



THE SEDONA CONFERENCE JOURNAL®

Volume 22 ❖ 2021

ARTICLES

- The Sedona Conference Commentary on Rule 45 Subpoenas to Non-Parties, Second Edition.** The Sedona Conference
- The Sedona Conference Commentary on ESI Evidence & Admissibility, Second Edition.** The Sedona Conference
- The Sedona Conference Commentary on the Proper Identification of Asserted Trade Secrets in Misappropriation Cases** The Sedona Conference
- The Sedona Conference Commentary on the Enforceability in U.S. Courts of Orders and Judgments Entered Under GDPR** The Sedona Conference
- The Sedona Conference Commentary on a Reasonable Security Test** The Sedona Conference
- The Sedona Conference Commentary on Ephemeral Messaging** The Sedona Conference
- The Sedona Conference Commentary on Quantifying Violations under U.S. Privacy Laws** The Sedona Conference
- From the KKK to George Floyd: Three Judges Explore Qualified Immunity** Hon. Cathy Bissoon, Hon. Benita Y. Pearson & Hon. David A. Sanders
- Implicit Bias: The Science, Influence, and Impact on Justice** Hon. Bernice B. Donald



ANTITRUST LAW, COMPLEX LITIGATION, INTELLECTUAL
PROPERTY RIGHTS, AND DATA SECURITY AND PRIVACY LAW

THE SEDONA CONFERENCE JOURNAL[®]

VOLUME 22



2021



The Sedona Conference Journal® (ISSN 1530-4981) is published on an annual or semi-annual basis, containing selections from the preceding year's conferences and Working Group efforts. A PDF copy of The Journal is available on a complimentary basis and can be downloaded from the Publications page on The Sedona Conference website: www.thesedonaconference.org. Check our website for further information about our conferences, Working Groups, and publications.

Comments (strongly encouraged) and requests to reproduce all or portions of this issue should be directed to:

The Sedona Conference,
301 East Bethany Home Road, Suite C-297, Phoenix, AZ 85012 or
info@sedonaconference.org or call 1-602-258-4910.

The Sedona Conference Journal® cover designed by MargoBDesignLLC at
www.margobdesign.com.

Cite items in this volume to "22 Sedona Conf. J. ____ (2021)."

Copyright 2021, The Sedona Conference.
All Rights Reserved.

PUBLISHER'S NOTE

Welcome to Volume 22 of The Sedona Conference Journal (ISSN 1530-4981), published by The Sedona Conference, a nonprofit 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, intellectual property rights, and data security and privacy law. The mission of The Sedona Conference is to move the law forward in a reasoned and just way through the creation and publication of nonpartisan consensus commentaries and advanced legal education for the bench and bar.

The various Working Groups in The Sedona Conference Working Group Series (WGS) pursue in-depth study of tipping-point issues, with the goal of producing high-quality, nonpartisan consensus commentaries that provide guidance of immediate and practical benefit to the bench and bar. The Sedona Conference conducts a “regular season” of limited-attendance conferences that are mini-sabbaticals for the nation’s leading jurists, lawyers, academics, and experts to examine cutting-edge issues of law and policy. The Sedona Conference also conducts continuing legal education programs under The Sedona Conference Institute (TSCI) banner, an annual International Programme on Cross-Border Data Transfers and Data Protection Laws, and webinars on a variety of topics.

Volume 22 of the Journal contains three nonpartisan consensus commentaries from The Sedona Conference Working Group on Data Security and Privacy Liability (WG11), two nonpartisan consensus commentaries from the Working Group on Electronic Document Retention and Production (WG1), one nonpartisan consensus commentary from the Working Group on International Electronic Information Management, Discovery, and Disclosure (WG6), and one nonpartisan consensus commentary from the Working Group on Trade Secrets (WG12). Additionally, this issue contains timely new articles on Qualified Immunity and Implicit Bias, both authored by members of the federal judiciary. I hope you find the commentaries and articles to be thought-provoking pieces that stimulate further dialogue and ultimately serve to move the law forward.

For more information about The Sedona Conference and its activities, please visit our website at www.thesedonaconference.org.

Craig Weinlein
Executive Director
The Sedona Conference
August 2021

The Sedona Conference gratefully acknowledges the contributions of its Working Group Series annual sponsors, event sponsors, members, and participants whose volunteer efforts and financial support make participation in The Sedona Conference and its activities a thought-provoking and inspiring experience.

JOURNAL EDITORIAL BOARD

Editor-in-Chief

Craig Weinlein

Managing Editor

David Lumia

Review Staff

Jim W. Ko

Michael Pomarico

Kenneth J. Withers

THE SEDONA CONFERENCE ADVISORY BOARD

Kevin F. Brady, Esq., Volkswagen Group of America, Herndon, VA

Prof. Stephen Calkins, Esq., Wayne State University Law School, Detroit, MI

Michael V. Ciresi, Esq., Ciresi Conlin LLP, Minneapolis, MN

The Hon. John Facciola (ret.), Washington, DC

Prof. Steven S. Gensler, University of Oklahoma College of Law, Norman, OK

Prof. George A. Hay, Cornell Law School, Ithaca, NY

Ronald J. Hedges, Esq., Dentons US LLP, New York, NY

Allan Kanner, Esq., Kanner & Whiteley, L.L.C., New Orleans, LA

The Hon. Paul R. Michel (ret.), Alexandria, VA

Dianne M. Nast, Esq., NastLaw LLC, Philadelphia, PA

The Hon. Nan R. Nolan (ret.), Redgrave LLP, Chicago, IL

The Hon. Andrew J. Peck (ret.), DLA Piper, New York, NY

Jonathan M. Redgrave, Esq., Redgrave LLP, Washington, DC

The Hon. James M. Rosenbaum (ret.), JAMS, Minneapolis, MN

Prof. Stephen A. Saltzburg, George Washington University Law School, Washington, DC

The Hon. Shira A. Scheindlin (ret.), Stroock & Stroock & Lavan LLP, New York, NY

Daniel R. Shulman, Esq., Shulman & Buske PLLC, Minneapolis, MN

Dennis R. Suplee, Esq., Schnader Harrison Segal & Lewis LLP, Philadelphia, PA

Prof. Jay Tidmarsh, University of Notre Dame Law School, Notre Dame, IN

The Hon. Tom I. Vanaskie (ret.), Stevens & Lee, Philadelphia, PA

The Hon. Patrick J. Walsh (ret.), Signature Resolution, Los Angeles, CA

The Hon. Ira B. Warshawsky (ret.), Meyer, Suozzi, English & Klein, P.C., Garden City, NY

JUDICIAL ADVISORY BOARD

The Hon. Jerome B. Abrams, Minnesota District Court Judge, First Judicial District

The Hon. Michael M. Baylson, Senior U.S. District Judge, Eastern District of Pennsylvania

The Hon. Laurel Beeler, U.S. Magistrate Judge, Northern District of California

The Hon. Cathy A. Bencivengo, U.S. District Judge, Southern District of California

The Hon. Cathy Bissoon, U.S. District Judge, Western District of Pennsylvania

The Hon. Hildy Bowbeer, U.S. Magistrate Judge, District of Minnesota

The Hon. Ron Clark, Senior U.S. District Judge, Eastern District of Texas

The Hon. Joy Flowers Conti, Chief U.S. District Judge, Western District of Pennsylvania

The Hon. Mitchell D. Dembin, U.S. Magistrate Judge, Southern District of California

The Hon. James L. Gale, Senior Judge, North Carolina Business Court

The Hon. George C. Hanks, U.S. District Judge, Southern District of Texas

The Hon. Susan Illston, U.S. District Court, Northern District of California, San Francisco, CA

The Hon. Kent A. Jordan, U.S. Appellate Judge, Third Circuit

The Hon. Barbara M.G. Lynn, Chief U.S. District Judge, Northern District of Texas

The Hon. Kristen L. Mix, U.S. Magistrate Judge, District of Colorado

The Hon. Kathleen McDonald O'Malley, U.S. Appellate Judge, Federal Circuit

The Hon. Katharine H. Parker, U.S. Magistrate Judge, Southern District of New York

The Hon. Anthony E. Porcelli, U.S. Magistrate Judge, Middle District of Florida

The Hon. Xavier Rodriguez, U.S. District Judge, Western District of Texas

The Hon. Lee H. Rosenthal, Chief U.S. District Judge, Southern District of Texas

The Hon. Elizabeth A. Stafford, U.S. Magistrate Judge, Eastern District of Michigan

The Hon. Gail J. Standish, U.S. Magistrate Judge, Central District of California

The Hon. Leda Dunn Wettre, U.S. Magistrate Judge, District of New Jersey

TABLE OF CONTENTS

Publisher’s Note	i
Journal Editorial Board	ii
The Sedona Conference Advisory Board	iii
The Sedona Conference Judicial Advisory Board	iv
The Sedona Conference Commentary on Rule 45 Subpoenas to Non-Parties, Second Edition	
The Sedona Conference	1
The Sedona Conference Commentary on ESI Evidence & Admissibility, Second Edition	
The Sedona Conference	83
The Sedona Conference Commentary on the Proper Identification of Asserted Trade Secrets in Misappropriation Cases	
The Sedona Conference	223
The Sedona Conference Commentary on the Enforceability in U.S. Courts of Orders and Judgments Entered under GDPR	
The Sedona Conference	277
The Sedona Conference Commentary on a Reasonable Security Test	
The Sedona Conference	345
The Sedona Conference Commentary on Ephemeral Messaging	
The Sedona Conference	435
The Sedona Conference Commentary on Quantifying Violations under U.S. Privacy Laws	
The Sedona Conference	489
From the KKK to George Floyd: Three Judges Explore Qualified Immunity	
Hon. Cathy Bissoon, Hon. Benita Y. Pearson & Hon. David A. Sanders	533
Implicit Bias: The Science, Influence, and Impact on Justice	
Hon. Bernice B. Donald	583

THIS PAGE INTENTIONALLY LEFT BLANK

THE SEDONA CONFERENCE COMMENTARY ON
RULE 45 SUBPOENAS TO NON-PARTIES, SECOND EDITION

*A Project of The Sedona Conference Working Group on
Electronic Document Retention and Production (WG1)*

Author:

The Sedona Conference

Editors-in-Chief:

Tessa K. Jacob

The Hon. Andrew J. Peck (ret.)

Drafting Team Leaders:

Tessa K. Jacob

Eric Schwarz

Drafting Team:

John Baker

Bryan Bleichner

Andrew Diana

Arthur Fahlbusch

Nathaniel C. Giddings

Ross Gotler

Beth Leland

Glenn Melcher

Sandra Metallo-Barragan

Joshua Schonauer

Ronnie Spiegel

Deric Yoakley

Steering Committee Liaisons:

Andrea L. D'Ambra

The Hon. Andrew J. Peck (ret.)

Peter Pepiton

Judicial Participant:

The Hon. David Horan

Staff editors:

David Lumia

Susan McClain

Copyright 2020, The Sedona Conference.
All Rights Reserved.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

We thank all of our Working Group Series Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors, just click on the “Sponsors” navigation bar on the homepage of our website.

This publication may be cited as follows:

The Sedona Conference, *Commentary on Rule 45 Subpoenas to Non-Parties, Second Edition*, 22 SEDONA CONF. J. 1 (2021).

PREFACE

Welcome to the final, October 2020, version of *The Sedona Conference Commentary on Rule 45 Subpoenas to Non-Parties, Second Edition*, a project of The Sedona Conference Working Group on Electronic Document Retention and Production (WG1).

This is one of a series of Working Group commentaries published by The Sedona Conference, a 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, intellectual property rights, and data security and privacy law. The mission of The Sedona Conference is to move the law forward in a reasoned and just way.

In 2008, The Sedona Conference published its first edition of the *Commentary on Non-Party Production & Rule 45 Subpoenas*. This Second Edition, now titled the *Commentary on Rule 45 Subpoenas to Non-Parties*, accounts for the 2013 amendments to Rule 45, the December 2015 amendments to other discovery rules, publication of *The Sedona Principles, Third Edition*, and significant case law development since 2008. Updating the 2008 Commentary was a topic of dialogue at the WG1 2018 Midyear Meeting, and the subsequently formed drafting team presented redlined drafts to the WG1 membership and entertained feedback at the 2018 Annual Meeting.

This Second Edition was first published for public comment in January 2020 and underwent significant revision after the public comment period, resulting in publication of a second public comment version in August 2020. Where appropriate, the comments received during the public comment periods have now been incorporated into this final version of the *Commentary*.

The Sedona Conference acknowledges the efforts of Drafting Team Leaders Tessa K. Jacob and Eric Schwarz, who were invaluable in driving this project forward. We thank Drafting

Team members John Baker, Bryan Bleichner, Anthony Diana, Arthur Fahlbusch, Nathaniel Giddings, Ross Gotler, Beth Leland, Glenn Melcher, Sandra Metallo-Barragan, Joshua Schonauer, Ronnie Spiegel, Deric Yoakley, and The Honorable David Horan for their commitments in time and attention to this project, and we also thank Katelyn Flynn and Kelly Warner for their contributions. Finally, we thank Andrea D’Ambra, Peter Pepiton, and The Honorable Andrew J. Peck (ret.) for their guidance and input as the WG1 Steering Committee Liaisons to the drafting team. Ms. Jacob and Judge Peck served as the Editors-in-Chief guiding this Commentary to publication.

Membership in The Sedona Conference Working Group Series is open to all. The Series includes WG1 and several other Working Groups in the areas of international electronic information management, discovery, and disclosure; patent damages and patent litigation best practices; data security and privacy liability; trade secrets; and other “tipping point” issues in the law. The Sedona Conference hopes and anticipates that the output of its Working Groups will evolve into authoritative statements of law, both as it is and as it should be. Information on membership and a description of current Working Group activities is available at <https://thesedonaconference.org/wgs>.

Craig Weinlein
Executive Director
The Sedona Conference
October 2020

TABLE OF CONTENTS

I.	INTRODUCTION.....	7
II.	RULE CHANGES AND THEIR IMPACT ON NON-PARTY DISCOVERY.....	9
	A. Introduction.....	9
	B. Rule Changes	9
	1. 2013 Amendments to Rule 45.....	9
	2. 2015 Amendments to the Federal Rules of Civil Procedure.....	12
	3. Federal Rule of Evidence 502	16
III.	THE POSSESSION, CUSTODY, AND CONTROL FRAMEWORK AND ITS IMPACT ON RULE 45 OBLIGATIONS	17
	A. Introduction.....	17
	B. Rule 45 Rights and Obligations Where a Party Has Possession, Custody, or Control	19
	1. Requesting Discovery When a Party to the Litigation Has Control Over ESI or Documents in a Non-Party’s Custody or Possession.....	21
	C. Rule 45 Rights and Obligations Where a Party Does Not Have Possession, Custody, or Control	28
	1. Subpoenaing a Non-Party with Sole Possession, Custody, and Control	28
	2. Subpoenaing a Non-Party That Has a Relationship to a Party	30
IV.	PRESERVATION	33
	A. Prior to Receipt of a Subpoena	33
	B. After Receipt of a Subpoena	38
	C. Remedies for Spoliation.....	40

V.	RULE 45(D) COSTS, SANCTIONS, AND MOTION PRACTICE	42
A.	Rule 45(d)(1)—Avoiding Undue Burden or Expense and Sanctions	42
B.	Rule 45(d)(2)(B)(ii)—Mandatory Cost Shifting When the Non-Party is Ordered to Produce with Significant Expense Over Objection	45
1.	Prerequisites for Seeking Cost Shifting.....	47
2.	Significant Expense and Cost Shifting	56
3.	When and How to Request Cost Shifting.....	56
C.	Rule 45(d)(3)(A)—Quashing or Modifying a Subpoena	71
D.	Rule 45(e)(1)(D)—Non-Party ESI That Is Not Reasonably Accessible Due to Undue Burden or Cost.....	77
VI.	THE SEDONA CONFERENCE RULE 45 PRACTICE POINTERS ...	78

I. INTRODUCTION

Developments since the 2008 edition of *The Sedona Conference Commentary on Non-Party Production and Rule 45 Subpoenas*¹ have led to significant revisions and additions now included in this *Second Edition*. Federal Rule of Civil Procedure 45 (Rule 45)² was revised substantially in 2013. The 2015 amendments to the Federal Rules of Civil Procedure also impact Rule 45. The rise of cloud computing has put appreciable amounts of party data into the hands of non-parties.

Like the 2008 edition, the scope of the current *Commentary* is limited to the use of Rule 45 subpoenas to obtain discovery from a non-party custodian of documents or electronically stored information (ESI). The *Commentary* does not address the use of Rule 45 subpoenas to (1) compel any person to appear and give testimony at a trial, hearing, or deposition, or (2) compel any person to appear and bring documents or ESI to a trial, hearing, or deposition.

Section II of this *Commentary* briefly explains the major revisions to Rule 45 made by the 2013 Rules amendments, as well as the effect of the 2015 Rules amendments.

Section III of this *Commentary* proposes an approach for analyzing whether a party has possession, custody, or control of information that may make a non-party subpoena inappropriate. In other words, if the non-party has possession or custody of ESI but a party retains control, the *Commentary* recommends that the information should be obtained from the party under Rule 34, not from the non-party under Rule 45.

Section IV of this *Commentary* deals with preservation. A letter or similar request for the preservation of evidence

1. See generally 9 SEDONA CONF. J. 197 (2008).

2. As used in this *Commentary*, the term “Rule(s)” refers to the Federal Rule(s) of Civil Procedure, unless otherwise specified.

generally does not create a non-party preservation obligation. In most cases, receipt of a properly served subpoena only obligates a non-party to take reasonable steps to produce the requested materials and does not obligate the non-party to initiate a formal legal hold process. Rather, the non-party's obligation is to ensure that the requested information is not destroyed during the compliance period. However, once a non-party has complied with a subpoena by producing responsive documents and ESI, the non-party has no duty to preserve them. Because Rule 45(d)(2)(B)(ii) places no time limit for a party to move to compel production of information sought by a subpoena, this *Commentary* encourages a non-party to provide a specific date after which it will no longer retain the documents or ESI that it objects to producing. Such a step thereby places the requesting party on notice of the date by which the requesting party needs to determine the completeness of the production and make a motion to compel.

The longest section of this *Commentary*, Section V, deals with the related concepts of sanctions under Rule 45(d)(1), cost shifting under Rule 45(d)(2)(B)(ii), and quashing or limiting the scope of a subpoena under Rule 45(d)(3). Section V analyzes the now extensive case law under each of these approaches. This *Commentary* focuses on case law discussing the importance of properly objecting within the required 14 days in order to benefit from the rule's mandatory cost-shifting component. There is also detailed discussion on what constitutes undue burden under the various subsections of Rule 45, including when and how courts have relieved non-parties of their obligations under a subpoena due to undue burden under these subsections.

Finally, Section VI sets forth recommended "Practice Pointers" for both parties and non-parties dealing with a Rule 45 subpoena.

II. RULE CHANGES AND THEIR IMPACT ON NON-PARTY DISCOVERY

A. Introduction

Since publication of the 2008 edition of this *Commentary*, there have been several revisions to the Federal Rules of Civil Procedure and Federal Rules of Evidence. In 2013, Rule 45 was substantially revised with the intent of decreasing disputes and streamlining the practice of non-party discovery. The 2015 amendments to the Federal Rules of Civil Procedure did not alter Rule 45, but revisions to other rules—such as the scope of discovery under Rule 26—relate to non-party subpoena practice. The enactment of Federal Rule of Evidence 502, governing the production of privileged or work-product protected material, also has implications for the issuance of and response to subpoenas.

B. Rule Changes

1. 2013 Amendments to Rule 45

The Advisory Committee Notes to the 2013 amendments to Rule 45 state that the revisions were intended to “clarify and simplify the rule.” The following revisions were implemented to accomplish this intent.

(a) Subpoenas Issued from Court in Which Action Is Pending

Pursuant to revised Rule 45(a)(2), all subpoenas, whether for documents, deposition, or trial testimony, must be issued from the court in which the action is pending. This revision addressed confusion resulting from the prior Rule’s differentiation based on the purpose of the subpoena. In addition to the clerk of the issuing court, an attorney who is authorized to practice in the

issuing court, including an attorney admitted *pro hac vice*, may issue the subpoena (Rule 45(a)(3)).

(b) Nationwide Service of Subpoenas

Pursuant to revised Rule 45(b)(2), subpoenas may be served at any place in the United States. Previously, the location for service was limited to the district or state where the issuing court was located, with some exceptions based on distance outside the district. These distinctions were eliminated in favor of simplicity. This revision is consistent with the requirement that all subpoenas be issued from the court in which the action is pending.

(c) Service on Opposing Parties Prior to Service on Recipients

Pursuant to Rule 45(a)(4), for subpoenas seeking the production of documents or ESI, a notice and a copy of the subpoena must be served on each party prior to serving the subpoena on the person or entity subject to the subpoena (subpoena recipient or recipient). The 2013 Advisory Committee Notes state that this requirement was emphasized and enhanced to require providing other parties with a copy of the subpoena due to “frequent fail[ure] to give the required notice to the other parties.” The amendment is “intended to achieve the original purpose of enabling the other parties to object or to serve a subpoena for additional materials.”³ The amended rule does not specify how far in advance the notice and copy of the subpoena must be served on the opposing party.

3. FED. R. CIV. P. 45 advisory committee’s note to 2013 amendment.

(d) Geographic Limitations to Place of Compliance

Revised Rule 45(c) “simplifies” where compliance with a subpoena can be required and abandons the prior version’s focus on location of service. Under revised Rule 45(c), the court may require documents or ESI to be produced within 100 miles of where the recipient resides, is employed, or regularly conducts business in person. For testimony at trial, hearing, or deposition, the court may require the recipient to appear in person within 100 miles of where the recipient resides, is employed, or regularly conducts business in person. However, this 100-mile travel limit is expanded to anywhere within the state where the recipient resides, is employed, or regularly conducts business in person, if the recipient: (1) is a party or party officer, or (2) is commanded to attend a trial unless doing so would cause the recipient to incur “substantial expense.”

(e) Jurisdiction to Enforce or Quash Subpoenas

Revised Rule 45(d) requires that the court for the district where compliance of the subpoena is required (not necessarily the court from which the subpoena was issued) must ensure that the party issuing the subpoena took reasonable steps to avoid imposing undue burden or expense on the person subject to the subpoena.⁴ This same court is given the responsibility to enforce, quash, or modify the subpoena.

A party’s motion to enforce a subpoena or a non-party motion to quash or modify the subpoena should be filed as a

4. *See, e.g.,* *Merch. Consulting Grp., Inc. v. Beckpat, LLC*, No. 17-11405, 2018 WL 4510269, at *3 (D. Mass. July 11, 2018) (discussing that the court or district “where compliance is required” is determined by the location or “place” for compliance identified on the subpoena); *Raap v. Brier & Thorn, Inc.*, No. 17-MC-3001, 2017 WL 2462823, at *3 (C.D. Ill. July 7, 2017) (“[T]he better approach is to tie the place of compliance to the location of the subpoenaed person or entity.”).

separate action in the court for the district where compliance is required. This action will be assigned a “miscellaneous” case number.

(f) Transfer of Disputes Related to Subpoenas

Rule 45(f) is a new section that allows the court where compliance is required to transfer a subpoena dispute to the court where the action is pending (a) if the subpoena recipient consents, or (b) upon a finding of “exceptional circumstances,” for which the party seeking the transfer has the burden of showing such circumstances exist.⁵

2. 2015 Amendments to the Federal Rules of Civil Procedure

Although the 2015 amendments to the Federal Rules of Civil Procedure did not make changes to Rule 45 itself, several revised provisions impact subpoenas.

(a) Rule 1

Rule 1 was amended to include that the rules are to be “employed by the court and the parties” to “secure the just, speedy, and inexpensive determination of every action and proceeding.” The 2015 Advisory Committee Notes emphasize that the parties share the court’s obligation to “construe and administer” the rules, thus creating an expectation for all parties

5. *See* Warkins v. Piercy, No. 4:16-MC-00324, 2016 WL 3683010, at *2 (E.D. Mo. July 12, 2016) (quoting FED. R. CIV. P. 45 advisory committee’s note to 2013 amendment) (noting that the court’s primary concern should be to avoid burdens on local non-parties, and it should not be assumed that the issuing court is in a superior position to resolve subpoena-related motions); *see also* Lima LS PLC v. Nassau Reinsurance Grp. Holdings, L.P., 160 F. Supp. 3d 574, 579–80 (S.D.N.Y. 2015) (noting that Rule 45 permits transfer but does not require it).

to apply the rules in a manner that makes litigation efficient. While the amendment refers to “the parties,” the spirit of Rule 1 should be applied to a party’s dealings with non-party subpoena recipients.⁶ Cooperation is helpful in reducing disputes and saving costs.

(b) Rule 26

Rule 26 was modified in two material ways that impact the scope of discovery. These revisions are applicable to all discovery, including the issuance of and response to subpoenas.

First, the drafters emphasized the concept of proportionality, which was already part of Rule 26(b)(2)(c), by moving the factors into the section regarding scope of discovery. In revised Rule 26(b)(1), the factors to be considered when assessing proportionality are the same as those which have been included in the rule for years, with the addition of the explicit instruction that courts consider “the parties’ relative access to relevant information.” The elevation of the proportionality factors reinforces obligations of the parties to consider these factors when propounding and responding to discovery requests. The few courts that have considered this issue in the context of Rule 45 subpoenas have found that Rule 26’s proportionality factors are applicable to Rule 45 subpoenas.⁷

Second, Rule 26 was amended to exclude the phrase “reasonably calculated to lead to the discovery of admissible

6. See also The Sedona Conference, *The Sedona Conference Cooperation Proclamation*, 10 SEDONA CONF. J. 331 (2009).

7. See generally *MetroPCS v. Thomas*, 327 F.R.D. 600 (N.D. Tex. 2018) (applying Rule 26’s proportionality factors to assess whether a non-party subpoena issued pursuant to Rule 45 sought information within the proper scope of discovery); *Walker v. H & M Henner & Mauritz, LP*, 16 Civ. 3818, 2016 WL 4742334 (S.D.N.Y. Sept. 12, 2016) (quashing subpoenas to non-party witnesses for failure to meet proportionality requirements).

evidence,” requiring only that “materials need not be admissible in order to be discoverable,” because the “reasonably calculated” language was misused to extend the scope of discovery. The 2015 Advisory Committee Notes to amended Rule 26 stress the role of the parties and, if necessary, the court in ensuring the principles of effective and cooperative case management are followed.

(c) Rule 34

The 2015 amendments to Rule 34 focus on requiring particularity in the objections asserted by a party responding to a Rule 34 request for production. The Rule requires the responding party to “state with specificity the grounds for objecting to a request” and clearly note if documents or ESI are being withheld on the basis of an objection. The 2015 Advisory Committee Notes state the amendments are “aimed at the potential to impose unreasonable burdens by objections to requests to produce,” and the cases interpreting this change underscore the focus on whether the parties are complying with the spirit of the rule.⁸

There is uncertainty regarding whether these revisions to Rule 34 are applicable to subpoenas issued pursuant to Rule 45. Rule 45 contemplates the recipient objecting to the request, but the drafters did not include additional language regarding requirements for those objections. Most courts have held that Rule 34 requirements regarding reasonable particularity and objections apply with equal force to a non-party’s responses to Rule 45 subpoenas.⁹ It may be beneficial for requesting parties

8. See The Sedona Conference, *Federal Rule of Civil Procedure 34(b)(2) Primer: Practice Pointers for Responding to Discovery Requests*, 19 SEDONA CONF. J. 447 (2018).

9. See, e.g., *Nasufi v. King Cable, Inc.*, No. 3:15-cv-3273, 2017 WL 3334110, at *4 (N.D. Tex. Aug. 4, 2017) (“[N]on-party’s Rule 45(d)(2)(B) objections to

to state their requests with specificity and for the responding non-party to object with specificity.

(d) Rule 37(e)

The 2015 amendments significantly changed Rule 37(e) regarding when a party may be sanctioned for the failure to preserve information. The 2015 Advisory Committee Notes explain that Rule 37 now “authorizes and specifies measures a court may employ if information that should have been preserved is lost, and specifies the findings necessary to justify these measures.” Rule 37(e) on its face does not apply to non-party subpoena recipients because all of its provisions specifically refer to a “party.”

Instead, Rule 45(g) provides that “[t]he court for the district where compliance is required—and also, after a motion is transferred, the issuing court—may hold in contempt a person who, having been served, fails without adequate excuse to obey the subpoena or an order related to it.”¹⁰

discovery requests in a subpoena are subject to the same prohibition on general or boiler-plate or unsupported objections and requirements that the objections must be made with specificity and that the responding party must explain and support its objections.”); *Am. Fed’n of Musicians of the U.S. & Can. v. Skodam Films, LLC*, 313 F.R.D. 39, 46 (N.D. Tex. 2015) (“[J]ust as Rule 34(b)(1)’s reasonable particularity requirement should apply with no less force to a subpoena’s document requests to a non-party, a non-party’s Rule 45(d)(2)(B) objections to those requests should be subject to the same requirements facing a party objecting to discovery under Rule 34.”).

10. *See, e.g., Jalayer v. Stigliano*, No. CV102285, 2016 WL 5477600, at *3 (E.D.N.Y. Sept. 29, 2016) (discussing how moving for sanctions against non-party subpoena recipient under Rule 37 was improper and instead finding that Rule 45(g) was the appropriate avenue for seeking such relief).

3. Federal Rule of Evidence 502

Federal Rule of Evidence (Fed. R. Evid.) 502, effective September 2008, was enacted subsequent to the April 2008 edition of this *Commentary*. Fed. R. Evid. 502 addresses the production of documents or ESI protected by the attorney-client privilege or work-product doctrine.

Fed. R. Evid. 502 was enacted for two primary purposes. First, it was intended to address inconsistent case law regarding the effect of a production of protected material. The Advisory Committee Notes state the rule was intended to “provide a predictable, uniform set of standards under which parties can determine the consequences of a disclosure of a communication or information covered by the attorney-client privilege or work-product protection.” Second, the rule was intended to respond to the “widespread complaint that litigation costs necessary to protect against waiver of attorney-client privilege or work product have become prohibitive.”

Many litigants now request Fed. R. Evid. 502(d) non-waiver orders to ensure that the production of privileged materials will not result in the waiver of attorney-client privilege or work product. Subpoena recipients should request a copy of the Fed. R. Evid. 502(d) order entered in the case and ensure that the language of the order applies to non-party productions as well. If such an order has not been entered or the non-party feels that the order does not provide adequate protection, the non-party should seek to have one entered.¹¹

11. See generally The Sedona Conference, *Commentary of the Protection of Privileged ESI*, 17 SEDONA CONF. J. 95 (2016); see also The Sedona Conference, *The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 SEDONA CONF. J. 1, 147 (2018) [hereinafter *The Sedona Principles, Third Edition*].

III. THE POSSESSION, CUSTODY, AND CONTROL FRAMEWORK AND ITS IMPACT ON RULE 45 OBLIGATIONS

A. Introduction

While Rule 45 does not require the parties to confer with each other or with a non-party prior to serving a subpoena, this *Commentary* recommends certain practices to reduce the burden and expense of litigation to parties and non-parties.

Whether a non-party subpoena or a Rule 34 document request should be used is dependent on the concept of possession, custody, or control. Particularly with ESI, a party may have legal control even when the ESI itself is in the possession or custody of a non-party. This *Commentary* will refer to such a non-party as a “custodial non-party.” Where a party has possession, custody, or control of documents or ESI, that party should have the burden of producing its own information (via Rule 34) rather than requiring the requesting party to seek it through a subpoena to the custodial non-party. Similarly, a requesting party should seek documents or ESI from the party that controls the information through a Rule 34 request before issuing a Rule 45 subpoena to the custodial non-party. Accordingly, this *Commentary* recommends that before any requests for documents or ESI from a custodial non-party are issued or enforced, the threshold analysis should be whether a party to the litigation has possession, custody, or control of the documents or ESI.

Importantly and generally, where a party to the litigation has control over the requested documents and ESI that are in the possession or custody of a non-party, document requests to a party, rather than subpoenas to a custodial non-party, are the appropriate method to obtain discovery of those documents and ESI. In such a situation, the party’s interests in as well as rights and obligations regarding the requested documents and ESI—

including retention, production costs, and management of the risks associated with privilege, privacy, data security, and confidentiality—are determinative of the obligations imposed upon and protections afforded to that party and non-party. This *Commentary* discusses this concept further in section III.B.

It might be beneficial for the parties to discuss, at the Rule 26(f) conference or other appropriate point, whether a party believes a non-party has documents or ESI responsive to the requesting party's discovery requests, and whether the responding party asserts that it does or does not have possession, custody, or control of such documents or ESI. The parties should work to reach stipulations concerning authenticity and admissibility to avoid the need to subpoena a non-party custodian to prove up documents or ESI. If after receipt of a notice of a subpoena to a custodial non-party the party is willing to produce all or some of the requested information, it should notify the non-party and the party issuing the subpoena.

In addition, where a non-party is related to a party to the litigation and the party to the litigation does not have possession, custody, or control of the requested information, the non-party may share such interests in as well as rights and obligations regarding the requested documents and ESI. This *Commentary* discusses this further in section III.C.2 below. Where a non-party has sole possession, custody, or control and does not share any interest in the litigation, the non-party is afforded the full protections of Rule 45, including cost-shifting mechanisms or quashing or modifying of the subpoena. This *Commentary* discusses this further in section III.C.1 below.

B. Rule 45 Rights and Obligations Where a Party Has Possession, Custody, or Control

Prior to the issuance or enforcement of a non-party subpoena, there initially should be an analysis of whether any party to the litigation has possession, custody, or control of the requested documents or ESI. When defining “possession, custody, or control,”¹² this *Commentary* follows The Sedona Conference’s prior publication on this issue, the *Commentary on*

12. Courts have applied inconsistent and varying standards to construe the meaning of “possession, custody, or control.” See The Sedona Conference, *Commentary on Rule 34 and Rule 45 “Possession, Custody, or Control,”* 17 SEDONA CONF. J. 467, 483–98 (2016) (surveying the applicable case law and observing that it provides three broad interpretations of “control”: the Legal Right Standard, the Legal Right Plus Notification Standard, and the Practical Ability Standard). These standards largely fall within three broad categories. The Legal Right Standard evaluates a party’s possession, custody, or control based on its legal right to obtain the documents or ESI in question. The Legal Right Plus Notification Standard builds on the Legal Right Standard by further obligating a party that does not have a legal right to the documents or ESI to notify the requesting party of the identities of non-parties that have possession, custody, or control of the documents or ESI requested. The Practical Ability Standard evaluates possession, custody, or control based on whether the party has the practical ability to obtain the documents or ESI, regardless of whether it has the legal right to do so. The analysis proposed by this *Commentary* can be applied in jurisdictions using any of these standards. Practitioners should be familiar with the standard that applies in the relevant jurisdiction and should review the *Commentary on Rule 34 and Rule 45 “Possession, Custody, or Control.”* Note that in courts using the Legal Right Plus Notification Standard, the party needs to inform opposing parties of non-parties that possess requested documents or ESI. Additionally, where a party controls information or documents in the hands of a non-party, the party has an independent obligation to “preserve, collect, search, and produce the Documents and ESI in the hands” of a non-party, “even though the producing party does not actually possess or have actual custody of the Documents and ESI at issue.” See *Commentary on Rule 34 and Rule 45 “Possession, Custody, or Control,”* 17 SEDONA CONF. J. at 483.

Rule 34 and Rule 45 "Possession, Custody, or Control," advocating that the Legal Right Standard is the proper standard for defining control. Put another way, if a party has: "(1) actual possession of Documents and ESI; or (2) the legal right to obtain Documents and ESI," then a party should be deemed to have "possession, custody, or control" of those documents and ESI.¹³

If a party to the litigation has possession, custody, or control of the requested documents or ESI, generally it is unnecessary to issue a non-party subpoena. It is a well-established principle that the burdens of discovery should fall on the parties to the litigation instead of on any non-party.¹⁴ A party to the litigation may also be best positioned and have an incentive to properly address and manage issues concerning privilege, data privacy, and confidentiality, while a non-party often has no capability nor incentive to do so. The Rule 45 subpoena process is not intended to circumvent the requirements and protections of Rule 34.¹⁵ Similarly, where discoverable information is not in the possession, custody, or control of a party, Rule 34 does not prevent a party from obtaining discoverable information from a non-party through Rule 45. Should the requesting party have any doubt as to whether the responding party has possession, custody, or control, the requesting party should confer, in good faith, with the responding party to determine if it has possession, custody, or control and could produce the requested documents or ESI without a non-party subpoena.¹⁶ If the responding party states that it lacks possession, custody, or control of the requested documents or ESI, or portions thereof, or does not respond to such an inquiry within a reasonable time,

13. *See id.* at 529.

14. *See* discussion in Sections IV & V, *infra*.

15. *See* Section II.B.2.c., *supra*.

16. *See The Sedona Conference Cooperation Proclamation, supra* note 6; *see also* FED. R. CIV. P. 26(f).

the requesting party may seek to have a court determine whether the responding party has possession, custody, or control of the documents or ESI, or issue a non-party subpoena, or both.

1. Requesting Discovery When a Party to the Litigation Has Control Over ESI or Documents in a Non-Party's Custody or Possession

The evolving nature of data management solutions has resulted in many organizations outsourcing the storage of ESI to third-party service providers, sometimes with or without contractual language ensuring the organization a legal right to the information. Therefore, circumstances where non-parties hold documents or ESI over which a party to the litigation has control (i.e., a legal right to obtain the requested documents or ESI from the non-party) have become increasingly common with the rise of cloud computing services.¹⁷ This ESI storage revolution has profound implications on confidentiality, privacy, and privilege, which are fundamental considerations in the discovery process. In such situations, the obligations and burdens to produce those documents and ESI, as well as the associated rights and protections regarding those documents and ESI, should be borne or exercised by the party to the litigation—rather than the custodial non-party.

Therefore, the request for documents and ESI should be made pursuant to Rule 34, not Rule 45, and directed to the party to the litigation in the first instance. In other words, although the custodial non-party has actual possession or custody of the requested ESI and documents, the party that *controls* the ESI and documents should be responsible for responding to the request,

17. *Commentary on Rule 34 and Rule 45 "Possession, Custody, or Control," supra* note 13, at 521.

which may include coordination with the custodial non-party. This custodial non-party should be protected from responding to a subpoena where the primary rights and obligations associated with the requested documents and ESI lie with the party to the litigation. Arguments that the party does not have the ability to obtain the documents and ESI in the custody of a custodial non-party may be tenuous, as the party—by definition—has a legal right to obtain its own documents and ESI.¹⁸ Where there is a lack of cooperation by the custodial non-party, a Rule 45 subpoena may appropriately be considered. (The requesting party should not have to wait, beyond a reasonable time, until any dispute between the responding party and the custodial non-party is resolved.) This is distinguishable from the context described below, where a party to the litigation does not have control of the requested documents or ESI, and the only mechanism for obtaining the documents or ESI is from the non-party by way of a subpoena.

This framework is consistent with the Stored Communications Act,¹⁹ which imposes specific limitations regarding subpoenas that seek the contents of communications served on providers of remote computing services and electronic communication services. Under the Stored Communications Act, electronic communication service

18. See *Brown v. Tellermate Holdings Ltd.*, No. 2:11-cv-1122, 2014 WL 2987051, at *3–5 (S.D. Ohio July 1, 2014), *aff'd in part*, 2015 WL 4742686, at *2 (S.D. Ohio Aug. 11, 2015) (rejecting defendants' suggestion that it did not have possession, custody, or control of relevant information reflecting plaintiffs' sales activities held by defendants' enterprise cloud provider, Salesforce.com).

19. The Stored Communications Act was enacted in 1986 under Title II of the Electronic Communications Privacy Act., 18 U.S.C §§ 2701 *et seq.*

providers²⁰ are prohibited from divulging the contents of communications that are in “electronic storage”²¹ by that service, and remote computing service providers²² may not divulge communications that are carried or maintained on that service (absent the customer’s consent).²³

Illustration 1: A loan servicing systems provider that hosts ESI for a financial institution is subpoenaed by the defaulting loan party in a lawsuit with the financial institution. The requested ESI about the loan is not in the actual possession of the financial institution, but

20. Electronic communication service “means any service which provides to users thereof the ability to send or receive wire or electronic communications.” 18 U.S.C. § 2510(15) (2002).

21. Electronic storage means: “(A) any temporary, intermediate storage of a wire or electronic communication incidental to the electronic transmission thereof; and (B) any storage of such communication by an electronic communication service for purposes of backup protection of such communication.” *Id.* § 2510(17).

22. Remote computing service means “the provision to the public of computer storage of processing services by means of an electronic communications system.” *Id.* § 2711(2).

23. However, the Stored Communications Act does not provide third parties with absolute immunity to Rule 45 subpoenas. *See, e.g.,* UN4 Prods., Inc. v. Doe-173.68.177.95, No. 17-CV-3278, 2017 WL 2589328 (E.D.N.Y. June 14, 2017) (permitting Rule 45 subpoenas on internet service providers to discover the true name, postal address, and email address of each subscriber associated with identified internet protocol (IP) addresses, asserting “ISP subscribers have a minimal expectation of privacy in the sharing of copyrighted material” (quoting *Malibu Media, LLC v. John Does 1-11*, No. 12 Civ. 3810, 2013 WL 3732839 (S.D.N.Y. July 16, 2013))).

The Stored Communications Act does allow disclosure under some circumstances. For example, a provider may disclose a customer record or other subscriber information with the lawful consent of the customer or subscriber. 18 U.S.C. § 2702(c)(2) (2018).

pursuant to existing contractual rights, the financial institution has the legal right to obtain that ESI from the loan servicing systems provider. The financial institution, and not the loan servicing systems provider that hosts the ESI, should provide that ESI to the defaulting party via Rule 34.

Illustration 2: An online human resources (HR) and payroll solutions provider that hosts HR and payroll information for a manufacturing company is subpoenaed for HR information in an employment class action against the manufacturing company. The requested documents and ESI are not in the actual possession of the manufacturing company, but the manufacturing company has a contractual legal right to obtain those documents and ESI from the online HR and payroll solutions provider. The manufacturing company, and not the online HR and payroll solutions provider, should provide the documents and ESI to the plaintiffs via Rule 34.

In addition, non-party service providers or vendors in possession of information may have their own terms of service and use, contractual obligations, or policies that govern the level of protection they afford to the documents or ESI being held. Depending upon the nature of those provisions, the party whose data is retained by the non-party may be in a better position than the non-party to address those issues, as well as the confidentiality and privilege of the party's information. Where there is a dispute between the service provider (or vendor) and the responding party (such as, for example, over nonpayment of fees), that should not delay the requesting

party's ability to serve or enforce a non-party subpoena against the service provider (or vendor).

However, a Rule 45 subpoena directed to a custodial non-party may be necessary in certain limited circumstances even where a party has a legal right to obtain the documents. Instances may include where: (i) a party to litigation has engaged in misconduct and failed to produce or destroyed certain documents or ESI; (ii) the non-party is likely to have nonduplicative documents or ESI in its sole possession, custody, or control that do not fall within the legal right of the party; or (iii) there are extenuating circumstances that necessitate timely compliance with a document request or a need to authenticate documents (e.g., temporary restraining order, preliminary injunction). For instance, following discovery sanctions, obtaining requested documents, or ESI directly from the custodial non-party instead of the sanctioned party may be reasonable. On the other hand, the position of this *Commentary* is that a Rule 45 subpoena should not be used simply to validate that the party to the litigation properly produced all documents or ESI that are also in the possession of the custodial non-party.²⁴

24. While the case law is not uniform, many courts agree with this view, holding that parties should not use Rule 45 subpoenas as a means to evade the requirements of Rule 34, and instead should only resort to non-party subpoenas in the exceptional circumstances outlined above where the responding party has possession, custody, or control of the requested information. *See, e.g.,* *McCall v. State Farm Mut. Auto. Ins. Co.*, No. 216CV01058JADGWF, 2017 WL 3174914, at *6 (D. Nev. July 26, 2017) (“Although most courts hold that a subpoena duces tecum may be served on another party, it cannot be used to circumvent Rule 34 or the other discovery rules . . . The court also has an obligation to protect non-parties from being burdened with subpoenas for documents that can more easily and inexpensively be obtained from the opposing party.”); *Layman v. Junior Players Golf Acad.*, 314 F.R.D. 379, 385 (D.S.C. 2016) (“[R]esort to Rule 45 should not be allowed when it circumvents the requirements and protections of Rule 34 for the production of documents belonging to a party If

documents are available from a party, it has been thought preferable to have them obtained pursuant to Rule 34 rather than subpoenaing them from a nonparty witness.”) (internal citations and quotation marks omitted); *DeGeer v. Gillis*, 755 F.Supp.2d 909, 924 (N.D. Ill. 2010) (“Although the [third party] Cravath law firm has possession and custody of the database, [party] Huron does not contend that it lacks a legal right to obtain its documents from Cravath. . . . Huron documents in Cravath’s possession are subject to Huron’s control and thus, not exempt from Defendants’ subpoena to Huron.”); *Stokes v. Xerox Corp.*, No. 05-CV-71683-DT, 2006 WL 6686584, at *3 (E.D. Mich. Oct. 5, 2006) (“Accordingly, the Court finds that the majority view is that a party should not be permitted to circumvent the requirements and protections of Rule 34 by proceeding under Rule 45 for the production of documents belonging to a party.”); *Morrow v. Air Ride Techs., Inc.*, No. 05-113, 2006 U.S. 99651 WL 559288, at *2 (S.D. Ind. Mar. 6, 2006) (despite claims of no response from party to discovery requests, absent motion to compel production, court “is reluctant to allow the Plaintiffs to jettison the burden of production on a non-party”).

However, while some decisions contain language that appears contradictory to the cases cited immediately above, a close reading of these cases shows they generally align with the position taken in this *Commentary* that absent some evidence of misconduct in party discovery, subpoenas should not be used to get documents from a non-party that are more easily obtained from a party in the case. *See, e.g.*, *Trustees of Boston Univ. v. Everlight Electronics Co.*, No. 12-cv-11935-PBS, 2014 WL 12792497 at *3 (D. Mass. Sept. 10, 2014) (quashing the subpoena as overly burdensome on a non-party when the information could be obtained from the party, while noting “[T]here is no absolute rule prohibiting a party from seeking to obtain the same documents from a non-party as can be obtained from a party. ‘In many cases, tell-tale differences may appear between [the document collections]; and in many cases when a party obtains what should be the same set of documents from two different sources a critical fact in the litigation turns out to be that one set omitted a document that was in the other set.’”) (quoting *Coffeyville Resources Refining & Mktg., LLC v. Liberty Surplus Ins. Corp.*, No. 4:08MC00017, 2008 WL 4853620 at *2 (E.D. Ark. Nov. 6, 2008)); *In re Mushroom Direct Purchaser Antitrust Litigation*, No. 06-620, 2012 WL 298480, at *4 (E.D. Pa. Jan. 31, 2012) (“This is not a case where plaintiffs have made no attempt to obtain the requested information from other sources. . . . A plaintiff seeking to discover information from a third-party is not required to compel defendants to produce potentially overlapping information before

Illustration 1: A law firm that represented the defendant in the negotiation of a contract is subpoenaed by the plaintiff in a breach of contract case to produce the relevant non-privileged transaction documents, including drafts. Based on the engagement agreement between the defendant and the law firm, the documents are under the defendant's control. The plaintiff should obtain the requested documents directly from the defendant via Rule 34.

Illustration 2: A law firm that represented the defendant in the negotiation of a contract is subpoenaed by the plaintiff to produce relevant internal and external non-privileged communications regarding the contract negotiations. Based on the engagement agreement, such documents and ESI are not under the party's "possession, custody, or control." In this example, the plaintiff should obtain the

seeking any third-party discovery. This is particularly true where, as here, the defendant with potentially overlapping information has already produced documents in response to the overlapping discovery requests. . . . Further, it is likely that [the non-party] possesses relevant documents that [Defendant] does not."); *Med Tech., Inc. v. Breg, Inc.*, No. 10-MC-00100, 2010 WL 3734719, at *4 (E.D. Pa. Sept. 21, 2010) (rejecting non-party argument that subpoena was cumulative of party discovery where the party had already responded to discovery without producing the documents sought from the non-party); *Davis v. City of Springfield*, No. 04-3168, 2009 WL 910204, at *4 (C.D. Ill. Apr. 1, 2009), *aff'd sub nom. Davis v. City of Springfield*, No. 04-3168, 2009 WL 1161619 (C.D. Ill. Apr. 28, 2009) (quashing the subpoena because the information could be obtained more easily from the party in the case, while stating "Certainly, Rule 45(c) does not require [a party] to exhaust other means of securing information before seeking it from [a non-party]; however, the Court will consider the availability of the information from other sources in balancing the relative hardships.").

requested documents and ESI from the law firm via Rule 45.

C. Rule 45 Rights and Obligations Where a Party Does Not Have Possession, Custody, or Control

Parties that do not have “control” over documents and ESI in the possession or custody of a non-party may still have significant interests at stake concerning the production of those documents and ESI and should generally be included in the management of the scope of such productions and limitations thereto. They also potentially bear some or all of the burden of production under certain circumstances.²⁵ This section explores procedures that should be considered to protect a party’s interests in documents or ESI of which it has no possession, custody, or control, as well as the protections afforded to the non-party.

1. Subpoenaing a Non-Party with Sole Possession, Custody, and Control

Some non-parties have sole possession, custody, or control of requested documents or ESI—i.e., a party to the litigation has no control over these documents or ESI. In such situations, requesting parties should issue non-party subpoenas. However, these non-parties should be afforded protections under Rule 45, including cost shifting, or quashing or modifying the subpoena as appropriate.²⁶

25. The analysis is slightly different where the non-party is related to a party to the litigation. *See* Section III.C.2., *infra*.

26. *See* *United States v. Columbia Broad. Sys., Inc.*, 666 F.2d 364, 371–72 (9th Cir. 1982) (“Nonparty witnesses are powerless to control the scope of litigation and discovery, and should not be forced to subsidize an unreasonable share of the costs of a litigation to which they are not a party.... [W]e... emphasize that a witness’s nonparty status is an

Illustration 1: An internet search company is subpoenaed by a plaintiff in a defamation case for search results relating to any statements made by the defendant about the plaintiff during a specific time period. The internet search company has no interest in the litigation and has sole possession, custody, or control of the requested documents and ESI. A non-party subpoena should be used, and the protections and cost-shifting mechanisms under Rule 45 should be fully available to the internet search company. To the extent cost shifting is appropriate, the court may allocate those costs to the plaintiff.

Illustration 2: A city traffic department is subpoenaed by a plaintiff in a personal injury case for its internally maintained video footage of an intersection where plaintiff claims defendant caused him injury. The city traffic department has no interest in the litigation and has sole possession, custody, or control of the requested documents and ESI. A non-party subpoena should

important factor to be considered in determining whether to allocate discovery costs on the demanding or the producing party.”); *Nitsch v. DreamWorks Animation SKG Inc.*, No. 14-cv-04062-, 2017 WL 930809, at *3 (N.D. Cal. Mar. 9, 2017) (finding non-party’s efforts to protect the confidential information to be reasonable and compensable and shifting the costs to the plaintiff); *Spears v. First Am. eAppraiseIT*, No. 5:08-cv-00868, 2014 WL 11369809, at *3–4 (N.D. Cal. July 3, 2014) (noting that a non-party subpoena recipient was not “substantially involved in underlying events [nor] had a significant relationship with the litigants” and thus was afforded full protection under Rule 45); *In re Honeywell, Int’l, Inc. Sec. Litig.*, 230 F.R.D. 293, 303 (S.D.N.Y. 2003) (finding that whether the non-party has an interest in the outcome of the case is an important factor in determining who should bear the costs of discovery).

be used, and the protections and cost-shifting mechanisms under Rule 45 should be fully available to the city traffic department. To the extent cost shifting is appropriate, the court may allocate those costs to the plaintiff.

2. Subpoenaing a Non-Party That Has a Relationship to a Party

Some non-parties have a prior or current relationship to a party to the litigation.²⁷ Subpoenas directed to these non-parties frequently involve complex possession, custody, or control issues as well as substantive issues of corporate separateness and veil piercing.²⁸ Here, where possession, custody, or control

27. See generally *St. Jude Med. S.C., Inc. v. Janssen-Counotte*, 305 F.R.D. 630 (D. Or. 2015) (finding there was a “sufficient indicia of effective control” to require European affiliates of the non-party to conduct a search for responsive documents and ESI, where European affiliates were acting as non-party’s agent-in-hiring and non-party itself did not have responsive documents and ESI); *Wells Fargo Bank, N.A. v. Konover*, 259 F.R.D. 206 (D. Conn. 2009) (finding that a non-party is not a “truly disinterested party,” due to its corporate structure and representation by the same law firm); *In re First Am. Corp.*, 184 F.R.D. 234, 242 (S.D.N.Y. 1998) (while Rule 45 protects non-parties from significant expense in producing documents and ESI, the non-party was not “the quintessential innocent, disinterested bystander,” as it should have reasonably anticipated being drawn into litigation resulting from the underlying alleged fraud and was therefore responsible for some of the production costs).

28. See *Flame S.A. v. Indus. Carriers, Inc.*, 39 F. Supp. 3d 752, 759 (E.D. Va. 2014) (“The analysis therefore applies with equal force, and for related non-parties, like parent, sister, or subsidiary corporations, courts examine (1) the corporate structure of the party/non-party; . . . (4) whether the related entities exchange documents in the ordinary course of business; . . . (6) common relationships between a party and its related non-party entity; (7) the ownership of the non-party; (8) the overlap of directors, officers, and employees; . . . and (11) agreements among the entities that may reflect the parties’ legal rights or authority to obtain certain documents.” (citing E.I.

does not exist but a relationship exists (which could implicate privacy, confidentiality, or privilege concerns), special considerations for coordination may still be appropriate for responses to the subpoena.

Where a non-party has a relationship to a party, that relationship may impact cost shifting and coordination among the parties and the non-party. Therefore, courts would need to balance the competing interests of the parties and the non-party in the requested documents and ESI.

Illustration: A non-party parent company is subpoenaed by the plaintiff in a patent infringement case against its subsidiary. The plaintiff requests specific relevant documents regarding the research and development of the allegedly infringing product. The requested documents are in the possession of the parent company, and the subsidiary does not have control over those documents. Therefore, the subpoena should be directed to the parent company, and if the nature of the relationship between the parent and the subsidiary is fully aligned and cooperative, special considerations for coordination in production between the parent and subsidiary may not be necessary because the parent may be incentivized and capable of

DuPont de Nemours & Co. v. Kolon Indus., Inc., 286 F.R.D. 288, 292 (E.D. Va. 2012)); *see also* Level One Techs., Inc. v. Penske Truck Leasing Co., L.P., No. 4:14 CV 1305, 2018 WL 3819042, at *1–2 (E.D. Mo. Aug. 10, 2018) (holding that defendant had “possession, custody, and control” of third-party service provider’s time records and monthly register, as the service provider was present at defendant’s offices, much of its work is stored on the defendant’s servers and systems, and “it is clear that [service provider] would comply with any demand from [defendant] for documents supporting the contracted projects”).

protecting the subsidiary's interests in the litigation context.

Were it not for the corporate relationship in the Illustration, the subpoena to the parent should still be used, but the unrelated non-party likely would be entitled to greater protection.

IV. PRESERVATION

A. *Prior to Receipt of a Subpoena*

Generally, a non-party has no obligation to preserve documents prior to receipt of a subpoena or after complying with a subpoena, absent a special relationship to a party to the litigation.²⁹ A written or oral preservation demand creates no duty to preserve materials.³⁰ In *Tassin v. Bob Barker Co.*, the court found that a written request for the preservation of evidence did not create a non-party preservation obligation.³¹ The plaintiff sent correspondence to a non-party's supervisory employee requesting that any video concerning his accident (the basis of his lawsuit against defendant) be preserved.³² When no response was received from the non-party, the plaintiff sought

29. See The Sedona Conference, *Commentary on Legal Holds, Second Edition: The Trigger & The Process*, 20 SEDONA CONF. J. 341, 365 (2019) (discussing circumstances where there is a "special, affiliated, or contractual relationship with a party" and where, after receipt of a subpoena, non-parties that "have actual or constructive control of discoverable information" should decide whether a duty to preserve discoverable information has been triggered).

30. Courts have specifically noted the distinction between a written request and the legal force of a general subpoena or preservation subpoena. See, e.g., *In re Heckmann Corp. Sec. Litig.*, No 10-378-LPS-MPT, 2011 WL 10636718, at *5 (D. Del. Feb. 28, 2011) (finding in Private Securities Litigation Reform Act (PSLRA) action, "[s]ending preservation letters . . . is distinct from serving preservation subpoenas because the latter imposes a legal obligation on third parties to take reasonable steps to preserve relevant documents") (relying in part upon *Koncelik v. Savient Pharms., Inc.*, No. 08 Civ. 10262, 2009 WL 2448029, at *2 (S.D.N.Y. Aug. 10, 2009) ("The only thing that is certain is that without preservation subpoenas, the third party corporations in possession of potentially relevant information are free to destroy that information.")).

31. *Tassin v. Bob Barker Co.*, No. 16-0382-JWD-EWD, 2017 WL 9963365, at *2 (M.D. La. Sept. 28, 2017).

32. *Id.* at *1.

an order from the court to compel the non-party to preserve the potential video evidence.³³ The court denied plaintiff's motion, explaining that it was not clear whether any such recordings existed, and there was no indication that the non-party had been dilatory.³⁴ The court advised the plaintiff that he could request potential video footage by subpoena and declined to address any issue of spoliation pending a response by the non-party to a subpoena.³⁵

A non-party, however, may have a preservation obligation prior to receipt of a subpoena where contractual obligations to a party exist.³⁶ In *Holmes v. Amerex Rent-A-Car*, a case brought by a motorist against the manufacturer of a rental car in which he was injured while driving, the court assumed "that a duty of care existed which was derived from a contractual relationship to transfer ownership of the car wreckage from the [car rental agency, which took possession of the wrecked automobile] to the [plaintiff] and that the duty was breached" by the rental

33. *Id.*

34. *Id.* at *2.

35. *Id.*

36. *Andra Grp., LP v. JDA Software Grp., Inc.* No. 3:15-MC-11, 2015 WL 12731762, at *15, *17 (N.D. Tex. Dec. 9, 2015); *Koplin v. Rosel Well Perforators, Inc.*, 734 P.2d 1177, 1179 (Kan. 1987) ("[S]ome special relationship or duty rising by reason of an agreement, contract, statute, or other special circumstance" is necessary to give rise to a "duty to preserve possible evidence for another party to aid that other party in some future legal action against a third party."); *Fletcher v. Dorchester Mut. Ins. Co.*, 773 N.E. 2d 420, 425 (Mass. 2002) ("A third-party witness may also agree to preserve an item of evidence and thereby enter into an enforceable contract." (citing *Koplin*, 734 P.2d at 1179)); see also *Tassin*, 2017 WL 9963365, at *1 ("[D]uty [to preserve] may extend to a non-party . . . when the non-party enters into an agreement to preserve the evidence sought to be obtained." (internal citation omitted)).

agency's destruction of the car.³⁷ Similarly, a non-party witness may agree to preserve an item of evidence and thereby enter into an enforceable contract with a party.³⁸

A non-party may also have a preservation obligation prior to a subpoena where it has a special relationship with a party.³⁹ However, very few cases discuss what kind of special relationship must exist to trigger a non-party's preservation obligation, and the cases that address this issue are fact-driven.

A close working relationship in and of itself does not rise to the level of the special relationship required to impose a retention obligation. In *Andra Group, LP v. JDA Software Group, Inc.*, the court examined whether a non-party project management limited liability company ("p202") for plaintiff Andra's software development project had a duty to preserve certain information prior to receipt of a subpoena from defendant JDA.⁴⁰ Andra had hired p202 to manage the software development project that was the basis for the dispute between Andra and JDA.⁴¹ In its subpoena, JDA sought a number of items from p202, including telephone recordings between Andra's CEO and p202's project manager.⁴² Prior to receipt of the subpoena, however, p202 had deleted those recordings.⁴³ p202's project manager testified at a deposition that he did not keep or archive the recordings once p202 had completed work on the project and been paid by Andra.⁴⁴ p202's project manager

37. *Holmes v. Amerex Rent-A-Car*, 710 A.2d 846, 850 (D.C. 1998).

38. *Fletcher*, 773 N.E. 2d at 425 (citing *Koplin*, 734 P.2d at 1177).

39. *See Andra*, 2015 WL 12731762, at *15.

40. *Id.* at *15–16.

41. *Id.* at *2.

42. *Id.* at *5.

43. *Id.* at *6.

44. *Id.*

also testified that p202 did not have a document or email retention policy, nor did it have, at the time the evidence was deleted, a practice or custom for storing digital, video, tape, or audio recordings.⁴⁵ p202's project manager further testified that prior to the subpoena from JDA, Andra's CEO informed him that she was considering litigation against JDA, and that despite this communication from Andra's CEO, p202 did not have a litigation hold in place until it was served with JDA's subpoena in the matter.⁴⁶ JDA filed a motion for civil contempt against p202 regarding its deletion of the telephone recordings and certain other items, claiming p202 had not been compliant.⁴⁷ The court determined that despite the close working relationship between Andra and p202, and p202 being informed of potential litigation between Andra and JDA, p202 did not have a duty to preserve the telephone recordings and certain other items prior to receipt of JDA's subpoena.⁴⁸

In some cases in which the commencement of discovery is delayed, generally due to a statutory stay or lengthy pre-discovery motion practice, such as securities actions subject to the Private Securities Litigation Reform Act of 1995 (PSLRA), courts have issued orders, based upon a specific evidentiary showing, for the issuance of so-called preservation subpoenas to a non-party requiring preservation of relevant documents of ESI.⁴⁹ Such court orders, however, should include Rule 45's

45. *Id.* at *6, *8.

46. *Id.* at *9.

47. *Id.* at *6.

48. *Id.* at *15.

49. *In re Smith Barney Transfer Agent Litig.*, No. 05 Civ. 7583(WHP), 2012 WL 1438241, at *3 (S.D.N.Y. Apr. 25, 2012); *see also*, e.g., *Gruber v. Gilbertson*, No. 16-cv-9727, 2017 WL 3891701, at *4 (S.D.N.Y. Sept. 5, 2017) (Where non-parties hold relevant documents to which plaintiff will be entitled if it prevails on the motion(s) to dismiss, "courts have generally permitted

protection against undue burden and expense by avoiding overbroad requests and properly tailoring preservation to the scope of discovery required by the circumstances, including proportionality. Leave of court to serve preservation subpoenas has also been granted in other litigation settings, including cases consolidated or coordinated through the Judicial Panel for Multidistrict Litigation and other complex cases in which the commencement of discovery may be delayed substantially.⁵⁰

plaintiffs in PSLRA actions to issue subpoenas that have given specified third parties notice of the action and impose upon them only a duty to preserve certain relevant evidence in their possession.” (quoting *In re Smith Barney*, 2012 WL 1438241, at *3)); *Avenue Capital Management II, LP v. Schaden*, No. 14-CV-02031-PAB-KLM, 2015 WL 758521, at *3–4 (D. Colo. Feb. 20, 2015) (same) (quoting *In re Smith Barney*, 2012 WL 1438241, at *3); *Caston v. Hoaglin*, No. Civil Action No. 2:08-CV-200, 2009 WL 1687927, at *2 (S.D. Ohio June 12, 2009) (holding plaintiff had good cause to serve preservation subpoenas in PSLRA action prior to Rule 26(f) discovery conference where information sought in subpoena request was narrow and the evidence was critical to defendants’ alleged breaches of fiduciary duties); *In re Refco, Inc. Sec. Litig.*, No. 05 Civ. 0826 (GEL), 2006 WL 2337212, at *5 (S.D.N.Y. Aug. 8, 2006) (“[C]ourts have generally permitted plaintiffs in PSLRA actions to ‘issue subpoenas that give specified third parties notice of the action and impose upon them only a duty to preserve certain relevant evidence in their possession.’” (internal citations omitted; collecting cases)); *Payne v. DeLuca*, CA No. 02-1927, 2005 WL 8152650, at *5 (W.D. Pa. Dec. 23, 2005) (granting defendants’ motion for order permitting issuance of preservation subpoenas); *In re Nat’l Century Fin. Enter., Inc. Fin. Inv. Litig.*, 347 F. Supp. 2d 538, 542 (S.D. Ohio 2004) (granting motion to issue document subpoena to debtor); *In re Tyco*, 2000 WL 33654141, at *3–4 (subpoenas authorized where, unlike the defendants, the non-parties had not necessarily received actual notice of the action, and plaintiff produced evidence that large corporations typically overwrite and thereby destroy electronic data in the course of performing routine backups).

50. *Johnson v. U.S. Bank Nat’l Ass’n*, No. 1:09-cv-492, 2009 WL 4682668, at *1–2 (S.D. Ohio Dec. 3, 2009) (Telemarketing Fraud: applying “good cause” standard and authorizing service of preservation subpoena prior to Rule 26(f) conference where non-party was a “critical link” in the alleged scheme,

B. After Receipt of a Subpoena

Rule 45 and its Advisory Committee Notes are devoid of any reference to preservation.⁵¹ The Rule does require that the issuing party take steps to avoid undue burden or expense to the subpoenaed non-party, and the subpoenaed non-party can either produce the subpoenaed documents, object to the subpoena, or move to quash. However, a non-party subpoena recipient should be careful not to destroy or discard information responsive to a subpoena, because the Rule provides for contempt sanctions if the non-party “fails without adequate excuse to obey the subpoena or an order related to it,”⁵² and some states have recognized spoliation as an independent tort.

Thus, although a subpoena imposes an obligation on the non-party to ensure documents responsive to the subpoena are not destroyed pending compliance with the subpoena, the nature and extent of the obligation varies depending on the facts and circumstances presented. In most cases, receipt of a properly served subpoena only obligates a non-party to take

and where preservation was necessary to ensure that records and databases were not destroyed, lost, or otherwise despoiled); *Tama Plastic Indus. v. Pritchett Twine & Net Wrap, LLC*, No. 8:12CV324, 2013 WL 275013, at *3–4 (D. Neb. Jan. 24, 2013) (Patent: excluding non-party discovery from stay: “Tama will likely have a more difficult time gathering information after a two-year wait because the third parties may dispose of documents and because memories tend to fade over the course of time. Accordingly, a granting of a complete stay of all discovery in this case will likely cause prejudice and tactical disadvantage to Tama with respect to information currently in the hands of third parties.”); *see also* *State Farm Mut. Auto. Ins. Co. v. Physiomatrix, Inc.*, Civil Action No. 12-cv-11500, 2013 WL 10936871, at *5 (E.D. Mich. Nov. 26, 2013) (RICO: ordering non-party to preserve emails from identified account for 180 days in order to permit plaintiff to subpoena emails if it learned that the account was controlled by defendants).

51. Although a reference to preservation was specifically added to Rule 26 in the 2015 amendments, it was not added to Rule 45.

52. *See* FED. R. CIV. P. 45(g).

reasonable steps to produce the requested materials. The subpoena does not obligate the non-party to initiate a formal legal hold process. What is required is to ensure that materials are retained until there is compliance. Absent a contractual or other special obligation, a non-party has no duty to preserve information after it has complied with the subpoena.

Since Rule 45(d)(2)(B)(ii) places no time limit on when a party must move to compel production of documents sought in a subpoena, a subpoenaing party may find itself in a more difficult position when the non-party objects to producing all or part of the information subpoenaed or otherwise fails to fully comply with the subpoena. In those circumstances, it is advisable for the requesting party to provide prompt notice of its intent to move to compel compliance. If the requesting party does not promptly move to compel, the non-party may be faced with a dilemma about how long it needs to preserve documents. To protect itself, a non-party should consider specifying a reasonable date after which it will no longer retain the documents or ESI, thereby placing the requesting party on notice of the date by which it needs to move to compel, if it plans to do so. The party issuing the subpoena should promptly move to compel in such a situation. In addition, while not currently required by the Federal Rules of Civil Procedure, the non-party and the party issuing the subpoena may wish to meet and confer, to discuss and try to resolve any disputes as to the scope of discovery and scope of the subpoena, or other matters including retention, before seeking to quash or enforce the subpoena.

C. Remedies for Spoliation

While certain states have recognized spoliation as an independent tort,⁵³ there does not exist an independent federal cause of action for spoliation of evidence.⁵⁴ Moreover, Rule 37(e) applies only to parties, not to a non-party.⁵⁵ A non-party's failure to produce documents or ESI responsive to a subpoena may result in a Rule 45(g) sanction of contempt—often a monetary fine—before the court in that action.⁵⁶

53. While it is beyond the scope of this paper to consider the potential tort liability of non-parties who destroy evidence relevant to others' disputes, the law in this area is developing and has been addressed by a majority of states. Several publications analyze and tally the states that recognize or reject spoliation as a separate tort. *See, e.g.*, 86 C.J.S. *Torts* § 78, Westlaw (database updated Dec. 2018); 1 STEVEN PLITT & JORDAN ROSS PLITT, *PRACTICAL TOOLS FOR HANDLING INSURANCE CASES* § 7:42 (2018); 40 ERIC M. LARSSON, *CAUSES OF ACTION 2D 1* (2018); AM. L. PROD. LIAB. 3D § 53:160, Westlaw (database updated Nov. 2018).

54. *R.C. Olmstead, Inc. v. CU Interface, L.L.C.*, 657 F. Supp. 2d 878, 887 (N.D. Ohio 2009) (citing *Lombard v. MCI Telecomm. Corp.*, 13 F. Supp. 2d 621, 628 (N.D. Ohio 1998)); *In re Elec. Mach. Enters., Inc.*, 416 B.R. 801, 872 (Bankr. M.D. Fla. 2009). In diversity actions, federal courts will apply local law with regard to substantive issues, but under the *Erie* doctrine, they will apply federal law to procedural issues. *Foster v. Lawrence Memorial Hosp.*, 809 F. Supp. 831, 835 n.1 (D. Kan. 1992).

55. *In re Correra*, 589 B.R. 76, 123–24 (N.D. Tex. 2018) (“Rule 37(e) applies only to parties.”).

56. Rule 45(g) provides that the court “may hold in contempt a person who, having been served, fails without adequate excuse to obey the subpoena or an order related to it.” FED. R. CIV. P. 45(g). Indeed, because properly served subpoenas have the effect of a court order, contempt sanctions are the logical remedy for the failure to comply, including failure to preserve documents. *See, e.g.*, *Washington v. Trump*, No. 17-0141JLR, 2017 WL 2172020, at *4 (W.D. Wash. May 17, 2017) (“The issuance of subpoenas to third parties . . . provide the force of a court order with respect to the preservation of . . . evidence.”); *In re Heckmann Corp. Sec. Litig.*, No 10-378-LPS-MPT, 2011 WL 10636718, at *34 (D. Del. Feb. 28, 2011) (deeming

preservation subpoenas to be court orders); *In re Napster, Inc. Copyright Litig.*, 462 F. Supp. 2d 1060, 1068–70 (N.D. Cal. 2006) (imposing sanctions for violation of legal obligation to preserve documents pursuant to receipt of subpoena where it determined non-party had reasonable cause to believe it would become a party); *Fletcher v. Dorchester Mut. Ins. Co.*, 773 N.E. 2d 420, 425 (Mass. 2002) (duty imposed by a subpoena is “enforced as needed by appropriate court orders, up to and including holding the witness in contempt”); *SonoMedica, Inc. v. Mohler*, No. 1:08-cv-230, 2009 WL 2371507, at *5 (E.D. Va. July 28, 2009) (finding non-parties in civil contempt for failing to produce documents pursuant to subpoena, destroying ESI, and lying at depositions; ordering non-parties to pay attorney’s fees; and referring matter to U.S. Attorney for criminal contempt proceedings).

V. RULE 45(D) COSTS, SANCTIONS, AND MOTION PRACTICE

Although many Rule 45 subpoenas are handled without any court intervention, Rule 45(d) provides three avenues by which a non-party subpoena recipient may be protected from the costs of compliance. This *Commentary* addresses these provision in the order they appear in the Rule—first, sanctions under Rule 45(d)(1); second, cost shifting under Rule 45(d)(2)(B)(ii)—but it recognizes that quashing or modifying a subpoena, which is discussed in the third subsection, is many times a more appropriate first step. Thus, practitioners should not give less consideration to quashing or limiting the scope of the subpoena under Rule 45(d)(3) to resolve issues.

A. *Rule 45(d)(1)—Avoiding Undue Burden or Expense and Sanctions*

Rule 45(d)(1) provides:

Avoiding Undue Burden or Expense; Sanctions. A party or attorney responsible for issuing and serving a subpoena must take reasonable steps to avoid imposing undue burden or expense on a person subject to the subpoena. The court for the district where compliance is required must enforce this duty and impose an appropriate sanction—which may include lost earnings and reasonable attorney’s fees—on a party or attorney who fails to comply.

The first mechanism for protecting subpoena recipients is squarely in the hands of the court. Rule 45(d)(1) requires that a party or attorney responsible for issuing a subpoena take *reasonable steps* to avoid imposing undue burden or expense on a non-party. The court for the district where compliance is sought must enforce this duty and impose an “appropriate” sanction on a party or attorney who fails to meet this

requirement. Although the rule provides that “appropriate” sanctions may include lost earnings and reasonable attorney’s fees, courts have discretion over the type and degree of sanctions imposed. In determining whether sanctions should be imposed under Rule 45(d)(1), courts consider a number of factors, “including the person’s status as a non-party, the relevance of the discovery sought, the subpoenaing party’s need for the documents, the breadth of the request, and the burden imposed on the subpoenaed party.”⁵⁷

Undue burden is assessed in a case-specific manner considering “such factors as relevance, the need of the party for the documents, the breadth of the document request, the time period covered by it, the particularity with which the documents are described and the burden imposed.”⁵⁸ “Courts are required to balance the need for discovery against the burden imposed on the person ordered to produce documents, and the status of that person as a non-party is a factor that weighs against disclosure.”⁵⁹ Thus, as this *Commentary* suggests, the status of a non-party as related to a party is a factor.

Courts are also required to evaluate the reasonableness of the steps the issuing party took to avoid undue burden. Where an issuing party attempted to engage in good-faith negotiations to either reduce the burden or narrow the scope of the subpoena, courts have declined to impose sanctions.⁶⁰ The plain language

57. *Parker v. Four Seasons Hotels, Ltd.*, 291 F.R.D. 181, 188 (N.D. Ill. 2013).

58. *Am. Elec. Power Co. v. U.S.*, 191 F.R.D. 132, 136 (S.D. Ohio 1999) (quoting *Concord Boat Corp. v. Brunswick Corp.*, 169 F.R.D. 44, 53 (S.D.N.Y. 1996)).

59. *Id.*

60. *See In re Am. Kidney Fund, Inc.* 2019 WL 1894248, at *6 (good-faith negotiations to limit the scope of the subpoena and the fact that requesting party refrained from serving the subpoena until obtaining party discovery

of the provision, however, does suggest that sanctions may be imposed when a subpoenaing attorney or party unfairly harms a subpoena recipient by acting carelessly or in bad faith when issuing a subpoena. However, a finding of bad faith is not required for sanctions to be imposed under Rule 45(d)(1).⁶¹

Merely losing a motion to compel does not in and of itself expose a requesting party to Rule 45(d)(1) sanctions.⁶² While failure to narrowly tailor a subpoena may be a ground for sanctions, the court need not impose sanctions every time it finds a subpoena overbroad; such overbreadth may sometimes result from normal advocacy and does not necessarily give rise to sanctions.

The history of Rule 45 provides guidance on how this section should be interpreted in the event of a misuse of the subpoena process. Rule 45 was amended in 1991 to bring the protections for subpoenaed non-parties under a single subdivision. But the 1991 Advisory Committee Notes suggest that the amendment did not effect a “change in existing law” and was designed to codify existing practice, including to give “specific application” to the principles stated in Rule 26(g).⁶³ Federal Rule of Civil Procedure 26(g)(1)(B) requires parties seeking discovery to act “(i) consistent with these rules and warranted by existing law or

demonstrates the requesting party took reasonable steps to limit undue burden; sanctions under Rule 45 (d)(1) not warranted).

61. *Legal Voice v. Stormans Inc.*, 738 F.3d 1178, 1185 (9th Cir. 2013); *see also* *Mount Hope Church v. Bash Back!*, 705 F.3d 418, 429 (9th Cir. 2012) (holding that bad faith is sufficient but not necessary to impose sanctions if Rule 45(d)(1) otherwise is violated).

62. *See Mount Hope Church*, 705 F.3d at 425–27 ; *Mattel Inc. v. Walking Mountain Prods.*, 353 F.3d 792, 814 (9th Cir. 2003). *Mount Hope Church*, 705 F.3d at 425–27.

63. *See* FED. R. CIV. P. 45(d) advisory committee’s note to 1991 amendment. The 1991 Advisory Committee Notes refer to subdivision (c), which became subdivision (d) in the 2013 amendments.

by a non-frivolous argument for extending, modifying, or reversing existing law, or for establishing new law; (ii) not interposed for any improper purpose, such as to harass, cause unnecessary delay, or needlessly increase the cost of litigation; and (iii) neither unreasonable nor unduly burdensome or expensive, considering the needs of the case, prior discovery in the case, the amount in controversy, and the importance of the issue at stake in the action.” Violation of any one of these duties without substantial justification may result in sanctions.⁶⁴ Because Rule 45(d)(1) gives “specific application” to Rule 26(g),⁶⁵ it follows that a violation of any one of the Rule 26(g) duties will be relevant to assessing the propriety of sanctions under Rule 45(d)(1)’s “undue burden” language.

B. Rule 45(d)(2)(B)(ii)—Mandatory Cost Shifting When the Non-Party is Ordered to Produce with Significant Expense Over Objection

Rule 45(d)(2)(B) provides:

Objections. A person commanded to produce documents or tangible things or to permit inspection may serve on the party or attorney designated in the subpoena a written objection to inspecting, copying, testing or sampling any or all of the materials or to inspecting the premises—or to producing electronically stored information in the form or forms requested. The objection must be served before the earlier of the time specified for compliance or 14 days after the subpoena is served. If an objection is made, the following rules apply:

64. FED. R. CIV. P. 26(g)(3).

65. *Id.*

- i. At any time, on notice to the commanded person, the serving party may move the court for the district where compliance is required for an order compelling production or inspection.
- ii. These acts may be required only as directed in the order, and the order must protect a person who is neither a party nor a party's officer from significant expense resulting from compliance. (emphasis added.)

Part 2 of Rule 45(d) gives the non-party the ability to protect itself from significant expense if complying with a subpoena over its objection. Under Rule 45(d)(2)(B)(ii), when a court orders compliance with a subpoena over a non-party's objection, the court should protect the non-party from significant expense resulting from compliance upon a showing by the non-party that they would incur significant expenses in responding to the subpoena.⁶⁶ If the non-party would be subjected to significant expense, this protection shifts as much of the compliance expense as necessary to the requestor to render the remaining expenses non-significant.⁶⁷

Before Rule 45 was amended in 1991, cost shifting was within the court's discretion.⁶⁸ As the Advisory Committee

66. FED. R. CIV. P. 45(d)(2)(B)(ii).

67. *Id.*; *Koopmann v. Robert Bosch LLC*, No. 18-CV-4065, 2018 WL 9917679, at *1 (S.D.N.Y. May 25, 2018) (holding that Petitioners "should bear some of the Respondent's costs for complying with the Subpoena. Courts have deemed Rule 45(d)(2)(B)(ii) 'to make cost shifting mandatory in all instances in which a non-party incurs significant expense from compliance with a subpoena.'" (quoting *Sands Harbor Marina Corp. v. Wells Fargo Ins. Servs. of Or., Inc.*, No. 09-CV-3855, 2018 WL 1701944, at *3 (E.D.N.Y. Mar. 31, 2018) (internal quotation marks & brackets omitted))).

68. *Linder v. Calero-Portocarrero*, 251 F.3d 178, 182 (D.C. Cir. 2001).

Notes to the 1991 amendment explain and courts have found, this section is now mandatory.⁶⁹ The changes were intended “to enlarge the protections afforded [non-parties] who are required to assist the court.”⁷⁰

1. Prerequisites for Seeking Cost Shifting

Before a non-party can seek reimbursement for costs under Rule 45(d)(2)(B)(ii)’s cost-shifting provision, several requirements must be met. First, the non-party must file a *timely* and *specific* objection to the subpoena. Second, the requesting party must move the court under Rule 45(d)(2)(B)(i) to compel production over the non-party’s objection. Third, the court must enter an order compelling the non-party to comply with the subpoena and produce the requested documents or ESI at a significant expense to the non-party. Only after these prerequisites have been met can a non-party request reimbursement for “significant expenses” under Rule 45(d)(2)(B)(ii)’s cost-shifting provision.

(a) Non-Party Serves Objections

(1) Must be Timely

If the non-party chooses to serve a written objection to a subpoena rather than, or in addition to, moving to quash or

69. *Id.* (finding that under the revised Rule 45, the “rule is susceptible of no other interpretation” but that it is mandatory); *see also* *Voice v. Stormans Inc.*, 738 F.3d 1178, 1184 (9th Cir. 2013) (“This language leaves no room for doubt that the rule is mandatory”; *Iowa Pub. Emples. Ret. Sys. v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, No. 17-6221, 2019 WL 7283254 at *2 (S.D.N.Y. Dec. 26, 2019) (finding that the “plain text obligates, and not merely empowers, the Court to protect third parties from significant expenses resulting from compliance with subpoenas.”).

70. *Id.* (citing FED. R. CIV. P. 45 advisory committee’s notes to 1991 amendment); *United States v. CBS, Inc.*, 666 F.2d 364, 371 n.9 (9th Cir. 1982).

modify it under Rule 45(d)(3)(A), the objection “must be served [on the issuing party] before the earlier of the time specified for compliance or 14 days after the subpoena is served.”⁷¹ Although some courts have held that in unusual circumstances the failure to submit timely objections is not an automatic waiver and the objections still may be considered,⁷² untimely service “typically constitutes a waiver of such objections.”⁷³ If necessary to ensure that the timeliness requirement is met, a non-party can request

71. See FED. R. CIV. P. 45(d)(2)(B).

72. *In re Ex Parte Application of Grupo Mexico SAB de CV for an Order to Obtain Discovery for Use in a Foreign Proceeding*, No. 3:14-mc-73, 2015 WL 12916415, at *3 (N.D. Tex. Mar. 10, 2015), *aff'd sub nom.* *Grupo Mexico SAB de CV v. SAS Asset Recovery, Ltd.*, 821 F.3d 573 (5th Cir. 2016) (“However, in unusual circumstances and for good cause . . . the failure to act timely will not bar consideration of objections.” (quoting *In re Denture Cream Prods. Liab. Litig.*, 292 F.R.D. 120, 124 (D.D.C. 2013) (internal quotations & citation omitted))). “Unusual circumstances” exist when “(1) the subpoena is overbroad on its face and exceeds the bounds of fair discovery; . . . (2) the subpoenaed witness is a nonparty acting in good faith; . . . and (3) counsel for witness and counsel for subpoenaing party were in contact concerning the witness’ compliance prior to the time the witness challenged [the] legal basis for the subpoena.” *In re Denture Cream Prods.*, 292 F.R.D. at 124 (alterations in original) (citation omitted); see also *Piazza’s Seafood World, L.L.C. v. Odom*, No. 07-413, 2011 WL 3664437, at *3 (M.D. La. Aug. 19, 2011) (“[I]f aspects of a subpoena are overbroad on their face and exceed the bounds of fair discovery and the subpoenaed witness is a non-party acting in good faith, waiver of the non-party’s untimely objections is not automatic, and the objections may be considered.”).

73. *In re Ex Parte Application of Grupo Mexico SAB de CV*, 2015 WL 12916415, at *3 (citing *Isenberg v. Chase Bank USA*, 661 F. Supp. 2d 627, 629 (N.D. Tex. 2009) (“The failure to serve written objections to a subpoena within the time specified by [Rule 45(d)(2)(B)] typically constitutes a waiver of such objections.” (quoting *Concord Boat Corp. v. Brunswick Corp.*, 169 F.R.D. 44, 48 (S.D.N.Y. 1996))); *Piazza’s Seafood World, L.L.C.*, 2011 WL 3664437, at *3 (“Failure of a nonparty to serve timely objections to a Rule 45 subpoena generally results in a waiver of all grounds for objection.” (citing *Moon v. SCP Pool Corp.*, 232 F.R.D. 633 (C.D. Cal. 2005))).

that the issuing party extend the non-party's deadline to serve written objections.⁷⁴ Some circuits have interpreted Rule 45 to require that all objections be "raised at once, rather than in staggered batches."⁷⁵ In addition to or rather than serving a written objection, the non-party may move to quash or modify the subpoena under Rule 45(d)(3)(A).⁷⁶

Timely service of written objections suspends the non-party's obligation to comply with the subpoena until there is a court order compelling compliance.⁷⁷

74. *Am. Fed'n of Musicians of the U.S. & Can. v. Skodam Films, LLC*, 313 F.R.D. 39, 43 (N.D. Tex. 2015) ("The serving party may agree to extend the deadline to respond to a subpoena, including the deadline to serve written objections." (citing *Shaw Grp., Inc. v. Zurich Am. Ins. Co.*, No. 12-257, 2014 WL 1816494, at *8 (M.D. La. May 7, 2014))).

75. *Young v. City of Chicago*, No. 13 C 5651, 2017 WL 25170, at *8 (N.D. Ill. Jan. 3, 2017) ("Rule 45 'require[s] the recipient of a subpoena to raise all objections at once, rather than in staggered batches, so that discovery does not become a game.'" (quoting *Ott v. City of Milwaukee*, 682 F.3d 552, 558 (7th Cir. 2012) (quoting *In re DG Acquisition Corp.*, 151 F.3d 75, 81 (2d Cir. 1998)))).

76. *MetroPCS v. Thomas*, 327 F.R.D. 600, 607 (N.D. Tex. 2018) ("Under Federal Rule of Civil Procedure 45(d), '[e]ither in lieu of or in addition to serving objections on the party seeking discovery, a person can 'timely' file a motion to quash or modify the subpoena' under Federal Rule of Civil Procedure 45(d)(3)(A)." (citing *In re Ex Parte Application of Grupo Mexico SAB de CV*, 2015 WL 12916415, at *3)).

77. *MetroPCS*, 327 F.R.D. at 607; *see also* FED. R. CIV. P. 45(d)(2)(B)(ii); *Pennwalt Corp. v. Durand-Wayland, Inc.*, 708 F.2d 492, 494 (9th Cir. 1983) (once a non-party objects to a subpoena *duces tecum*, the non-party is no longer "obligated to produce the subpoenaed documents"); *Ctr. for Individual Rights v. Chevaldina*, No. 16-20905-Civ, 2017 WL 5905191, at *4 (S.D. Fla. Nov. 29, 2017) ("If a non-party timely serves written objections, the non-party's objection to comply with the subpoena is suspended pending a court's order." (citations omitted)); *Am. Fed'n of Musicians*, 313 F.R.D. at 44 (denying subpoena issuer's request for fees for filing a motion to compel because the non-party was not required to produce documents unless and

(2) Must be Specific

In 2015, Rule 34 was amended to focus on specificity of objections. The pre-2015 case law held non-parties to similar standards when it came to objections.⁷⁸ Pre-2015 cases noted that a non-party's objection should be free from general or boilerplate language and should be made with enough specificity to allow the parties to understand the scope of the

until the subpoena issuer obtained a court order); *Forsythe v. Brown*, 281 F.R.D. 577, 587 (D. Nev. 2012). This benefit does not extend to subpoenas seeking deposition testimony, objections to which do not suspend a non-party's duty to appear and testify. *See BNSF Ry. Co. v. Alere, Inc.*, 18-CV-291-BEN-WVG, 2018 WL 2267144, at *7 (S.D. Cal. May 17, 2018) (finding that "the only way for a nonparty to seek excusal from a subpoenaed deposition is to file a motion seeking to quash or modify the subpoena"); *Abbott v. Kidder, Peabody & Co.*, No. 97 C 3251, 1997 WL 337228, at *3 (N.D. Ill. June 16, 1997) (finding that written objections to subpoena did not excuse non-party from attending deposition absent a motion to quash or a protective order).

78. *MetroPCS*, 327 F.R.D. at 607 ("And 'a non-party's Rule 45(d)(2)(B) objections to discovery requests in a subpoena are subject to the same prohibition on general or boiler-plate [or unsupported] objections and requirements that the objections must be made with specificity and that the responding party must explain and support its objections.' *Am. Fed'n of Musicians*, 313 F.R.D. at 46 (citing *Heller v. City of Dallas*, 303 F.R.D. 466, 483 (N.D. Tex. 2014), and adopting 'the explanations in *Heller* of what is required to make proper objections and how to properly respond to discovery requests'."); *Orix USA Corp. v. Armentrout*, No. 3:16-mc-63, 2016 WL 4095603, at *2 (N.D. Tex. Aug. 1, 2016); *Am. Fed'n of Musicians*, 313 F.R.D. at 46 ("Rule 34(b)(1)'s reasonable particularity requirement should apply with no less force to a subpoena's document requests to a non-party," so too "a non-party's Rule 45(d)(2)(B) objections to those requests should be subject to the same requirements facing a party objecting to discovery under Rule 34."); *but see Ctr. for Individual Rights*, 2017 WL 5905191, at *4 ("In the Eleventh Circuit, objections should be plain enough and specific enough so that the court can understand in what way the [discovery sought is] alleged to be objectionable." (citations omitted)).

objection and the court to determine if the objection has merit.⁷⁹ These cases also note that the objection must specify the part of the request to which the objection pertains, the grounds for objecting, and whether any responsive materials are being withheld on the basis of that objection; and the non-party must produce documents or ESI responsive to the remainder of the request.⁸⁰ Failure to provide adequate specificity may result in a waiver of the objection.⁸¹ Thus, the addition of specificity language in Rule 34 but not in Rule 45 should not diminish the importance of the prior Rule 45 cases on this issue.

Non-parties may not be familiar enough with the details of the underlying litigation to object on any grounds other than undue burden or expense and privilege. The specificity that courts require related to Rule 34 objections must be reconciled

79. See generally *Federal Rule of Civil Procedure 34(b)(2) Primer: Practice Pointers for Responding to Discovery Requests*, *supra* note 8.

80. *MetroPCS*, 327 F.R.D. at 607 (For each item or category, the non-party must “state with specificity the grounds for objecting to the request, including the reasons, and must state whether any responsive materials are being withheld on the basis of that objection; that an objection to part of a request must specify the part and permit inspection of the rest; that ‘general or so-called boilerplate or unsupported objections are improper under Rule 45(d)(2)(B)’; and that the explanations in *Heller*, 303 F.R.D. at 466, of what is required to make proper objections and how to properly respond to discovery requests apply equally to non-parties subject to a Rule 45 subpoena.” (citing FED. R. CIV. P. 34(b)(2)(B)–(C); *Am. Fed’n of Musicians*, 313 F.R.D. at 46)).

81. *Sabol v. Brooks*, 469 F. Supp. 2d 324, 328 (D. Md. 2006) (finding that a non-party is “subject to the same obligations and scope of discovery under Rule 45 as if it were a party proceeding under Rule 34” and that a “failure to make particularized objections to document requests constitutes a waiver of those objections.” (citing FED. R. CIV. P. 45 advisory committee’s notes to 1991 amendment; *Hall v. Sullivan*, 231 F.R.D. 468 (D. Md. 2005); *Thompson v. HUD*, 199 F.R.D. 168 (D. Md. 2001); *Marens v. Carrabba’s Italian Grill*, 196 F.R.D. 35 (D. Md. 2000)).

with this limitation. Courts should not hold a non-party's lack of specificity as to proportionality factors of which it is not aware against the non-party. The non-party, however, should be specific as to its burdens and costs, should refrain from boilerplate objections, and should clearly identify what it is providing and what it is withholding on the basis of objections. The requesting party should cooperate with the non-party in an effort to narrow the scope of the request, if needed, to what the requestor truly needs to litigate its case.⁸² And if brought to the court's attention, the court should temper its expectation that a non-party comply with Rule 34 objection specificity standards and should tailor its assessment of the non-party's objections to the circumstances of the case.

As indicated earlier, non-parties should comply with the sections of the subpoena to which there are no objections. Conversely, complying with portions of a subpoena that have been objected to *before* a court has ordered compliance may, absent certain precautions,⁸³ lessen the likelihood that costs will

82. See discussion Section V.A., *supra* (violation of any Rule 26(g) duties by the requestor will be relevant to assessing the propriety of sanctions under Rule 45(d)(1)'s "undue burden" language).

83. *New Prods. Corp. v. Dickinson Wright, PLLC (In re Modern Plastics Corp.)*, 890 F.3d 244, 252–53 (6th Cir. 2018) (distinguishing *Angell v. Kelly*, 234 F.R.D. 135, 138 (M.D.N.C. 2006) (holding that expenses incurred before issuance of an order to compel are compensable where production itself did not precede such order, particularly where the requesting party is apprised of the non-party's intention to seek reimbursement and where the requesting party would, absent reimbursement, unfairly benefit from the non-party's efforts)); see also *In re Modern Plastics Corp.*, 577 B.R. 690, 707 (W.D. Mich. 2017) (affirming the lower court's order requiring requestor to pay the non-party's reasonable costs of compliance, "including costs that were incurred before the [] court ordered [the nonparty] to turn over the documents" and noting that the "rule does not distinguish compliance costs incurred prior to the court's order from costs incurred after the order. It might be argued that the term 'compliance' in 45(d)(2)(B)(ii) specifically refers to compliance with

shift to the requestor.⁸⁴ To minimize that risk, the non-party should notify the requesting party as early as possible that it intends to pursue reimbursement and should seek the requesting party's cooperation to limit expenses and avoid

the court's order, but this interpretation is inconsistent with the rest of Rule. When the term 'compliance' is used in other parts of Rule 45(d)(2), it always means compliance with the subpoena." (citations omitted)).

84. See *In re Aggrenox Antitrust Litig.*, No. 3:14-md-02516, 2017 WL 4679228, at *13–14 (D. Conn. Oct. 18, 2017); *Wellin v. Wellin*, No. 2:13-CV-1831, 2016 WL 7613663, at *8 (D.S.C. July 1, 2016) (recounting that "courts are reluctant to shift costs where the subpoenaed party has not provided the procuring party with sufficient notice of available cost information prior to incurring the expense to allow the procuring party an opportunity to re-evaluate its request and seek less costly alternatives"); *Sun Capital Partners, Inc. v. Twin City Fire Ins. Co.*, No. 12-CIV-81397, 2016 WL 1658765, at *5 (S.D. Fla. Apr. 26, 2016) ("The Non-Parties should have notified the Court . . . that production of the documents listed in the subpoena was becoming excessively burdensome and expensive to produce so that the Court could have worked with the parties and the Non-Parties on the front end of this discovery issue to try to minimize the costs incurred. . . . The Non-Parties' failure to notify the Court and Twin City of the significant expenses the Non-Parties were incurring prevented the Court from further protecting the Non-Parties from significant expense and prevented Twin City from further taking steps to try and reduce the expense. The Court will not allow the Non-Parties to sit back, fail to respond to the Court's Order, and then later assert they require reimbursement This is akin to sandbagging, which the Court will not permit."); *Spears v. First Am. eAppraiseIT*, 2014 WL 6901808, at *3 (N.D. Cal. December 8, 2014) (noting that "costs may be shifted under Rule 45(d)(2)(B)(ii) if the requesting party is on notice that the non-party will seek reimbursement of costs," but finding that the non-party did not provide clear notice to requestor that it would seek reimbursement of costs (of over \$450,000) until after production was underway or complete); *but see Miller v. Ghirardelli Chocolate Co.*, No. C 12-4936 LB, 2013 WL 6774072, at *5 (N.D. Cal. December 20, 2013) ("One good insight that a meet-and-confer process gives is how much it might cost to get the discovery, which in turn will guide Plaintiff's decision about what to ask for (knowing that costs can be shifted) and the court's inquiry about whether to shift costs. The court will not order that cost-shifting without a record.").

delays. The non-party should also outline for the requesting party its anticipated efforts and expenses so the requesting party can understand those efforts and, if appropriate, limit its discovery requests to reduce the burden. The requesting party should, in turn, engage with the non-party in these efforts to avoid unnecessary expenses that it may be required to pay.⁸⁵ It may be beneficial for the non-party and the requesting party to confer on these issues.

85. See *Modern Plastics Corp. v. Tibble*, No. 13-80252, 534 B.R. 723 (W.D. Mich. 2015) (“To accept New Products’ argument based on Rule 45(d)(2)—*i.e.*, that the [non-party] must now absorb all compliance costs incurred after they served their Objections and that [the requestor] is entitled to the documents at no charge—would reward gamesmanship and punish cooperation. The court cannot countenance such a windfall on this record, and will not construe Rule 45 in this way.”); *In re Modern Plastics Corp.*, 890 F.3d at 252–53 (distinguishing *Angell*, 234 F.R.D. at 138); see also *In re Modern Plastics Corp.*, 577 B.R. at 706–07 (“Appellants contend that after serving objections, Recipients were required to cease all efforts toward complying with the subpoena until ordered to comply by the court. Then, and only then, would Recipients be entitled to protection from significant expense. [This court sees] ‘no point in penalizing a cooperative [non-party] who gathers documents while reaching out to the requesting party in an effort to limit the expense and delay for all concerned.’ . . . Recipients repeatedly [voiced] their concerns with the subpoenas and [] their intent to seek reimbursement of the costs and expenses for compliance, but [requesting party] turned a deaf ear. Rather than work with Recipients to reduce the burden and expense of the subpoenas, or even inquire what those expenses might be, he encouraged them to continue working by extending the deadline for compliance. . . . Allowing Appellants to obtain the benefit of production without payment of Recipients’ reasonable fees and expenses would reward inaction by Appellants and is inconsistent with Appellants’ duty to take reasonable steps to avoid imposing an undue burden or expense on Recipients. Moreover, Appellants’ position would encourage non-compliance with subpoenas and unnecessary court intervention rather than communication, cooperation and expedient discovery. . . .”).

(b) Requesting Party Files Motion to Compel

The second condition before a non-party can seek reimbursement for costs is met when the requesting party files a motion to compel under Rule 45(d)(2)(B)(i).⁸⁶ Although a non-party must serve objections before the earlier of the time specified for compliance or 14 days after the subpoena is served,⁸⁷ there is no time limit under Rule 45(d)(2)(B)(i) for when a requesting party has to move to compel in response to non-party objections.

(c) Court Orders Compliance

A court order compelling the non-party to comply with the subpoena and produce the requested documents or ESI at a significant expense to the non-party satisfies the third condition before a non-party can seek reimbursement for costs under Rule 45(d)(2)(B)(ii)'s mandatory cost-shifting provision.⁸⁸ Without a motion to compel and a court order granting the motion, this mandatory cost-shifting mechanism is unavailable. Similarly, Rule 45(d)(2)(B)(ii) mandatory cost shifting may not apply where the party and non-party have entered into an agreement that governs reimbursement for subpoena compliance costs.⁸⁹

86. *Williams v. City of Dallas*, 178 F.R.D. 103, 113 (N.D. Tex. 1998) (relief provided by Rule 45(d)(2)(B) applies only when a motion to compel is filed in response to an objection to a subpoena).

87. *See* FED. R. CIV. P. 45(d)(2)(B).

88. *See* FED. R. CIV. P. 45(d)(2)(B)(i)–(ii).

89. *See* *FDIC v. LSI Appraisal LLC*, No. SACV 11-00706, 2014 WL 12561102, at *3 (C.D. Cal. July 21, 2014) (“[P]rivate agreements should be considered and honored by the courts. . . . *Legal Voice* does state that cost-shifting is ‘mandatory,’ but does not address whether the parties may alter the requirements of Rule 45 through agreement. This is not a situation in which the Court is exercising any discretion to decide whether fees are owed. Instead, the Court finds only that the parties entered into a separate binding agreement that addresses the substance of Rule 45(d)(2)(B)(ii)’s

2. Significant Expense and Cost Shifting⁹⁰

If all prerequisites above have been met and compliance will impose “significant expense” on the non-party, the court must order mandatory cost shifting. Courts consider several factors when determining if compliance has imposed significant expense on the non-party to warrant mandatory full or partial cost shifting. A non-party has the burden of presenting these factors (including its incurred or anticipated costs) to the court during the motion-to-compel briefing or as soon as it becomes evident to the non-party that compliance will result in “significant cost.”

3. When and How to Request Cost Shifting

In response to the requesting party’s motion to compel, the non-party should describe with particularity and provide a detailed affidavit or declaration describing its anticipated costs to comply with the subpoena,⁹¹ and should specifically request

requirements. Because their arrangement covers costs of subpoena compliance, Rule 45(d)(2)(B)(ii) is simply inapposite.” (citing *Angell v. Shawmut Bank Conn. Nat’l Ass’n*, 153 F.R.D. 585, 590 (M.D.N.C. 1994)).

90. Under Rule 45(d)(2)(B)(ii), when a court orders compliance with a subpoena over an objection, “the order must protect a person who is neither a party nor a party’s officer from significant expense resulting from compliance.” FED. R. CIV. P. 45(d)(2)(B)(ii).

91. *In re EpiPen (Epinephrine Injection, USP) Mktg., Sales Practices & Antitrust Litig.*, No. 17-md-2785, 2018 WL 3240981, at *4 (D. Kan. July 3, 2018) (“Express Scripts asks the Court to order Plaintiffs to pay the costs of compliance if the Court grants the motion to compel. Express Scripts has submitted an affidavit showing it has spent more than \$20,000 in legal fees and costs to serve objections, produce documents, negotiate and otherwise respond to the subpoena. Express Scripts also projects costs in the range of \$75,000 to \$250,000 to search for and produce additional documents. . . . Plaintiffs object that Express Script’s declaration is speculative, premature, and does not address the reasonableness of its projected costs. While the Court finds it appropriate for Class Plaintiffs to share in the cost of

that if the court orders production, it also should shift costs to the requesting party to the extent necessary to render costs insignificant. If, at that time, the non-party cannot detail its expenses or does not know whether the cost of compliance will be significant so as to trigger cost shifting under the rule, the non-party should notify the court and the requesting party as soon as it becomes apparent that continued compliance will necessitate a request for reimbursement. A response to a motion to compel or a subsequent cost shifting/reimbursement motion should “include a careful accounting of all expenses, how they ‘resulted from compliance,’ and an explanation as to their reasonableness;”⁹² and it should focus on whether the cost of compliance was “significant,” not on whether compliance was “unduly burdensome.”⁹³

production, such payment will of course be limited to Express Script’s actual and reasonable costs in producing documents pursuant to this order. Accordingly, the Court will require Class Plaintiffs to bear 50% of the reasonable costs Express Scripts incurs in timely producing documents responsive to the subpoena as ordered herein.”).

92. *United States v. McGraw-Hill Cos.*, 302 F.R.D. 532, 536 (C.D. Cal. 2014); *see also* *Bridgestone Ams., Inc. v. Int’l Bus. Machs. Corp.*, 2016 WL 11683327 (N.D. Ga. Aug. 5, 2016); *Hyundai Motor Am., Inc. v. Pinnacle Grp., LLC*, No. SAVC 14-0576-CJC, 2016 WL 6208313, at *1 (C.D. Cal. Apr. 20, 2016) (“[Non-party] [h]as made almost no factual showing in support of its request.”); *Callwave Commc’ns, LLC v. Wavemarket, Inc.*, No. C 14-80112 JSW (LB), 2014 WL 2918218, at *6 (N.D. Cal. June 26, 2014) (“The problem here is that Location Labs did not even tell the court how much it estimates it will spend to comply with the subpoena, let alone provide any evidence to support that amount. Without a specific dollar amount, the court cannot say whether Location Labs’ costs are significant.”).

93. *Legal Voice v. Stormans Inc.*, 738 F.3d 1178, 1185 (9th Cir. 2013).

(a) Factors to Consider

Rule 45(d)(2)(B)(ii) requires a two-step inquiry: (1) whether the costs are considered “expenses,” and, if so, (2) whether the expenses are “significant.”⁹⁴

(1) “Expense” Under the Rule

As previously stated, a non-party seeking compensation must demonstrate that the expense for which it seeks reimbursement is reasonable.⁹⁵ That determination is within the court’s discretion.⁹⁶ Courts have clarified that “an unreasonably incurred expense is not an expense ‘resulting from compliance.’”⁹⁷ Thus, “‘services provided by an attorney to a non-party for the non-party’s sole benefit and peace of mind’ [likely cannot] be counted as ‘expenses.’ . . . In other words, unnecessary or unduly expensive services do not ‘result from

94. *McGraw-Hill Cos.*, 302 F.R.D. at 536 (citing *Legal Voice*, 738 F.3d at 1184 (adopting the rule set out by *Linder v. Calero-Portocarrero*, 251 F.3d 178, 182 (D.C. Cir. 2001)).

95. *See In re Modern Plastics Corp.*, 577 B.R. 690, 707–08 (W.D. Mich. 2017) (“Rule 45(d)(2) does not expressly limit the compensable expenses to those that are reasonable, but courts have read it to do so.”); *Sands Harbor Marina Corp. v. Wells Fargo Ins. Servs. of Or., Inc.*, No. 09-CV-3855, 2018 WL 1701944, at *4 (E.D.N.Y. Mar. 31, 2018) (“A non-party who moves for costs and fees bears the burden of demonstrating that those costs and fees are reasonable” (internal citations omitted)).

96. *See In re Aggrenox Antitrust Litig.*, No. 3:14-MD-02516, 2017 WL 4679228, at *1–2 (D. Conn. Oct. 18, 2017).

97. *McGraw-Hill Cos.*, 302 F.R.D. at 536 (citing *Michael Wilson & Partners, Ltd. v. Sokol Holdings, Inc.* (*In re Michael Wilson & Partners, Ltd.*), 520 Fed. App’x. 736 (10th Cir. 2013)); *see also Steward Health Care Sys. LLC v. Blue Cross & Blue Shield of R.I.*, No. 15-272, 2016 WL 8716426, at *4–5 (E.D. Pa. Nov. 4, 2016).

compliance’ and, therefore, do not count as ‘expenses.’”⁹⁸ When opposing a motion to compel, a non-party should inform the court, prior to incurring any costs, the type of expenses it will undertake to comply with the subpoena and for which it will seek to shift costs. Once the court determines the types of expenses subject to cost shifting, the non-party can move forward with both an understanding of what expenses it may need to cover in full and the ability to determine the risk it is willing to take to forgo any such expenses.

In determining what counts as an “expense,” “[t]he touchstone is whether the expense ‘result[s] from compliance’ with the court’s order compelling production.”⁹⁹ Expenses allowed in the context of non-party subpoenas are broader than those allowed for party discovery.¹⁰⁰ Courts may allow

98. *McGraw-Hill Cos.*, 302 F.R.D. at 536 (citing *O’Cheskey v. Koehler (In re Am. Hous. Found.)*, No. 12-cv-00222, 2013 WL 2422706, at *3 (N.D. Tex. June 4, 2013)); see *Steward Health Care Sys. LLC*, 2016 WL 8716426, at *6 (non-party’s vendor costs deemed excessive and not resulting from compliance with subpoena where requesting party had offered a less expensive vendor that non-party failed to even contact and where non-party chose its vendor due to a relationship of trust that inured only to its own benefit, not to the benefit of requesting party); *United States v. Cardinal Growth, L.P.*, No. 11 C 4071, 2015 WL 850230, at *3 (N.D. Ill. Feb. 23, 2015) (non-party’s selected method of storing e-mails drove the need for an outside vendor, resulting in non-compensable overhead expenses); *In re Michael Wilson & Partners, Ltd.*, 520 Fed. App’x. at 741 (cutting shifted costs by fifty percent on the grounds that the non-parties ‘assume[d], rather than demonstrate[d], that all of their requested attorney’s fees are reasonable’); see also *In re Am. Hous. Found.*, 2013 WL 2422706, at *3 (expressing skepticism that “services provided by an attorney to a non-party for the non-party’s sole benefit and peace of mind” can be counted as “expenses”).

99. *McGraw-Hill Cos.*, 302 F.R.D. at 536 (quoting the text of Rule 45(d)(2)(B)(ii)); see also *Steward Health Care Sys. LLC*, 2016 WL 8716426, at *4–7.

100. Many courts have held that costs for responsiveness, privilege, and confidentiality review costs are non-compensable. See, e.g., *Lefta Assocs. v.*

non-party expenses to include printing costs and technology consulting fees¹⁰¹ as well as costs associated with collection, database creation, and, under certain circumstances, document review¹⁰² and privilege log preparation.¹⁰³ Given that electronic discovery is often the most costly part of compliance, it follows that courts consider these types of items as expenses so that, if significant, the cost to comply shifts to the requestor. Notably, attorneys' fees may count as costs resulting from compliance if incurred for "production-related legal tasks,"¹⁰⁴ but a court

Hurley, No. 1:09-CV-2487, 2011 WL 1793265, at *4 (M.D. Pa. May 11, 2011) (declining to award costs for conducting responsiveness and privilege review); *Cahoo v. SAS Inst. Inc.*, No. 17-10657, 2019 WL 4139152, at *4 (E.D. Mich. Aug. 30, 2019) ("It has been well recognized that a subpoenaed party cannot seek reimbursement for costs of privilege review."); *Sands Harbor Marina Corp.*, 2018 WL 1701944, at *6 (E.D.N.Y. Mar. 31, 2018) (declining to award costs for privilege review).

101. *Stormans Inc. v. Selecky*, No. C07-5374 RBL, 2015 WL 224914, at *5 (W.D. Wash. Jan. 15, 2015).

102. *G&E Real Estate, Inc. v. Avison Young-Washington, D.C., LLC*, 317 F.R.D. 313, 318–20 (D.D.C. 2016) (excluding costs related to extensive subpoena litigation because it believed a collaborative approach was more appropriate, but permitting costs related to document review where costs were based on adequately explained estimated hourly rates for reviewers (distinguishing *W. Convenience Stores, Inc. v. Suncor Energy (U.S.A.), Inc.*, No. 11-cv-01611, 2014 WL 1257762 at *24 (D. Colo. Mar. 27, 2014), where the task descriptions were vague)).

103. *Selecky*, 2015 WL 224914, at *5 ("There is no doubt that Rule 45 expenses resulting from compliance may include some attorneys' fees. Complying with a subpoena will almost always require some production-related legal tasks like document review, creating a privilege log, and drafting protective orders. Attorneys' fees for those production-related legal tasks are 'expenses resulting from compliance,' whether they are completed by in-house counsel or outside attorneys.").

104. *Id.*; see also *Steward Health Care Sys. LLC*, 2016 WL 8716426, at *4 ("[A] nonparty's legal fees, especially where the work benefits the requesting party, have been considered a cost of compliance and may be subject to reimbursement." (citations omitted)).

generally will exclude “attorneys’ fees for litigating a subpoena”¹⁰⁵ or, as with other costs, those that are unnecessary and incurred only for the benefit of the producing non-party.¹⁰⁶ At least one court, however, has held that even attorneys’ fees incurred in litigating fee disputes are compensable.¹⁰⁷

(2) “Significant” Under the Rule

Before the 1991 amendment to Rule 45, courts considering whether to shift costs could consider at least seven factors related to “significant expense.”¹⁰⁸ Some courts continued to

105. *Selecky*, 2015 WL 224914, at *5 (“It is a tenuous proposition, at best, that attorneys’ fees incurred resisting a subpoena are expenses resulting from compliance. . . . [Without] this interpretation . . . , when a party abuses its subpoena power or files frivolous or vexatious motions to compel, a non-party could contend that attorneys’ fees for litigating a subpoena are expenses resulting from compliance. But that situation is exactly what Rule 45(d)(1) is meant to address.”); see also *In re Aggrenox Antitrust Litig.*, No. 3:14-MD-02516, 2017 WL 4679228, at *9–10 (D. Conn. Oct. 18, 2017) (“[M]any of the expenses that [non-party] Gyma incurred . . . appear to have been related—directly or indirectly—to its efforts to resist the subpoena. Gyma is not entitled to recoup those costs. . . . Attributing [those] costs to the DPPs is particularly unwarranted because Gyma’s efforts to resist the subpoena were largely unsuccessful.”).

106. *Steward Health Care Sys. LLC*, 2016 WL 8716426, at *4–5 (a “tailored production,” which the non-party claimed it had put together in order to avoid a document dump, was deemed excessive and not compensable where: the requesting party had not asked for attorney review on relevance or other grounds; the non-party had conducted the review due to its own desire to check for privileged and confidential documents; and the review ultimately did not benefit the requesting party).

107. See *Linglong Americas Inc. v. Horizon Tire, Inc.*, No. 1:15CV1240, 2018 WL 1631341, at *3–4 (N.D. Ohio Apr. 4, 2018).

108. Those factors, as detailed in *United States v. McGraw-Hill Cos.*, 302 F.R.D. 532, 534 (C.D. Cal. 2014), are: (1) the non-party’s interest in the case; (2) the parties’ relative abilities to bear the costs; (3) the public importance of the litigation; (4) the scope of discovery; (5) the invasiveness of the request;

analyze cost shifting as though the original factors survived the 1991 amendment.¹⁰⁹ In *United States v. McGraw-Hill Cos.*, however, the U.S. District Court for the Central District of California, upon a thorough review of the relevant authority and statutory background, deemed the original factors “obsolete.”¹¹⁰ The Court found that many of the factors “d[id] not bear” on the question of whether the subpoena imposes significant expense on the non-party but instead were developed to guide the court’s exercise of discretion on whether cost shifting was appropriate, which the 1991 amendment eliminated.¹¹¹ It is the non-party’s obligation to adequately document the costs it seeks.

It is still within the court’s discretion, however, to determine which costs are “significant”¹¹²—“a term that readily lends itself to myriad interpretations depending on the circumstances of a particular case.”¹¹³ The main factors, derived from the seven equitable factors used in the pre-1991 analysis, by which courts

(6) the extent to which the producing party must conduct a privilege or responsiveness review; and (7) the reasonableness of the costs of production.

109. *Id.* at 534–35 (noting support for the notion that the original factors survived the amendment appears in *In re Exxon Valdez*, 142 F.R.D. 380 (D.D.C. 1992)).

110. 302 F.R.D. at 534–36 (citations omitted); *see also* *Cornell v. Columbus McKinnon Corp.*, No. 13-CV-02188-SI, 2015 WL 4747260, at *3 (N.D. Cal. Aug. 11, 2015) (describing *McGraw-Hill* analysis as “compelling”).

111. *McGraw-Hill Cos.*, 302 F.R.D. at 534–36.

112. *See* *Callwave Commc’ns, LLC v. Wavemarket, Inc.*, No. C 14-80112 JSW (LB), 2014 WL 2918218, at *3 (N.D. Cal. June 26, 2014) (citing *Sound Sec., Inc. v. Sonitrol Corp.*, No. CIV.3:08-CV-05350-RB, 2009 WL 1835653, at *1 (W.D. Wash. June 26, 2009)).

113. *McGraw-Hill Cos.*, 302 F.R.D. at 536; *see also* *Balfour Beatty Infrastructure, Inc. v. PB&A, Inc.*, 319 F.R.D. 277, 281 (N.D. Cal. 2017); *Cornell*, 2015 WL 4747260, at *4 (“[T]he weight of the case law makes clear that determining what constitutes a ‘significant cost’ is a relative, not an absolute, inquiry.”).

determine whether expenses are “significant” are: (a) whether the non-party actually has an interest in the outcome of the case; (b) whether the non-party can bear the costs; and, in some courts, (c) whether the underlying litigation is of public importance.¹¹⁴

a. Non-Party’s Interest in Outcome

Cost shifting is less appropriate where the non-party “was substantially involved in the underlying transaction,” could have anticipated that the “transaction would reasonably spawn some litigation,” and “[had] an interest in the outcome of the litigation.”¹¹⁵ Additionally, cost shifting is not appropriate where a non-party stands to recoup money from the underlying

114. See *Wellin v. Wellin*, No. 2:13-CV-1831, 2016 WL 7613663, at *11 (D.S.C. July 1, 2016); *Sun Capital Partners, Inc. v. Twin City Fire Ins. Co.*, No. 12-CIV-81397, 2016 WL 1658765, at *7 (S.D. Fla. Apr. 26, 2016); *United States v. Cardinal Growth, L.P.*, No. 11 C 4071, 2015 WL 850230, at *3 (N.D. Ill. Feb. 23, 2015); *Stormans Inc. v. Selecky*, No. C07-5374 RBL, 2015 WL 224914, at *6 (W.D. Wash. Jan. 15, 2015) (citing *Linder & Wells Fargo Bank, N.A. v. Konover*, 259 F.R.D. 206 (D. Conn. 2009)); *In re Application of Michael Wilson & Partners, Ltd.*, No. 06-CV-02575-MSK-KMT, 2012 WL 1901217, at *3 (D. Colo. May 24, 2012); *In re Exxon Valdez*, 142 F.R.D. at 383–84.

115. *W. Convenience Stores, Inc. v. Suncor Energy (U.S.A.) Inc.*, No. 11-CV-01611-MSK-CBS, 2014 WL 1257762, at *23 (D. Colo. Mar. 27, 2014) (citing *United States v. Blue Cross Blue Shield of Michigan*, No. 10-CV-14155, 2012 WL 4838987, at *3 (E.D. Mich. Oct. 11, 2012)).

judgment¹¹⁶ or, more broadly, where it is intimately involved with or has already financially benefited from a party.¹¹⁷

116. See, e.g., *Cornell*, 2015 WL 4747260, at *3, *5 (“FedEx, as plaintiff’s employer at the time of the accident, has filed a lien against any judgment or settlement in plaintiff’s favor in order to recoup worker’s compensation benefits it has paid to him. . . . While slightly more attenuated than its direct financial interest, the outcome of this case could also affect FedEx’s employee training, safety policies, and future exposure to liability. . . . The Rule . . . was not intended as a mechanism for entities which stand to benefit from certain litigation outcomes to evade discovery costs arising from their involvement in the underlying acts that gave rise to the lawsuit.”).

117. See *Balfour Beatty Infrastructure, Inc.*, 319 F.R.D. at 282 (“[Non-party] URS perhaps is not in the typical position of a completely uninterested nonparty, as it was purportedly involved in the underlying acts that gave rise to the lawsuit.” (citation omitted)); *Ala. Aircraft Indus. v. Boeing Co.*, No. 2:11-cv-03577-RDP, 2016 WL 6892113, at *5 (N.D. Ala. Oct. 17, 2016) (“TCP is an interested non-party, as it has ‘a significant, underlying connection to the case,’ namely its intimate involvement in AAI’s affairs during the periods relevant to the instant case. . . . [W]hile TCP is successful in showing that it lacks a financial or reputational stake in this case’s outcome, it cannot explain away its significant connection to the underlying events. . . . [I]t also exercised a certain level of influence over AAI’s decisions and actions due to its control of numerous seats on AAI’s Board of Directors.”); *Wellin*, 2016 WL 7613663, at *12 (non-party trust beneficiaries had an interest in the outcome of the case where they were aligned in interest with their parents, who were parties to the litigation); *Sun Capital Partners, Inc.*, 2016 WL 1658765, at *5 (non-parties were either co-defendants in separate, underlying litigation or were created to facilitate settlement of that litigation, coordinated with party counsel in motion practice, and stood to be reimbursed by a party); *Hyundai Motor Am., Inc. v. Pinnacle Grp., LL*, No. SAVC 14-0576-CJC, 2016 WL 6208313, at *1 (C.D. Cal. Apr. 20, 2016) (“[Non-party] Mobis . . . is affiliated with Plaintiff Hyundai—they share the same parent company—and serves as its parts distributor. As such, it is a competitor of Defendant’s and has a strong interest in the outcome of this litigation.”); *Am. Fed’n of Musicians of the U.S. & Can. v. Skodam Films, LLC*, 313 F.R.D. 39, 58 (N.D. Tex. 2015) (non-party film producer’s “level of involvement with the production—including the scoring—of the Movie at issue in the underlying Litigation” deemed a sufficient “interest in the case” to weigh against an award of costs);

b. Non-Party's Ability to Bear the Cost

When determining whether a subpoena imposes a significant expense, courts also consider a non-party's "financial ability to bear the costs of production."¹¹⁸ Specifically, courts consider whether the non-party can "more readily bear the costs than the requesting party."¹¹⁹ When assessing the non-party's financial means, the court should note that burden is relative and fact-specific.¹²⁰ Some of the factors the court may consider

Cardinal Growth, L.P., 2015 WL 850230, at *3 ("P&H served as Cardinal's counsel [and] derived substantial income from Cardinal, drafted and prepared hundreds of transactional documents, and participated in the design of numerous complex transactions. Thus, P&H 'is not a classic disinterested non-party.'").

118. *Balfour Betty Infrastructure, Inc.*, 319 F.R.D. at 281; *Cedar Rapids Lodge & Suites, LLC v. Seibert*, No. 0:14-cv-04839, 2018 WL 3019899, at *2 (D. Minn. June 18, 2018).

119. *Koopmann v. Robert Bosch LLC*, No. 18-CV-4065, 2018 WL 9917679, at *1 (S.D.N.Y. May 25, 2018) (citing *Sands Harbor Marina Corp. v. Wells Fargo Ins. Servs. of Or., Inc.*, No. 09-CV-3855, 2018 WL 1701944, at *3 (E.D.N.Y. Mar. 31, 2018) (internal quotation marks & brackets omitted)).

120. *See United States v. McGraw-Hill Cos.*, 302 F.R.D. 532, 536 (C.D. Cal 2014) ("This consideration makes practical sense—an expense might be 'significant,' for instance, to a small family-run business, while being 'insignificant' to a global financial institution." (citing *Linder v. Calero-Portocarrero*, 251 F.3d 178, 182 (D.C. Cir. 2001))); *Ala. Aircraft Indus.*, 2016WL 6892113, at *6 ("[G]iven TCP's admitted ability to pay for its production costs, its apparent reluctance to provide a workable sense of its financial condition, and its established interest in the case, the undersigned is comfortable concluding that this element weighs against shifting costs."); *Cornell*, 2015 WL 4747260, at *4 ("FedEx is correct to argue that this factor is not dispositive in every instance. However, in this particular case, the discovery costs are dwarfed by FedEx's profit figures, and therefore weigh in favor of finding them insignificant."); *Cardinal Growth, L.P.*, 2015 WL 850230, at *3 ("Relative to the substantial income that P&H collected from Cardinal, the expenses incurred by P&H in complying with the Court's order do not constitute a 'significant expense.'"); *Seibert*, 2018 WL 3019899, at *2 ("[Non-party] did not present any argument or proof to demonstrate that he cannot bear the costs

include the (i) cost of compliance as a percentage of the non-party's total value/yearly revenue; (ii) cost of compliance as a percentage of the total that a party has contributed to the non-party in a business relationship; and (iii) size of the non-party company.¹²¹ In some circumstances, the court may also take into account the financial status of the requesting party.¹²²

These factors protect individuals and smaller companies, who may have a more limited ability to bear the cost of compliance when facing significant expense. This is the intent of Rule 45(d)(2)(B)(ii), which seeks to “protect a person who is neither a party nor a party’s officer from significant expense resulting from compliance.”¹²³

of production. Given that he bills his time at \$450 per hour, it seems unlikely that the costs associated with three hours of gathering and reviewing documents for production really amounts to a significant expense.”).

121. *Shasta Linen Supply, Inc. v. Applied Underwriters Inc.*, No. 2:16-CV-00158, 2018 WL 2981827, at *5 (E.D. Cal. June 14, 2018) (finding that \$15,000 is not a “significant” cost such that fee shifting is appropriate because the non-party did not provide the court with information regarding its gross revenues or “indicating that \$15,000 is significant with respect to its total value as a company,” and noting that because the non-party was a “national company with multiple offices,” it had “the financial ability to bear the costs of production”); *Stormans Inc. v. Selecky*, No. C07-5374 RBL, 2015 WL 224914, at *7 (W.D. Wash. Jan. 15, 2015) (stating that because a non-party nonprofit received over \$700,000 in contributions in one year, it was capable of paying some of its own expenses).

122. *See, e.g., Pitts v. Davis*, No. 212CV0823TLNACP, 2015 WL 6689856, at *5 (E.D. Cal. Oct. 30, 2015) (“Rule 45 does not preclude post-compliance reimbursement of costs. . . . Since plaintiff is proceeding in forma pauperis . . . and has made no indication that he is capable of covering such costs, the motion to compel will be denied to the extent it seeks further production of non-staff complaint grievances.”).

123. *See* FED. R. CIV. P. 45(d)(2)(B)(ii).

Large or very profitable organizations may deem the ability of the non-party to bear the cost as unjust, particularly if they regularly receive a large volume of subpoenas. Entities such as large national banks or car rental companies may receive thousands of non-party subpoenas a year pertaining to customers involved in litigation. Even a \$500-per-subpoena cost could result in substantial aggregate costs. In most of these matters, the non-party is a custodial non-party, i.e., the non-party corporation is a repository of customer information and does not have an interest in the litigation. The language of Rule 45, however, does not support cost shifting unless the cost of responding to the specific subpoena over objections is significant. Entities that face a large volume of low-cost subpoenas for their customers cannot usually claim significant costs, so they should anticipate and look for other ways to defray these costs. One way is to contractually obligate customers to be liable for costs related to non-party subpoenas of records if customers become parties to a litigation. Another option is to negotiate the cost issue with the requesting party prior to production. Here, too, it could be beneficial for the non-party and requesting party to confer to keep costs down and reduce the burden on the non-party, no matter its size or ability to pay.

c. Public Importance

When determining whether a subpoena imposes a significant expense, some courts also consider the underlying litigation's public importance. These courts have noted that this factor "is very much in the eye of the beholder"¹²⁴ and can turn, in part, on the nature of the parties themselves and the functions

124. *W. Convenience Stores, Inc. v. Suncor Energy (U.S.A.) Inc.*, No. 11-CV-01611-MSK-CBS, 2014 WL 1257762, at *24 (D. Colo. Mar. 27, 2014).

they perform.¹²⁵ Other courts, however, have refused to consider this factor, noting that a “non-party’s expenses are not made less significant by the fact that the litigation is important to the general public.”¹²⁶

(b) Allocation of Costs

Even if the court determines that a non-party bears “significant expense” in complying with a subpoena, “this does not mean that the requesting party must necessarily bear the *entire* cost of compliance.”¹²⁷ A non-party can be required to bear some or all of its expenses “where the equities of the particular case demand it.”¹²⁸ Courts also are not inclined to award cost

125. See, e.g., *Cardinal Growth, L.P.*, 2015 WL 850230, at *3 (“[Plaintiff,] [t]he SBA[,] is a public agency that regulates the operations of publically [sic] financed SBICs. . . . [T]he SBA has a duty to responsibly liquidate Cardinal’s assets, pay its creditors, and preserve its claims in furtherance of the public interest. To properly execute those duties, the SBA needed documents that were in the possession of P&H. Under these circumstances, the SBA should not have to bear the cost of production.”); *Selecky*, 2015 WL 224914, at *7 (noting that there was “no doubt” that the underlying lawsuit regarding a challenge to Washington’s State Board of Pharmacy regulations that compelled pharmacies and pharmacists to dispense lawfully prescribed emergency contraceptives over sincere religious objections was of “great public importance”).

126. *Cornell v. Columbus McKinnon Corp.*, No. 13-CV-02188-SI, 2015 WL 4747260, at *2 (N.D. Cal. Aug. 11, 2015) (citing *United States v. McGraw-Hill Cos.*, 302 F.R.D. 532, 534 (C.D. Cal. 2014)).

127. *Callwave Commc’ns, LLC v. Wavemarket, Inc.*, No. C 14-80112 JSW (LB), 2014 WL 2918218, at *3 (N.D. Cal. June 26, 2014) (citing *Legal Voice v. Stormans Inc.*, 738 F.3d 1178, 1185 (9th Cir. 2013)).

128. *Legal Voice*, 738 F.3d at 1184 (“[If] the subpoena imposes significant expense on the non-party . . . the district court must order the party seeking discovery to bear at least enough of the cost of compliance to render the remainder ‘non-significant.’” (quoting *Linder v. Calero-Portocarrero*, 251 F.3d 178, 182 (D.C. Cir. 2001))); *Wellin v. Wellin*, No. 2:13-CV-1831, 2016 WL 7613663, at *8 (D.S.C. July 1, 2016); *Sound Sec., Inc. v. Sonitrol Corp.*, No.

shifting to a non-party that has engaged in needlessly litigious, obstructionist behavior.¹²⁹

Rule 45(d)(2)(B)(ii) mandates cost shifting sufficient to render the non-party's subpoena expenses non-significant. As the D.C. Circuit noted in *Linder v. Calero-Portocarrero*:

Under the revised Rule 45, the questions before the district court are whether the subpoena imposes expenses on the non-party, and whether those expenses are "significant." If they are, the court must protect the non-party by requiring the party seeking discovery to bear at least enough of the expense to render the remainder "non-significant."¹³⁰

CIV.3:08-CV-05350-RB, 2009 WL 1835653, at *2 (W.D. Wash. June 26, 2009) ("[A] non-party can be required to bear some or all of its expenses where the equities of a particular case demand it." (quoting *In re Exxon Valdez*, 142 F.R.D. 380, 383 (D.D.C. 1992))); *In re EpiPen* (Epinephrine Injection, USP) Mktg., Sales Practices & Antitrust Litig., No. 17-md-2785, 2018 WL 3240981, at *3 (D. Kan. July 3, 2018) ("If Plaintiffs maintain their interest in these documents to the extent they are willing to pay a share of the actual reasonable costs Express Scripts incurs in producing them [approximately \$75,000], the Court orders Express Scripts to search for and produce the documents according to Plaintiffs' proposal.").

129. See *In re Aggrenox Antitrust Litig.*, No. 3:14-MD-02516, 2017 WL 4679228, at *11 (D. Conn. Oct. 18, 2017) ("Gyma was notably intransigent and dilatory in its response to the subpoena, taking a full year and necessitating three interventions by the court to complete a review of 5,545 pages of documents. Gyma also appears to have repeatedly exaggerated its costs, claiming in its latest motion that it spent nearly \$20 per page in document review. . . . Considering the Second Circuit's admonition that courts 'not endors[e] scorched earth tactics' or 'hardball litigation strategy,' . . . Gyma should bear the . . . balance of its costs." (citations omitted)).

130. *Linder*, 251 F.3d at 182 ("The rule is susceptible of no other interpretation."); see also *McGraw-Hill Cos.*, 302 F.R.D. at 534; *CallWave Commc'ns, LLC*, 2014 WL 2918218, at *3; *Legal Voice*, 738 F.3d at 1184 ("[O]nly

After considering the factors and determining that the non-party will bear “significant expense” resulting from compliance,¹³¹ a court will then allocate responsibility for these expenses between the non-party and the requesting party to ensure that the costs incurred by the non-party are *non-significant*.¹³²

If the non-party has served objections, the requesting party could consider offering to pay most or all of the non-party’s compliance costs up front to expedite production and avoid motion practice. This approach limits the ability of the non-party to argue “significant expense” and delay compliance.

two considerations are relevant under the rule: ‘[1] whether the subpoena imposes expenses on the non-party, and [2] whether those expenses are significant.’ . . . The plain language of the rule dictates our conclusion. . . . [If] the subpoena imposes significant expense on the non-party . . . the district court must order the party seeking discovery to bear at least enough of the cost of compliance to render the remainder ‘non-significant.’” (quoting *Linder*, 251 F.3d at 182)).

131. See, e.g., *Legal Voice*, 738 F.3d at 1185 (\$20,000 in expenses held “significant”); *Linder*, 251 F.3d at 182 (finding \$9,000 in costs “significant”); *G&E Real Estate, Inc. v. Avison Young-Washington, D.C., LLC*, 317 F.R.D. 313, 320–21 (D.D.C. 2016) (holding \$3,148.44 in expenses “significant”); see also *Broussard v. Lemons*, 186 F.R.D. 396, 398 (W.D. La. 1999) (finding that \$43 to copy and mail 11 sheets of paper was a “significant” expense).

132. *In re Aggrenox*, 2017 WL 4679228, at *11 (“Rule 45 only protects non-parties from ‘significant expense resulting from compliance,’ *Legal Voice*, 738 F.3d at 1184 (quoting FED. R. CIV. P. 45(d)(2)(B)(ii)), and ‘[a] non-party may be required to absorb a non-significant portion of its expenses,’ particularly ‘where the equities of the particular case demand it.’”); see *In re Subpoena of American Nurses Ass’n*, 924 F. Supp. 2d 607 (D. Md. 2013) (“Although the . . . Plaintiffs shall bear the majority of the costs of production, there are some costs the ANA should absorb.”); *Koopmann v. Robert Bosch LLC*, No. 18-CV-4065, 2018 WL 9917679, at *1 (S.D.N.Y. May 25, 2018) (deeming it reasonable that the requestor bear half the cost of compliance, up to a maximum of \$30,000).

C. *Rule 45(d)(3)(A)—Quashing or Modifying a Subpoena*

Rule 45 also provides that a court may quash or modify a subpoena under certain circumstances. Specifically, Rule 45(d)(3)(A) states:

When Required. On timely motion, the court for the district where compliance is required must quash or modify a subpoena that:

- i. fails to allow a reasonable time to comply;
- ii. requires a person to comply beyond the geographical limits specified in Rule 45(c);
- iii. requires disclosure of privileged or other protected matter, if no exception or waiver applies; or
- iv. subjects a person to undue burden.

This provision is noteworthy as it provides the court the procedural authority to alter the scope of—or quash altogether—the requesting party’s subpoena. For example, courts have used this provision to prevent a foreign witness from being required to appear for a deposition;¹³³ permit a non-party to withhold clearly privileged documents called for by a subpoena;¹³⁴ and revise a subpoena to provide the responding non-party enough time to produce documents and seek appropriate protection for sensitive materials.¹³⁵

133. See, e.g., *In re Donald Edwin May*, 2014 WL 12923988 (Bankr. N.D. Ind. Jul. 9, 2014) (quashing a subpoena that would have required a deponent to travel outside the state and more than 100 miles from where he resided, was employed, and regularly transacted business in person).

134. *Cones v. Parexel Int’l Corp.*, No.: 16cv3084, 2018 WL 3046424, at *1–2 & n.2 (S.D. Cal. Jun. 20, 2018).

135. *Verisign, Inc. v. XYZ.com*, No. 15-mc-175, 2015 WL 7960976, at *3 (D. Del. Dec. 4, 2015) (noting that several courts have found 14 days to be

The provision that has created the greatest source of conflict with other parts of Rule 45 is the authority to quash or modify a subpoena where it subjects a non-party to undue burden. Even before the inception of this provision, courts have attempted to build a framework to guide litigants in their analysis of whether information requested by a Rule 45 subpoena constitutes an undue burden for the non-party. However, the rule (and resulting case law) lacks clarity for how the analysis of undue burden under this section is related to or impacted by other “undue burden” provisions in the rule—particularly those that afford non-parties the right to seek costs associated with their burden.

When confronted with the question of “undue burden,” the 2008 edition of this *Commentary* noted that “[o]nly a few reported cases address the acquisition of ESI from non-parties.”¹³⁶ In the decade since, the volume of subpoenas seeking ESI production has skyrocketed. As a result, courts have continued to refine the contours of what imposes an undue burden on a responding non-party.

Courts routinely note that the movant bears the responsibility of establishing that a subpoena imposes an undue burden.¹³⁷ Although the court has discretion to determine whether a subpoena’s request constitutes an undue burden on the non-party, it is tasked with weighing the requesting party’s

presumptively reasonable and that others have found that seven days is “clearly unreasonable” (citations omitted)).

136. The Sedona Conference, *Commentary on Non-Party Production and Rule 45 Subpoenas*, 9 SEDONA CONF. J. 197 (2008).

137. *Stokes v. Cenveo Corp.*, No. 2:16cv886, 2017 WL 3648327, at *2 (W.D. Penn. Aug. 24, 2017) (“[T]he burden of establishing that a subpoena *duces tecum* imposes an undue burden is on the party moving to quash the subpoena. This burden is a heavy one. . . . A successful demonstration of undue burden requires more than ‘generalized and unsupported allegations.’” (citation omitted)).

need for the requested documents against the hardship imposed on the non-party. In making these determinations, courts have relied on a case-specific balancing test that typically includes some combination of the following six factors:

1. the relevance of the information requested;
2. the requesting party's need for the documents or ESI;
3. the breadth of the document or ESI request;
4. the time period covered by the request;
5. the particularity with which the requesting party describes the requested documents or ESI;
6. the burden imposed upon the responding non-party.¹³⁸

While case law dealing with obtaining ESI from non-parties has increased, the courts' concerns about a subpoena's burden placed on a non-party have largely remained the same. In particular, courts emphasize that non-parties should not be required to subsidize litigation in which they have no stake in the outcome.¹³⁹

138. *Wiwa v. Royal Dutch Petroleum Co.*, 392 F.3d 812, 818 (5th Cir. 2004); *see also* *New Prods. Corp. v. Dickinson Wright, PLLC (In re Modern Plastics Corp.)*, 890 F.3d 244, 251 (6th Cir. 2018) (quoting *Am. Elec. Power Co., Inc. v. United States*, 191 F.R.D. 132, 136 (S.D. Ohio 1999) (citation omitted)); *Koch v. Pechota*, No. 10 Civ. 9152, 2012 WL 4876784, at *3 (S.D.N.Y. Oct. 12, 2012) (quoting *Night Hawk Ltd. v. Briarpatch Ltd., L.P.*, No. 03 CIV.1382, 2003 WL 23018833, at *8 (S.D.N.Y. Dec. 23, 2003)); *Call of the Wild Movie, LLC v. Does 1-1,062*, 770 F. Supp. 2d 332, 354 (D.D.C. 2011). The same factors are also applied to subpoenas for testimony. *See, e.g.*, *Black Knight Fin. Servs. v. Powell*, No. 3:14-mc-42, 2014 WL 10742619, at *2 (M.D. Fla. Dec. 11, 2014).

139. *Cusumano v. Microsoft Corp.*, 162 F.3d 708, 717 (1st Cir. 1998) ("Although discovery is by definition invasive, parties to a law suit must accept its travails as a natural concomitant of modern civil litigation. Non-

With these factors in mind, courts have expanded their analysis of undue burden as they examine the growing universe of ESI. For instance, courts have granted motions to quash subpoenas demanding the forensic imaging of a non-party's cell phone where a party failed to narrowly tailor the request to ESI relevant to the matter.¹⁴⁰ Courts also have used the undue-burden framework to quash subpoenas seeking ESI from a non-party where the volume and scope of the requested ESI is unlikely to lead to the discovery of relevant information.¹⁴¹

Given the rule's language requiring a court to quash or modify a subpoena that imposes an undue burden on a non-party recipient, there is some tension between this provision's protections to the non-party and a party's need for discovery necessary to afford a fair opportunity to develop and prepare its case. Mindful of this tension, courts tend to take a pragmatic, measured approach to motions to quash. Several courts have noted that "[m]odification of a subpoena is

parties have a different set of expectations."); *see also* *Butler v. Christian Island Food Serv.*, No. 4:15-CV-1118, 2016 WL 11683326 (E.D. Mo. May 9, 2016) ("[C]oncern for the unwanted burden thrust upon non-parties is a factor entitled to special weight in evaluating the balance of competing needs." (quoting *Cusumano*, 162 F.3d at 717)); *Pugh v. Junqing*, No. 4:16-CV-1881, 2018 WL 10733633 (E.D. Mo. Mar. 29, 2018) ("Where, as here, discovery is sought from a non-party, courts have wide latitude in deciding motions regarding non-party subpoenas, and courts are directed to give special consideration in assessing whether the subpoena subjects a non-party to annoyance or an undue burden or expense." (citation omitted)).

140. *Charles Schwab & Co. v. Highwater Wealth Mgmt., LLC*, No. 17-cv-00803, 2017 WL 4278494, at *6-7 (D. Colo. Sep. 27, 2017).

141. *Hock Foods, Inc. v. William Blair & Co.*, No. 09-2588-KHV, 2011 WL 884446, at *7 (D. Kan. Mar. 11, 2011) (denying the portion of a motion to compel that would have required a non-party to search through 12,786 boxes of hard copy data and 12 terabytes of ESI to find "a needle in a haystack—an irrelevant needle").

generally preferable to quashing it outright.”¹⁴² Frequently, courts achieve this goal by ordering production of a narrowed set of the requested documents or by imposing other limiting factors on the subpoena’s requests.¹⁴³

Of course, not every request to a non-party to produce ESI constitutes an undue burden on the non-party. For instance, courts have held that there is no undue burden where the requesting party agrees to cover the expenses of the responding non-party.¹⁴⁴ Courts also have found no undue burden where the burden is of the non-party’s own making.¹⁴⁵

142. *Andra Grp., LP v. JDA Software Grp., Inc.*, 312 F.R.D. 444, 449 (N.D. Tex. 2015) (quoting *Witwa*, 392 F.3d at 818); *see also* *Fernandez v. Cal. Dep’t of Corr. & Rehab.*, No. 2:11-cv-01125, 2014 WL 794332, at *2 (E.D. Cal. Feb. 27, 2014) (“Quashing subpoenas goes against the court’s general preference for a broad scope of discovery, [but] limiting discovery is appropriate when the burden of providing the documents outweighs the need for it.”).

143. *Sams v. GA West Gate, LLC*, 316 F.R.D. 693, 697 (N.D. Ga. 2016) (modifying a subpoena to “provide for an initial production” of ESI and permitting supplemental production “if, and only if, the electronic documents point to additional, relevant documents”).

144. *In re Domestic Drywall Antitrust Litig.*, 300 F.R.D. 234, 252 (E.D. Pa. 2014) (“Although it is true that compliance with the subpoena will require [the non-party] to review and redact numerous reports and investigative files, this burden is not undue because Plaintiffs will compensate [the non-party].”); *see also* *Wood v. Town of Warsaw*, No. 7:10-CV-00219, 2011 WL 6748797, at *3 (E.D.N.C. Dec. 22, 2011) (rejecting a non-party’s undue burden objection where, *inter alia*, the plaintiff “agreed to pay the cost of a forensic expert to copy and search [the non-party’s] hard drive”); FED. R. CIV. P. 45(d)(3)(C) (“[T]he court may, instead of quashing or modifying a subpoena, order appearance or production under specified conditions if the serving party shows a substantial need for the . . . material that cannot be otherwise met without undue hardship; and ensures that the subpoenaed person will be reasonably compensated.”).

145. *See* *W. Convenience Stores, Inc. v. Suncor Energy (U.S.A.) Inc.*, No. 11-CV-01611-MSK-CBS, 2014 WL 1257762, at *25 (D. Colo. Mar. 27, 2014) (“I understand that the court should be especially vigilant to protect the non-

Although courts have continued to refine jurisprudence under this provision, the fact-specific nature of whether a request constitutes an undue burden on a subpoenaed non-party will continue to change with advances in technology and new ways of creating and retaining ESI.

Perhaps the most significant outstanding questions concern how this provision should operate in the context of other Rule 45 provisions dealing with undue burden. Arguably, Rule 45(d)(3)(A)(iv) can be best understood as a purely procedural mechanism that affords courts the ability to quash or modify a subpoena when other provisions of Rule 45 are not an option—such as Rule 45(e)(1)(D) (which outlines the undue burden of inaccessible ESI) and Rule 45(d) (which puts requesting parties on notice that they may be subject to cost shifting if they request documents or ESI that may impose an undue burden).

party from undue burden and expense. However, this principle should not be invoked to excuse the non-party's own evasive or obstructive conduct. It strains logic to suggest that the court should hold a party or attorney issuing a subpoena to a standard of reasonableness, but then turn a blind eye to a non-party's unreasonable behavior Counsel for a non-party subpoena recipient, however, should be expected to 'stop and think' before taking actions that will almost certainly result in unnecessary delay and burden an already congested court docket. Rule 45(d)(1) correctly focuses on the burdens imposed upon the subpoena recipient. However, Rule 45(d)(1) should not be construed or applied in a way that ignores the subpoena recipient's own conduct or confers a right to obfuscation or obstinacy." (internal quotations & citations omitted); *Morgan Hill Concerned Parents Ass'n v. California Dep't of Educ.*, No. 2:11-CV-3471, 2017 WL 445722, at *7 (E.D. Cal. Feb. 2, 2017), *reconsideration denied*, No. 2:11-CV-03471, 2017 WL 1382483 (E.D. Cal. Apr. 18, 2017) (rejecting an undue burden argument where the party making the argument "created the problem it now complains about").

D. Rule 45(e)(1)(D)—Non-Party ESI That Is Not Reasonably Accessible Due to Undue Burden or Cost

Rule 45(e)(1)(D) protects non-parties from the production of ESI that is not reasonably accessible due to undue burden or cost. It provides as follows:

Inaccessible Electronically Stored Information. The person responding need not provide discovery of electronically stored information from sources that the person identifies as not reasonably accessible because of undue burden or cost. On motion to compel discovery or for a protective order, the person responding must show that the information is not reasonably accessible because of undue burden or cost. If that showing is made, the court may nonetheless order discovery from such sources if the requesting party shows good cause, considering the limitations of Rule 26(b)(2)(C). The court may specify conditions for the discovery.¹⁴⁶

The not-reasonably-accessible issue is the same under Rule 34 and Rule 45.

146. FED. R. CIV. P. 45(e)(1)(D). Because this Rule is materially the same as Federal Rule of Civil Procedure 26(b)(2)(B), case law interpreting that Rule may be applicable to Rule 45(e)(1)(D).

VI. THE SEDONA CONFERENCE RULE 45 PRACTICE POINTERS

- Practice Pointer 1.** Timely disclosure by the parties is helpful to prevent over reliance on Rule 45 subpoenas. Early in litigation (possibly during the Rule 26(f) conference), each party should endeavor to identify relevant documents and ESI held by non-parties.
- Practice Pointer 2.** When the parties confer about discovery, they should work to reach stipulations concerning authenticity and admissibility to avoid the need to subpoena a non-party custodian to prove up documents or ESI.
- Practice Pointer 3.** Prior to issuing a subpoena to a non-party, it may be beneficial for a party to confirm that the information cannot be obtained through discovery from a party. The party issuing a subpoena generally should avoid seeking information from a non-party that likely is duplicative of information in a party's possession, custody, or control. If the non-party has possession or custody of ESI but a party retains control, the *Commentary* recommends that the information should be obtained from the party under Rule 34, not from the non-party under Rule 45.
- Practice Pointer 4.** It may be beneficial, before service on the non-party, for the party issuing the subpoena to give the other parties time to raise relevance, proportionality, confidentiality, and privilege concerns.
- Practice Pointer 5.** If the party issuing a subpoena does not have sufficient information to tailor the subpoena, the party should seek to confer with the non-party promptly after issuance of the subpoena, or possibly before

issuance, in order to properly tailor the scope of the subpoena and to reduce the burden and expense on the non-party. The subpoena recipient should meet and confer in good faith with the issuing party to explain any objections it may have and work collaboratively to resolve them without need for court intervention. The parties should be mindful of local rules that may require parties and non-parties alike to meet and confer before bringing motions.

Practice Pointer 6. The party issuing a subpoena should be mindful of its obligations under Rule 45(d)(1) and Rule 26(g) to avoid imposing undue burden and expense on a non-party subpoena recipient. Accordingly, subpoenas should include, as applicable, limitations regarding time periods, individuals involved, and scope. The party issuing a subpoena should consider and incorporate the concept of proportionality. That would include whether the information sought is proportional to the needs of the case, including whether the burden and cost of preserving or producing such information outweigh the potential value and uniqueness of the information.

Practice Pointer 7. If not clear, a subpoena should seek to explain the non-party's relationship to the lawsuit or a party, in order to provide context to the non-party recipient and facilitate identifying responsive information. The party issuing the subpoena should consider enclosing a copy of the complaint and the answer to assist the non-party.

Practice Pointer 8. It may be beneficial for the parties to ensure that the protective order and Fed. R. Evid. 502(d) order in place protect the non-party. The party issuing a subpoena should include a copy of any protective order

and Fed. R. Evid. 502(d) order that were entered in the action.

Practice Pointer 9. The party issuing a subpoena should specify a form of production and, if applicable, attach any ESI order addressing the form of production that may have been entered in the action if the issuing party seeks the non-party's compliance with that format. If the subpoena specifies a form of production, the non-party subpoena recipient can object to the requested form of production and specify a different form of production. The non-party subpoena recipient usually will want to specify a form of production, regardless of whether the subpoena specifies one. It may be beneficial for the requesting party to consider agreeing to an alternative production format or to pay some or all of such additional cost and expense necessary to comply with the requested format that is less costly or seek other solutions that reduce the costs of compliance. For example, a party may wish to consider offering less costly means of processing and production to reduce the non-party's processing and production costs.

Practice Pointer 10. It may be beneficial for a non-party recipient to initiate discussions with the issuing party soon after receiving a subpoena (or vice versa), due to the relatively short period for serving objections and responses under Rule 45(d)(2)(B).

Practice Pointer 11. Whenever feasible, the party issuing a subpoena and the non-party recipient should agree to a reasonable extension of time for the non-party to serve objections and to respond to the subpoena. Meaningful dialogue regarding issues concerning the subpoena is more likely to occur when an extension has been

provided, and the dialogue may reduce or eliminate the need for objections and subsequent unnecessary motion practice.

Practice Pointer 12. When an extension is not feasible, the non-party recipient should assert objections prior to compliance or 14 days of service of a subpoena (whichever is earlier) to ensure that mandatory cost-shifting provisions for significant expenses are available.

Practice Pointer 13. Subpoenas should be written with reasonable particularity based on the issuing party's then-knowledge of the non-party's documents and custodians. The non-party should be as specific as possible under the circumstances in its objections.

Practice Pointer 14. This *Commentary* encourages a non-party to provide a specific date after which it will no longer retain the documents or ESI that it objects to producing. Such a step thereby places the requesting party on notice of the date by which it needs to determine the completeness of the production and move to compel.

Practice Pointer 15. It may be beneficial for the party issuing a subpoena and the non-party recipient to confer in an effort to resolve any disputes regarding the scope of the subpoena before seeking to quash or enforce a subpoena. If appropriate, other parties should be given the opportunity to participate in such discussions.

Practice Pointer 16. The party issuing a subpoena and the non-party recipient should consider, where appropriate, a tiered or staged production, particularly if requested by the non-party.

Practice Pointer 17. Rule 45(e)(2)(A) and (B) require a non-party subpoena recipient to, among other things, expressly make a “claim [of privilege] and the basis for it” and set forth a process for the handling of the inadvertent production of such information. The party issuing a subpoena should seek to minimize the burden of privilege claims on the non-party. For example, the issuing party and the non-party may agree to exclude some potentially privileged and protected information from the subpoena based upon dates, general topics, or subjects. To minimize the burden on the non-party, the subpoenaing party should consider alternatives to the traditional privilege log.

Practice Pointer 18. The parties should work together and with the non-party, as appropriate, to facilitate the authentication of material received through non-party subpoenas. To avoid the necessity for a non-party’s appearance at trial, the parties should obtain and utilize, when possible, non-party certifications under Fed. R. Evid. 902. The parties should stipulate to the authenticity and the business records hearsay exception, when possible, to minimize the burden and expense imposed on a non-party subpoena recipient, including any need to testify for foundational matters.

THE SEDONA CONFERENCE COMMENTARY ON ESI
EVIDENCE & ADMISSIBILITY, SECOND EDITION

*A Project of The Sedona Conference Working Group on
Electronic Document Retention and Production (WG1)*

Author:

The Sedona Conference

Editors-in-Chief & WG1 Steering Committee Liaisons:

Kevin F. Brady

Heather Kolasinsky

Philip Favro

Drafting Team:

Carey Busen

Gita Radhakrishna

Holly Dyer

Kristin Walinski

Endel (Del) Kolde

Martin Wolf

Jonathan Le

Judicial Participants:

Judicial Advisor:

Hon. Ralph Artigliere (ret.)

Hon. Paul W. Grimm

Hon. Thomas Vanaskie (ret.)

Staff editors:

David Lumia

Susan McClain

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The

Copyright 2020, The Sedona Conference.
All Rights Reserved.

Sedona Conference Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

We thank all of our Working Group Series Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors, just click on the “Sponsors” navigation bar on the homepage of our website.

This publication may be cited as follows:

The Sedona Conference, *Commentary on ESI Evidence & Admissibility, Second Edition*, 22 SEDONA CONF. J. 83 (2021).

PREFACE

Welcome to the final, October 2020, version of The Sedona Conference *Commentary on ESI Evidence & Admissibility, Second Edition*, a project of the Sedona Conference Working Group 1 on Electronic Document Retention and Production (WG1). This is one of a series of Working Group commentaries published by The Sedona Conference, a 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, intellectual property rights, and data security and privacy law. The mission of The Sedona Conference is to move the law forward in a reasoned and just way.

This is the second iteration of The Sedona Conference *Commentary on ESI Evidence & Admissibility*. The first edition was published in March 2008 to address a concern at that time about whether and how electronically stored information (ESI), once produced, can actually be authenticated and used as evidence at trial or in motion practice. The 2008 edition provided a framework, practical guidance, and a checklist for authenticating ESI and getting it admitted into evidence. That 2008 *Commentary* focused primarily on the applicability and application of the Federal Rules of Evidence and case law to the existing data sources at that time, as well as addressing the potential issues and pitfalls for data sources that were looming on the horizon. Much has changed in the past 12 years, and this second edition reflects those changes.

In 2017 and 2019, the Federal Rules of Evidence were amended. In contrast to the fanfare accompanying the changes to the Federal Rules of Civil Procedure in 2006 and 2015, little attention was paid to the 2017 changes to Federal Rules of Evidence 803(16), 807, and 902(13) and (14). Those changes are significant and intended to influence how parties manage ESI. For example, the changes to Rule 803(16) address authentication of

digital information that has been stored for more than 20 years, eliminating the concern that factual assertions made in massive volumes of ESI will be admissible for the truth simply because of their age. The concurrent addition of new subsections (13) and (14) to Rule 902 provide for streamlined authentication of ESI and potentially eliminate the need to call a witness at trial to authenticate the evidence. As we note at the end of this *Commentary*, future developments in the law and ever-changing landscape of technology may warrant another iteration.

An update to the 2008 edition of the *Commentary on ESI Evidence & Admissibility* was a topic of dialogue at the WG1 2018 Annual and 2019 Midyear meetings, and drafts of this *Commentary* were circulated for member comment at the 2019 Midyear Meeting and again in early 2020. This second edition was published for public comment in July 2020. Where appropriate, the comments received during the public-comment period have been incorporated into this final version of the *Commentary*. The Sedona Conference acknowledges the efforts of Editors-in-Chief and Steering Committee Liaisons Kevin F. Brady, Philip Favro, and Heather Kolasinsky, who were invaluable in driving this project forward. We also thank Drafting Team members Carey Busen, Holly Dyer, Del Kolde, Jonathan Le, Gita Radhakrishna, Kristin Walinski, and Martin Wolf, as well as The Honorable Ralph Artigliere (ret.), The Honorable Thomas Vanaskie (ret.), and The Honorable Paul Grimm for their efforts and commitments in time and attention to this project.

We encourage your active engagement in the dialogue. Membership in The Sedona Conference Working Group Series is open to all. The Series includes WG1 and several other Working Groups in the areas of international electronic information management, discovery, and disclosure; patent damages and patent litigation best practices; data security and privacy liability; trade secrets; and other “tipping point” issues in the law.

The Sedona Conference hopes and anticipates that the output of its Working Groups will evolve into authoritative statements of law, both as it is and as it should be. Information on membership and a description of current Working Group activities is available at <https://thesedonaconference.org/wgs>.

Craig Weinlein
Executive Director
The Sedona Conference
October 2020

TABLE OF CONTENTS

I.	INTRODUCTION.....	91
II.	APPLYING EXISTING RULES AND CASE LAW TO ESI EVIDENCE	93
A.	Early Focus on Authentication and Evidentiary Issues	93
B.	Summary Judgment Motions and ESI Evidence	93
C.	Authentication Tools: Rules 104, 901, and 902.....	95
1.	Rule 104	95
2.	Rules 901 and 902.....	96
3.	Rules 902(13) and (14).....	98
4.	Rule 902(13) and (14) Certifications	102
D.	Various Types of ESI Require Different Approaches	104
1.	Email	105
2.	Text Messages.....	111
3.	Websites.....	113
4.	Social Media Sites.....	120
5.	Internet of Things.....	132
6.	Ephemeral or Self-Destructing Photographs/Messages.....	133
7.	Digitally Stored Data	136
8.	Digital Photographs.....	137
9.	Group Collaboration Tools.....	139
10.	Computer Processes, Animations, Audio/ Video, Virtual Reality, and Simulations	140
11.	Cloud Computing	141
12.	Emoji	142
E.	Hard Copies	148

F.	Potential Challenges to Using Rule 902(14).....	149
1.	The Requirement of a Process of Digital Identification.....	149
2.	Certification Hazard: The Potential Exposure of Electronic Discovery Protocols	151
G.	Recent Changes to Rule 807 (Residual Exception to Hearsay Rule)	152
III.	EMERGING ESI EVIDENTIARY ISSUES	156
A.	Determining the Owner/Creator of ESI	156
B.	Understanding the Limits of Technology	156
1.	Hashing	158
2.	Encryption.....	160
3.	System Metadata	162
4.	Computer Forensics and Anti-Forensics	163
5.	Blockchain	164
C.	Application of Federal Rules and Cases in State Court and Vice Versa	168
1.	Federal law application in state cases	168
2.	State law application in federal cases.....	169
IV.	PRACTICAL GUIDANCE ON THE USE OF ESI IN COURT	173
A.	Use of ESI in Static vs. Native/Live Format.....	174
B.	Evidence to Assist the Jury on the Permissive Spoliation Inference	175
C.	Practical Tips for Administration of ESI as Evidence.....	179
D.	Practical Tips for Seeking Authority on Admission of ESI as Evidence	179
V.	ARTIFICIAL INTELLIGENCE USES IN BUSINESS AND LAW	183
	Appendix A: Summary Federal Rules of Evidence	
	901 and 902 Rules for Authentication	192

Appendix B: Committee Note on Rule 807	210
Appendix C: 12 V.S.A. § 1913. Blockchain enabling	214
Appendix D: Checklist of Potential Authentication Methods	218

I. INTRODUCTION

The ability to present admissible evidence is an essential skill for successful litigators. At its core, admissibility is about what evidence may be considered by the decision-maker. Many civil cases settle, but they settle at different stages of the litigation process. Summary judgment proceedings and pretrial motion practice often, if not always, require a party to offer admissible evidence for a proposition, claim, or defense. If a civil case is not resolved before trial, a judge or jury will decide the merits of the case, which will also require the presentation of admissible evidence. Criminal cases, on the other hand, which are more likely to go to trial, may result in a higher number of reported decisions regarding electronically stored information (ESI) evidence,¹ primarily due to the lack of pre-trial discovery of devices in criminal cases.²

The growth of electronic discovery reflects the increasing digitization of information in society, which also results in more relevant evidence being sourced from ESI. This phenomenon means that successful litigators must understand how to get ESI admitted into evidence, which is a different question than preserving or gathering it for discovery. As U.S. District Judge Paul W. Grimm noted in the seminal case *Lorraine v. Markel American Insurance Co.*, “it makes little sense to go to all the bother and

1. As used in this *Commentary*, evidence means “material presented to a competent legal tribunal to prove or disprove a fact.” See BRYAN A. GARNER, GARNER’S DICTIONARY OF LEGAL USAGE 34 (1987). Most ESI that exists, or is collected and produced in discovery, will never be promoted to the status of evidence submitted before a tribunal. The focus of this *Commentary* is on the small subset of ESI that will be offered as evidence.

2. Criminal cases may also lead to more reported decisions because many defense counsel may perceive an ethical duty not to stipulate to the admissibility of ESI evidence that will be used to attempt to convict their client. Different incentives to cooperate may prevail in civil cases.

expense” of electronic discovery only to have that evidence excluded when it really matters.³ This *Commentary* focuses specifically on that concern.

This *Commentary* is divided into three parts. First, there is a survey of the application of existing evidentiary rules and case law addressing the authenticity of ESI. Second, there are discussions about new issues and pitfalls that are looming on the horizon such as ephemeral data, blockchain, and artificial intelligence. Finally, there is practical guidance on admissibility and the use of ESI in depositions and in court.

While this *Commentary* primarily addresses the Federal Rules of Evidence, the overwhelming volume and the widest diversity in types and size of cases occur in state courts, where the subject-matter jurisdiction is much broader. Space prohibits state-by-state coverage, but this *Commentary* compares the federal law and principles to rules of evidence and admissibility arising in state court. Guidance for addressing state court admissibility occurs throughout this *Commentary*.

3. 241 F.R.D. 534, 538 (D. Md. 2007). *Lorraine* remains a frequently cited case on ESI admissibility, with nearly 1,600 citing references on WestlawNext.

II. APPLYING EXISTING RULES AND CASE LAW TO ESI EVIDENCE

A. *Early Focus on Authentication and Evidentiary Issues*

Judge Grimm's discussion in *Lorraine* makes it clear that parties should start to think about evidentiary issues much earlier than was the practice when dealing only with hard-copy evidentiary materials. Consideration should be given to how potential ESI evidence is handled by records management programs, and parties should be mindful of authentication possibilities throughout the discovery process. For example, under the pretrial disclosure provisions of Rule 26(a)(3), a party has 14 days to object to the admissibility of an opponent's proposed documents of other trial exhibits, and the failure to do so results in a waiver. Additionally, given the extent to which summary judgment has displaced trial as a procedure for resolving legal disputes, parties should be prepared to deal with evidentiary issues at the summary judgment stage.

B. *Summary Judgment Motions and ESI Evidence*

Summary judgment is a critical stage in any litigation and is likely the first time that issues of evidence admissibility, including authenticity, will be considered, because the court is only allowed to consider evidence that is admissible.⁴

This point was made clear in *Lorraine*, where the court rejected unsworn, unauthenticated documents from both parties. As the Judge Grimm explained, the court could only consider evidence at summary judgment that would be admissible at trial.⁵ Judge Grimm also detailed how the Rules:

4. *Celotex Corp. v. Catrett*, 477 U.S. 317, 322 (1986); FED. R. CIV. P. 56(c); see also *Gannon Int'l, Ltd. v. Blocker*, 684 F.3d 785, 793 (8th Cir. 2012).

5. The Court in *Celotex* noted that under Rule 56(e), a party can oppose summary judgment using any of the evidentiary materials identified in Rule

present themselves like a series of hurdles to be cleared by the proponent of the evidence. Failure to clear any of these evidentiary hurdles means that the evidence will not be admissible. Whenever ESI is offered as evidence, either at trial or in summary judgment, the following evidence rules must be considered: (1) is the ESI relevant as determined by Rule 401 (does it have any tendency to make some fact that is of consequence to the litigation more or less probable than it otherwise would be); (2) if relevant under 401, is it authentic as required by Rule 901(a) (can the proponent show that the ESI is what it purports to be); (3) if the ESI is offered for its substantive truth, is it hearsay as defined by Rule 801, and if so, is it covered by an applicable exception (Rules 803, 804 and 807); (4) is the form of the ESI that is being offered as evidence an original or duplicate under the original writing rule, or if not, is there admissible secondary evidence to prove the content of the ESI (Rules 1001- 1008); and (5) is the probative value of the ESI substantially outweighed by the danger of unfair prejudice or one of the other factors identified by Rule 403, such that it should be excluded despite its relevance.⁶

56(c), except for the pleadings themselves, and it is from that list (which includes affidavits) that one would normally expect the nonmoving party to make that showing. *Celotex*, 477 U.S. at 324. However, that is not always the case. If the content of the affidavit would not be admissible if it is offered into evidence at trial by a live witness, then it is not considered admissible evidence for summary judgment purposes notwithstanding the fact that it is in an acceptable form for Rule 56(c) purposes. FED. R. CIV. P. 56(c).

6. *Lorraine*, 241 F.R.D. at 538.

C. *Authentication Tools: Rules 104, 901, and 902*

Authenticity is one part of admissibility, requiring that the proponent of an exhibit “make a prima facie showing that it is what he or she claims it to be.”⁷ The comparatively recent additions of Federal Rule of Evidence 902(13) and (14) provide additional tools for the authentication of ESI, including system metadata and files such as an email or an Excel spreadsheet.

1. Rule 104

There is a complex interplay between “preliminary rulings” on admissibility, governed by Rules 104(a) and (b), and the authenticity determination, governed by Rules 901 and 902. Rule 104(a) governs the admissibility of matters such as whether an expert is qualified and, if so, whether the expert’s opinions are admissible; whether the evidence is privileged; and whether evidence is hearsay, and, if so, whether any recognized exception applies.⁸ As explained in *Lorraine*, under Rule 104(a), the court, not the fact finder, makes the admissibility determination. In making that determination, the court is not bound by the restrictions of the rules of evidence except those concerning privileges.⁹

On the other hand, the authenticity of ESI and other evidence is governed by Rule 104(b), which affords the court a much narrower role. Under this rule, the court addresses only a threshold question of law: does the evidence have sufficient probative value to sustain a rational jury finding that the evidence is what the proponent claims it to be? The fact finder makes the ultimate determination of whether the evidence is authentic.

7. *Id.* at 542.

8. *See id.* at 539.

9. *Id.*

For example, if an email is offered into evidence, the jury makes the authenticity determination under Rule 104(b) using only admissible evidence.¹⁰

2. Rules 901 and 902

Examples of methods a proponent may use to authenticate ESI are set forth in Rules 901 and 902. Just as with hard-copy evidence, a “party seeking to admit an exhibit need only make a prima facie showing that it is what he or she claims it to be.”¹¹ This is not a particularly high barrier to overcome.

In *United States v. Safavian*, the court analyzed the admissibility of email, noting that:

[t]he question for the Court under Rule 901 is whether the proponent of the evidence has “offered a foundation from which the jury could reasonably find that the evidence is what the proponent says it is.” The Court need not find that the evidence is necessarily what the proponent claims, but only that there is sufficient evidence that the *jury* ultimately might do so.¹²

10. *Id.* at 540.

11. *Id.* at 542.

12. *United States v. Safavian*, 435 F. Supp. 2d 36, 38 (D.D.C. 2006) (internal citations omitted) (emphasis in original); *see also* *Dunn v. Hunting Energy Servs.*, 288 F. Supp. 3d 749, 764 (S.D. Tex. 2017) (citing *Lorraine* and admitting emails); *United States v. Bertram*, 259 F. Supp. 3d 638, 640, 642–43 (E.D. Ky. 2017) (citing *Lorraine* and *Safavian* and admitting emails).

The first edition of this *Commentary* included a discussion of an eleven-factor authentication test for computerized records adopted by the U.S. Bankruptcy Appellate Panel of the Ninth Circuit in *In re Vinhnee*, 336 B.R. 437, 446–47 (B.A.P. 9th Cir. 2005). The more stringent test applied in that case has been omitted from this edition of the *Commentary* because it has rarely been cited outside the Ninth Circuit, and the analysis is discussed in only a

It is important to note that the methods for authentication listed in Rules 901 and 902 are non-exhaustive and can be used in combination with each other, although, as discussed below, courts have identified particular provisions of 901 and 902 that are appropriate or most useful for specific types of ESI.

Rule 902¹³ identifies evidence that is “self-authenticating,” that is, information that can be admitted at trial without being authenticated by a witness. Self-authenticating evidence may be admissible without extrinsic evidence of authenticity “sometimes for reasons of policy but perhaps more often because practical considerations reduce the possibility of unauthenticity to a very small dimension.”¹⁴ Most, but not all, of the items listed in Rule 902 are self-authenticating on their face, thus requiring no extrinsic evidence of authenticity for the document to be admitted. There are sections of Rule 902—such as Rule 902(11) and Rule 902(12) (for records of regularly conducted activity, domestic and foreign, respectively), 902(13) (records generated by an electronic process or system), and 902(14) (data copied from an electronic device)—that are self-authenticating *only* to the extent the party seeking to introduce them into evidence submits a proper certification to their authenticity and provides notice to the opposing party to give it a fair opportunity to challenge the certification.

few reported decisions. Cautious practitioners may nevertheless want to be aware that *In re Vinhnee* can be cited to support a more stringent authentication standard, including proving the existence of access control and an audit trail. In general, however, the courts have become more comfortable with authenticating ESI over the past decade.

13. The following discussion (up to Section D) is taken with permission from Hon. Paul W. Grimm & Kevin F. Brady, *Recent Changes to Federal Rules of Evidence: Will They Make It Easier to Authenticate ESI?*, 19 SEDONA CONF. J. 707, 711–21 (2018).

14. FED. R. EVID. 902 advisory committee’s notes to 1972 proposed rules.

3. Rules 902(13) and (14)

In 2017, the Advisory Committee supplemented Rule 902 by adding two subsections permitting similar certifications to authenticate electronic evidence. The amendments are intended to eliminate the need for a live witness to testify as to the authenticity of certain ESI, thereby streamlining the process at trial.

The new subsections to Rule 902 are:

(13) Certified Records Generated by an Electronic Process or System. A record generated by an electronic process or system that produces an accurate result, as shown by a certification by a qualified person that complies with the certification requirements of Rule 902(11) or Rule 902(12). The proponent must also meet the notice requirements of Rule 902(11).

(14) Certified Data Copied from an Electronic Device, Storage Medium, or File. Data copied from an electronic device, storage medium, or file, if authenticated by a process of digital identification, as shown by a certification by a qualified person that complies with the certification requirements of Rule 902(11) or (12). The proponent also must meet the notice requirements of Rule 902(11).

As with the provisions on business records in Rules 902(11) and 902(12), the Advisory Committee noted that the expense and inconvenience of producing a witness to authenticate an item of electronic evidence is often unnecessary because the adversary either stipulates to authenticity before the witness is called or fails to challenge the authentication testimony once it

is presented.¹⁵ Under the amendments to Rule 902, the parties are now able to determine in advance of trial whether a real challenge to authenticity will be made.

Note that Rule 902(11) relates “only to the procedural requirements” of authentication.¹⁶ Likewise, new subsections 902(13) and (14) are designed to do “nothing more than authenticate” ESI.¹⁷ Therefore, the proponent of the evidence sought to be admitted still must prove the requirements of Rule 803(6) after clearing the authenticity hurdle. What is important to note from Rules 902(13) and (14) is that the references to Rules 902(11) and (12) are simply to the form of the declaration: the affidavit the party wishes to introduce must have the same formality and style as the certifications referred to in Rules 902(11) and (12). Rules 902(13) and (14) do not require that the certification for subsections (13) and (14) to include the substantive certification of Rule 902(11), which is tied to Rule 803(6)(A)(B)(C) elements for the business-record exception.

New subsections 13 and 14, like Rules 902(11) and (12), permit a foundation witness or “qualified person” to establish the authenticity of information by way of certification.¹⁸ Subsection 902(13) provides for self-authentication of machine-generated information—such as system metadata—upon the submission

15. FED. R. EVID. 902 advisory committee’s notes to 2017 amendments ¶¶ 13 & 14.

16. *Id.*

17. ADVISORY COMMITTEE ON EVIDENCE RULES, MINUTES OF THE MEETING OF APRIL 29, 2016, https://www.uscourts.gov/sites/default/files/2016-04-evidence-minutes_0.pdf.

18. Pursuant to Rule 901(11) and 901(12), a “qualified person” is a custodian or other individual who has the ability to establish the authenticity of the ESI as if that person would have testified at trial such as under FED. R. EVID. 901(b)(1) (Testimony of a Witness with Knowledge) or 901(b)(4) (Distinctive Characteristics and the Like).

of a certification prepared by a qualified person. Subsection 902(14) provides for authentication of data copied from an electronic device, medium, or file—such as an email or Excel spreadsheet that was stored on a computer—through digital identification.

The Advisory Committee noted that in most instances, digital identification involves authentication of data copied from electronic devices by comparing the “hash value” of the proffered copy to that of the original document. A message-digest hash value is a unique alphanumeric sequence of characters that an algorithm determines based upon the digital contents of the device.¹⁹ The hash value serves as the digital fingerprint that a qualified person uses to compare the numeric value of the proffered item with the numeric value of the original item. If the hash values for the original and copy are identical, the information can be proffered, and the court can rely on them as authentic copies.²⁰ The Advisory Committee also noted that “[t]he rule is flexible enough to allow certifications through processes other than comparison of hash value, including by other reliable means of identification provided by future technology.”²¹

New Rules 902(13) and 902(14) have the same effect as other Rule 902 provisions of shifting to the opponent the burden of going forward—but not the burden of proof—on authenticity disputes regarding the electronic evidence at issue. Shifting the burden of questioning the authenticity of such records to the opponent who has a fair opportunity to challenge both the certification and the records streamlines the process by which these items can be authenticated, thereby reducing the time, cost, and

19. FED. R. EVID. 902 advisory committee’s notes to 2017 amendments ¶ 14. See Section III.B.1, *infra*, for a more detailed definition of “hashing.”

20. *Id.*

21. *Id.*

inconvenience of presenting this evidence at trial or summary judgment.

Rule 902(13) is designed to permit the proponent to show that the evidence in question is authentic by attaching an affidavit under oath by the person or people with the technical or specialized knowledge of how the system or process works, certifying that the evidence is reliable and accurate.²²

Rule 902(14) allows for a certification that would explain the process by which that person took a forensic copy of the evidence such as a hard drive of a laptop, hashed it, and then compared the hash value of the forensic copy with the hash value of the original hard drive. Certification is an affidavit or declaration by someone with firsthand, personal knowledge or with qualified expertise under Rule 702. If the original hash value and the hash value of the forensic copy are the same, then the information in the copy is identical to the information in the original.

For example, if an individual takes a picture with a smartphone, embedded within the electronic metadata of that photograph are global positioning system (GPS) coordinates of the location where that photograph was taken. In a criminal case, where the prosecution must prove that the defendant was in a specific location by virtue of photographs taken from that defendant's mobile phone, the metadata from that electronic photograph that shows the GPS coordinates is evidence of where the smartphone and (by extension) the person were located when the picture was taken.

22. See *United States v. Forty-Febrs*, No. 16-330, 2018 WL 2182653, at *2 (D.P.R. May 11, 2018) *appeal docketed*, No. 18-2106 (1st Cir. Nov. 17, 2018) (granting motion *in limine* to admit electronic records of the Puerto Rico Department of Transportation based upon a certification from the custodian of the records).

Under the Rule, the prosecutor can put that information in an affidavit and offer the affidavit to the defendant with the request to voice any objection regarding authenticity. If the defendant objects, the prosecutor must actually prove the authenticity and will need to bring one or more witnesses—persons with the scientific, technical, or specialized knowledge—to testify at trial how the system and processes produce reliable results.²³ If the defendant does not object, the prosecutor has established authenticity and no authenticating witness would be needed at trial. Unless qualified as an expert under Rule 702, the affiant must provide information based on direct personal knowledge. The affiant's testimony cannot be based on hearsay. Moreover, if the proponent has a system or process that requires explanation by multiple persons in order to be complete, affidavits are needed from each of those persons.

4. Rule 902(13) and (14) Certifications

A Rule 902 certification is intended to take the place of the testimony traditionally required to establish the authenticity of the ESI sought to be admitted; therefore, it should follow the same pattern as the testimony it is intended to replace.²⁴ The certification should start by establishing the background, education, training, and expertise of the affiant in order to establish

23. Criminal cases involving such certifications can also raise Confrontation Clause issues. *Compare* *Melendez-Diaz v. Massachusetts*, 557 U.S. 305, 329 (2009) (“The Sixth Amendment does not permit the prosecution to prove its case via *ex parte* out-of-court affidavits. . .”) *with* *United States v. Yeley-Davis*, 632 F.3d 673, 681 (10th Cir. 2011) (Rule 902(11) certifications of authenticity concerning certified copies of telephone toll records are not testimonial and therefore do not violate the Sixth Amendment Confrontation Clause). Thus, there may be a distinction between records generated specifically for a prosecution and historic records that pre-existed a legal dispute.

24. *See* Grimm & Brady, *supra* note 13, at 740 for sample certifications under Rules 902(13) and 902(14).

that the affiant is a “qualified person” as required by Rules 902(11) and (12). Although Rules 902(13) and (14) do not refer to Rule 702, counsel would be wise to ensure that the affiant providing the certificate meets the requirements of an expert witness under Rule 702 if the underlying facts to be authenticated involve scientific, technical, or specialized knowledge. The added benefit of showing that the affiant meets these Rule 702 requirements is that the affiant may base the certification on information beyond personal knowledge, provided it is reliable, as described in Rule 703. The certification should then describe the affiant’s role in the case, that is, that the affiant was retained by the party as a computer forensics expert to assist the party and its counsel in the identification, preservation, collection, and production of ESI. The certification should describe in detail the evidence in question and establish its authenticity consistent with the formality requirements of Rules 901(11) and (12). The certification need not meet the requirements of Rule 803(6)(A–C), unless the proponent also seeks to qualify the evidence as a business record. Instead, the certification must provide the information required by Rules 902(13) and (14), as discussed below.

If the certificate seeks to authenticate evidence under Rule 902(13), the affiant should describe in detail the “electronic process or system” that was used to generate the information in question. For example, if the information in question is a series of monthly sales reports, the affiant should describe: (i) the system from which the reports were generated; (ii) the process by which the data that was used to generate the statements was gathered, processed, and stored; and (iii) the process by which the statements or reports sought to be admitted were generated and produced for the litigation. The Rule 902(13) certificate should establish that the information sought to be admitted has not been altered from the form in which it was maintained in the ordinary course of business. While the process of preparing

a certification under Rule 902 is seemingly straightforward, the affiant must be careful to describe the “electronic process or system” with enough specificity to satisfy the court and the opponent of the evidence’s authenticity. Doing so can help avoid a hearing during which the opponent of the evidence may cross-examine the affiant.²⁵

If the certificate seeks to authenticate evidence under Rule 902(14), the affiant also should describe in detail the ESI that was copied from its original location and now offered into evidence. The affiant should additionally detail the steps taken by the affiant at the time of duplication (including recording the date, time, surrounding circumstances, and hardware and software tools as well as versions utilized). For example, if the information sought to be admitted comprises a series of Excel and PowerPoint files that were stored on the departmental file share for the client’s accounting department, the affiant should list the files in question and include the hash value of each of the files as they existed on the file share. The affiant should also describe the hash value for the copy of each of the files sought to be admitted to establish that the files are authentic copies of the files as they were maintained in the ordinary course of business. The identical hash values will attest that the information sought to be admitted into evidence is a true and correct copy of the information as it existed in its original state.

D. Various Types of ESI Require Different Approaches

All ESI shares certain common characteristics, but some types of ESI present unique challenges to authentication, necessitating different approaches. For example, the creator of certain

25. See *La Force v. Gosmith, Inc.*, No. 17-cv-05101-YGR, 2017 WL 9938681, at *3 (N.D. Cal. Dec. 12, 2017) (deeming an attorney’s declaration submitted in support of printouts of web pages insufficient to meet the requirements of Rule 902(13)).

ESI types may be unidentifiable, and the ESI may be stored in various systems with different security measures. Some ESI may contain clues about its history, while other types are completely lacking in provenance. It is thus useful to quickly survey some representative categories of ESI.

1. Email

For many organizations, email remains the primary form of business communication.²⁶ Other forms of electronic communication, including various forms of instant messaging, are also increasingly part of the mix, but email is still predominant.

There are many ways in which email evidence may be authenticated:

- a witness with personal knowledge—Rule 901(b)(1)
- expert testimony or comparison with authenticated examples—Rule 901(b)(3)
- distinctive characteristics, including circumstantial evidence—Rule 901(b)(4)
- a system or process capable of proving a reliable and dependable result—Rule 901(b)(9)
- trade inscriptions—Rule 902(7)
- certified copies of a business record—Rule 902(11)
- certified records generated by an electronic process or system—Rule 902(13)

26. “The total number of business and consumer emails sent and received per day will exceed 306 billion in 2020, and is forecast to grow to over 361 billion by year-end 2024.” THE RADICATI GROUP, INC., EMAIL STATISTICS REPORT, 2020-2024 EXECUTIVE SUMMARY 2 (FEB. 2020), <https://www.radicati.com/wp/wp-content/uploads/2019/12/Email-Statistics-Report-2020-2024-Executive-Summary.pdf>.

- certified data copied from an electronic device, storage medium, or file—Rule 902(14)²⁷

The addition of two new subsections to Rule 902 gives practitioners additional options for authenticating emails or metadata associated with emails, although admissibility will still need to be established.²⁸ For example, under Rule 902(13), an email could qualify as data copied from a storage medium, which could be digitally authenticated by a qualified person. Similarly, under 902(14), system metadata could be used to authenticate an attachment to an email as a record generated by an electronic process or system.

(a) Email as a business record

In litigation involving business entities or government agencies, many emails will potentially qualify as business records, allowing a proponent to establish both authenticity and admissibility by meeting a single test. But it is insufficient to “simply [] say that since a business keeps and receives emails, then *ergo* all those e-mails are business records falling within the ambit of [the business records exception].”²⁹

Longstanding Rule 902(11) is particularly “helpful in establishing the foundation elements for a business record without the need to call a sponsoring witness to authenticate the document and establish the elements of the hearsay exception.” This, in turn, allows a proponent to establish both authenticity and a

27. See Appendix D: Checklist of Potential Authentication methods, *infra*.

28. See Section II.C, *supra*.

29. *United States v. Cone*, 714 F.3d 197, 220 (4th Cir. 2013) (ruling that emails concerning counterfeit goods were improperly admitted). *But see* *Alig v. Quicken Loans Inc.*, No. 5:12-CV-114, 2017 WL 5054287, at *8 (N.D.W. Va. July 11, 2017) (finding that executives’ emails qualified as business records).

major component of admissibility.³⁰ Rule 902(11) allows the self-authentication of a business record. The proponent must produce an original or duplicate of a domestic record of regularly conducted activity that would be admissible under Rule 803(6) if accompanied by a written declaration of its custodian or other qualified person, in a manner complying with any Act of Congress or rule prescribed by the Supreme Court pursuant to statutory authority, certifying that the record:

- (a) was made at or near the time of the occurrence of the matters set forth by, or from information transmitted by, a person with knowledge of those matters;
- (b) was kept in the course of the regularly conducted activity; and
- (c) was made by the regularly conducted activity as a regular practice.³¹

Because the elements for Rules 902(11) and 803(6) are essentially the same, they frequently are analyzed together when Rule 902(11) is the proffered means by which a party seeks to admit a business record.³²

With respect to the “personal knowledge” component of Rule 803(6) (that there be personal knowledge of the entrant or of an informant who had a business duty to transmit the information to the entrant), it is relatively simple to prove personal knowledge if the author of the email is available to testify and

30. *Lorraine v. Markel Am. Ins. Co.*, 241 F.R.D. 534, 571 (D. Md. 2007). Rule 803(6) is often referred to as the business-records exception to the hearsay rule and presents a common way for gaining admissibility of ESI evidence in civil cases involving companies and other organizations that maintain business records.

31. *Id.*

32. *Id.* at 572.

had personal knowledge of the contents. But in many instances, the email contains information from a source outside the business of the maker of the business record, which presents special evidentiary problems.

In *Lorraine*, the court noted that the majority view for meeting the requirements of the business-record exception in that situation is that the supplier or source of the information memorialized in the email must have had “a business duty to transmit the information to the maker of the record, if the maker, him or herself lacks personal knowledge of the facts or events.”³³ “However, some courts have held that it may be possible to meet the requirements of the business-record exception even if the source of the information had no business duty to provide it to the maker of the record, provided the recipient of the information has a business duty to verify the accuracy of the information provided.”³⁴

In addition, it may be useful for litigants to establish the elements of the business-records exception for high-value emails during depositions, prior to offering them as evidence in a court. If a manager or party representative admits, in a deposition, to having sent or received an email in the course of regularly conducted business activity, that manager’s employer will be hard-pressed to challenge authenticity at a later stage in the lawsuit.

33. *See id.* at 571 n.52 (citing FED. R. EVID. 803(3) advisory committee’s note (“Sources of information presented no substantial problem with ordinary business records. All participants, including the observer or participant furnishing the information to be recorded, were acting routinely, under a duty of accuracy, with employer reliance on the result, or in short ‘in the regular course of business.’ If, however, the supplier of the information does not act in the regular course, an essential link is broken; the assurance of accuracy does not extend to the information itself, and the fact that it may be recorded with scrupulous accuracy is of no avail.”)).

34. *Id.* (citing *Rambus, Inc. v. Infineon Techs. AG*, 348 F. Supp. 2d 698, 706–07 (E.D. Va. 2004)).

Depositions, however, may not always have taken place, and they would not ordinarily be available in criminal cases.

Finally, in civil cases, a party may be precluded from challenging the authenticity of ESI that it produced during discovery. Some courts have held that “[parties] cannot voluntarily produce documents and implicitly represent their authenticity and then contend they cannot be used by the [opposing party] because the authenticity is lacking.”³⁵ In practice, however, this rule may not always apply, especially if a party is in possession of records it did not generate. For example, an email received from an outside entity might be subject to discovery and production, but it would not necessarily be appropriate to imply that the producing party had a definitive position on the identity of the sender or the authenticity of the document. Similarly, if a party originally received the ESI from an opposing party and then subsequently produced it back to the opposing party in accordance with a new discovery request or a duty to supplement, it would not necessarily follow that the party was claiming that the ESI was authentic.

(b) Authenticating emails using circumstantial evidence

In a nonbusiness context or other situations where an email does not qualify as a business record, practitioners can often authenticate emails with circumstantial evidence by reference to distinctive characteristics in the contents of the email.³⁶ For

35. *Indianapolis Minority Contractors Ass’n, Inc. v. Wiley*, IP 94-1175-C-T/G, 1998 WL 1988826, at *6 (S.D. Ind. May 13, 1998), *aff’d sub nom.* *Indianapolis Minority Contractors Ass’n, Inc. v. Wiley*, 187 F.3d 743 (7th Cir. 1999). *See also* *Radiance Capital Receivables Eighteen, LLC v. MBO Investments, LLC*, 4:16-CV-1921-SPM, 2019 WL 330463, at *4 (E.D. Mo. Jan. 25, 2019).

36. FED. R. EVID. 901(b)(4).

example, an email might contain “details known only to the sender and the person receiving the message.”³⁷

Thus, in *United States v. Safavian*, emails between the defendant and a lobbyist were sufficiently authenticated because both persons’ names were part of the respective email user names. In addition, the contents of the emails referred to matters the lobbyist or defendant were known to be working on.³⁸

Similarly, when it comes to the next step, admissibility, there are numerous options for nonbusiness records. Frequently, an email may be the statement of a party opponent, which is not hearsay.³⁹ Even where an email contains non-party statements, they might not be hearsay at all. For example, in *Safavian*, the court held that email content from a lobbyist was non-hearsay because the lobbyist asked questions, sought favors, or made requests for assistance rather than making declarative statements about the truth of a matter.⁴⁰ Likewise, in *United States v. Fluker*,

37. *Lorraine*, 241 F.R.D. at 554.

38. *United States v. Safavian*, 435 F. Supp. 2d 36, 40–41 (D.D.C. 2006) (emails admissible as admissions of a party opponent and non-hearsay); *see also* *United States v. Fluker*, 698 F.3d 988, 998–1000 (7th Cir. 2012) (email addresses were consistent with purported senders and contents showed sender had knowledge of relevant issues); *United States v. Bertram*, 259 F. Supp. 3d 638, 642–43 (E.D. Ky. 2017) (witness with history of email exchanges with defendants could authenticate emails based on distinctive characteristics); *Johnson v. State*, 137 A.3d 253, 271–74 (Md. Ct. Spec. App. 2016), *cert. denied*, 146 A.3d 471 (Md. 2016) (email contents referred to personal and family circumstances specific to defendant).

39. FED. R. EVID. 801(d)(2); *see also* *Lorraine*, 241 F.R.D. at 568 (noting the universality of electronic communication and the application of the party opponent rule); *Safavian*, 435 F. Supp. 2d at 43–44 (admitting emails containing statements directly attributed to defendant and forwarded emails where context showed they were adoptive admissions).

40. *Safavian*, 435 F. Supp. 2d at 44–45.

the emails contained fraudulent statements that, by definition, were not offered for the truth of the matter asserted.⁴¹

2. Text Messages

Text messages are frequently used to communicate in business and nonbusiness settings but occupy a less formal space than email. This is because the communications are often shorter, may be sent and received on personally owned devices, and may exist outside of formal information governance policies. As a result, text messages may not be considered business records even if they relate to the business of a particular organization.

There are many ways in which text messages may be authenticated:

- a witness with personal knowledge—Rule 901(b)(1)
- expert testimony or comparison with authenticated examples—Rule 901(b)(3)
- distinctive characteristics, including circumstantial evidence—Rule 901(b)(4)
- a system or process capable of proving a reliable and dependable result—Rule 901(b)(9)
- trade inscriptions—Rule 902(7)
- certified copies of a business record—Rule 902(11)
- certified records generated by an electronic process or system—Rule 902(13)
- certified data copied from an electronic device, storage medium, or file—Rule 902(14)

41. 698 F.3d at 998–1000.

In practice, the authentication and admissibility of text messages are handled just like email. A key question is often whether the purported sender actually sent the text, which is a subset of authentication. In other words, is the text what its proponent claims: a message sent by a named person to another person at a specific date and time. Absolute certainty is not required. For example, in a criminal prosecution for gun running, the government used circumstantial evidence to authenticate texts that were taken off an iPhone, which was in the defendant's possession at the time of his arrest, and a Samsung device found in his room.⁴² One phone listed the defendant's nickname—"Big Dave"—in the properties section, and both phones contained information in the contacts directory associated with the defendant, including the defendant's mother under the heading "Mom."⁴³ Moreover, the texts sent by him were non-hearsay admissions of a party opponent.⁴⁴ Similarly, in another case, the government authenticated text messages where a witness testified that although she was not certain that the defendant authored the messages, she had talked to him at the phone number that was the source of the texts, and the content indicated that they were from the defendant.⁴⁵

Texts can also present unique questions of collection and preservation. Unlike emails, texts do not ordinarily reside on an enterprise server, nor are they typically foldered or archived for long-term retention. Often the simplest way to facilitate preservation of messages is for users to harvest or collect them from their own smartphones. Recipients wishing to retain texts in a legal dispute have resorted to various means of preservation,

42. *United States v. Lewisbey*, 843 F.3d 653, 657–58 (7th Cir. 2016).

43. *Id.* at 658.

44. *Id.*

45. *United States v. Barnes*, 803 F.3d 209, 217 (5th Cir. 2015).

including cutting and pasting screenshots into emails or word-processing files that are then offered into evidence. These methods predictably elicit an authentication objection. As long as a witness with personal knowledge can testify as to the process used to generate the secondary document or image and assert that it accurately reflects the content of the text messages, courts have tended to find that authenticity was sufficiently established for the issue to go the jury.⁴⁶ Similarly, courts in these situations have not usually required the presentation of reliable chain-of-custody procedures or elaborate forensic processes.

3. Websites

“Websites are inherently changeable,” which can make them difficult to authenticate.⁴⁷ The most well-known approach to preserving web pages is the screen capture or variations on it, such as creating a PDF (portable document format) image or preserving a site through application programming interfaces (APIs). For static web pages—those that lack any interactive features or features personalized to the viewer, these methods might suffice; they do, at least, provide a view of what the web page looked like at that moment on that browser. However, it is easy to manually alter hypertext before capture or to manipulate PDF files and other screenshots after capture using software like Photoshop.⁴⁸ Moreover, API captures may miss significant

46. See *United States v. Arnold*, 696 F. App'x 903, 906–07 (10th Cir. 2017) (reflecting testimony from the witness who explained that he copied text messages into another document); *United States v. Ramirez*, 658 F. App'x 949, 952 (11th Cir. 2016) (memorializing testimony from a witness who indicated the photographs of text messages were pictures from her phone).

47. *Supermedia LLC v. Law Firm of Asherson*, No. 2:12-CV-03834, 2013 WL 12113386, at *3 (C.D. Cal. Feb. 13, 2013).

48. See, e.g., *Leidig v. BuzzFeed, Inc.*, No. 16 Civ. 542, 2017 WL 6512353, at *2 (S.D.N.Y. Dec. 19, 2017) (finding that the plaintiffs produced “documents

chunks of data, and many companies have withdrawn their APIs in response to data security threats and breaches.⁴⁹ Even so, if the court and parties can access the current version of the web page and it has not changed, then there is no authenticity issue.⁵⁰ But this is rarely the case given the dynamic nature of today's websites.⁵¹

Modern websites pose complicated authentication problems because no longer are they static pages of images and text. Today, 95 percent of websites incorporate JavaScript,⁵² a tool that developers use to create interactive web elements such as chat boxes, dropdown menus, and other personalized content. To ensure that this interactive website evidence remains admissible, something more than screenshots or PDF captures is required to view, preserve, and authenticate it.

Authentication issues typically include what the actual content of the web page was at a particular point in time, whether the exhibit or testimony accurately reflects this content and, if

bearing no metadata, including manually manipulated PDFs, summaries of underlying documents not produced, and screenshots and other text files").

49. See, e.g., Mike Schroepfer, *An Update on Our Plans to Restrict Data Access on Facebook*, FACEBOOK, (Apr. 4, 2018), <https://newsroom.fb.com/news/2018/04/restricting-data-access/>.

50. See *United States v. Bari*, 599 F.3d 176, 180 (2d Cir. 2010) (noting that a judge can conduct a "basic internet search" to confirm the authenticity of current website content).

51. See, e.g., *Adobe Sys. Inc. v. Christenson*, No. 2:10-cv-00422-LRH-GWF, 2011 WL 540278, at *9 (D. Nev. Feb. 7, 2011) ("Although Defendants can probably determine, with little difficulty, whether a *current* Google search for the search terms 'software surplus' provides links on the first page for the 'resellerratings.com' and 'Eopinions.com' websites, this would not prove that such a search would have resulted in such a link at a prior point in time.").

52. *Usage of JavaScript for Websites as client-side programming language on websites*, W3TECHS, <https://w3techs.com/technologies/details/cp-javascript> (last visited May 5, 2020).

so, whether the content is attributable to the site owner.⁵³ Alternatively, parties can authenticate a web page through the personal knowledge of a person who created or who maintains the website.⁵⁴

In addressing these evidentiary problems, the authentication rules most likely to apply include the following:

- a witness with personal knowledge—Rule 901(b)(1)
- expert testimony or comparison with authenticated examples—Rule 901(b)(3)
- distinctive characteristics, including circumstantial evidence—Rule 901(b)(4)
- a system or process capable of proving a reliable and dependable result—Rule 901(b)(9)
- certified records generated by an electronic process or system—Rule 902(13)
- certified data copied from an electronic device, storage medium, or file—Rule 902(14)

Typically, the witness will need to testify or certify that the witness typed in the web address at the date and time on an exhibit, that the witness reviewed the contents of the web page, and that the exhibit is a fair and accurate reflection of what the

53. *See* *Supermedia LLC v. Law Firm of Asherson*, No. 2:12-CV-03834, 2013 WL 12113386, at *3 (C.D. Cal. Feb. 13, 2013) (“A purported printout of the content of a website on a past date requires proof from someone with actual knowledge that the printout is in fact what would have been viewed if the website had been accessed at the stated time period.”).

54. *St. Luke’s Cataract & Laser Inst., P.A. v. Sanderson*, No. 8:06-cv-223-T-MSS, 2006 WL 1320242, at *2 (M.D. Fla. May 12, 2006) (finding that a webmaster’s testimony can authenticate a website printout).

witness saw.⁵⁵ The exhibit should include two things: the web page's internet address and the date and time the web page contents were downloaded.⁵⁶

A point of contention is "whether a website's owner, webmaster, or author is necessary to authenticate a web posting when its relevancy depends on its accuracy or its author."⁵⁷ In determining authenticity, courts may consider circumstantial evidence in determining whether the content of the website was posted by the site's owner under Rule 901(b)(4).⁵⁸ This evidence can include whether the website has a distinctive design or specific logos, photos, or images that are linked to the website or its

55. *See, e.g.,* SMS Audio, LLC v. Belson, No. 16-81308-CIV, 2107 WL 1533971, at *3 (S.D. Fla. Mar. 20, 2017) ("[C]ourts generally permit the authentication of web postings, bearing a web address and the date printed, by a witness who saw and printed the postings 'for the limited purpose of proving that the postings had appeared on the world wide web on the days that [the witness] personally saw the postings and printed them off the computer.'" (quoting Saadi v. Maroun, No. 8:07-cv-1976-T-24 MAP, 2009 WL 3736121, at *4 (M.D. Fla. Nov. 4, 2009)); Estate of Konell v. Allied Prop. & Cas. Ins. Co., No. 3:10-cv-955-ST, 2014 WL 11072219, at *1 (D. Or. Jan. 28, 2014) ("To authenticate a printout of a web page, the proponent must offer evidence that: (1) the printout accurately reflects the computer image of the web page as of a specified date; (2) the website where the posting appears is owned or controlled by a particular person or entity; and (3) the authorship of the web posting is reasonably attributable to that person or entity.").

56. *See, e.g.,* Foreword Magazine, Inc. v. OverDrive, Inc., No. 1:10-cv-1144, 2011 WL 5169384, at *3 (W.D. Mich. Oct. 31, 2011) (admitting website screenshots based on an attorney's sworn affidavit plus "other indicia of reliability (such as the Internet domain address and the date of printout)").

57. *SMS Audio, LLC*, 2017 WL 1533971, at *4; *see also* United States v. Browne, 834 F.3d 403, 413–15 (3d Cir. 2016) (ruling that Facebook chats are sufficiently authenticated by circumstantial evidence that the defendant was the author), *cert. denied*, 137 S. Ct. 695 (2017).

58. *See* Hon. Paul W. Grimm, et al., *Authenticating Digital Evidence*, 69 BAYLOR L. REV. 1, 26 (2017).

owner.⁵⁹ Courts may also evaluate whether the contents of the proffered web pages are of the kind typically posted on similar websites, whether the site owner wholly or partially published the website content elsewhere, whether the contents have been otherwise republished elsewhere and attributed to the proffered website, or the length of time that the website content was posted.⁶⁰

Another popular—if limited—method of authentication is the Wayback Machine. Launched in 2001 by the nonprofit Internet Archive, the Wayback Machine is a digital archive of the web. Courts have occasionally taken judicial notice of the contents of these archived sites.⁶¹ Some courts have permitted an Internet Archive witness to testify about the reliability of the Wayback Machine’s results under 901(b)(9).⁶² Now, the reliability of the Wayback Machine process may be established by a certificate of an Internet Archive official under Rule 902(13).

Although the Wayback Machine captures information, what it actually memorializes is inconsistent. The archive may not

59. *See, e.g.,* Metcalf v. Blue Cross Blue Shield of Mich., No. 3:11-cv-1305-ST, 2013 WL 4012726, at *10 (D. Or. Aug. 5, 2013) (finding that authenticity of website information of an organization’s purported website was established by logos or headers matching those of the organization), *cited in* Grimm, et al., *supra* note 58, at 26.

60. *See* Grimm, et al., *supra* note 58, at 26.

61. *See, e.g.,* Under a Foot Plant, Co. v. Exterior Design, Inc., No. 6:14-cv-01371-AA, 2015 WL 1401697, at *2 (D. Or. Mar. 24, 2015) (“District courts have routinely taken judicial notice of content from The Internet Archive . . .”).

62. *See, e.g.,* Specht v. Google Inc., 747 F.3d 929, 933 (7th Cir. 2014) (“[T]he district court reasonably required . . . authentication by someone with personal knowledge of reliability of the archive service from which the screenshots were retrieved.”); Open Text S.A. v. Box, Inc., No. 13-cv-04910-JD, 2015 WL 428365, at *2 (N.D. Cal. Jan. 30, 2015) (refusing to admit a Wayback Machine screenshot into evidence without testimony from an Internet Archive representative confirming its authenticity).

capture all of a website's content. Moreover, users can ask that the archive delete or change information. This led at least one court to find that a party could not show that data from the archive was "reliable, complete, and admissible in court."⁶³ As a result, the Wayback Machine is not accepted as a forensic evidence collection method.⁶⁴

The ISO 28500 WARC (Web ARChive) standard, established by the International Internet Preservation Consortium, addresses authentication issues by making it possible to obtain an exact native file of the collected content of a website.⁶⁵ A WARC file is a container for all accessed web resources and metadata; it is a collection of records, each of which relates to an element of a web page. A web crawler or similar program captures the data, stores the data in a WARC file, and generates relevant metadata about the capture that confirms the data's integrity. The saved data is an identical replica of the website, with working links, graphics, and other dynamic content. The saved website also records every possible server request and the answer to that request, along with all of the supporting metadata to establish the authenticity of its information. Some software timestamps and hashes each event in the collection, simplifying the process of establishing a chain of custody and facilitating authentication.⁶⁶

63. See *Leidig v. BuzzFeed, Inc.*, No. 16 Civ. 542, 2017 WL 6512353, at *13 (S.D.N.Y. Dec. 19, 2017).

64. *Id.*

65. International Organization for Standardization, *ISO 28500:2017: Information and Documentation— WARC File Format*, <https://www.iso.org/standard/68004.html> (last visited May 5, 2020).

66. For example, Hanzo Archives offers a WARC native file copy of web content with its Preserve service. See Hanzo Archives, *eDiscovery for the Interactive Age*, <https://www.hanzo.co/ediscovery-software-0> (last visited May 9, 2020).

For certain websites, authentication is a simpler matter. Three types of website evidence are self-authenticating under Rule 902. Under Rule 902(5), federal, state, local, and international government websites are self-authenticating, and courts typically take judicial notice of these sites.⁶⁷ Under Rule 902(6), online newspapers and periodicals are self-authenticating.⁶⁸ Finally, business records kept in the ordinary course of business that satisfy Rule 803(6) are self-authenticating.⁶⁹

Courts may also take judicial notice of other reputable websites, such as internet maps,⁷⁰ calendars,⁷¹ the publication of articles in newspapers and periodicals,⁷² and online versions of textbooks, dictionaries, rules, and charters.⁷³ Note that courts

67. *See, e.g., Williams v. Long*, 585 F. Supp. 2d 679, 686–88 & n.4 (D. Md. 2008) (collecting cases indicating that posts on government websites are self-authenticating).

68. *See, e.g., White v. City of Birmingham*, 96 F. Supp. 3d 1260, 1274 (N.D. Ala. 2015) (noting that online news articles are “analogous to traditional newspaper articles and could be found self-authenticating at trial”).

69. *See, e.g., United States v. Hassan*, 742 F.3d 104, 132–34 (4th Cir. 2014) (finding social media posts, including links to videos, were self-authenticating under Rule 902(11) where accompanied by “certifications of records custodians of Facebook and Google, verifying that the Facebook pages and YouTube videos had been maintained as business records in the course of regularly conducted business activities.”). *See* Section II.D.1.a, *supra*.

70. *See, e.g., United States v. Burroughs*, 810 F.3d 833, 835 n.1 (D.C. Cir. 2016) (granting a motion to take judicial notice of a Google map).

71. *See, e.g., Tyler v. United States*, No. 1:08-CR-165-CC & No. 1:11-LV-4592-CC, 2012 WL 6808525, at *3 n.6 (N.D. Ga. Dec. 6, 2012).

72. *See, e.g., Ford v. Artiga*, No. 2:12-CV-02370, 2013 WL 3941335, at *7 n.5 (E.D. Cal. July 30, 2013) (taking judicial notice of the fact of publication but not of the articles’ content).

73. *See, e.g., Williams v. Emp’rs Mut. Cas. Co.*, 845 F.3d 891, 905 (8th Cir. 2017) (taking judicial notice of a dictionary); *Morgan Stanley Smith Barney LLC v. Monaco*, No. 14-cv-00275-RM-MJW, 2014 WL 5353628, at *2 (D. Colo. Aug. 26, 2014) (taking judicial notice of FINRA rules).

have declined to accord the same courtesy to the crowdsourced Wikipedia, finding it “not sufficiently reliable.”⁷⁴

4. Social Media Sites

(a) What is social media?

“Social media” is a broad and imprecise term encompassing a range of platforms, applications, and tools that permit users to share information with others, typically in an internet-based environment.⁷⁵ Since their introduction in the early 2000s, social media applications and platforms have been constantly changing and expanding. Although even the traditional platforms differ from site to site, their basic feature is social networking—the ability to connect with other people and share content.⁷⁶ Platforms like Facebook, Twitter, and LinkedIn allow people to “friend,” “follow,” or “retweet” each other and to share comments, photos, videos, and events. YouTube, Snapchat, and Instagram provide for similar social interaction, with the focus on sharing photos and videos. Dating apps like Tinder, Bumble, and Grindr also provide opportunities for online (and real life) social connection.

Social media has expanded into territory previously occupied by SMS text messaging. Over-the-top (OTT) messaging applications use the internet and travel directly from device to device instead of going through servers belonging to SMS

74. See, e.g., *Blanks v. Cate*, No. 2:11-cv-0171 WBS CKD P., 2013 WL 322881, at *3 n.3 (E.D. Cal Jan. 28, 2013). *But see* *United States v. Bazaldua*, 506 F. 3d 671, 673 n.2 (8th Cir. 2007) (court took judicial notice of an article in Wikipedia).

75. See *The Sedona Conference, Primer on Social Media, Second Edition*, 20 SEDONA CONF. J. 1, 10 (2019); Hon. Paul W. Grimm et al., *Authentication of Social Media Evidence*, 36 AM. J. TRIAL ADVOC. 433, 434 (Spring 2013).

76. See *Primer on Social Media*, *supra* note 75, at 10.

providers. Examples of OTT messaging applications include WhatsApp, Facebook Messenger, iMessage, Snapchat, and Kik.⁷⁷ Some messaging applications also give the user the ability to be anonymous or to send messages that will self-destruct.⁷⁸

More recent additions to social media include applications for cloud-based messaging, collaboration applications, live-streaming video, health information sharing, wearable technologies, and location-based platforms.⁷⁹

(b) Social Media Content as Evidence

It was not long after the advent of social media that participants in the justice system recognized it as a source of evidence. A Facebook comment could be an admission of a crime. A photo of a criminal defendant with known gang members could tend to show gang affiliation. A video of someone dancing exuberantly at his daughter's wedding reception could undermine a personal injury claim, the need for workers compensation, or long-term disability payments.⁸⁰

The recognition of social media's evidentiary value also gave rise to admissibility challenges. These issues have arisen mostly in the authentication arena: whether the social media post, photo, video, message, or comment is what the proponent claims it to be.

77. *Id.* at 13–14.

78. *Id.* at 14–15; *see* Sect. II.D.8 (Digital Photographs), *infra*.

79. *Primer on Social Media*, *supra* note 75, at 15–20.

80. It is worth noting, however, that the vast majority of cases dealing with the admissibility of social media evidence are criminal in nature.

Social media evidence can come in a variety of forms. Often it will be presented in the form of screenshots or printouts.⁸¹ Photos and videos can be downloaded in their native formats.⁸² Content available through websites can be preserved through APIs.⁸³ Social media evidence can also be gathered using individual platform download tools.⁸⁴ Social media content also may contain metadata that might be relevant in legal disputes.⁸⁵

81. See, e.g., *Hawkins v. State*, No. S18A0886, 2018 WL 3965665, at *4 (Ga. Aug. 20, 2018); *State v. Jones*, No. 109,027, 2014 WL 802022, at *4 (Kan. Ct. App. Feb. 28, 2014).

82. See, e.g., *United States v. Farrad*, 895 F.3d 859, 875–76 (6th Cir. 2018); *Lamb v. State*, 246 So. 3d 400, 404–05 (Fla. Dist. Ct. App. 2018).

83. See Sect. II.D.3 (Websites), *supra*.

84. See *How to Access Your Twitter Data*, TWITTER, <https://help.twitter.com/en/managing-your-account/accessing-your-twitter-data> (last visited May 5, 2020); *Accessing & Downloading Your Information*, FACEBOOK, https://www.facebook.com/help/1701730696756992/?helpref=hc_fnav (last visited May 5, 2020); see also Katie Canales, *Instagram is rolling out a feature that will let you download all of your photos and past searches in one fell swoop*, BUS. INSIDER (Apr. 24, 2018, 5:48 PM), <https://www.businessinsider.com/instagram-data-download-feature-gdpr-privacy-photos-searches-2018-4>; Abby Ohlheiser, *Here's how to download all your data from Facebook. It might be a wake-up call*, WASH. POST (Mar. 27, 2018, 9:23 a.m.), https://www.washingtonpost.com/news/the-intersect/wp/2018/03/27/heres-how-to-download-all-your-data-from-facebook-it-might-be-a-wake-up-call/?utm_term=.1b84ec6553f2; see, e.g., *Ehrenberg v. State Farm Mut. Auto. Ins. Co.*, No. 16-17269, 2017 WL 3582487, at *3 n.2 (E.D. La. Aug. 18, 2017) (refusing to decide whether request seeking plaintiff's Facebook, Twitter, and Instagram accounts via "data link" was appropriate).

85. See *In re Adoption of Nash*, No. 15-P-1302, 2016 WL 2755864, at *3 (Mass. App. Ct. May 12, 2016) (holding Facebook messages were not authenticated based on metadata review that could not link them to mother).

(c) Authentication of Social Media Evidence

Generally, the standard for authentication of evidence, whether under Rule 901 and or its state counterparts, is low.⁸⁶ To authenticate evidence, “the proponent must produce evidence sufficient to support a finding that the item is what the proponent claims it is.”⁸⁷ This may be shown by either direct or circumstantial evidence.⁸⁸ A *prima facie* case is all that is necessary.⁸⁹

In addressing these evidentiary problems, the authentication rules most likely to apply include the following:

- a witness with personal knowledge—Rule 901(b)(1)
- expert testimony or comparison with authenticated examples—Rule 901(b)(3)
- distinctive characteristics, including circumstantial evidence—Rule 901(b)(4)

86. *United States v. Barnes*, 803 F.3d 209, 217 (5th Cir. 2015) (stating that the authentication standard is not a burdensome one); *United States v. Vayner*, 769 F.3d 125, 130 (2d Cir. 2014) (“bar for authentication of evidence is not particularly high”); *United States v. Hassan*, 742 F.3d 104, 133 (4th Cir. 2014) (“the burden to authenticate under Rule 901 is not high”); *United States v. Ortiz*, 966 F.2d 707, 716 (1st Cir. 1992) (901(a) “does not erect a particularly high hurdle”); *State v. Newman*, 916 N.W.2d 393, 409 (Neb. 2018) (authentication statute “does not impose a high hurdle for authentication or identification”); *State v. Adams*, 161 A.3d 1182, 1199 (R.I. 2017) (“authentication is not a high hurdle to clear”); *see also* Grimm et al., *supra* note 75, at 458.

87. FED. R. EVID. 901(a).

88. *Vayner*, 769 F.3d at 130; *Tienda v. State*, 358 S.W.3d 633, 638 (Tex. Crim. App. 2012) (“Evidence may be authenticated in a number of ways, including by direct testimony from a witness with personal knowledge, by comparison with other authenticated evidence, or by circumstantial evidence.”).

89. *Stout v. Jefferson Cty. Bd. of Educ.*, 882 F.3d 988, 1008 (11th Cir. 2018); *Hassan*, 742 F.3d at 133.

- a system or process capable of proving a reliable and dependable result—Rule 901(b)(9)
- certified records generated by an electronic process or system—Rule 902(13)
- certified data copied from an electronic device, storage medium, or file—Rule 902(14)

While these basic authentication standards have never changed, social media evidence nevertheless appeared to drive some courts to raise the evidentiary bar.⁹⁰ Commentators noted that courts appeared to fall into two camps.⁹¹ In the beginning, courts were openly skeptical of social media and concerned about the possibility of forgery, falsification, and impersonation.⁹² Other courts did not appear to share this skepticism and kept the bar low.⁹³ The low-bar approach was exemplified by courts that articulated a “reasonable jury” standard—authentication was shown if there was sufficient direct or circumstantial evidence to allow a reasonable jury to find that the evidence is what it is purported to be.⁹⁴

More recently, some courts in the high-bar camp appear to have softened.⁹⁵ This is in line with other cases that show a growing comfort level among attorneys, litigants, and judges

90. See *Primer on Social Media*, *supra* note 75.

91. *Id.* See generally Grimm et al., *supra* note 75; Wendy Angus-Anderson, *Authenticity and Admissibility of Social Media Website Printouts*, 14 DUKE L. & TECH. REV. 33 (2015).

92. See, e.g., *Griffin v. State*, 19 A.3d 415, 422 (Md. 2011); *Smith v. State*, 136 So. 3d 424, 432 (Miss. 2014); see also Grimm et al., *supra* note 75, at 441–49.

93. See *id.* at 449–54.

94. See, e.g., *Tienda v. State*, 358 S.W.3d 633, 638 (Tex. Crim. App. 2012).

95. See *Sublet v. State*, 113 A.3d 695, 712–18 (Md. 2015) (distinguishing *Griffin* and applying a “reasonable juror” standard articulated in *United States v. Vayner*, 769 F.3d 125 (2d Cir. 2014), *Tienda*, and *United States v. Hassan*, 742 F.3d 104 (4th Cir. 2014)).

with the use of social media evidence.⁹⁶ The picture today is not so much one of division among courts based on different legal standards, but one of different outcomes based on different facts.⁹⁷

Turning to the examples of authentication evidence in Rule 901(b), the typical or most likely to be used, whether alone or in combination, are 901(b)(1) (testimony of a witness with knowledge) and 901(b)(4) (distinctive characteristics).⁹⁸ Authentication can also be satisfied under 901(b)(3) by comparison to an already authenticated specimen by either an expert or the trier of fact.⁹⁹

The issue of authorship and identity is usually critical because the identity of the author, creator, or owner of social media evidence is often essential to its relevance and its admissibility. It is in this context that judicial suspicions about the integrity of social media evidence are most evident, driven by the

96. See KENNETH S. BROUN ET AL., MCCORMICK ON EVIDENCE § 227 (Robert P. Mosteller ed., 7th ed. 2016) (“[T]he approach by courts imposing a heavier burden on social networking evidence is reminiscent of the conservative response many courts had to the advent of other technologies such as the telegraph, the computer, and the internet. With time the trend may well shift towards the second category of cases as courts become more familiar with the social networking medium and the perceived dangers of this evidence dissipate. Given that many of the cases taking a lenient approach to social networking evidence have arisen in only the last two to three years, this shift may already be occurring.”).

97. See *id.* (“Despite the seeming novelty of social network-generated documents, courts have applied the existing concepts of authentication under Federal Rule 901 to them.”).

98. See *id.* at 545–47; *People v. Glover*, 363 P.3d 736, 741 (Colo. App. 2015).

99. See *Lorraine v. Markel Am. Ins. Co.*, 241 F.R.D. 534, 556; Patrick Marshall, What You Say on Facebook May Be Used Against You in a Court of Family Law: Analysis of This New Form of Electronic Evidence and Why It Should Be on Every Matrimonial Attorney’s Radar, 63 ALA. L. REV. 1115, 1129 (2012).

perception that social media is more susceptible to forgery or falsification than hard-copy evidence.¹⁰⁰ The Mississippi Supreme Court described the issue this way:

Not only can anyone create a profile and masquerade as another person, but such a risk is amplified when a person creates a real profile without the realization that third parties can mine their personal data. . . . Thus, concern over authentication arises because anyone can create a fictitious account and masquerade under another person's name or can gain access to another's account by obtaining the user's username and password, and, consequently, the potential for fabricating or tampering with electronically stored information on a social networking [website] is high and poses challenges to authenticating printouts from the website.¹⁰¹

When authorship is critical to the admissibility of social media evidence, courts have required "direct or circumstantial evidence that tends to corroborate the identity of the author of the communication in question."¹⁰² This may include "testimony

100. See *Commonwealth v. Mangel*, 181 A.3d 1154, 1162–64 (Pa. Super. Ct. 2018) (trial court did not abuse its discretion in denying Commonwealth's motion *in limine* to admit social media posts and messages based in part on the concern about the ease with which social media accounts may be falsified or a legitimate account accessed by an imposter).

101. *Smith v. State*, 136 So. 3d 424, 432 (Miss. 2014) (internal citations and quotations omitted); see also *Mangel*, 181 A.3d at 1162 (raising similar concerns). For further discussion and cases see Section V.C.2 (State law application in federal cases), *infra*.

102. *Mangel*, 181 A.3d at 1162; see also *Glover*, 363 P.3d at 742; *United States v. Recio*, 884 F.3d 230, 236–37 (4th Cir. 2018) (authenticating Facebook posts through circumstantial evidence).

from the person who sent or received the communication, or contextual clues in the communication tending to reveal the identity of the sender.”¹⁰³ Authorship of social media evidence is subject to authentication by the same “wide range of extrinsic evidence”¹⁰⁴ as traditional hard-copy evidence. But courts have still held that the proponent need not absolutely prove authorship.¹⁰⁵

Not all social media evidence, however, presents an issue of identity or authorship. In some cases, courts have appeared to require either a lesser quantum of evidence, or no evidence, pertaining to the authorship or identity.¹⁰⁶ This is often seen in the admission of photos and videos posted to social media.¹⁰⁷

In *Lamb v. State*, the Florida court permitted the introduction of a Facebook live video that purported to show the defendant driving the stolen vehicles.¹⁰⁸ The video had been posted to a co-defendant’s public Facebook page and downloaded by a “digital forensic examiner” who simply visited the page. Beyond the examiner’s testimony as to how he downloaded the video, the only other evidence was the testimony from two witnesses who

103. *Mangel*, 181 A.3d at 1162.

104. *United States v. Browne*, 834 F.3d 403, 411–12 (3d Cir. 2016), *cert. denied*, 137 S. Ct. 695 (2017).

105. *See Gagliardi v. Comm’r of Children & Families*, 110 A.3d 512, 518 (Conn. App. Ct. 2015) (only need to make a *prima facie* showing of authenticity and “once a *prima facie* showing of authorship is made to the court, the evidence, as long as it is otherwise admissible, goes to the [finder of fact], which ultimately will determine its authenticity.”).

106. *Beaty v. State*, No. 03-16-00856-CR, 2017 WL 5560078 at *4–5 (Tex. App. Nov. 15, 2017).

107. *See, e.g., United States v. Broomfield*, 591 F. App’x 847, 852 (11th Cir. 2014); *Lamb v. State*, 246 So. 3d 400, 409 (Fla. Dist. Ct. App. 2018); *State v. Gray*, No. 2016-KA-1195, 2017 WL 3426021, at *15–16 (La. Ct. App. June 28, 2017).

108. 246 So. 3d at 409 (Fla. Dist. Ct. App. 2018).

watched the video and identified the defendant as being in the video. This was a sufficient *prima facie* showing of authenticity.

The court cited the Eleventh Circuit for not requiring more authentication evidence:

[T]he Eleventh Circuit and other courts . . . have permitted the admission of social media videos in criminal cases based on sufficient evidence that the video depicts what the government claims, even though the government did not: (1) call the creator of the videos; (2) search the device which was used to create the videos; or (3) obtain information directly from the social media website. *See, e.g., U.S. v. Washington*, 2017 WL 3642112, *2 (N.D. Ill. Aug. 24, 2017) (YouTube video which the government contended showed the defendant and several other men pointing firearms at the camera was sufficiently authenticated where law enforcement witness would testify that he watched this video on YouTube, recognized the defendant, and downloaded the video); *State v. Gray*, — So.3d — —, —, —, 2017 WL 3426021, *16 (La. Ct. App. June 28, 2017) (YouTube videos were sufficiently authenticated where the investigating officer's testimony provided sufficient support that the videos were what the state claimed them to be, that is, videos depicting the defendant and other gang members in a park and surrounding area). As the *Washington* court stated, "[w]hile a witness with [knowledge of the video's creation] could authenticate [the] video, Rule 901 does not require it." 2017 WL 3642112 at *2.¹⁰⁹

109. *Id.* at 409–10.

The relevance of the video did not depend on who created the video or even who posted the video, even though it purportedly came from a co-defendant's Facebook page. Its relevance was in its content—that it depicted someone identified as the defendant with the stolen vehicle. In this respect, the Facebook Live video in *Lamb* was essentially no different than any other video.

In *Commonwealth v. Martin*, the Pennsylvania Superior Court distinguished the *Mangel* decision discussed above (which required evidence to tie social media messages to an individual) and held that Instagram posts depicting the defendant did not require evidence that he had made the posts.¹¹⁰ In addition, the issue did not depend on whether the defendant made the posts, but on whether they accurately portrayed the defendant.¹¹¹

Similarly, in *United States v. Thomas*, the Sixth Circuit upheld the admission of photos downloaded by law enforcement from Facebook and Instagram pages using a version of the name “Jabron Thomas,” the same name as the defendant.¹¹² Thomas argued the photos were inadmissible because there was no evidence of who created the Facebook page or whether the page itself was authentic.

The court set out some hypotheticals to illustrate the authentication issue posed:

In many contexts, the question could conceivably be quite interesting: what if, for example, the owner of a social-media profile (let's call him Alex) used a picture of someone else (say, Bob) as his profile picture? If Bob robbed a bank, Alex

110. No. 1962 MDA 2016, 2018 WL 3121766, at *9 (Pa. Super. Ct. June 26, 2018) (non-precedential decision).

111. *Id.*

112. 701 F. App'x 414, 419 (6th Cir. 2017).

would not want to be implicated as the robber simply because he had Bob's picture on his social-media profile. Or, what if Bob fabricated a social-media profile under Alex's name, but with Bob's picture—and then Bob robbed a bank? Or, less convolutedly, what if there were allegations that the online photographs had been digitally manipulated or hacked in some way?¹¹³

But the court concluded that those questions weren't before it. Instead, the court saw "no reason to depart from the ordinary rule that photographs, including social-media photographs, are authenticated by 'evidence sufficient to support a finding that the [photograph] is what the proponent claims it is,' Fed. R. Evid. 901(a)."¹¹⁴ As with *Lamb*, it was what was depicted in the photos, not necessarily who took them or to what social media site they were posted, that was relevant. The photos were offered to identify Thomas—they showed his distinctive tattoos on his hands and arms and that he was wearing Detroit Tigers gear similar to the hat worn by the robber.¹¹⁵

(d) Business Records

When social media posts or profiles are offered into evidence, Rule 902(11) may be unavailable because the evidence may not qualify as a business record.¹¹⁶ Posts by users or user profiles are often not business activities—they are not records

113. *Id.*

114. *Id.*

115. *Id.*; see also *Beaty v. State*, No. 03-16-00856-CR, 2017 WL 5560078, at *4 (Tex. App. Nov. 15, 2017) (holding that Facebook photos offered to show defendant's clothing and appearance at the time of the shooting did not demand proof of identify of person who created the photos or the social media post).

116. *People v. Glover*, 363 P.3d 736, 741-42 (Colo. App. 2015).

that the social media site would use or rely on for a business purpose. Instead, they are declarations from the individuals who posted the information. As such, they are not usually admissible business records.¹¹⁷

(e) Other Social Media Admissibility Challenges

Authentication, however, does not guarantee admissibility. As with all evidence, to be admissible, social media evidence must also be relevant,¹¹⁸ not inadmissible hearsay,¹¹⁹ and not unduly prejudicial, confusing, cumulative, or misleading.¹²⁰ Some

117. See *United States v. Farrad*, 895 F.3d 859, 878–79 (6th Cir. 2018); *United States v. Browne*, 834 F.3d 403, 434–35 (3d Cir. 2016), *cert. denied*, 137 S. Ct. 695 (2017). *But see* *United States v. Recio*, 884 F.3d 230, 237–38 (4th Cir. 2018). In *Recio*, the Fourth Circuit found that authentication was achieved through a certification of a Facebook records custodian showing that the Facebook user in question had made the post at or near the time showed by the post. This was in addition to other (strong) evidence tying the defendant to the account, including that the name on the account was the same as the defendant, “Larry Recio”; an email address associated with the account was larryrecio20@yahoo.com; the defendant appeared in over 100 photos posted to the account; and one photo included the caption “Happy Birthday Larry Recio.” *Id.* at 237.

118. FED. R. EVID. 402; *Recio*, 884 F.3d at 235–36 (holding that a lyric posted on Facebook was relevant because it matched the details of the alleged crime and illustrated the defendant’s motive).

119. FED. R. EVID. 802; *Recio*, 884 F.3d at 234–35 (holding that a lyric posted on Facebook was admissible as an adoptive admission under Fed. R. Evid. 801(d)).

120. FED. R. EVID. 403; *Recio*, 884 F.3d at 236 (holding that the probative value of admitting a lyric posted on Facebook outweighed the risk of undue prejudice); *United States v. Khoa*, No. 17-4518, 2018 WL 2905432, at *3 (4th Cir. 2018) (holding that photos of victim posted to social media were not unduly prejudicial under Fed. R. Evid. 403).

courts have also applied the “best evidence” rule to social media evidence.¹²¹

5. Internet of Things

The Internet of Things (IoT) is a network of computing devices and sensors embedded in everyday objects that create, collect, and share data through the internet. Some examples include wearables that track our steps and sleep, appliances that track our consumption, and thermostats that adjust to our habits. The data that these devices create is often stored in structured databases and may be stored in multiple locations in the cloud.

IoT data is already playing a significant role in cases. For example, in one murder case, data indicating movement from a wife’s fitness wearable convinced the police that her husband killed her.¹²² In another, prosecutors used Fitbit data to show that a victim falsely accused a man of raping her.¹²³

The risk that IoT data could be manipulated should not bar this evidence entirely. In the best-case scenario, the wearer or owner of an IoT device can testify to authenticate the device and its data (and metadata) as a witness with personal knowledge under Rule 901(b)(1). Any analysis of the data would need to undergo a separate process to authenticate the data produced and its accuracy using 901(b)(3) (expert testimony), 901(b)(4)

121. See, e.g., *Woods v. State*, No. 11-15-00134-CR, 2017 WL 3711104, at *6 (Tex. App. Aug. 25, 2017) (holding that Facebook posts satisfied best evidence rule).

122. Christine Hauser, *In Connecticut Murder Case, a Fitbit Is a Silent Witness*, N.Y. TIMES (Apr. 27, 2017), <https://www.nytimes.com/2017/04/27/nyregion/in-connecticut-murder-case-a-fitbit-is-a-silent-witness.html>.

123. Jacob Gershman, *Prosecutors Say Fitbit Device Exposed Fibbing in Rape Case*, WALL ST. J.: L. BLOG (Apr. 21, 2016, 1:53 PM), <https://blogs.wsj.com/law/2016/04/21/prosecutors-say-fitbit-device-exposed-fibbing-in-rape-case/>.

(distinctive characteristics, including circumstantial evidence), 901(b)(9) (system or process capable of proving a reliable and dependable result), 902(13) (certified records generated by an electronic process or system), or 902(14) (certified data copied from an electronic device, storage medium, or file).

6. Ephemeral or Self-Destructing Photographs/Messages

Since the release of Snapchat in September 2011, the use of self-destructing messaging (also referred to as “ephemeral messaging”) has increased exponentially. In 2019, over 200 million people were using Snapchat, creating over 3.5 billion snaps each day.¹²⁴ Additional ephemeral messaging providers have emerged, including Wickr,¹²⁵ Telegram,¹²⁶ Confide,¹²⁷ and Signal.¹²⁸ The default setting in ephemeral messaging applications is for messages and images to self-destruct after a limited amount of time.¹²⁹ Some applications claim to be “screen-shot

124. SnapChat Revenue and Usage Statistics (2020), BUSINESS OF APPS (Apr. 24, 2020), <https://www.businessofapps.com/data/snapchat-statistics/>.

125. WICKR, <https://wickr.com/> (last visited May 6, 2020).

126. TELEGRAM, <https://telegram.org/> (last visited May 6, 2020).

127. CONFIDE, <https://getconfide.com/> (last visited May 6, 2020).

128. SIGNAL, <https://signal.org/en/> (last visited May 6, 2020).

129. *When does Snapchat delete Snaps and Chats*, SNAPCHAT, <https://support.snapchat.com/en-US/a/when-are-snaps-chats-deleted> (last visited May 6, 2020); *see also Features*, CONFIDE, <https://getconfide.com/> (“Messages disappear forever after they are read once, making them as private and secure as the spoken word.”) (last visited May 6, 2020); *Set and manage disappearing messages*, SIGNAL, <https://support.signal.org/hc/en-us/articles/360007320771-Set-and-manage-disappearing-messages> (“Use disappearing messages to keep your message history tidy. The message will disappear from your devices after the timer has elapsed.”) (last visited May 6, 2020). What sets these applications apart from SMS text messaging or OTT messaging applications is their ability to automate the destruction of content on the sender’s and the recipient’s devices. Another key aspect of ephemeral messaging is endpoint encryption of messages, which ostensibly prevents third parties from gaining

proof,” and one even requires the receiver to scroll over redacted text with a finger to briefly unredact the text before it is permanently deleted.¹³⁰ Although not in the context of authentication or admissibility, ephemeral communications figured prominently in discovery disputes in recent trade secret matters.¹³¹

Given that Snapchat is currently one of the most prevalent ephemeral messaging applications, this *Commentary* analyzes authentication issues through Snapchat. In 2020, 78 percent of internet users aged 18 to 24 used Snapchat, with 71 percent of those users accessing the platform daily.¹³²

Over time, Snapchat has evolved to allow users to save “snaps” as memories so that they do not self-destruct.¹³³ In those

access to message content. Philip Favro, *Ephemeral Messaging: Balancing the Benefits and Risks*, PRACTICAL LAW THE JOURNAL: LITIGATION (June/July 2019).

130. See *Features: Screenshot-Proof*, CONFIDE, <https://getconfide.com/> (“For extra privacy on iOS and Android, our patented reading experience ensures that only one line of the message is unveiled at a time and that the sender’s name is not simultaneously visible.”) (last visited May 6, 2020). Use of such technology would present some interesting authentication challenges in court. Message recipients could film the temporary unredaction of a message with a second device while scrolling their finger over the text, avoiding the first layer of screen-shot protection, but with the sender’s name invisible, there would be one less piece of information tying the message to the sender. But if the recipient was able to authenticate the video of the message, it might still be authenticated under the right facts, much like other electronic messages.

131. *WeRide Corp. v. Kun Huang*, No. 5:18-cv-07233, 2020 WL 1967209 (N.D. Cal. Apr. 24, 2020); *Waymo LLC v. Uber Techs., Inc.*, No. 3:17-cv-00939, 2018 WL 646701, at *21 (N.D. Cal. Jan. 30, 2018); *Waymo LLC v. Uber Techs., Inc.*, 3:17-cv-00939, 2018 WL 6501798, at *6–8 (N.D. Cal. Dec. 15, 2017).

132. See *Snapchat by the Numbers: Stats, Demographics & Fun Facts*, OMNICORE (Feb. 7, 2020), <https://www.omnicoreagency.com/snapchat-statistics/>.

133. See, e.g., *How to Use Memories*, SNAPCHAT, <https://support.snapchat.com/en-US/a/about-memories> (last visited May 6, 2020). Snapchat is

situations, the “memories” are like any other social media posts. Thus, parties would need to authenticate snaps or analogous content from other ephemeral messaging applications in the same way.

Self-destructing snaps may need to be handled differently. Snaps that disappear have not necessarily been erased once Snapchat deletes them. A receiver of a snap can save the snap by taking a screenshot of the snap, taking a photograph of the screen, or using image-capture software or apps. A Snapchat user can adjust the privacy settings to determine who can send snaps to the user and who can view the user’s “story” (other saved content on a user’s application). If a recipient chooses to “screenshot” or “screen capture” a photo before it disappears, Snapchat will notify the sender that the recipient took a screenshot of the snap.¹³⁴ These types of saved snaps are likely to be authenticated using 901(b)(1) (personal knowledge) or 902(14) (certified data copied from device). Snaps saved in this manner are likely to be treated similarly to digital photographs or videos.

There is limited case law discussing the authentication of Snapchat messages. In one criminal matter, a defendant sought appellate review of a trial court order that admitted a video shared through Snapchat.¹³⁵ During the trial, two witnesses who had contemporaneously viewed the snaps testified that the videos played in the courtroom were the same videos posted to the defendant’s account. One of the witnesses also remembered a

used as an example. The technology evolves rapidly and changes quicker than articles about technology.

134. Henry T. Casey & David Murphy, *How to Use the New Snapchat Like a Pro*, TOM’S GUIDE (Sept. 25, 2018), <https://www.tomsguide.com/us/snapchat-tutorial,news-21216.html>.

135. *Schaffer v. State*, No. 238, 2017, 2018 WL 1747793, at *1 (Del. Apr. 10, 2018).

caption on the video referencing the victim being scared. The defendant argued such testimony was insufficient to authenticate the video because the witnesses could not remember exactly when they watched the video and that the video apparently did not have a time stamp. The Delaware Supreme Court rejected the defendant's argument.¹³⁶

7. Digitally Stored Data

The mere fact that information has been created and stored within a computer system does not make that information reliable or authentic. Electronic records are most frequently authenticated under Rule 901(b)(4), which permits authentication by “[a]pppearance, contents, substance, internal patterns, or other distinctive characteristics, taken in conjunction with circumstances.”¹³⁷ The primary authenticity issue in the context of computer-stored records and databases is chain of custody.

The methods of authentication most likely to be appropriate for computerized records are as follows:

- a witness with personal knowledge—Rule 901(b)(1)
- expert testimony or comparison with authenticated examples—Rule 901(b)(3)
- distinctive characteristics, including circumstantial evidence—Rule 901(b)(4)
- a system or process capable of proving a reliable and dependable result—Rule 901(b)(9)
- certified records generated by an electronic process or system—Rule 902(13)

136. *Id.* at *6 (observing as well that the defendant's arguments went “to the appropriate weight to be given the evidence, not its admissibility.”).

137. *Lorraine v. Markel Am. Ins. Co.*, 241 F.R.D. 534, 546 (D. Md. 2007).

- certified data copied from an electronic device, storage medium, or file—Rule 902(14)

8. Digital Photographs

Historically, photographs were authenticated by the person taking the photograph or the person who witnessed the event who can show that a photograph is a fair and accurate representation of the scene depicted.¹³⁸ However, when photographs were captured on film, there were fewer photographs, and it was much more difficult to alter or manipulate the photographs. Today, digital photographs are ubiquitous—both through cell phone and camera usage.¹³⁹

Addressing the authenticity of photographs is not limited to the content of the photograph itself. The potential for altering or enhancing of the photograph must be considered.¹⁴⁰ In addition, the metadata of photographs could have an abundance of information relevant to a case, including date, time, location, and GPS coordinates. Additional issues may arise when a film photograph is converted to digital.

When authenticating digital photographs, the most likely Rules to apply are as follows:

138. *People v. Goldsmith*, 326 P.3d 239, 246 (Cal. 2014).

139. It is estimated that over one trillion digital photographs are now taken every year. Stephen Heyman, *Photos, Photos Everywhere*, N.Y. TIMES (July 29, 2015), <https://www.nytimes.com/2015/07/23/arts/international/photos-photos-everywhere.html>.

140. *See Hines v. Carpenter*, No. 3:05-0002, 2015 WL 1208684, at *19 (M.D. Tenn. Mar. 16, 2015) (quoting *Lorraine*, 241 F.R.D. at 561–62) (“enhancement consists of removing, inserting, or highlighting an aspect of the photograph that the technician wants to change.”); *Guarisco v. Boh Brothers Construction Co., LLC*, No. 18-7514, 2019 WL 4881272 (E.D. La. Oct. 3, 2019) (imposing sanctions against the plaintiff for modifying a digital photograph to enhance her negligence claims against defendant and observing that the original unmodified photograph was still available on the plaintiff’s Facebook page).

- a witness with personal knowledge—Rule 901(b)(1)
- a system or process capable of providing a reliable and dependable result—Rule 901(b)(9)
- official publications—Rule 902(5)
- certified records generated by an electronic process or system—Rule 902(13)
- certified data copied from an electronic device, storage medium, or file—Rule 902(14)

The leading authority on authenticating digital photographs remains *Lorraine*,¹⁴¹ which considered the authentication issues surrounding digital photographs, digitally enhanced images, digitally converted images, and photograph metadata. As with film photographs, Rule 901(b)(1) is a viable option for authenticating a digital photograph if a witness with personal knowledge of the scene in the photograph is available. If such a person is not available, a digitally converted image requires testimony by someone knowledgeable about the film-to-digital conversion process.

Authentication of a digitally enhanced photograph likely implicates Rule 901(b)(9) because of the unlikelihood that a witness can testify regarding subtle differences in the original photograph as compared to the enhanced image.¹⁴² Metadata of a photograph was not considered in depth a decade ago. Photographs taken with cell phones have information that may be important for multiple reasons. The metadata from a photograph

141. 241 F.R.D. at 561–62.

142. *Id.* at 560 (discussing *State v. Swinton*, 847 A.2d 921, 942 (Conn. 2004)).

taken with a cell phone may automatically capture the geographic coordinates of where a picture was taken.¹⁴³

9. Group Collaboration Tools

Collaboration applications, such as Slack, Jive, Confluence, Microsoft Teams, Salesforce Chatter, and others, facilitate group discussions as well as message exchanges between users and in private channels.¹⁴⁴ These applications often store shared content in the cloud, though some are deployed on a company's servers.¹⁴⁵

Bases for authentication will typically include the following:

- a witness with personal knowledge—Rule 901(b)(1)
- expert testimony or comparison with authenticated examples—Rule 901(b)(3)
- distinctive characteristics, including circumstantial evidence—Rule 901(b)(4)
- a system or process capable of proving a reliable and dependable result—Rule 901(b)(9)
- certified records generated by an electronic process or system—Rule 902(13)
- certified data copied from an electronic device, storage medium, or file—Rule 902(14)

143. See *United States v. Post*, 997 F. Supp. 2d 602, 603–04 (S.D. Tex. 2014) (discussing how image metadata can reveal the location where a digital photograph was taken).

144. See *Primer on Social Media*, *supra* note 75, at 16.

145. *Id.*

Collaboration tools typically offer programs that use APIs to access and share information with the application.¹⁴⁶ Using the API, some discovery review platforms can import machine-readable, searchable data that includes content and its metadata; some even collect metadata that can authenticate the content and may provide a message-digest hash for verification of the extracted data.

As noted with website collections, collecting data through an API can be problematic. An API collection lacks perfect synchronicity with the original content—it may change its context, format, or appearance—and it may be difficult to access. Moreover, provider restrictions may limit the amount of data that can be collected through an API.¹⁴⁷

10. Computer Processes, Animations, Audio/Video, Virtual Reality, and Simulations

When machines are responsible for recording audio or video or implementing processes, authentication will be relatively simple, presuming that the recording device was in good working order, under 902(13) (certified records generated by an electronic process or system).

146. *Guide to Slack import and export tools*, SLACK, <https://get.slack.help/hc/en-us/articles/204897248-Guide-to-Slack-import-and-export-tools> (last visited May 6, 2020).

147. *Id.* For example, Slack only permits “Enterprise Grid” plan users to export all data from their accounts. *A guide to Slack’s Discovery APIs*, <https://slack.com/help/articles/360002079527> (last visited May 6, 2020). In contrast, Slack places restrictions on “Free,” “Standard,” and “Plus” plan users to export messages from “private channels” and “direct messages.” Slack also forbids such plans from exporting files attached to user messages. *Guide to Slack Import and Export Tools*, SLACK, <https://get.slack.help/hc/en-us/articles/204897248-Guide-to-Slack-import-and-export-tools> (last visited May 6, 2020).

However, where a person is creating audio or video, virtual reality scenarios, or simulations, authentication becomes more complex. It may require testimony regarding the operation of the equipment, the accuracy of the data, and the motion and sound. Typical methods for authenticating this evidence are as follows:

- a witness with personal knowledge—Rule 901(b)(1)
- expert testimony or comparison with authenticated examples—Rule 901(b)(3)
- a system or process capable of proving a reliable and dependable result—Rule 901(b)(9)¹⁴⁸

Computer simulations, which are based on scientific principles and data and offered as substantive evidence, face a stiffer test for authentication, wrapped up in an analysis of their reliability.¹⁴⁹

11. Cloud Computing

Cloud computing services often transfer ESI to servers other than the “original” server (i.e., the server on which it was stored in the first instance). The cloud computing service’s servers may be located in various locations across the country or even around the world. It may be difficult, if not virtually impossible,

148. *Lorraine v. Markel Am. Ins. Co.*, 241 F.R.D. 534, 559 (D. Md. 2007) (stating that computer animations offered to illustrate testimony must be “authenticated by testimony of a witness with personal knowledge of the content of the animation, upon a showing that it fairly and adequately portrays the facts and that it will help to illustrate the testimony given in the case.”).

149. *Id.* at 560–61 (“[T]he most frequent methods of authenticating computer simulations are 901(b)(1) (witness with personal knowledge); and 901(b)(3) (expert witness). Use of an expert witness to authenticate a computer simulation likely will also involve Federal Rules of Evidence 702 and 703.”).

to establish a chain of custody of a file, for example, that has been moved multiple times. Moreover, a single file may be disassembled and its parts stored on multiple servers. By analogy, this would be similar to cutting paper document into pieces, putting each piece in a separate file cabinet, and distributing the file cabinets to various warehouses around the world. To an end user, the service is seamless. When retrieved, the document pieces are reassembled from their various locations. How does a party establish that the reassembled document is identical to the “original” file before disassembly? Possible answers may be matching hash values or expert testimony about a process.

In addition, cloud computing services must duplicate and store copies of ESI on various servers to protect against loss from some catastrophic failure (e.g., fire, flood, etc.). It will be difficult, if not impossible, to know whether a particular file is the “original.” This issue, however, may be more theoretical than practical. In any event, matching hash values may once again provide a sufficient basis to authenticate the “original” or “copy.”

12. Emoji

Emoji, from the Japanese word meaning “picture character,” are small pictographs.¹⁵⁰ These images are often used in text messages, social media, emails, and chat apps “to express the emotional attitude of the writer, convey information succinctly, [and] communicate a message playfully without using words, etc.”¹⁵¹ They are distinct from *emoticons*, which are letters, numbers, and other standard ASCII characters grouped into a

150. *Frequently Asked Questions: Emoji and Pictographs*, UNICODE, https://unicode.org/faq/emoji_dingbats.html#1.05 (last visited May 6, 2020).

151. *Commonwealth v. Castano*, 82 N.E.3d 974, 978 n.2 (Mass. 2017) (citing MERRIAM-WEBSTER ONLINE DICTIONARY, <https://www.merriam-webster.com/dictionary/emoji>).

pictograph, like a smiley face :-) or a heart <3, and are used to “represent[] a facial expression or suggest[] an attitude or emotion and that is used especially in computerized communications (such as e-mail).”¹⁵²

Emoji have typically been used in consumer correspondence and have been increasingly a subject of evidence in criminal cases.¹⁵³ With emoji showing up now in business communications, they are also becoming a source of evidence in civil litigation. Despite their seemingly straightforward cartoonish appearance, emoji can be fraught with difficulty for the unwary practitioner given the rapid growth in emoji variety and depictions, together with the challenges of interpreting their meaning.¹⁵⁴

First, the variety of emoji is continually expanding—and with it, the multiplicity of ways they are depicted. Over 3,000 emoji are now listed in the Unicode Standard.¹⁵⁵ Unicode is a computer-industry standard that assigns each letter, digit, and symbol, including emoji, a unique numeric value that will apply across different operating systems, devices, applications, and languages. Its purpose is to ensure the consistent encoding, handling, and representation of characters and emoji symbols. However, though a single code is assigned to Unicode emoji, that does not mean that there is a single depiction or meaning of each Unicode emoji. Instead, a platform can render emoji using

152. Emoticon, MERRIAM-WEBSTER ONLINE DICTIONARY, <https://www.merriam-webster.com/dictionary/emoticon>.

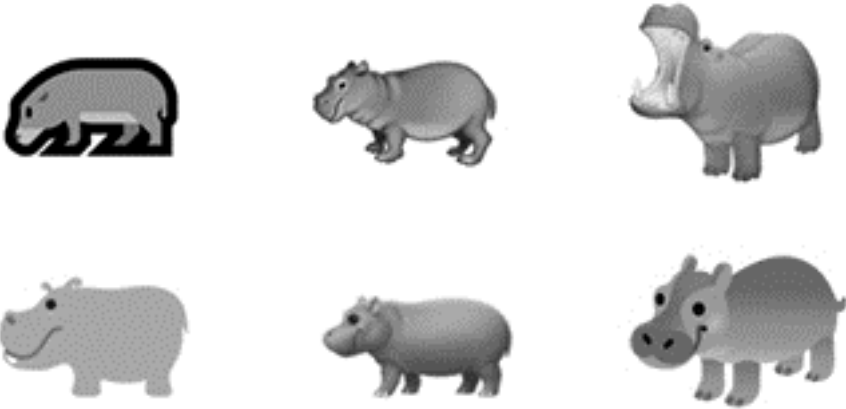
153. See, e.g., *In re JP*, No. 344812, 2019 WL 4648450 (Mich. Ct. App. 2019) (memorializing in the court’s opinion emoji the appellant exchanged with friends through Snapchat).

154. See Eric Goldman, *Emojis and the Law*, 93 WASH. L. REV. 1227, 1230 (2018).

155. *Full Emoji List, v 13.0*, UNICODE, <http://unicode.org/emoji/charts/full-emoji-list.html> (last visited May 6, 2020).

its own idiosyncratic color and shape choices.¹⁵⁶ Complicating this further is that the emoji intended by a sender may appear differently on the recipient's device.¹⁵⁷

Take, for example, the hippopotamus emoji, which was approved as part of Unicode 11.0 in 2018. Here are some renderings of the hippo across various platforms (Microsoft, Samsung, Facebook, Twitter, Apple, and Google, respectively):



156. Hannah Miller et al., “Blissfully Happy” or “Ready to Fight”: Varying Interpretations of Emoji, in PROCEEDINGS OF THE TENTH INTERNATIONAL AAAI CONFERENCE ON WEB AND SOCIAL MEDIA 259, 267 (2016) (“Unlike plain text where people view the same characters in their exchange, platforms effectively *translate* emoji: the emoji that the sender chose is translated to the receiver’s platform’s rendering.”).

157. Further, since emoji render differently on different platforms, the emoji sent by one person from one device may differ markedly from what a recipient using a different device sees. *Id.* at 259. Such a phenomenon is apparent in the *In re JP* matter where the court inserted what appear to be Gmail emoji into its opinion to reflect the emoji exchanged by the appellant and her friends on Snapchat. *In re JP*, 2019 WL 4648450 at *2. See Eric Goldman, *More Teenagers Mistakenly Think “Private” Chat Conversations Will Remain Private—People v. JP*, TECHNOLOGY & MARKETING LAW BLOG (Oct. 7, 2019), <https://blog.ericgoldman.org/archives/2019/10/more-teenagers-mistakenly-think-private-chat-conversations-will-remain-private-people-v-jp.htm>.

Problematically, Unicode is not the only type of emoji. There are many more non-Unicode emoji that are idiosyncratic to different platforms. Often called “bespoke emoji” or “stickers,” these are available on platforms like Facebook and Snapchat. Other apps also let users create their own custom emoji, such as avatars from the Bitmoji app. Since these emoji lack Unicode codes, they may not be compatible with other platforms, so they may not display properly—or at all—to recipients who are not using the same platform as the sender.


The differences in renderings have implications for discovery as well. When emoji are collected and processed, the image may very well appear differently—or as an empty rectangular box or space—for review.

A second hurdle to using emoji as evidence is the issue of interpretation. Emoji can be difficult to interpret on their own. Emoji are small and many appear similar with minor differences. For example, the Unicode crying face has a tear, but the Unicode downcast face has a similar shape indicating a bead of sweat (both shown in Apple renderings). Only the eyes and placement of the water drop clue the reader in as to the meaning.



Finally, while “a picture is worth a thousand words,” those words may be different in the eye of the beholder. Face emoji can be particularly problematic because people have difficulty interpreting facial expressions and given the different ways that

platforms choose to depict those faces.¹⁵⁸ Moreover, facial expressions may be used to indicate irony: for example, a winking emoji may indicate a joking tone, but a recipient may perceive the joke differently—or more maliciously—than the sender intended.¹⁵⁹

Additionally, some emoji have  multiple meanings. For example, the alien emoji may mean that something is out of this world or strange. Alternatively, it may be used to refer to someone who is an illegal alien. Meanings can also depend on the cultural background of the sender and recipient (as well as a judge or jury).¹⁶⁰ For instance, the angel emoji may denote innocence, but a Chinese reader may perceive an angel as a threatening sign of death.¹⁶¹ As a result, it can be difficult from an evidentiary point of view for parties, courts, and juries to give proper meaning to emoji. Meanings can become especially muddled when emoji are grouped together: it may be unclear whether the emoji are

158. Miller et al., *supra* note 156, at 261, 263–67.

159. NEXUS Servs., Inc. v. Moran, No. 5:16-cv-00035, 2018 WL 1461750, at *4 (W.D. Va. Mar. 23, 2018) (interpreting a Hitler emoji as ironic, finding that “[w]hile any image evoking Hitler obviously can be offensive, the emoji was contained in an internal email between two work colleagues in which, taken in context, one was jokingly calling the other a ‘meanie’ and a taskmaster.”); United States v. Christensen, No. CR 06-085-BLG-RFC, 2013 WL 1498950, at *2 (D. Mont. Apr. 11, 2013) (“Christensen claims Neuhardt violated attorney-client privilege and the Sixth Amendment by offering, in an e-mail to the prosecutor accompanied by an emoticon, to ‘stipulate that my client is guilty. :)’ No one took Neuhardt’s frivolous e-mail as an actual stipulation.”).

160. VYVYAN EVANS, *THE EMOJI CODE: THE LINGUISTICS BEHIND SMILEY FACES AND SCAREDY CATS* 102, 123 (2017).

161. Alex Rawlings, *Why emoji mean different things in different cultures*, BBC (Dec. 11, 2018), <http://www.bbc.com/future/story/20181211-why-emoji-mean-different-things-in-different-cultures>.

independent of each other, modify each other, or are lined up to tell a story.

Emoji are already finding their way into judicial opinions. In one criminal case involving allegations of drug trafficking, firearms offenses, and racketeering, the defendants argued that there was no probable cause to search their Facebook accounts.¹⁶² The investigating ATF agent testified, using his investigative experience, that the emoji referred to illicit activity: namely, a cloud emoji referred to drugs, while a cloud-of-gas emoji symbolized a gang. The court permitted the agent to use his training to interpret the emoji and establish probable cause. In another criminal case, the jury used emoji in a text message to conclude that a killing was not accidental.¹⁶³ The defendant had texted a friend the victim's nickname along with an emoji face showing Xs instead of eyes. The prosecution argued that the text indicated the shooting had already occurred.

Presenting emoji as evidence presents several challenges for authentication and admissibility. Parties will need to consider the context of the emoji in the sequence of communications to help define their meaning as well as the platforms used to depict those emoji. In addition, because emoji evolve over time, parties will need to determine how the emoji was rendered on a particular platform and operating system at a particular time for both the sender and recipient.

To authenticate emoji, expert testimony may be particularly important. The authentication rules most likely to play a role are as follows:

- A witness with personal knowledge—Rule 901(b)(1)

162. *United States v. Westley*, No. 3:17-CR-171, 2018 WL 3448161 (D. Conn. July 17, 2018).

163. *Commonwealth v. Castano*, 82 N.E.3d 974, 982–83 (Mass. 2017).

- expert testimony or comparison with authenticated examples—Rule 901(b)(3)
- distinctive characteristics, including circumstantial evidence—Rule 901(b)(4)
- a system or process capable of proving a reliable and dependable result—Rule 901(b)(9)
- certified data copied from an electronic device, storage medium, or file—Rule 902(14)

E. Hard Copies

Lorraine contains numerous points of comparison between ESI and hard-copy record systems in resolving authentication and admissibility issues.¹⁶⁴ While comparisons to the familiar world of tangible evidence are a useful starting point in many legal analyses, it is important to note some key differences between the two systems.

With hard-copy record systems, the mechanics of creating, storing, managing, organizing, controlling, and securing records and the systems that maintain them are generally simple and easily understood. Control largely depends on physical access to the records, which are basically stable and durable; one would need to be physically present to manipulate, mutilate, or destroy a hard-copy record. Moreover, manipulation or mutilation of documents has the potential for leaving indications of the tampering. Control systems can be designed to take advantage of physical realities such as the contiguous nature of the environment in which the records persisted, including known points of ingress and egress and singularity (uniqueness, originality, and the fact that a hard-copy record cannot simultaneously be physically present in more than one location at the same time).

164. *Lorraine v. Markel Am. Ins. Co.*, 241 F.R.D. 534, 537, 542, 561 (D. Md. 2007).

Further, a physical or hard-copy record cannot be accessed and used simultaneously by multiple people without those people also being physically present and aware that access and use are shared.

This is not the case with ESI, particularly with regard to the issues of controlling and securing records. Unlike paper documents, access to ESI is not naturally constrained. Most computers are members of networks (or are intermittently on and off networks), and these networks generally are internetworked. With the advent of cloud storage, ESI may no longer reside on a local hard drive or server but may be accessed by a user half a world away. Moreover, scarcely a month goes by without another serious data breach being reported.¹⁶⁵

F. Potential Challenges to Using Rule 902(14)

1. The Requirement of a Process of Digital Identification

To take advantage of Rule 902(14), litigants should think ahead, as the rule requires proof of “a process of digital identification.” Any counsel who waits until the eve of trial to ponder hash values may be out of luck—the benefits of self-authentication cannot be applied to electronic evidence retroactively. The time to consider Rule 902(14) begins at the collection phase.

The most common method for authenticating electronic evidence under Rule 902(14) is hash-value verification. This involves comparing the hash value of an original, native version of an electronic file to the hash value of the collected, copied version. If both hash values are identical, then the copied version

165. See The Sedona Conference, *Commentary on Privacy and Information Security*, 17 SEDONA CONF. J. 1, 5 (2015) (“Personal identities, privacy, confidential client information, work product, and even attorney-client communications have never been more vulnerable to unauthorized disclosures, breaches, loss, or theft than they are today.”).

proffered at trial is self-authenticating, assuming that a qualified person explains the process by which the original and copied hash values were generated and compared.¹⁶⁶

The challenge that litigants are most likely to encounter with Rule 902(14) will be their failure to prepare for the first step—that is, generating an original hash value for each native file they intend to collect. This is because many litigants “self-collect” by either copying and pasting or dragging and dropping ESI onto a storage device or platform. It is often the most cost-effective way to preserve or collect information, but depending on how this is done, it may preclude reliance on Rule 902(14) for authentication.

Litigants should consider that original hash values do not self-generate. Currently, only specialized, third-party software can assign the unique alphanumeric identifiers for the authenticity of original ESI. While these programs are widely available, the practical reality is that given time limits and other reasons, most litigants, including large organizations with sophisticated Information Technology (IT) departments, do not use hash values with regularity for certain types of collections; they simply collect the files without collecting hash values. However, other avenues of authentication may be available. For example, ESI may still be authenticated as a business record or by a sender or recipient with the requisite personal knowledge.

166. FED. R. EVID. 902 advisory committee’s notes to 2017 amendments ¶ 14 (“If the hash values for the original and copy are different, then the copy is not identical to the original. If the hash values for the original and copy are the same, it is highly improbable that the original and copy are not identical. Thus, identical hash values for the original and copy reliably attest to the fact that they are exact duplicates. This amendment allows self-authentication by a certification of a qualified person that she checked the hash value of the proffered item and that it was identical to the original.”).

Those entities wishing to rely on Rule 902(14) should consider developing their own hashing policies and procedures. Whether responsibility falls to outside counsel, a third-party vendor, in-house counsel, or internal IT specialists, such litigants will benefit from having given their teams clear direction on how ESI is to be collected and digitally identified.

Even if litigants are diligent about assigning original hash values, they should consider how they will prove compliance with Rule 902(14) and should consider generating, maintaining, and preserving hash-value logs. This approach regarding original and copied hash values is a new concept—one unlikely to be on litigants' radar—but it is now key to admissibility under Rule 902(14). Creating these logs is not difficult; the software that generates the hash values also generates the logs. But maintaining them could be a challenge for some. With many years passing between the collection of documents and the admission of evidence, counsel should consider this issue early in the discovery process.

2. Certification Hazard: The Potential Exposure of Electronic Discovery Protocols

While careful adherence to Rule 902(14)'s requirements can streamline authentication, litigants should be alert to one potential drawback: exposing their electronic discovery protocols to adversaries. Typically, the details of a litigant's preservation, collection, and processing methods fall outside the scope of permissible discovery under Rule 26(b)(1) as being unrelated to the parties' "claims or defenses."¹⁶⁷ But the best supported Rule 902(14) declarations will be based on thorough ESI-collection

167. The Sedona Conference, *The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 SEDONA CONF. J. 1, 118–30 (2018) [hereinafter *The Sedona Principles, Third Edition*].

documentation. This could mean having to explain a litigant's electronic discovery procedures.

In preparing the certification, litigants may want to refer to documentation confirming the chain of custody, which might encompass a range of sensitive details about the evidence and the collection process. This may very well include a description of the ESI source, custodian information, identification of the party performing the collection, collection date, and the storage or transfer means for the copy. It could also identify the copying tools and methods.

G. Recent Changes to Rule 807 (Residual Exception to Hearsay Rule)

Federal Rule of Evidence 807, also known as the residual exception, provides that certain hearsay statements may be admissible, even if they do not fall into one of the other hearsay exceptions in Rules 803 or 804. A revised version of Rule 807 adding a totality-of-the-circumstances standard took effect on December 1, 2019.

Amended Rule 807 eliminates the requirement that the evidence must be material and the requirement that the proffered evidence must serve the interests of justice. Before the amendment, Rule 807 allowed admission only when notice of an intent to use was made before trial. Under amended Rule 807, the out-of-court statement must be trustworthy and be more probative than other reasonably available evidence. It also expands the procedure for admission of such evidence by permitting the trial court to admit hearsay "during the trial or hearing if the court, for good causes, excuses a lack of earlier notice."

In 2016 and 2017, the Advisory Committee on the Rules of Evidence debated whether to expand the Rule 807 exception to allow the admission of reliable hearsay even absent "exceptional circumstances." Ultimately, the committee decided

against expanding the exception; instead, it opted for an amendment to cure several problems with the current rule.¹⁶⁸

The problems that the committee identified included the following:

- The requirement that the court find trustworthiness “equivalent” to the circumstantial guarantees in the Rules 803 and 804 exceptions is difficult to apply because these exceptions offer no single trustworthiness standard.
- The requirements in Rule 807 that the residual hearsay must prove a “material fact” and that admission of residual hearsay be in “the interests of justice” are superfluous because these issues are addressed in Rules 102 and 401.
- The requirement that the hearsay statement must be “more probative than any other evidence that the proponent can obtain through reasonable efforts” is unnecessary.¹⁶⁹

After receiving public comments, the Advisory Committee approved and then submitted the proposed amendment to the Standing Committee for final approval. Under the amended rule, the proponent of the evidence must still establish that the hearsay statement is not otherwise admissible under Rule 803 or 804. Instead of equivalence, the new rule requires the court to analyze the totality of the circumstances surrounding the making of the statement, including any corroborating evidence, to

168. COMM. ON RULES OF PRACTICE AND PROCEDURE, JUDICIAL CONFERENCE OF THE U.S., AGENDA BOOK 99–100 (Jan. 4, 2018), <http://www.uscourts.gov/sites/default/files/2018-01-standing-agenda-book.pdf>.

169. COMM. ON RULES OF PRACTICE AND PROCEDURE, JUDICIAL CONFERENCE OF THE U.S., AGENDA BOOK 736–37 (June 12–13, 2017), http://www.uscourts.gov/sites/default/files/2017-06-standing-agenda_book_0.pdf.

assess whether there are sufficient guarantees of trustworthiness.

The following is the language of the amended Rule 807 (Committee Notes to amended Rule 807 are in Appendix B):

Rule 807. Residual Exception

(a) In General. Under the following ~~circumstances~~ conditions, a hearsay statement is not excluded by the rule against hearsay even if the statement is not ~~specifically covered by~~ admissible under a hearsay exception in Rule 803 or 804:

(1) the statement ~~has equivalent circumstantial~~ is supported by sufficient guarantees of trustworthiness—after considering the totality of the circumstances under which it was made and evidence, if any, corroborating the statement; and

(2) it is offered as evidence of a material fact;

(3) it is more probative on the point for which it is offered than any other evidence that the proponent can obtain through reasonable efforts; ~~and.~~

(4) admitting it will best serve the purposes of these rules and the interests of justice.

(b) Notice. The statement is admissible only if, ~~before the trial or hearing,~~ the proponent gives an adverse party reasonable notice of the intent to offer the statement ~~and its particulars, including the declarant's name and address,~~ including its substance and the declarant's name—so that the party has a fair opportunity to meet it. The notice must be provided in writing before the trial or

hearing—or in any form during the trial or hearing if the court, for good cause, excuses a lack of earlier notice.¹⁷⁰

170. COMM. ON RULES OF PRACTICE AND PROCEDURE, JUDICIAL CONFERENCE OF THE U.S., AGENDA BOOK 409–10 (June 12, 2018), <https://www.uscourts.gov/rules-policies/archives/agenda-books/committee-rules-practice-and-procedure-june-2018>. (new material is underlined; matter to be omitted is struck).

III. EMERGING ESI EVIDENTIARY ISSUES

A. *Determining the Owner/Creator of ESI*

ESI may be created by aggregating data from various sources, with various owners. With increasingly more complex interconnected systems, determining the actual owner or creator of ESI becomes more challenging. However, a custodian or other qualified witness must be able to testify as to the source of the information, circumstances associated with the record's creation, and the degree of regularity of the organization's practice and its record making and keeping. Therefore, it becomes imperative to determine who or what created the content to be able to authenticate it.

An individual may create various electronic documents that are in turn passed to others through various electronic mediums such as emails, collaborative environments, and other shared networks. These individuals may in turn modify the document either on the shared space or on their individual devices.

B. *Understanding the Limits of Technology*

The proliferation of technology has transformed the nature of "documents." What was once primarily in hard-copy, ink-and-paper format is now often in ESI format but is no less a document.¹⁷¹ The overwhelming majority of documents generated today are ESI, including not only digital versions of those that are analogous to documents of the past (e.g., word processing and spreadsheets) but also an entirely new class of digital documents consisting of what were formerly verbal conversations:

171. Indeed, one of the most ubiquitous word-processing applications refers to individual files as "documents." *Create a document in Word*, MICROSOFT, https://support.office.com/en-us/article/create-a-document-in-word-aafc163a-3a06-45a9-b451-cb7250dcbaa1?wt.mc_id=fsn_word_quick_start (last visited May 6, 2020).

text messages, Skype, Voice over Internet Protocol (VoIP) calls, video conferences, and social media postings, to name a few.¹⁷² Moreover, some technology—the IoT—has created an entire class of ESI that otherwise wouldn't exist, such as GPS location data and biological data from wearable devices.¹⁷³

Given the proliferation in the volume of ESI and the changing nature of such “documents,” actors in the legal system have and will continue to turn to technology for assistance in identifying, analyzing, and ultimately authenticating ESI for use as evidence in both civil and criminal cases. Such technology may also be important in establishing the closely related chain of custody.¹⁷⁴ While deficiencies in the chain of custody do not destroy the admissibility of the proffered evidence, they go to the weight that the jury may give to the evidence. In light of the interplay between Rule 104(a) and (b), however, deficiencies in either authentication or chain of custody may produce the same result.¹⁷⁵

172. 2 RAYMOND T. NIMMER & HOLLY K. TOWLE, *THE LAW OF ELECTRONIC COMMERCIAL TRANSACTIONS, E-Mails and Evidence in E-Commerce Contexts* § 13.09, pt. C (2d ed. 2018).

173. See Section II.D.5, *supra*.

174. *United States v. Blank*, No. WDQ-14-10448, 2015 WL 4041408, at *8 (D. Md. June 30, 2015), *aff'd*, 659 F. App'x 727 (4th Cir. 2016) (quoting *United States v. Howard-Arias*, 679 F.2d 363, 366 (4th Cir. 1982)) (finding that, as a practical matter, chain of custody is a variation of the authenticity requirement).

175. See U.S. COURT OF APPEALS FOR THE THIRD CIRCUIT, MODEL CIVIL JURY INSTRUCTIONS 1.5 (2015) (“Consider it in light of your everyday experience with people and events, and give it whatever weight you believe it deserves.”); U.S. COURT OF APPEALS FOR THE SEVENTH CIRCUIT, FEDERAL CRIMINAL JURY INSTRUCTIONS 2.02 (2012) (“Give the evidence whatever weight you decide it deserves.”); Pattern Instruction No. 2.02 (“It is up to you to decide how much weight to give to any evidence, whether direct or circumstantial.”); U.S. COURT OF APPEALS FOR THE NINTH CIRCUIT, MODEL CIVIL JURY INSTRUCTIONS 1.12 (2017) (“It is for you to decide how much weight to give to any evidence.”); *Flores v. City of Westminster*, 873 F.3d 739, 758 (9th

Although technology can provide many tools to assist in the process of authentication (including establishing the chain of custody), it is important to understand these tools and their potential role, including their limitations.

1. Hashing

One of the most important ways of authenticating ESI is through hash values:

A hash value is a unique numerical identifier that can be assigned to a file, a group of files, or a portion of a file, based on a standard mathematical algorithm applied to the characteristics of a data set. The most commonly used algorithms, known as MD5 and SHA, will generate numerical values so distinctive that the chance that any two data sets will have the same hash value, no matter how similar they appear, is less than one in one billion. “Hashing” is used to guarantee the authenticity of an original data set and can be used as a digital equivalent of the Bates stamp used in paper document production.¹⁷⁶

Cir. 2017), *cert. denied sub nom.*, Hall v. Flores, 138 S. Ct. 1551 (2018) (quoting Tortu v. Las Vegas Metro. Police Dep’t, 556 F.3d 1075, 1084 (9th Cir. 2009)); United States v. Vidacak, 553 F.3d 344, 350 (4th Cir. 2009); United States v. Pantic, 308 F. App’x 731, 733 (4th Cir. 2009); United States v. Cardenas, 864 F.2d 1528, 1531 (10th Cir. 1989) (“[D]eficiencies in the chain of custody go to the weight of the evidence, not its admissibility; once admitted, the jury evaluates the defects and, based on its evaluation, may accept or disregard the evidence.”).

176. See Grimm et al., *supra* note 58, at 17 n.47 (quoting BARBARA J. ROTHSTEIN ET AL., *MANAGING DISCOVERY OF ELECTRONIC INFORMATION: A POCKET GUIDE FOR JUDGES* 38 (2d. ed. 2007)).

There are three areas of concern regarding the use of hash algorithms: (i) encryption; (ii) known file identification; and (iii) file and or data authentication.¹⁷⁷ This *Commentary* focuses on the latter two concerns.

Hashing is based on algorithms that are essentially a set of rules for a mathematical process.¹⁷⁸ Herein lies its inherent weakness, because a mathematical process is based on rules that are predictable and repeatable.¹⁷⁹ Such predictability can lend itself to manipulation and cause either a “collision attack” of algorithms or result in a matching value. Such manipulation, however, is a complex process and has only been successfully accomplished in a laboratory setting where the manipulator must have physical possession of the target file and be able to alter the file before the hash algorithm is run. Outside the laboratory, for purposes of identifying and authenticating ESI (item iii, above), such a collision is statistically nearly impossible.¹⁸⁰ Nevertheless, a strict protocol for the chain of custody of files should be implemented to eliminate the opportunity to manipulate the target file. Further, for purposes of known file identification,¹⁸¹ known file hash sets (known as Secure Hash

177. Don L. Lewis, *The Hash Algorithm Dilemma—Hash Value Collisions*, FORENSIC MAG. (Dec. 2008).

178. *Id.*

179. *Id.*; see FED. R. EVID. 902 advisory committee’s notes to 2017 amendments ¶ 14.

180. Lewis, *supra* note 177 (“For use in file identification and authentication, there is a greater probability that [a] single individual, from a twelve member jury, will win the Power Ball Lottery sixty days in a row, than an accidental occurrence of two matching MD5 hash values from files that have not been manipulated to collide.”).

181. See Orin S. Kerr, *Searches and Seizures in a Digital World*, 119 HARV. L. REV. 531, 541 (2005); *United States v. Reddick*, 900 F.3d 636, 637 (5th Cir. 2018), *petition for cert. filed*, No. 18-6734 (U.S. Nov. 19, 2018); *United States v. Borowy*, 595 F.3d 1045, 1048 (9th Cir. 2010); *United States v. Cartier*, 543 F.3d

Algorithm) have been created independently by the National Institute of Standards and Technology and the National Software Reference Library. Although it is virtually impossible to create a hash value of a contraband image, even if it were possible, it would be traceable in the independent known data sets.¹⁸² One important caveat: the research is based on current technology. It is possible that use of artificial intelligence and other advanced computing capabilities may produce tools to manipulate hashes in the future. There is likely to be a continuing technology race to further strengthen on the one hand, and manipulate on the other hand, the hashing algorithms.

Regardless of future possibilities of compromise, hashing can be a means of efficiently determining whether two files are exact duplicates of each other or whether a single file has been altered. The reliability and usefulness of hashing depends on a trustworthy reference. Either the subject file or the copy (or its hash value) must be preserved in a way that ensures there has been no tampering with that reference file.

2. Encryption

The use of encryption and digital signatures can also provide a basis for trust. At a simple level, encryption uses a secret key to scramble the contents of a file so that only those with access to the key may read the file. A digital signature uses the same technology to enable a party to use its secret key to indicate that

442, 445 (8th Cir. 2008); *United States v. Miller*, No. CV 16-47-DLB-CJS, 2017 WL 2705963, at *1 (E.D. Ky. June 23, 2017); *United States v. Noden*, No. 8:16-cr-00283-LSC-MDN, 2017 WL 1406377, at *1 (D. Neb. Apr. 20, 2017); *United States v. Feldman*, No. 13-CR-155, 2014 WL 7653617, at *4, (E.D. Wis. July 7, 2014); *United States v. Woods*, 730 F. Supp. 2d 1354, 1362 (S.D. Ga. 2010); *United States v. Cartier*, No. 2:06-cr-73, 2007 WL 319648, at *1 (D.N.D. Jan. 30, 2007).

182. See Lewis, *supra* note 177.

it has “signed” an electronic document. Well-established products enable these processes to work fairly seamlessly, although managing the keys used for encryption can become an issue, especially at an enterprise level.

Using these technologies, it is possible to assert that a person signing an electronic document has viewed and approved the document, much as someone would indicate their acceptance of a document (or indicate their authorship of a letter) by signing their name in ink. In legal circles, this is commonly referred to as “non-repudiation.”

However, a digital signature actually indicates something slightly different: that someone with access to the key has signed the document. Keys can be stolen or borrowed (copied), frequently without the knowledge of the owner of the key. Similarly, one must link a key back to a specific individual, which generally requires an inquiry to the party that issued the key and an assessment of the veracity of the key issuer. And, even assuming a reputable issuer, that party may distribute keys under varying levels of scrutiny, requiring only an email address at the lower end all the way to requiring a passport or other official identification at the higher end.

For example, it may easily be proven that a key issued to John Smith by KeyCorp was used to sign an important document. However, upon inquiry to KeyCorp, it may be determined that the key was sent by email to JohnSmith@yahoo.com without any verification of John Smith’s identity.

Additionally, there is nothing about a plain digital signature that can be used to prove when it was created. It is possible for a party in control of the digital certificate (cryptographic key) to falsify the value/appearance of time in conjunction with manipulated data and force a signing event that would be technically impossible to identify or distinguish from a legitimate digital signature. In such a scenario, the resulting data/signature

combination would be mathematically true but semantically false. However, digital signatures can be used in combination with alternative methods for establishing authenticity.

3. System Metadata

Metadata can be another useful checkpoint for determining authenticity.¹⁸³ For example, email messages generally contain a substantial amount of metadata information, including a unique message ID as well as information on the unique internet locations (IP addresses) where the message originated and was handled along the way to its destination. Similarly, operating system metadata can be a useful tool. Most operating systems maintain information about individual files: the dates that a file was created, last modified, and last accessed. For example, in a case where an individual claims that she did not create a document until July 1 but the system metadata shows that the document was created on May 1, this data may be helpful.

However, metadata can be unreliable and may be subject to manipulation and nonobvious deletion. A moderately sophisticated user may be able to manipulate system dates. Although traces of this manipulation may be left behind, detecting such traces can be extremely difficult and expensive or simply impossible. Worse, use of files after the fact, such as an investigator opening a file for review, can modify metadata and make it useless or misleading for authenticity purposes. Accordingly, careful attention should be paid to the methods used to authenticate metadata.

183. For a detailed discussion about metadata, see *The Sedona Principles, Third Edition*, *supra* note 167, Principle 12 at 169 and *The Sedona Conference, Commentary on Ethics & Metadata*, 14 SEDONA CONF. J. 169, 173–75 (2013).

4. Computer Forensics and Anti-Forensics

Computer forensics “is the art and science of applying computer science to aid the legal process. Although plenty of science is attributable to computer forensics, most successful investigators possess a nose for investigations and a skill for solving puzzles, which is where the art comes in.”¹⁸⁴ Computer forensics involves the location, examination, identification, collection, preservation, and analysis of computer systems and ESI and often includes the rendering of a qualified expert opinion regarding those systems and ESI.

Computer forensics typically involves the employment of specialized and sophisticated computer-based tools to aid in the performance of the various investigation and documentation activities, which can be costly and time consuming. Use of forensic software to identify, acquire, analyze, and store ESI can generally be divided into two processes: (1) static environment and (2) live environment. In a static environment, a mirror image copy is made of the system or storage device (e.g., a hard drive). The accuracy of the copy is established by matching the hash values of the target drive, and each file on the drive, with the hash values of the copy. Then, forensic software can be used to extract evidence from the copy. In a live environment, the forensic software runs in the target system’s environment, which in itself affects the system (e.g., changing system metadata, directories, etc.). While evidence from both processes has been admitted in court, evidence acquired from a live system can be vulnerable to attack, particularly if there is a break in the digital chain of custody.

Anti-forensics is the employment of sophisticated tools and methods used for the intentional fabrication and/or

184. CHRISTOPHER L.T. BROWN, *COMPUTER EVIDENCE: COLLECTION & PRESERVATION* 4 (2d ed. 2010).

manipulation of ESI on a computer system intended to thwart forensic examination. In short, anti-forensics is digital forgery.

The sophistication of anti-forensics may soon overtake (if it has not already) the ability to detect or defend against it. For example, in *United States v. Tippens*, the defendant proffered exhibits that he had acquired from Wikileaks that documented that agencies of the United States have the:

ability to hack into a computer without leaving any trace that it had been hacked or that an exploit had been placed on it . . . [such] that even if Defendant completed a thorough forensic examination of Defendant's computer and devices, Defendant would not be able to determine whether child pornography had been planted or whether security settings had been modified.¹⁸⁵

Such capabilities to thwart forensic detection of infiltration and tampering threaten the veracity of expert testimony regarding the results from a forensic examination. There will almost certainly be a race between forensic and anti-forensic capabilities as technology continues to advance.

5. Blockchain

Blockchain is a distributed digital ledger that maintains a continuously growing list of ordered records, called "blocks." It uses algorithms to encrypt data that is shared widely across numerous computers known as "nodes," so that no single person or organization controls that data. A hash is created to ensure trust on the network. Each signature is combined with others to form an unbreakable cryptographic chain that can be

185. No. CR 16-5110 RJB, 2017 WL 11511726, at *2 (W.D. Wash. Mar. 16, 2017).

independently tracked and its authenticity verified.¹⁸⁶ Transactions using a blockchain cannot be changed; they can only be reversed with another transaction. A block generally contains four pieces of information: (1) the hash of the previous block, (2) a summary of the included transaction, (3) a time stamp, and (4) the proof of work that went into creating the secure block.¹⁸⁷

To authenticate the data stored in the blockchain, the veracity of the data must be established before it is added to the blockchain. Therefore, the electronic devices (e.g., IoT) capturing the data must each be certified and authenticated independently.¹⁸⁸ The human element involved in these processes means that authenticating the link between the physical data and the digital data cannot be done by the blockchain technology itself as yet.¹⁸⁹ However, once the link is established, the evidence from the blockchain will establish the chain of custody. The blockchain will reveal whether a document has been manipulated, whether it is what it purports to be, and whether all data that is supposed to come with the document is actually there.

A blockchain network lacks a centralized point of vulnerability, making it extremely difficult for hackers to exploit. Further, as each block includes the previous block's hash, any attempts to alter any transaction within the blockchain will be detectable. Because the blockchain is a decentralized network that connects multiple parties, it would act as a single digital

186. John McKinlay et al., *Blockchain: background, challenges and legal issues*, DLA PIPER, (Feb. 2, 2018), <https://www.dlapiper.com/en/oman/insights/publications/2017/06/blockchain-background-challenges-legal-issues/>.

187. *Id.*

188. Adrian Clarke, *The Blockchain Can Finally Secure Supply Chains Against Cyberattacks*, (Dec. 26, 2018, 7:00 AM), <https://www.law.com/legaltech-news/2018/12/26/the-blockchain-can-finally-secure-supply-chains-against-cyberattacks/>.

189. *Id.*

master ledger for an entire financial system, enabling any transaction to be tracked from beginning to end.

Reported opinions in which ESI derived from blockchain ledgers was admitted into evidence include: *United States v Ulbricht* and *Alibaba Group Holding Limited v. Alibabacoin Foundation*.¹⁹⁰ In *Ulbricht*, the Department of Justice was able to identify Ulbricht as “Dread Pirate Roberts,” the operator of the online drug distribution system known as Silk Road. This was accomplished, in part, by tracing \$18 million worth of Bitcoin on Ulbricht’s computer to transactions on Silk Road servers using transaction history on Silk Road’s blockchain ledger.¹⁹¹ In *Alibaba*, the defendant attempted unsuccessfully to escape the reach of New York’s long-arm statute by introducing evidence that the subject transactions linked to New York were found to be on blockchain servers outside the United States in Minsk, Belarus.¹⁹²

Though blockchain can by itself be comparatively secure, it is not entirely invulnerable. It is only as secure as the system that it works on, the application that interacts with it, and the protocol that supports it (i.e., private and public keys), which are all vulnerable to attack resulting from human interaction. For example, blockchain is famously associated with Bitcoin and other cryptocurrency trading, which have been the subject of various reported scams. In February 2018, a complaint was filed in the Delaware Superior Court by Elizabeth White,¹⁹³ who was the

190. *United States v Ulbricht*, 858 F.3d 71 (2d. Cir. 2017); *Alibaba Grp. Holding Ltd. v. Alibabacoin Found.*, No. 18-CV-2897 (JPO), 2018 WL 5118638 (S.D.N.Y. Oct. 22, 2018).

191. *Ulbricht*, 858 F.3d at 87–88.

192. *Alibaba Grp. Holding Ltd.*, 2018 WL 5118638, at *3–4.

193. Rhys Dipshan, *Successful Fraud Case Breaks New Ground: Assistance from a Cryptocurrency Exchange*, LEGALTECH NEWS, (June 29, 2018 11:10 AM),

victim of cryptocurrency fraud in December 2017 by an anonymous man who contracted to trade Bitcoin for her XRP.¹⁹⁴ Instead, he manipulated the escrow and exchange platform Cointal to steal White's cryptocurrency. White was eventually able to trace her XRP to a digital wallet on the Delaware-registered cryptocurrency exchange platform Bittrex. An application was filed requiring Bittrex to disclose the identity of the anonymous fraudster and turn over White's stolen assets from his account. Default judgment was obtained against the anonymous fraudster and Cointal. With Bittrex's cooperation, she was able to recover her XRP.¹⁹⁵

The admissibility of blockchain receipts as evidence of some underlying activity that was recorded on a blockchain could raise hearsay issues. It could probably be admitted through certification by a qualified person under a combination of the "business records" exception to hearsay under Rule 803(6) and Rule 902(13) on the reliability of the system or process that produced it. To qualify as a "business record," testimony would be required from a programmer-custodian or similarly knowledgeable person that the blockchain receipt was generated at the time of the transaction and kept in the course of a regularly conducted business activity.¹⁹⁶

<https://www.law.com/legaltechnews/2018/06/29/successful-fraud-case-breaks-new-ground-assistance-from-a-cryptocurrency-exchange/>.

194. Jake Frankenfield, *Ripple (Cryptocurrency)*, INVESTOPEDIA (Aug. 11, 2019), <https://www.investopedia.com/terms/r/ripple-cryptocurrency.asp> ("Ripple is a technology that acts as both a cryptocurrency and a digital payment network for financial transactions. Ripple was released in 2012 and co-founded by Chris Larsen and Jed McCaleb. The coin for the cryptocurrency is premined and labeled XRP.").

195. Dipshan, *supra* note 193.

196. See 12 VT. STAT. ANN. § 1913. Blockchain enabling (2018) (providing rules for authentication, admissibility, and presumptions for blockchain records including that a blockchain digital record "shall be self-authenticating

Vermont recently implemented a statute to facilitate the authentication and admissibility of blockchain evidence.¹⁹⁷ The rule recognizes that blockchain can be self-authenticated under Vermont's version of Rule 902 if accompanied by the declaration of a qualified person. The text of the rule is provided in Appendix C, *infra*.

C. *Application of Federal Rules and Cases in State Court and Vice Versa*

1. Federal law application in state cases

Many states model their rules of evidence and procedure as much as possible after federal rules for many good reasons. The most prominent is that where a state and federal rule of evidence or procedure are the same or similar, most state court judges may use federal cases applying the equivalent rule in similar circumstances as guidance or persuasive authority.¹⁹⁸ In the case of electronic evidence, federal cases on discovery and admissibility issues are far more abundant than state cases, the latter of which normally remain unpublished unless a case is appealed. Federal district and magistrate judges also address ESI evidence and discovery issues far more often than state court judges, which adds to the quality and persuasiveness of federal decisions as a whole.

pursuant to Vermont Rule of Evidence 902, if it is accompanied by a written declaration of a qualified person"). *See also* Illinois Blockchain Technology Act, 205 ILL. COMP. STAT. 730/10 (2020) (permitted use of blockchain in a proceeding).

197. *Id.*

198. *Ellis v. Toshiba Am. Info. Sys., Inc.*, 218 Cal. App. 4th 853, 861, n.6 (Cal. 2013) ("There is little California case law regarding discovery of electronically stored information under the act. 'Because of the similarity of California and federal discovery law, federal decisions have historically been considered persuasive absent contrary California decisions.'").

In addressing an admissibility issue involving ESI evidence, if there is no binding state authority on the issue, a comparison of the applicable Federal Rule of Evidence with the analogous state rule is the first step. If the rules are identical or similar in all respects material to the case at hand, the applicable principles and guidance in this *Commentary* as well as any relevant federal cases applying the rule may serve as persuasive authority.

2. State law application in federal cases

Given that new ESI admissibility issues emerge frequently as technology and the culture of information creation and communication evolve, finding binding precedent for the application of evidentiary rules can be difficult. Many regard state courts as a suitable laboratory for developing federal rules and case law, especially when the state courts are addressing issues frequently and in systematic fashion. While federal courts are not bound by state court precedent, there is no reason litigators should not identify and cite state court cases in the absence of direct federal authority. A federal court may accept or reject the reasoning of the state court cases, but, because many state court rules of evidence are identical or similar to their federal counterparts, guidance from state courts may be useful. This is especially true for cases from the same state in which the federal court sits.

Some admissibility issues are especially common in state court, where unique jurisdiction establishes common issues. One such example is foreclosure cases, in which state court judges and judicial officers frequently encounter the issue of ESI evidence of promissory notes that pass from entity to successor entity. When the lender forecloses, proving ownership of the note at the time the foreclosure is filed can be problematic when challenged by the debtor. This raises issues of authentication and hearsay. It also implicates the business-records exception to the hearsay rule.

Admissibility of bank records in an industry that frequently assigns mortgages and notes can be challenging. For example, in Florida foreclosure cases where a successor corporation takes custody of business records created by a predecessor organization and integrates them within its own records, the acquired records are treated as having been “made” by the successor business, such that both records constitute the successor business’s singular “business record.”¹⁹⁹ When introducing such records, a successor business may establish the trustworthiness of records under the business-records exception by independently confirming the accuracy of the third party’s business records upon receipt and providing testimony setting forth the procedures used to independently verify the accuracy of the payment history records from the prior organization.²⁰⁰

Foreclosure cases and hearsay objections to documents presented in court play out in lower and appellate state courts. For example, *Jackson v. Household Financial Corporation III* held that introducing bank records through an employee who regularly reviewed home loans and claimed to be familiar with the bank’s loan servicing practices was sufficient foundation under the business-records exception for the initial foundation burden, thus shifting the burden to the opposing party. In doing so, Florida’s Supreme Court held that a qualified witness who testifies as to each element of the business-records exception for the admission of a business record lays sufficient predicate for admission of the document such that the document should be admitted unless the opponent establishes it to be untrustworthy.²⁰¹ However, *Knight v. GTE Federal Credit Union* held that the witness proffering a record was not competent to provide

199. See *Deutsche Bank Nat’l Trust Co. v. Sheward*, 245 So. 3d 890, 892–93 (Fla. Dist. Ct. App. 2018).

200. See *id.*

201. *Jackson v. Household Fin. Corp. III*, 298 So. 3d 531 (Fla. 2020).

foundation where he did not demonstrate that he was well enough acquainted with the entity's business practices to authenticate the letter. *Knight* premised its holding on the fact that the witness did not work for the servicing agent, never visited its facility, never spoke with its employee, and had no documents other than the servicer's letter log to support his testimony.²⁰²

In the context of a Florida foreclosure action, a representative of a loan servicer testifying at trial was not required to have personal knowledge of the documents being authenticated but did have to be familiar with and know how the company's data was produced.²⁰³ The witness must ultimately be well enough acquainted with the activity to provide testimony.²⁰⁴ *Wells Fargo Bank, N.A. v. Balkissoon* describes the qualifications needed for a witness qualifying records under the business-records exception to the hearsay rule.²⁰⁵ If the witness is sufficiently familiar with the records to be admitted, the witness need not be familiar with the mechanics of actually typing the data into the system because there is no requirement that the witness have such knowledge.²⁰⁶ However, in *Maslak v. Wells Fargo Bank, N.A.*, the opposite result occurred where a bank's witness did not know whether someone at outside counsel's office changed or modified a document; she failed to testify about how payments were received and processed or the bank's procedures for inputting

202. *Knight v. GTE Fed. Credit Union*, No. 2D16-3241, 2018 WL 844352, at *2-3 (Fla. Dist. Ct. App. Feb. 14, 2018).

203. *See Sanchez v. Suntrust Bank*, 179 So. 3d 538, 541 (Fla. Dist. Ct. App. 2015); *Glarum v. LaSalle Bank Nat'l Ass'n*, 83 So. 3d 780, 783 (Fla. Dist. Ct. App. 2011).

204. *Cayea v. CitiMortgage, Inc.*, 138 So. 3d 1214, 1217 (Fla. Dist. Ct. App. 2014); *Cooper v. State*, 45 So. 3d 490, 493 (Fla. Dist. Ct. App. 2010).

205. 183 So. 3d 1272, 1275-77 (Fla. Dist. Ct. App. 2016).

206. FLA. STAT. § 90.803(6)(a) (2014).

payment information or the computer system the bank used.²⁰⁷ Similarly, in *Cassell v. Green Planet Servicing, LLC*, testimony on the business-records exception was inadequate when the witness testified that she had no personal knowledge of the policies and procedures used by the entities that created the payment history and notice letters.²⁰⁸ Published authority making close distinctions in such cases may provide guidance to federal courts and other state courts looking at similar admissibility issues.

Foreclosure cases have raised admissibility issues relating to ownership of e-notes. In *Rivera v. Wells Fargo Bank, N.A.*, the borrowers in a foreclosure case challenged the ownership and admissibility of an e-note, which was the only original, signed evidence of indebtedness in the case.²⁰⁹ The appellate court affirmed the foreclosure, holding that the bank proved foundation for admissibility and ownership of the electronic document.

In *DiGiovanni v. Deutsche Bank National Trust Company*, a printout produced from the trial judge's own internet research during a foreclosure trial was held to be not properly authenticated.²¹⁰ Because websites are not self-authenticating, the party proffering the evidence had to produce some statement or affidavit from someone with knowledge of the website. The appellate court also held that it was improper for the judge to do *ex parte* fact research on the internet.

207. 190 So. 3d 656, 659–60 (Fla. Dist. Ct. App. 2016).

208. 188 So. 3d 104, 105 (Fla. Dist. Ct. App. 2016).

209. 189 So. 3d 323, 327–29 (Fla. Dist. Ct. App. 2016).

210. 226 So. 3d 984, 988–89 (Fla. Dist. Ct. App. 2017).

IV. PRACTICAL GUIDANCE ON THE USE OF ESI IN COURT

Judge Grimm's discussion in *Lorraine* makes it clear that parties should start to think about evidentiary issues much earlier than was the practice when dealing only with hard-copy materials. This is especially critical because parties will need to ensure they have defensible preservation and collection protocols in place to maintain the information that the amended Federal Rules of Evidence require in the certification. Thus, parties should approach the discovery of ESI by always keeping the end goal—the successful admission of evidence—in mind.

The first step is to assess what potentially discoverable information is available. Only with that understanding can parties determine the appropriate scope of discovery, the proper tools and resources required to harvest the ESI, and the proportionality—or lack thereof—of the cost of discovery compared to the needs of the case. To the extent possible, parties should strive to collect only that data that is necessary for the case, narrowing the scope of the collection as much as possible by using relevant file types, date ranges, and the like. The prerequisite steps here include identifying and interviewing custodians and determining where data is stored. Another is determining who owns that information. For example, if a social media platform owns information, or if an individual has potentially relevant information on a personal cell phone, special permission and methods may be needed to preserve and collect that data.

As parties collect data, they should take steps to ensure they maintain its integrity. To this end, they should use the appropriate approach, which could include using a write-blocking solution that avoids data alteration. The improper collection of data, including metadata, can lead to data loss, alteration, or manipulation.

Before and after collection, parties should engage in quality assurance to validate that the data's integrity is intact. One way

to do this is to perform a hash analysis, both before and after collection, to ensure that the collection process did not alter any files.

In assessing whether to self-collect or to outsource data collection entirely, a key consideration is how much cost and risk the organization is willing to bear in collecting the data. That may vary from case to case. Self-collection, which comes in different forms, is often the fastest and least expensive way to collect data, but the individuals doing the collecting may lack specialized training and tools. Outsourcing offers the benefit of allowing trained forensic data professionals with the proper tools to perform collections.

No matter the method of collection, an essential step is to document the chronology of the ESI, including details about its custody, control, transfer, and disposition, in a chain of custody that can be used to authenticate the evidence later in the case. The documentation should also log who collected and handled the data at each stage.

A. Use of ESI in Static vs. Native/Live Format

In the past, parties were limited to sharing exhibits in hard copy. Today, parties can instead choose between static format and native (or live) format—the format in which the ESI was created and maintained—when presenting ESI. Parties should evaluate the advantages and disadvantages of different formats early in discovery, as these decisions can later affect the evidence they are able to present at trial.²¹¹

Static ESI, often presented in TIFF (tagged image file format) or PDF file format, may be simpler and less expensive to produce than native images, because it does not require any special know-how or tools. Its simplicity also makes it easier to copy,

211. See *Primer on Social Media*, *supra* note 75, at 44.

share, and authenticate. But it has several drawbacks that can make it inferior to native format ESI in many cases, particularly when the ESI is dynamic and complex.

One clear advantage of native format ESI is that it maintains the characteristics of data that would be lost if we reduced the data to static form, such as by playing a video or sound recording, revealing the formulas behind spreadsheet cells, or running a process. Another advantage is that native format files allow parties to manipulate data for demonstrative purposes without destroying the underlying data. A static form of ESI may also lack metadata that may be helpful to interpreting the ESI's origin. Of course, with these benefits comes the hardship of ensuring that data does not become corrupted and the potential requirements for additional hardware or software as well as technical expertise.

B. Evidence to Assist the Jury on the Permissive Spoliation Inference

Spoliation occurs where “the evidence was in the party’s control; the evidence is relevant to the claims or defenses in the case; there has been actual suppression or withholding of evidence; and, the duty to preserve the evidence was reasonably foreseeable to the party.”²¹² A range of sanctions is available when a party destroys ESI “with the intent to deprive another party of the information’s use in the litigation.”²¹³ The trial court

212. *Pace v. Wal-Mart Stores East, LP*, 799 F. App’x 127, 130 (3d Cir. 2020).

213. FED. R. CIV. P. 37(e)(2). The admission of relevant evidence of spoliation is also an option under Rule 37(e)(1) to address prejudice and without a finding of intent to deprive. Courts exercising that option have tried to explain why the evidence is admissible. *See EPAC Techs., Inc. v. Thomas Nelson, Inc.*, No. 3:12-cv-00463, 2018 WL 3322305, at *3 (M.D. Tenn. May 14, 2018) and *Karsch v. Blink Health Ltd.*, 17-CV-3880 (VM) (BCM), 2019 WL 2708125, at *27–28 (S.D.N.Y. June 20, 2019). The degree to which it makes a

may, for example, dismiss the action or impose default judgment. It may instead, however, instruct the jury that it may or must presume that the lost ESI was unfavorable to the spoliator.²¹⁴

If the court elects to give a permissive inference instruction to the jury, evidence may be presented to the jury to aid in the determination of whether to draw the adverse inference while at the same time avoiding unfair prejudice confusion of the issues, misleading the jury, or undue delay.²¹⁵ This issue was addressed in *GN Netcom, Inc. v. Plantronics, Inc.*²¹⁶

During the course of discovery in this antitrust action, plaintiff GN learned that defendant Plantronics had engaged in extensive destruction of ESI. GN moved for default judgment as a sanction. Following a hearing, the district court found that Plantronics had acted in bad faith with the intent to deprive GN of relevant evidence but declined to order default judgment.²¹⁷ Instead, the trial court opted to give the jury a permissive adverse inference instruction while fining Plantronics \$3 million and directing “it to pay GN’s spoliation-related fees.”²¹⁸

At trial, GN sought to introduce evidence of the spoliation, including testimony from an expert witness, Dan Gallivan, on

fact material to the claims or defenses “more or less probably” is crucial. See *Duran v. County of Clinton, NO. 4:14-CV-2047*, 2019 WL 2867273, at *5 (M.D.Pa. July 3, 2019).

214. FED. R. CIV. P. 37(e)(2).

215. Federal Rule of Evidence 403 provides that “[t]he court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.”

216. 930 F.3d 76 (3d Cir. 2019).

217. *Id.* at 81.

218. *Id.*

the extent of the spoliation. Concerned that the spoliation evidence would obscure the dispositive antitrust questions presented in the case, the court refused to allow Gallivan to testify. Instead, the court determined that the jury would consider 17 stipulations concerning the spoliation. The jury returned a verdict in favor of Plantronics.

On appeal, a divided Third Circuit concluded that the exclusion of the expert testimony was an abuse of discretion. Finding that the stipulation on the extent of the spoliation was extremely vague (“[I]t may be that several hundred or even up to 15,000 potentially responsive relevant emails were deleted or destroyed”²¹⁹), the majority explained that the expert’s proffered testimony was highly probative:

Gallivan’s expert testimony would have assisted the jury in narrowing that range, giving it evidence on which it could base an important decision: whether Plantronics engaged in a “massive cover-up.” Without Gallivan’s testimony, it is possible, if not entirely probable, that jurors concluded that only a few hundred emails were deleted, falling short of a massive cover-up; however, if they had evidence that fifteen, five, or even just one thousand emails had been deleted, they might have taken a very different view on whether to apply the adverse inference. . . . The “maximum reasonable probative force” of his testimony was high; therefore, the District Court could have properly excluded it only if that probative value was substantially outweighed by the

219. *Id.* at 87.

evidence's potential prejudice or by other risks outlined in Rule 403.²²⁰

Observing that "highly probative evidence is 'exceptionally difficult to exclude,'"²²¹ the majority ruled that Gallivan's testimony was not unfairly prejudicial, was likely to clarify the stipulations, would not mislead the jury, and would not have unduly prolonged the trial.

The dissenting judge believed that the majority had assigned too little weight to the spoliation stipulations, stating that "[t]hese stipulations gave the jury an adequate basis to decide whether to adopt the permissive adverse inference."²²² The dissent also determined that "the majority fail[ed] to give the required deference to the District Court's reasonable conclusions that Gallivan's spoliation testimony posed a substantial risk of distracting the jury from the antitrust merits of the case and that such risk eclipsed the testimony's probative value."²²³

GN Netcom illustrates the delicate balancing of interests that must be undertaken when a jury is being asked to decide whether to draw an adverse inference against a bad-faith spoliator. On the one hand, there is a strong preference to have cases adjudicated on their merits. On the other hand, there is an equally strong concern that the jury should have an adequate presentation of the facts underlying the trial court's decision to give the permissive inference instruction. In some cases, that adequate presentation cannot be made by way of stipulations.

220. *Id.*

221. *Id.* at 85.

222. *Id.* at 91 (Smith, C.J., dissenting).

223. *Id.*

C. *Practical Tips for Administration of ESI as Evidence*

ESI admissibility issues should be addressed as early as possible. Consideration should be given to incorporating agreements regarding admissibility into production stipulations or submitting these agreements to the court for approval. This may not be available in criminal cases.

As the degree to which ESI is static decreases, the difficulties of replicating it increase. Therefore, care should be taken to choose the most replicable form of ESI that provides the necessary probative information (including metadata).

D. *Practical Tips for Seeking Authority on Admission of ESI as Evidence*

Finding case support for admissibility of ESI can be challenging because so few civil cases are actually tried.²²⁴ However, the Federal Rules of Evidence are trans-substantive and apply in civil and criminal proceedings.²²⁵ The only exceptions to the applicability in criminal cases are grand-jury proceedings and “miscellaneous proceedings” such as extradition or rendition;

224. See *Civil Jury Project at NYU School of Law*, <https://civiljuryproject.law.nyu.edu/about/> (last visited May 7, 2020) (“[I]t is beyond dispute that the civil jury trial is a vanishing feature of the American legal landscape. In 2018 . . . 0.5 percent of federal civil cases were tried before juries—down from 5.5 percent in 1962. This amounted to an average of 2 civil jury trials per authorized federal judgeship in 2018—down from 10 in 1962. Similar trends are evident in states across the nation.”).

225. FED. R. EVID. 1101(b) (“These rules apply in: civil cases and proceedings, including bankruptcy, admiralty, and maritime cases; criminal cases and proceedings; and contempt proceedings”); see also Stephan Landsman, *Are the Federal Rules of Evidence Dynamite?* 33 B.U. INT’L L.J. 343, 351 (2015) (“A fourth characteristic that strongly colors the FRE is its commitment to a ‘trans-substantive’ approach to the rules of evidence While that approach is open to a variety of criticisms, it expresses important values. Chief among them is a democratic impulse that all litigants be treated alike.”).

issuing an arrest warrant, criminal summons, or search warrant; a preliminary examination in a criminal case; sentencing; granting or revoking probation or supervised release; and considering whether to release on bail or otherwise.²²⁶ Far more criminal cases end up being tried, and the nature of criminal practice necessarily involves frequent challenges to admissibility and less formal discovery pathways to resolution of authenticity, such as civil requests for admission. Thus, criminal cases should be included in legal research on admissibility issues for civil cases. Criminal cases are creating authority on admissibility of social media,²²⁷ digital security camera ESI,²²⁸ text messaging,²²⁹ emoji,²³⁰ and other forms of ESI.

State court criminal cases may provide helpful or persuasive authority on specific issues of admissibility. For example, authentication of a Facebook video may be accomplished under Rule 901(b)(3) (comparison with an authenticated specimen by an expert witness or the trier of fact) and 901(b)(4) (appearance,

226. FED. R. EVID. 1101(d)(2)–(3).

227. *See, e.g.*, *State v. Smith*, 181 A.3d 118, 134–36 (Conn. App. Ct. 2018) (authenticating Facebook messages using circumstantial evidence); *Lamb v. State*, 246 So. 3d 400, 409–10 (Fla. Dist. Ct. App. 2018) (authenticating and admitting a Facebook Live video); *State v. Hannah*, 151 A.3d 99, 107 (N.J. Super. Ct. App. Div. 2016) (authenticating Twitter posting using circumstantial evidence and reply doctrine).

228. *See, e.g.*, *People v. Taylor*, 956 N.E.2d 431, 438–43 (Ill. 2011) (copy of motion-activated video in non-native format).

229. *See, e.g.*, *State v. Papineau*, 190 A.3d 913, 935–36 (Conn. App. Ct. 2018) (allowing circumstantial evidence of authorship to authenticate text messages); *Pavlovich v. State*, 6 N.E.3d 969, 978–79 (Ind. Ct. App. 2014) (using circumstantial evidence to authenticate text messages); *State v. Young*, 369 P.3d 205, 208–09 (Wash. Ct. App. 2016) (using content to authenticate text messages).

230. *See* Section II.D.12., *supra*.

contents, substance, internal patterns, or other distinctive characteristics of the item, taken together with all the circumstances).

In *Lamb v. State*, a Florida criminal case, one of the defendant's phones contained a Facebook video posted twenty-one minutes after two crimes, showing the defendants with the two stolen vehicles and a stolen watch on a defendant's wrist.²³¹ Over objection, the appellate court applied equivalent Rule 901 principles and held that the prosecution sufficiently authenticated a social media video under Florida Statute § 90.901.²³²

Conversely, in a prosecution for aggravated assault, the Superior Court of Pennsylvania upheld the trial court's exclusion of Facebook postings that the prosecution attempted to use to link the defendant to the assault.²³³ The prosecution could show that the account bore defendant's name, high school, and hometown but was unable to show that the defendant authored ambiguous chat messages or posted a photo of bloody hands because there were no contextual clues, and third persons were posting some of the information in question. Thus, the trial court did not abuse its discretion in finding that the social media evidence lacked authentication.²³⁴

At least one state has gone so far as adopting an evidence rule specifically dealing with authentication of emails. Washington Evidence Rule 901(b)(10) sets forth the factors that may be used to authenticate email:

(b) Illustrations. By way of illustration only, and not by way of limitation, the following are

231. *Lamb*, 246 So. 3d at 408–10.

232. *Id.* at 410.

233. *Commonwealth v. Mangel*, 181 A.3d 1154, 1163–64 (Pa. Super. Ct. 2018).

234. *Id.* at 1164.

examples of authentication or identification conforming with the requirements of this Rule:

....

(10) *Electronic Mail (E-mail)*. Testimony by a person with knowledge that (i) the email purports to be authored or created by the particular sender or the sender's agent; (ii) the email purports to be sent from an e-mail address associated with the particular sender or the sender's agent; and (iii) the appearance, contents, substance, internal patterns, or other distinctive characteristics of the e-mail, taken in conjunction with the circumstances, are sufficient to support a finding that the e-mail in question is what the proponent claims.²³⁵

These factors have been applied by analogy to other forms of electronic communication.²³⁶

235. WASH. R. EVID. 901(b)(10).

236. See *State v. Young*, 369 P.3d 205, 208–09 (Wash. Ct. App. 2016) (text messaging); *In re Detention of H.N.*, 355 P.3d 294, 302 (Wash. Ct. App. 2015) (authenticating emailed screenshots of text messages by analogy to Wash. R. Evid. 910(b)(10)).

V. ARTIFICIAL INTELLIGENCE USES IN BUSINESS AND LAW²³⁷

Artificial intelligence (AI) is making major inroads into many industries such as health care, automotive, fitness, financial services, and even litigation. This is and will continue to present significant legal, technological, and ethical challenges for lawyers.²³⁸

In late 2019, before Covid-19 became a pandemic, a Canadian-based company, BlueDot, used AI to identify an emerging health risk in Wuhan, China. That AI subsequently predicted the global spread of the disease.²³⁹ Voice-controlled personal

237. The Editors wish to acknowledge the significant contribution of The Hon. Paul W. Grimm to this discussion and thank him for allowing us to borrow extensively from his forthcoming law review article on this topic. However, the final draft of this Commentary represents consensus of the drafting team and the Working Group 1 Steering Committee and should not be imputed to any individual contributor.

238. Under the Model Rules of Professional Conduct, lawyers must assess whether they have the requisite skill and knowledge, including *understanding the benefits and risks of the technology involved*, to perform the tasks (either by themselves or in collaboration with an experienced counsel or consultant) involving AI such as: (i) assisting their client in identifying sources (including custodians) of relevant ESI; (ii) engaging in meaningful meet-and-confer sessions with opposing counsel concerning an eDiscovery plan that targets AI as a data source; and (iii) advising a client about the proper method to collect responsive ESI in a manner that preserves the integrity of that ESI for evidentiary purposes when AI is the data source. These challenges will test a lawyer's ability to comply with, among others, Rule 1.1 (competence), Rule 1.3 (diligence), Rule 1.4 (Communications), Rules 5.1 and Rule 5.3 (Supervision), and Rule 5.4 (Professional Independence of a Lawyer).

239. Isaac I. Bogoch, et al., *Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel*, 27(2) J. TRAVEL MED. (Mar. 2020), <https://bluedot.global/>. See also Cory Stieg, *How this Canadian start-up spotted coronavirus before everyone else knew about it*, CNBC (Mar. 3, 2020, 10:27 a.m.), <https://www.cnbc.com/2020/03/03/bluedot-used-artificial-intelligence-to-predict-coronavirus-spread.html>. BlueDot also has used its

assistants with evolving “personality” traits allow the “assistant” to use machine-learning algorithms to learn how to pattern its behavior after the “voice-controller.” The assistant also has a visual component that allows it to use machine-learning algorithms to understand a voice-controller’s facial expressions, voice inflections, and verbal patterns from conversations. Robotic vacuums use AI to scan room size, identify obstacles, and remember the most efficient routes for cleaning.

AI also is making major inroads into law-related activities beyond technology-assisted review.²⁴⁰ A software program called COMPASS uses AI technology to assess the risk that defendants awaiting sentencing will re-offend, allowing sentencing judges to consider this risk in fashioning conditions of supervision. Similarly, facial recognition software, using AI algorithms, is being used by law enforcement agencies to identify suspects and fugitives in a crowd or captured on closed-circuit television videos (CCTV). Machine-learning algorithms can automatically analyze draft contracts and identify which portions of the contract are acceptable and which are problematic based on prior contracts. In addition, machine-learning models are being used to predict the outcomes of pending cases, using inputs from automated legal research and contextualization of the case’s particular fact pattern.

Technology that employs AI programming also will present significant evidentiary challenges when it is offered at hearings and trials. Although, to date, no reported court decision has

AI to make early predictions about where the Zika virus and the Ebola outbreak would spread.

240. Ellen M. Gregg, et al., *How Artificial Intelligence is Impacting Litigators*, ALAS LOSS PREVENTION JOURNAL 49 (Summer 2019); and Rob Toews, *AI Will Transform the Field of Law*, FORBES (Dec. 19, 2019 2:09 p.m.), <https://www.forbes.com/sites/robtoews/2019/12/19/ai-will-transform-the-field-of-law/#e1907ed7f01e>.

been found that comprehensively explores the many evidentiary issues that surround determinations of admissibility of AI, there are a number of rules of evidence that are likely to figure prominently in any such determination. Although there is no single rule in the Federal Rules of Evidence that specifically addresses admissibility of AI technology, Rule 102 encourages counsel and courts to employ the existing rules of evidence to “administer every proceeding fairly, eliminate unjustifiable expense and delay, and promote the development of evidence law, to the end of ascertaining the truth and securing a just determination.”²⁴¹ In essence, the existing rules of evidence are flexible enough to address novel evidentiary challenges not already directly covered in the rules. There are, however, several rules of evidence that offer great promise in connection with determining admissibility of AI evidence.

The starting place is Rule 401, which defines relevance.²⁴² Evidence is relevant if it has “*any* tendency to make a fact more or less probable than it would be without the evidence,” and “the fact is of consequence in determining the action.”²⁴³ Relatedly, irrelevant evidence is never admissible. But relevant, and therefore presumptively admissible, evidence may nonetheless be excluded if precluded by the U.S. Constitution, a federal statute, the rules of evidence, or other rules promulgated by the Supreme Court.²⁴⁴ Relevant evidence also is inadmissible if its probative value is substantially outweighed by the danger of unfair prejudice, confusing the issues, misleading the fact finder, wasting time, or is needlessly cumulative.²⁴⁵ Finally, even if relevant

241. FED. R. EVID. 102.

242. FED. R. EVID. 401.

243. *Id.* (emphasis added).

244. FED. R. EVID. 402.

245. FED. R. EVID. 403.

and not otherwise excluded, the fact that evidence is relevant (i.e. may be considered by the fact finder) is no guarantee that it will be deemed credible or given much weight by the fact finder.²⁴⁶

In framing this discussion, there are some “big picture” evidentiary concepts to keep in mind when considering the admissibility of AI evidence. First, if a foundation cannot be established to show that the AI-powered technology produces accurate results, the evidence is unreliable and therefore has no relevance. Unreliable evidence has no tendency to prove or disprove facts that are of consequence to resolving a case or issue. But the challenge for lawyers and judges alike is that determining the reliability of AI evidence depends on understanding how the applicable algorithm works. Given the countless applications for AI technology in connection with doing a seemingly endless number of technical chores, the proponent, opponent, and judge deciding whether to admit this evidence must have sufficient information to understand how the technology works. While individuals technically trained in the operation of AI applications may be able to explain *what* the algorithm did and the results the algorithm obtained, those individuals may have difficulty explaining the complexity as to how the algorithm was programmed, or how it produces accurate results. For example, the algorithm developed by the Canadian company Blue Dot (mentioned above) to predict the origins and transmission of the Covid-19 virus took a year to develop and involved input from an “eclectic mix of engineers, ecologists, geographers, and veterinarians, all under one roof”, and entailed “training” the computer to detect 150 deadly pathogens through use of thousands of examples.²⁴⁷ Imagine the challenge a lawyer might face when

246. FED. R. EVID. 104(e).

247. CBS 60 Minutes: *The Computer Algorithm That Was Among the First to Detect the Coronavirus Outbreak* (Apr. 27, 2020).

trying to establish the reliability for this AI application and have evidence of the results of the Blue Dot technology admitted in a trial. Fundamentally, this is an issue of authentication—showing that the technology produces the results that its proponents claim it produces.²⁴⁸

Rule 901(b) provides ten nonexclusive examples of how authentication of nontestimonial evidence can be accomplished. Two of them are most likely to be helpful in authenticating AI evidence. First, Rule 901(b)(1) permits the authentication of evidence through “[t]estimony that an item is what it is claimed to be.” If this rule is used, then the witness must either meet the conditions of Rule 602 (requiring that witnesses have personal knowledge of the matters they testify about) or meet the qualification requirements of Rule 702 (that the witness have sufficient expertise to testify to a matter requiring scientific, technical, or specialized knowledge, experience, or training, in which case the witness may testify in the form of an opinion, or otherwise). If the witness qualifies under Rule 702, then his opinion testimony may be based on information not personally known by the witness, so long as that information is of the type that similar experts reasonably would rely on.²⁴⁹ Using the Blue Dot AI technology as an illustration, it is easy to see why a qualified expert would be the most useful person to authenticate that the Blue Dot algorithm produces accurate results, given that it was developed by multiple individuals with different specialties. And a single expert may be sufficient to base his testimony on reliable information provided by the many other experts who helped to develop the algorithm. The time-consuming, and likely expensive, alternative would be to call multiple witnesses

248. FED. R. EVID. 901(a).

249. FED. R. EVID. 703.

to authenticate the algorithm if their testimony was limited to facts about which they have personal knowledge.

Rule 901(b)(9) is the second method of authentication that is likely to be most useful in authenticating AI evidence. It permits authentication by producing evidence “describing a process or system and showing that it produces an accurate result.”²⁵⁰ In this regard, authenticating AI evidence using Rule 901(b)(9) will usually, if not always, be done the same way described above for Rule 901(b)(1)—one or more witnesses with personal knowledge of the authenticating facts, or one or more witnesses meeting the qualifications of Rule 702.

There is an important feature of authentication that needs to be given careful consideration in connection with admitting AI evidence. Normally, a party has fulfilled its obligation to authenticate nontestimonial evidence by producing facts that are sufficient for a reasonable fact finder to conclude that the evidence *more likely than not* is what its proponent claims it is—by a mere preponderance.²⁵¹ This is a relatively low threshold—51 percent, slightly better than a coin toss. However, given the complexity of AI algorithms, and the tasks that they can accomplish that would be otherwise impossible, or nearly so, judges may be reluctant to allow a jury to consider AI evidence if its reliability has been established by little more than an

250. FED. R. EVID. 901(b)(9).

251. See 31A FEDERAL PRACTICE AND PROCEDURE, EVIDENCE, 2013 QUICK REFERENCE GUIDE, 383 (“Rule 901(a) prescribes that authentication or identification of an item requires only evidence sufficient to support a finding—a ‘prima facie case’—that the item is genuine. A bona fide dispute as to authenticity or identity is not to be decided by the judge, but rather is to go to the jury In other words, conflicting evidence on genuineness goes to weight, not admissibility, so long as some reasonable person could believe that the item is what it is claimed to be.”); *Ricketts v. City of Hartford*, 74 F. 3d 1397, 1411 (2d Cir. 1996); *United States v. Johnson*, 637 F. 2d 1224, 1247 (9th Cir. 1980).

evidentiary coin toss. Because the judge must act as the gatekeeper who determines whether the evidence that may be considered by the jury,²⁵² a party relying on AI evidence would be wise to provide as much evidence as possible to authenticate the AI.

One way a party can enhance the weight of the evidence it offers to authenticate AI applications is to clearly demonstrate how the results it produces are accurate. In this task, Rule 702 and the cases that have explored the criteria for admitting scientific or technical evidence provide helpful guidance. Rule 702 requires that expert testimony be based on sufficient facts and reliable methodology, reliably applied to the facts of the case.²⁵³ These factors were added to the evidence rules in 2000²⁵⁴ to bolster the rule in light of the Supreme Court's decisions in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*²⁵⁵ and *Kumho Tire Co. v. Carmichael*.²⁵⁶ Therefore, while Rule 702 was not intended to codify the decision in *Daubert*, the factors discussed in that decision relating to determining the reliability of scientific or technical evidence are quite informative in showing that Rule 702's reliability factor has been met. As described in the Advisory Committee Notes to Rule 702, the "*Daubert* Factors" are:

- (1) whether the expert's technique or theory can be or has been tested . . . ;
- (2) whether the technique or theory has been subject to peer review and publication;
- (3) the known or potential rate of error of the technique or theory when applied;
- (4) the

252. See FED. R. EVID. 104(a) ("The court must decide any preliminary question about whether . . . evidence is admissible").

253. FED. R. EVID. 702(b)-(d).

254. FED. R. EVID. 702 advisory committee's notes to 2000 amendment.

255. 509 U.S. 579 (1993).

256. 526 U.S. 137 (1999).

existence and maintenance of standards and controls; and (5) whether the technique or theory has been generally accepted in the scientific [or technical] community.

To authenticate AI technology, a proponent must show that the technology produces accurate, reliable results. When the accuracy of technical evidence has been verified by testing; the methodology used to develop it has been published and subject to review by others in the same field of science or technology; when the error rate associated with its use is not unacceptably high; when standard testing methods and protocols have been followed; and when the methodology used is generally accepted within the field of similar scientists or technologists; then it can be established as authentic because it does what its proponents say it does. Contrastingly, when the accuracy of evidence has not been tested; when its methodology has been treated as a trade secret by its developer, and not verified by others; when applied it produces an unacceptably high error rate; when standard procedures not followed when the methodology was developed or employed; or when the methodology is not accepted by others in the same field; then it would be challenging to maintain that the methodology does what its proponent claims it can do, which might render the evidence inadmissible. The bottom line is that if a proponent is going to rely on evidence produced by AI technology, he should consider these factors and marshal facts to show compliance with as many of factors as possible.

The final rule that is promising when authenticating AI technology is Rule 902(13), which permits the self-authentication of certified records generated by an electronic system or process shown to produce an accurate result.²⁵⁷ In lieu of calling one or

257. FED. R. EVID. 902(13).

more witnesses to establish the accuracy of the results of the AI technology, the party planning to introduce the AI evidence can prepare a certificate that meets the requirements of Rule 902(11). The records generated by the AI technology and the authenticating certificate are then produced in advance of the trial or hearing where the evidence will be introduced, and if there is no objection raised, the evidence is authenticated without the need to call live witnesses. This can significantly reduce the cost of authenticating AI evidence at a hearing or trial. But Rule 902(13) is no shortcut for completeness or accuracy in providing the facts necessary to show the accuracy of the AI technology. In fact, to succeed, the certificate must be as detailed and complete as live testimony by the witnesses with personal knowledge or technical expertise who would be called if the proponent of the AI evidence planned to authenticate it with witnesses. And the person or persons who provide the certificate must be similarly qualified (i.e., personal knowledge or scientific or technological expertise).

Given the rapid expansion of the use of AI in major industries and the evidentiary issues presented by AI, The Sedona Conference Working Group 1 will continue to watch this area of the law and update this *Commentary* as appropriate.

**APPENDIX A: SUMMARY FEDERAL RULES
OF EVIDENCE 901 AND 902
RULES FOR AUTHENTICATION**

Type of e-Evidence: Email	
FRE Rules	Methods
<p>Rule 901(b)(1): Testimony of a witness with knowledge that the document is what it purports to be.</p> <p>Rule 901(b)(4): Appearance, content.</p> <p>Rule 901(b)(9): System or process capable of proving a reliable and dependable result [902(13,14)].</p> <p>Rule 902(7): Trade inscriptions.</p> <p>Rule 902(11): Self authenticating.</p>	<p>Witness testifies on process of creation, acquisition, preservation etc.:</p> <ol style="list-style-type: none"> i. who sent: author, ii. who received, iii. someone who saw it being authored/received, iv. email chain recipient: accuracy of contents. <p>Business records: Rule 803(6) certificate by a qualified witness.</p> <p>Production of document in discovery.</p> <p>Circumstantial evidence: about authorship, content, writing style, etc.</p> <p>Forensic information, hash values, etc.</p>
Cases	
<p><i>Lorraine v. Markel Am. Ins. Co.</i>, 241 F.R.D. 534, 538–39, 547 (D. Md. 2007) (noting that “[h]ash values can be inserted into original electronic documents when they are created to provide them with distinctive characteristics that will permit their authentication under Rule 901(b)(4).”).</p>	

Type of e-Evidence: Email

United States v. White, 660 F. App'x 779, 783 (11th Cir. 2016) (allowing a witness to authenticate an email chain with many emails sent between the defendant and the witness and holding the “anomalies and inconsistencies” in the email insufficient to affect the admissibility of the documents).

United States v. Cone, 714 F.3d 197, 220 (4th Cir. 2013) (“While properly authenticated e-mails may be admitted into evidence under the business records exception, it would be insufficient to survive a hearsay challenge simply to say that since a business keeps and receives e-mails, then *ergo* all those e-mails are business records falling within the ambit of Rule 803(6)(B).”).

Broadspring, Inc. v. Congoo, LLC, No. 13-CV-1866 (JMF), 2014 WL 7392905, at * 3 (S.D.N.Y. Dec. 29, 2014) (holding that third-party emails sent to a party in the ordinary course of business and produced by the party in litigation are sufficiently authenticated).

Nola Fine Art, Inc. v. Ducks Unlimited, Inc., 88 F. Supp. 3d 602, 607 (E.D. La. 2015) (“[Defendant] produced the email to plaintiffs in discovery and therefore cannot seriously dispute the email’s authenticity.”).

United States v. Siddiqui, 235 F.3d 1318, 1322 (11th Cir. 2000) (holding that an email identified as originating from the defendant’s email address and that automatically included the defendant’s address when the reply function was selected was considered sufficiently authenticated).

FRE Rule	Method
Rule 901(b)(3): Comparison by trier or expert witness.	Expert witness may explain either the technology or the method.
FRE Rule	Method
Rule 901(b)(4): Distinctive characteristics and the like.	Appearance, content.

Type of e-Evidence: Text Messages	
FRE Rules	Methods
As above.	As above for 901(b)(1). Additionally: <ul style="list-style-type: none"> • the purported author’s ownership of the phone or other device from which the text was sent, • the author’s possession of the phone, • the author’s known phone number, • the author’s name, • the author’s name as stored on the recipient’s phone, and • the author’s customary use of emoji or emoticons.
Cases	
<p><i>United States v. Teran</i>, 496 F. App’x 287, 292 (4th Cir. 2012) (holding that threatening texts were authenticated where recipient testified to personal nature of messages and texts aligned with defendant’s knowledge of recipient’s family).</p> <p><i>United States v. Kilpatrick</i>, No. 10-20403, 2012 WL 3236727, at *3–6 (E.D. Mich. Aug. 7, 2012) (holding that texts were authenticated where SkyTel records-custodian verified that the texts had not been and could not be edited in any way because texts were automatically saved on SkyTel’s server with no capacity for editing).</p> <p><i>United States v. Ramirez</i>, 658 F. App’x 949, 952 (11th Cir. 2016) (admitting photos that were sent by text because the recipient testified she received them, an agent testified he was present when the</p>	

Type of e-Evidence: Text Messages

texts were sent, and the defendant was listed as the owner of the phone number sending the texts).

United States v. Barnes, 803 F.3d 209, 217 (5th Cir. 2015) (finding that government laid a proper foundation to authenticate Facebook and text messages as having been sent by the defendant where recipient testified she had seen the defendant use Facebook, she recognized his Facebook account, and the messages matched his manner of communicating; and further stating “[a]lthough she was not certain that [the defendant] authored the messages, conclusive proof of authenticity is not required for admission of disputed evidence”).

Type of e-Evidence: Mobile Devices, Voicemail

FRE Rules	Methods
<p>Rule 901 (b)(1): Testimony of a witness with knowledge that the document is what it purports to be.</p>	<p>A witness who overheard the person leaving the message and can say the message being offered into evidence is the same message, or use chain of custody.</p>
Cases	
<p><i>Furlev Sales & Assocs., Inc. v. N. Am. Auto. Warehouse, Inc.</i>, 325 N.W.2d 20, 27 n.9 (Minn. 1982) (noting seven foundational elements for admission of a tape recording that have the potential to apply to ESI).</p> <p><i>State v. Williams</i>, 150 P.3d 111, 118 n.7 (Wash. Ct. App. 2007) (“[i]dentification of a voice [whether firsthand or through mechanical or electronic transmission or recording] by opinion based upon hearing the voice at any time under circumstances connecting it with the alleged speaker”) (quoting Wash. R. Evid. 901(b)(5)).</p>	

Type of e-Evidence: Internet Websites/Web pages	
FRE Rules	Methods
<p>Rule 901(b)(1): Testimony of a witness with knowledge.</p> <p>Rules 902 (5), (7) and (11): Public authorities' websites: self-authenticating official publication.</p>	<p>Follow Rules 104(a) and (b):</p> <ol style="list-style-type: none"> i. What was actually on the website? ii. Does the exhibit or testimony accurately reflect it? iii. If so, is it attributable to the owner of the site? <p>Consider the totality of the circumstances, e.g.:</p> <ul style="list-style-type: none"> • "The length of time the data was posted on the site; • Whether others report having seen it; • Whether it remains on the website for the court to verify; • Whether the data is of a type ordinarily posted on that website or websites of similar entities (e.g., financial information from corporations); • Whether the owner of the site has elsewhere published the same data, in whole or in part;

Type of e-Evidence: Internet Websites/Web pages	
	<ul style="list-style-type: none"> • Whether others have published the same data, in whole or in part; • Whether the data has been republished by others who identify the source of the data as the website in question.”²⁵⁸
Cases	
<p><i>U.S. Equal Emp’t Opportunity Comm’n v. E.I. DuPont de Nemours & Co.</i>, No. Civ. A. 03-1605, 2004 WL 2347559, at *1–2 (E.D. La. Oct. 18, 2004) (denying motion to exclude government website printout where date and domain were shown).</p> <p><i>Telewizja Polska USA, Inc. v. EchoStar Satellite Corp.</i>, No. 02 C 3293, 2004 WL 2367740, at 6* (N.D. Ill. Oct. 15, 2004) (finding that Way-back Machine copies of website, verified by affidavit, met Rule 901’s threshold requirements).</p> <p><i>People v. Beckley</i>, 110 Cal. Rptr. 3d 362, 366–67 (Cal. Ct. App. 2010) (holding that prosecution failed to authenticate photograph downloaded from an internet website where “no expert testified that the picture was not a ‘composite’ or ‘faked’ photograph,” and noting that “digital photographs can be changed to produce false images”).</p> <p><i>United States v. Hassan</i>, 742 F.3d 104, 133 (4th Cir. 2014) (holding that Facebook posts, including YouTube videos, were self-authenticating under Rule 902(11) where accompanied by certificates from Facebook and Google custodians “verifying that the Facebook pages and YouTube videos had been maintained as business</p>	

258 Gregory P. Joseph, *Internet and Email Evidence (Part 1)*, THE PRACTICAL LAWYER 19, 21 (Feb. 2012); see also Hon. Alan Pendleton, *Admissibility of Electronic Evidence: A New Evidentiary Frontier*, BENCH & B. MINN. 14, 15 (Oct. 2014).

Type of e-Evidence: Internet Websites/Web pages	
<p>records in the course of regularly conducted business activities. According to those certifications, Facebook and Google create and retain such pages and videos when (or soon after) their users post them through use of the Facebook or Google servers.”).</p> <p><i>United States v. Jackson</i>, 208 F.3d 633, 638 (7th Cir. 2000) (holding website postings were not properly authenticated because the proponent needed to show that the website postings were actually posted by a particular group and not the proponent herself).</p>	
FRE Rule	Method
<p>Rule 901(b)(3): Comparison by trier or expert witness.</p>	<p>As above for 901(b)(3).</p> <p>Archived internet content could be obtained through the Internet Archive’s Wayback Machine.</p>
Cases	
<p><i>St. Luke’s Cataract & Laser Inst., P.A. v. Sanderson</i>, No. 8:06-cv-223-T-MSS, 2006 WL 1320242, at *2 (M.D. Fla. May 12, 2006).</p> <p><i>United States v. Gasperini</i>, No. 17-2479-cr, 2018 WL 3213005, at *5 (2d Cir. 2018).</p> <p><i>United States v. Bansal</i>, 663 F.3d 634, 667–68 (3rd Cir. 2011).</p> <p><i>Novak v. Tucows, Inc.</i>, No. 06-CV-1909(JFB)(ARL), 2007 WL 922306, at *5 (E.D.N.Y. Mar. 26, 2007) (holding that information about Wayback Machine was not properly authenticated pursuant to Fed. R. Evid. 901 because the plaintiff proffered neither testimony nor sworn statements attesting to the authenticity of the contested web-page exhibits by an employee of the companies hosting the sites from which the plaintiff printed the pages).</p>	
FRE Rule	Method
<p>Rule 901(b)(4): Distinctive characteristics and the like.</p>	<p>As above for 901(b)(4).</p>

Type of e-Evidence: Internet Websites/Web pages	
Cases	
<i>Premier Nutrition, Inc. v. Organic Food Bar, Inc.</i> , No. SACV 06-0827 AG (RNBx), 2008 WL 1913163, at *6 (C.D. Cal. Mar. 27, 2008) (noting that “[c]ourts consider the distinctive characteristics of a website in making a finding of authenticity,” i.e., printouts of web pages with web addresses and dates).	
FRE Rule	Method
Rule 901(b)(7): Public records or reports.	Proof of custody needed; proof of reliability of system not needed.
Cases	
<i>Williams v. Long</i> , 585 F. Supp. 2d 679, 686–88, & n.4 (D. Md. 2008) (collecting cases indicating that postings on government websites are self-authenticating).	
FRE Rule	Method
Rule 901(b)(9): Process or system.	Proof that the process or system is trustworthy.

Type of e-Evidence: Chat Room, Blogs, and Other Social Media	
FRE Rules	Methods
Rule 901(b)(1): Testimony of a witness with knowledge.	As above for 901(b)(1). As above for 901(b)(4).
Rule 901(b)(3): Comparison by trier or expert witness.	Showing that a posting appears on a particular user’s webpage is insufficient to authenticate as written by account holder.
Rule 901(b)(4): Distinctive characteristics and the like.	
Rule 902(b)(9): System or process.	Evidence: <ul style="list-style-type: none"> • testimony from a witness who identifies the social

**Type of e-Evidence: Chat Room, Blogs,
and Other Social Media**

Rule 902(5), (6): Official publications, newspapers etc.

Rule 902(13): Certified records generated by an electronic process.

Rule 902(14): Certified data copied from an electronic device, storage medium.

media account as that of the alleged author, on the basis that the witness on other occasions communicated with the account holder,

- testimony from a participant in the conversation based on firsthand knowledge that the transcript fairly and accurately captures the conversation,
- evidence from the hard drive of the purported author's computer reflecting that a user of the computer used the screen name in question, or
- evidence that the chat appears on the computer or other device of the account owner and purported author.

Social media as business records:

- time stamps, metadata, etc. maintained by the owner,
- testimony from the purported creator of the social network profile and related postings,

**Type of e-Evidence: Chat Room, Blogs,
and Other Social Media**

- testimony from persons who saw the purported creator establish or post to the page, or
- references or links to, or contact information about, loved ones, relatives, co-workers, others close to the purported author.

Cases

Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 538–39 (D. Md. 2007).

Griffin v. State, 19 A.3d 415, 427–28 (Md. 2011) (citing three methods of authentication)

- i. “[A]sk the purported creator if she indeed created the profile and also if she added the posting in question.”
- ii. Search the computer of the alleged person and “examine the computer’s internet history and hard drive to determine whether that computer was used to originate the social networking profile and posting in question.”
- iii. “[O]btain information directly from the social networking website that links the establishment of the profile to the person who allegedly created it and also links the posting sought to be introduced to the person who initiated it.”

State v. Eleck, 23 A.3d 818, 821–25 (Conn. App. Ct. 2011) (affirming exclusion of printouts of Facebook messages for lack of authentication where defendant did not provide enough circumstantial evidence to prove who sent the Facebook messages).

United States v. Browne, 834 F.3d 403, 410–14 (3d Cir. 2016), cert. denied, 137 S. Ct. 695 (2017) (holding that Facebook chats were sufficiently authenticated because witnesses testified they

**Type of e-Evidence: Chat Room, Blogs,
and Other Social Media**

communicated with the creator of the page through Facebook, they could identify the alleged creator of the page in court, and the available biographical data on Facebook matched the defendant).

United States v. Encarnacion-LaFontaine, 639 F. App'x 710, 713 (2d Cir. 2016) (finding that threatening Facebook posts were properly authenticated where "the Government introduced evidence that (1) the Facebook accounts used to send the messages were accessed from IP addresses connected to computers near Encarnacion's apartment; (2) patterns of access to the accounts show that they were controlled by the same person; (3) in addition to the Goris threats, the accounts were used to send messages to other individuals connected to Encarnacion; (4) Encarnacion had a motive to make the threats[;] and (5) a limited number of people, including Encarnacion, had information that was contained in the messages.").

Type of e-Evidence: Computerized Records or Data

FRE Rule	Methods
Rule 901(b)(1): Testimony of a witness with knowledge.	As above for 901(b)(1).
FRE Rule	Methods
Rule 901(b)(4): Distinctive characteristics and the like.	As above for 901(b)(4).
FRE Rule	Method
Rule 901(b)(9): Process or system.	As above for 901(b)(9).
Cases	
<i>Liser v. Smith</i> , 254 F. Supp. 2d 89, 94, 97–98 (D.D.C. 2003) (discussing significance of time lag between actual time and time indicated on surveillance tape in deciding summary judgment in false arrest case).	

Type of e-Evidence: Computerized Records or Data	
<p><i>State v. Chun</i>, 943 A.2d 114, 120–21 (N.J. 2008) (concluding, after extensive testing for scientific validity, that new breathalyzer, Alcotest using New Jersey Firmware version 3.11, is “generally scientifically reliable” but ordering modifications to enable it to be admitted into evidence because results of third test indicated inherent errors).</p>	
FRE Rule	Method
<p>Rule 902(13): Certification of records.</p>	<p>Affidavit by deponent:</p> <ol style="list-style-type: none"> i. with specialized or technical knowledge on how the system or process works. ESI was obtained from systems that produced reliable results. ii. detailed description of what was done. <p>Notice under Rule 902(11).</p>
Cases	
<p><i>Lamb v. State</i>, 246 So. 3d 400, 408–09 (Fla. Dist. Ct. App. 2018) (upholding trial court’s ruling that the Facebook Live video was properly authenticated and admissible to the jury). Authentication problem in the manner in which the prosecutor attempted to authenticate the Facebook Live video. FRE 902 (13) and (14) all provide parameters in which practitioners can easily present electronically stored information (ESI) as self-authenticating.</p>	
FRE Rule	Method
<p>Rule 902(14): Certification of data copied or stored (e.g., metadata).</p>	<p>As above for 902(14).</p> <p>By comparing the “hash value” of the proffered copy to that of the original document.</p> <p>Notice under Rule 902(11).</p>

Type of e-Evidence: Audios and Videos	
FRE Rule	Methods
Rule 901(b)(1): Testimony of a witness with knowledge	Testifies to personal observation of events Forensic expert
Cases	
<i>United States v. Broomfield</i> , 591 F. App'x 847, 848–49, 851–52 (11th Cir. 2014) (upholding trial court's admission, in possession of fire-arm case, of YouTube video showing defendant discharging an AR-15 rifle in front of Fowler Firearms where Fowler manager testified that: (i) defendant was a Fowler Firearms member; (ii) defendant purchased two boxes of PMC .223 ammunition at the time in question; (iii) he had not purchased the ammunition at any other time; and (iv) the only firearm Fowler rented that used PMC .223 ammunition was the AR-15).	
FRE Rule	Method
Rule 901(b)(9): Process or system. Digitally altered audios and videos.	Proof that the process or system is trustworthy. Integrity of the recording speaks for itself: <ul style="list-style-type: none"> • fidelity of equipment; • absence of modifications; • handling and storing procedure; • establishing the authenticity and correctness of the resulting recording; • time and date; • operating, testing and security procedures, chain of custody.

Type of e-Evidence: Audios and Videos	
	<ul style="list-style-type: none"> • Metadata should include time, date, geolocation, and device IDs of other devices in close proximity.
Cases	
<p><i>U.S. v. Chapman</i>, 804 F.3d 895, 902 (7th Cir. 2015).</p> <p><i>People v. Jackson</i>, 994 N.Y.S.2d 438, 440–41 (N.Y. App. Div. 2014).</p> <p>Julia Day, <i>Reuters Drops Photographer over ‘Doctored’ Image</i>, THE GUARDIAN (Aug. 7, 2006 7:05 AM), https://www.theguardian.com/media/2006/aug/07/reuters.pressandpublishing.</p> <p><i>Tillerson in Afghanistan: Photo of meeting apparently doctored</i>, BBC NEWS (Oct. 24, 2017), https://www.bbc.com/news/world-asia-41734559 (clock cropped out to conceal the true location of the meeting).</p>	
FRE Rule	Method
Rule 902(13): Certification of records.	As above for 902(13).

Type of e-Evidence: Computer Simulations and Computer Animations	
FRE Rule	Methods
Rule 901(b)(1): Testimony of a witness with knowledge.	As above for 901(b)(1).
FRE Rule	Methods
Rule 901(b)(3): Comparison by trier or expert witness.	As above for 901(b)(3).
FRE Rule	Method
Rule 901(b)(9): Process or system.	As above for 901(b)(9).

Type of e-Evidence: Digital Photographs	
FRE Rules	Methods
<p>Rule 901(b)(9): Process or system.</p> <p>Rule 902(13): Certification of records.</p>	<p>As above for 901(b)(9) and 902(13)</p> <p>Certification by a technician, metadata, GPS co-ordinates, camera log</p>
Cases	
<p><i>Rodd v. Raritan Radiologic Assocs., P.A.</i>, 860 A.2d 1003 (N.J. Super. Ct. App. Div. 2004) (computerized images of mammograms).</p>	

Type of e-Evidence: Cloud	
FRE Rules	Methods
<p>Rule 901(b)(1): Testimony of a witness with knowledge.</p> <p>Rule 901 (b)(9): Process or system.</p>	<p>Witness to testify on contractual service level agreements with cloud service providers that specify:</p> <ol style="list-style-type: none"> i. data ownership, ii. confidentiality and non-disclosure requirements, iii. notification about third-party requests for access, iv. trusted third-party security audit or verification procedures, and v. intrinsic data protective controls directly given by the data holder before uploading to the cloud.

Type of e-Evidence: Cloud	
	<p>Authenticate:</p> <ol style="list-style-type: none"> i. proof of its origin by identifying its creator or authorized signatory; ii. content integrity, i.e., that the document has not been altered since its creation; iii. time of its creation and attestation, including proof of the implementation of effective safeguards by a reliable or trustworthy source to ensure its integrity; and iv. recordkeeping system, allocation of operational control and responsibility, and access control. <p>Forensics can detect traces of the use of a cloud computing service stored in PCs and smartphones. (For example, Dropbox can be found in the Windows system. These traces can be located in the installation directory, registry changes on installation, network activity, database files, internet log files, and uninstallation data.)²⁵⁹</p>
Cases	
<p><i>Rearden LLC v. Rearden Commerce, Inc.</i>, 597 F. Supp. 2d 1006 (N.D. Cal. Jan. 27, 2009, <i>vacated</i>, 683 F.3d 1190 (9th Cir. 2012) (granting</p>	

Type of e-Evidence: Cloud

summary judgment (later vacated and remanded) involving claims of trademark infringement of personal-assistant device between parties involved in cloud computing).

International Business Machines Corp. v. Johnson, No. 09 Civ. 4826(SCR), 2009 WL 2356430 (S.D.N.Y. July 30, 2009) (noting, in noncompetition agreement case, requirement that vice president of corporate development advise on “enterprise services, servers, storage, so-called ‘Cloud’ computing and business analytics”).

State v. Bellar, 217 P.3d 1094, 1110–11 & n.10–11 (Or. Ct. App. 2009) (discussing defendant’s privacy rights relating to data stored in the cloud).

Type of e-Evidence: USB Device and Other Removable Storage Devices

FRE Rule	Methods
Rule 902(13): Certification from a forensic technician.	As above for 902(13).

Type of e-Evidence: IoT

FRE Rule	Methods
Rule 901(b)(1): Testimony of a witness with knowledge. Rule 901(b)(9): Process or system.	Expert witness: forensic analysis <ol style="list-style-type: none"> i. explain scope and nature of data collection and analysis; ii. security features; iii. devices: function, process, system; and

259. See Frank McClain, *Dropbox Forensics*, FORENSIC FOCUS (May 31, 2011), <https://www.forensicfocus.com/articles/dropbox-forensics/>.

Type of e-Evidence: IoT	
	iv. data stored in the cloud, as for cloud above.

APPENDIX B: COMMITTEE NOTE ON RULE 807²⁶⁰

Rule 807 has been amended to fix a number of problems that the courts have encountered in applying it.

Courts have had difficulty with the requirement that the proffered hearsay carry “equivalent” circumstantial guarantees of trustworthiness. The “equivalence” standard is difficult to apply, given the different types of guarantees of reliability, of varying strength, found among the categorical exceptions (as well as the fact that some hearsay exceptions, e.g., Rule 804(b)(6), are not based on reliability at all). The “equivalence” standard has not served to guide a court’s discretion to admit hearsay, because the court is free to choose among a spectrum of exceptions for comparison. Moreover, experience has shown that some statements offered as residual hearsay cannot be compared usefully to any of the categorical exceptions and yet might well be trustworthy. Thus the requirement of an equivalence analysis has been eliminated. Under the amendment, the court should proceed directly to a determination of whether the hearsay is supported by guarantees of trustworthiness. See Rule 104(a). As with any hearsay statement offered under an exception, the court’s threshold finding that admissibility requirements are met merely means that the jury may consider the statement and not that it must assume the statement to be true.

The amendment specifically requires the court to consider corroborating evidence in the trustworthiness enquiry. Most courts have required the consideration of corroborating evidence, though some courts have disagreed. The rule now provides for a uniform approach, and recognizes that the existence or absence of corroboration is relevant to, but not dispositive of,

260. COMM. ON RULES OF PRACTICE AND PROCEDURE, JUDICIAL CONFERENCE OF THE U.S., AGENDA BOOK 410–14 (June 12, 2018), http://www.uscourts.gov/sites/default/files/2018-06_standing_agenda_book_final.pdf.

whether a statement should be admissible under this exception. Of course, the court must consider not only the existence of corroborating evidence but also the strength and quality of that evidence.

The amendment does not alter the case law prohibiting parties from proceeding directly to the residual exception, without considering admissibility of the hearsay under Rules 803 and 804. A court is not required to make a finding that no other hearsay exception is applicable. But the opponent cannot seek admission under Rule 807 if it is apparent that the hearsay could be admitted under another exception.

The rule in its current form applies to hearsay “not specifically covered” by a Rule 803 or 804 exception. The amendment makes the rule applicable to hearsay “not admissible under” those exceptions. This clarifies that a court assessing guarantees of trustworthiness may consider whether the statement is a “near-miss” of one of the Rule 803 or 804 exceptions. If the court employs a “near-miss” analysis it should—in addition to evaluating all relevant guarantees of trustworthiness—take into account the reasons that the hearsay misses the admissibility requirements of the standard exception.

In deciding whether the statement is supported by sufficient guarantees of trustworthiness, the court should not consider the credibility of any witness who relates the declarant’s hearsay statement in court. The credibility of an in-court witness does not present a hearsay question. To base admission or exclusion of a hearsay statement on the witness’s credibility would usurp the jury’s role of determining the credibility of testifying witnesses. The rule provides that the focus for trustworthiness is on circumstantial guarantees surrounding the making of the statement itself, as well as any independent evidence corroborating the statement. The credibility of the witness relating the statement is not a part of either enquiry.

Of course, even if the court finds sufficient guarantees of trustworthiness, the independent requirements of the Confrontation Clause must be satisfied if the hearsay statement is offered against a defendant in a criminal case.

The Committee decided to retain the requirement that the proponent must show that the hearsay statement is more probative than any other evidence that the proponent can reasonably obtain. This necessity requirement will continue to serve to prevent the residual exception from being used as a device to erode the categorical exceptions.

The requirements that residual hearsay must be evidence of a material fact and that its admission will best serve the purposes of these rules and the interests of justice have been deleted. These requirements have proved to be superfluous in that they are already found in other rules. See Rules 102, 401.

The notice provision has been amended to make four changes in the operation of the rule:

- First, the amendment requires the proponent to disclose the “substance” of the statement. This term is intended to require a description that is sufficiently specific under the circumstances to allow the opponent a fair opportunity to meet the evidence. See Rule 103(a)(2) (requiring the party making an offer of proof to inform the court of the “substance” of the evidence).
- Second, the prior requirement that the declarant’s address must be disclosed has been deleted. That requirement was nonsensical when the declarant was unavailable, and unnecessary in the many cases in which the declarant’s address was known or easily obtainable. If prior disclosure of the declarant’s address is critical and cannot be obtained by the opponent

through other means, then the opponent can seek relief from the court.

- Third, the amendment requires that the pretrial notice be in writing—which is satisfied by notice in electronic form. See Rule 101(b)(6). Requiring the notice to be in writing provides certainty and reduces arguments about whether notice was actually provided.
- Finally, the pretrial notice provision has been amended to provide for a good cause exception. Most courts have applied a good cause exception under Rule 807 even though the rule in its current form does not provide for it, while some courts have read the rule as it was written. Experience under the residual exception has shown that a good cause exception is necessary in certain limited situations. For example, the proponent may not become aware of the existence of the hearsay statement until after the trial begins; or the proponent may plan to call a witness who without warning becomes unavailable during trial, and the proponent might then need to resort to residual hearsay.

The rule retains the requirement that the opponent receive notice in a way that provides a fair opportunity to meet the evidence. When notice is provided during trial after a finding of good cause, the court may need to consider protective measures, such as a continuance, to assure that the opponent is not prejudiced.

APPENDIX C: 12 V.S.A. § 1913. BLOCKCHAIN ENABLING

(a) As used in this section:

(1) “blockchain” means a cryptographically secured, chronological, and decentralized consensus ledger or consensus database maintained via Internet, peer-to-peer network, or other interaction.

(2) “Blockchain technology” means computer software or hardware or collections of computer software or hardware, or both, that utilize or enable a blockchain.

(b) Authentication, admissibility, and presumptions.

(1) A digital record electronically registered in a blockchain shall be self-authenticating pursuant to Vermont Rule of Evidence 902, if it is accompanied by a written declaration of a qualified person, made under oath, stating the qualification of the person to make the certification and:

(A) the date and time the record entered the blockchain;

(B) the date and time the record was received from the blockchain;

(C) that the record was maintained in the blockchain as a regular conducted activity; and

(D) that the record was made by the regularly conducted activity as a regular practice.

(2) A digital record electronically registered in a blockchain, if accompanied by a declaration that meets the requirements of subdivision (1) of this subsection, shall be considered a record of regularly conducted business activity pursuant to

Vermont Rule of Evidence 803(6) unless the source of information or the method or circumstance of preparation indicate lack of trustworthiness. For purposes of this subdivision (2), a record includes information or data.

(3) The following presumptions apply:

(A) A fact or record verified through a valid application of blockchain technology is authentic.

(B) The date and time of the recordation of the fact or record established through such a blockchain is the date and time that the fact or record was added to the blockchain.

(C) The person established through such a blockchain as the person who made such recordation is the person who made the recordation.

(D) If the parties before a court or other tribunal have agreed to a particular format or means of verification of a blockchain record, a certified presentation of a blockchain record consistent with this section to the court or other tribunal in the particular format or means agreed to by the parties demonstrates the contents of the record.

(4) A presumption does not extend to the truthfulness, validity, or legal status of the contents of the fact or record.

(5) A person against whom the fact operates has the burden of producing evidence sufficient to support a finding that the presumed fact, record,

time, or identity is not authentic as set forth on the date added to the blockchain, but the presumption does not shift to a person the burden of persuading the trier of fact that the underlying fact or record is itself accurate in what it purports to represent.

(c) Without limitation, the presumption established in this section shall apply to a fact or record maintained by blockchain technology to determine:

(1) contractual parties, provisions, execution, effective dates, and status;

(2) the ownership, assignment, negotiation, and transfer of money, property, contracts, instruments, and other legal rights and duties;

(3) identity, participation, and status in the formation, management, record keeping, and governance of any person;

(4) identity, participation, and status for interactions in private transactions and with a government or governmental subdivision, agency, or instrumentality;

(5) the authenticity or integrity of a record, whether publicly or privately relevant; and

(6) the authenticity or integrity of records of communication.

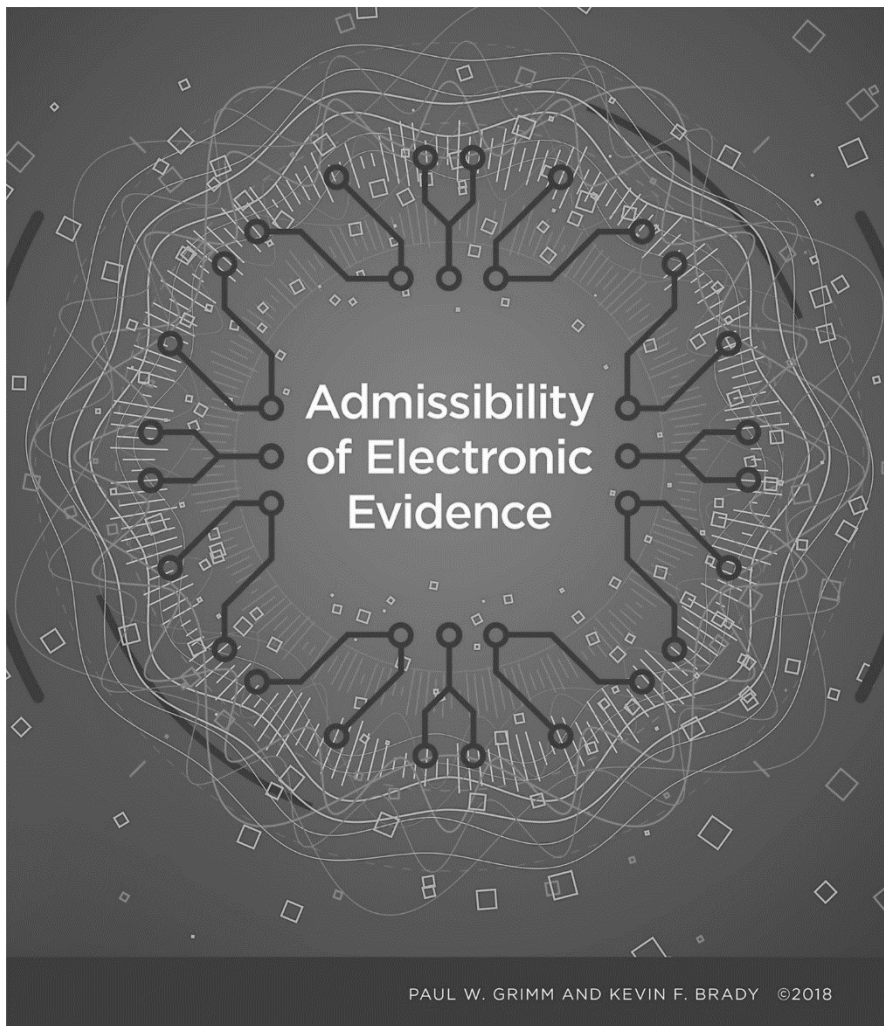
(d) The provisions of this section shall not create or negate:

(1) an obligation or duty for any person to adopt or otherwise implement blockchain technology for any purpose authorized in this section; or

(2) the legality or authorization for any particular underlying activity whose practices or data are verified through the application of blockchain technology. (Added 2015, No. 157 (Adj. Sess.), § I.1.)²⁶¹

261. *Id.*

**APPENDIX D: CHECKLIST OF POTENTIAL
AUTHENTICATION METHODS²⁶²**



262. Full-size PDF available at https://thesedonaconference.org/sites/default/files/Grimm_Brady_Evidence_Admissibility_Chart_2018.pdf.

Potential Authentication Methods



Email, Text Messages, and Instant Messages

- Witness with personal knowledge (901(b)(1))
- Expert testimony or comparison with authenticated examples (901(b)(3))
- Distinctive characteristics including circumstantial evidence (901(b)(4))
- System or process capable of proving a reliable and dependable result (901(b)(9))
- Trade inscriptions (902(7))
- Certified copies of business record (902(11))
- Certified records generated by an electronic process or system (902(13))
- Certified data copied from an electronic device, storage medium, or file (902(14))



Chat Room Postings, Blogs, Wikis, and Other Social Media Conversations

- Witness with personal knowledge (901(b)(1))
- Expert testimony or comparison with authenticated examples (901(b)(3))
- Distinctive characteristics including circumstantial evidence (901(b)(4))
- System or process capable of proving a reliable and dependable result (901(b)(9))
- Official publications (902(5))
- Newspapers and periodicals (902(6))
- Certified records generated by an electronic process or system (902(13))
- Certified data copied from an electronic device, storage medium, or file (902(14))



Digitally Stored Data and Internet of Things

- Witness with personal knowledge (901(b)(1))
- Expert testimony or comparison with authenticated examples (901(b)(3))
- Distinctive characteristics including circumstantial evidence (901(b)(4))
- System or process capable of proving a reliable and dependable result (901(b)(9))
- Certified records generated by an electronic process or system (902(13))
- Certified data copied from an electronic device, storage medium, or file (902(14))



Computer Processes, Animations, Virtual Reality, and Simulations

- Witness with personal knowledge (901(b)(1))
- Expert testimony or comparison with authenticated examples (901(b)(3))
- System or process capable of proving a reliable and dependable result (901(b)(9))
- Certified records generated by an electronic process or system (902(13))



Digital Photographs

- Witness with personal knowledge (901(b)(1))
- System or process capable of providing reliable and dependable result (901(b)(9))
- Official publications (902(5))
- Certified records generated by an electronic process or system (902(13))
- Certified data copied from an electronic device, storage medium, or file (902(14))



Social Media Sites (Facebook, LinkedIn, Twitter, Instagram, and Snapchat)

- Witness with personal knowledge (901(b)(1))
- Expert testimony or comparison with authenticated examples (901(b)(3))
- Distinctive characteristics including circumstantial evidence (901(b)(4))
- Public records (901(b)(7))
- System or process capable of proving a reliable and dependable result (901(b)(9))
- Official publications (902(5))
- Certified records generated by an electronic process or system (902(13))
- Certified data copied from an electronic device, storage medium, or file (902(14))

Know Which Approach Your Jurisdiction Follows

Maryland Approach to Rules 104 and 901:

A higher standard for authentication for social media evidence. In this approach, the burden is on the admitting party to show that the social media evidence was not falsified or created by another user through either:

- Testimony of the creator of the website page or the post
- Search of the internet history or hard drive of the purported creator's computer
- Information obtained directly from social media site

See, *Griffin v. State*, 19 A. 3d 415, 423 (Md. 2011).

Texas Approach to Rules 104 and 901:

A lower standard for authentication of social media evidence. In this approach, the burden is on the admitting party to show evidence sufficient to support a finding by a reasonable juror that the social media evidence is what its proponent claims it to be through either:

- Direct testimony of a witness with personal knowledge
- Expert testimony or comparison with authenticated evidence
- Circumstantial evidence

See, *Tienda v. State*, 358 S. W. 3d 633 (Tex. Crim. App. 2012).

Is Evidence Hearsay?

FRE 801 (a-c)

- Is it a statement? (written/ spoken assertion, non-verbal/ non-assertive verbal conduct intended to be assertive.)
- Is statement made by "Declarant?" (person, not generated by machine)
- Is statement offered for proving truth of assertion?
 - NOTE: Statement is not offered for substantive truth if offered to prove:
 - Communicative/ comprehension capacity of declarant
 - Effect on the hearer
 - Circumstantial evidence of state of mind of declarant
 - Verbal acts/parts of acts
 - Utterances of independent legal significance

Is statement excluded from definition of hearsay by 801(d)(1) and (2)?

Prior witness statements—801(d)(1)

- Prior testimonial statement 801(d)(1)(A)
- Prior consistent statement 801(d)(1)(B) to rebut allegations of recent fabrication or rehabilitate a witness that has been impeached
- Statement of identification 801(d)(1)(C)

Admission by party opponents—801(d)(2)*

- Individual admission 801(d)(2)(A)
- Adoptive admission 801(d)(2)(B)
- Admission by person with authority 802(d)(2)(C)
- Admission by agent/ employees 802(d)(2)(D)
- Co-conspirator statements 801(d)(2)(E)

** Documents produced in discovery by opposing party are presumed to be authentic under 801(d)(2). Certification of business records under 801(1) and (1E) must meet requirements of 803(6).*

If **HEARSAY**, then it is **INADMISSIBLE** unless covered by a recognized exception.

Hearsay Exception

Availability of Declarant Irrelevant—803

- Present sense impression 803(1)
- Excited utterance 803(2)
- State of mind exception 803(3)
- Statements for purposes of medical diagnosis or treatment 803(4)
- Past recollection recorded 803(5)
- Business records 803(6)
- Absence of an entry in records kept in the regular course of business 803(7)
- Public records or reports 803(8)
- Records of vital statistics 803(9)
- Absence of public record or entry 803(10)
- Records/ documents affecting interest in property 803(14) & (15)
- Statements in ancient documents 803(16)
- Market reports and commercial publications 803(17)
- Learned treatises 803(18)
- Character reputation testimony 803(21)
- Record of felony convictions 803(22)

Declarant Unavailable—804

- Unavailability – 804(a)(1-5) (privilege, refused to testify, lack of memory, death/illness, beyond subpoena power)
- Unavailability Exceptions—804(b):
 - Former Testimony 804(b)(1)
 - Dying Declaration 804(b)(2)
 - Statement Against Interest 804(b)(3)
 - Statement of personal or family history 804(b)(4)
 - Forfeiture by wrongdoing 804(b)(6)
- Residual "Catchall" Exception—807

A hearsay statement is not excluded by Rule 802 even if the statement is not specifically covered by Rule 803 or 804 under the following circumstances:

- Statement has equivalent circumstantial guarantees of trustworthiness
- Offered as evidence of a material fact
- More probative on the point for which it is offered than any other evidence that the proponent can obtain through reasonable efforts
- Admitting it will best serve the purposes of these rules and the interest of justice

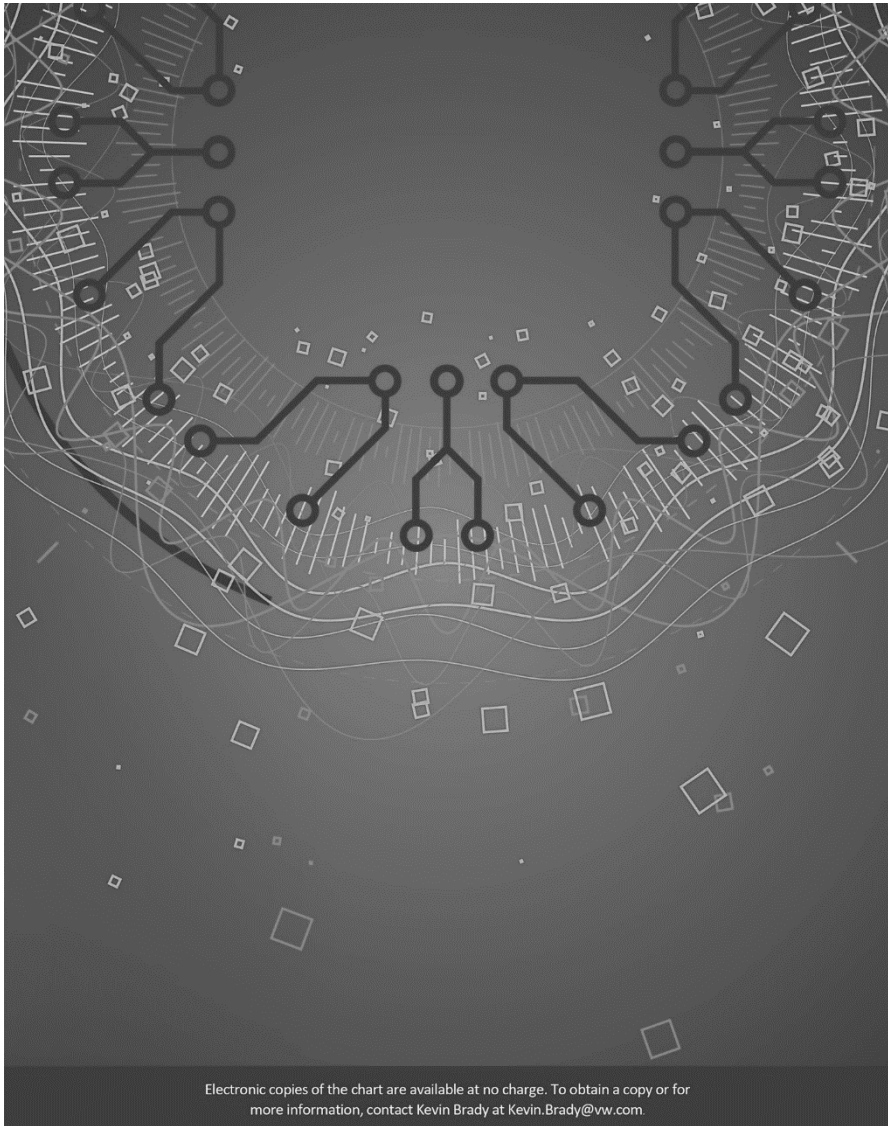
The statement is admissible only if, before the trial or hearing, the proponent gives reasonable notice of intent to offer the statement and its particulars, and the opposing party has a fair opportunity to meet it.

Original Writing Rule FRE 1001-1008

- Is the evidence "original," "duplicative," "writing," or "recording" (Rule 1001)
- Rule 1002 requires the original to prove the contents of a writing, recording, or photograph unless "secondary evidence" (any evidence other than original or duplicative) is admissible. (Rules 1004, 1005, 1006, and 1007)
- Duplicates are co-extensively admissible as originals, unless there is a genuine issue of authenticity of the original or circumstances indicate that it would be unfair to admit duplicate in lieu of original (Rule 1003)
- Permits proof of the contents of writing, recording or paragraph by use of "secondary evidence"—any proof of the contents of a writing, recording or photograph other than the original or duplicate (Rule 1004) if
 - Non-bad faith loss/destruction of original/duplicate
 - Inability to subpoena original/duplicate
 - Original/duplicate in possession, custody, or control of opposing party
 - "Collateral record" (i.e., not closely related to controlling issue in the case)
- Admission of summary of voluminous books, records, or documents (Rule 1006)
- Testimony or deposition of party against whom offered or by that party's written admission (FRCP 30, 33, 36) (Rule 1007)
- If admissibility depends on the fulfillment of a condition or fact, question of whether condition has been fulfilled is for fact finder to determine under Rule 104(b) (Rule 1008)
- But, the issue is for the trier of fact, if it is a question:
 - Whether the asserted writing ever existed
 - Whether another writing, recording, or photograph produced at trial is the original, or reflects the contents, the issue is for the trier of fact

Practice Tips

- Be prepared and start with a defensible and comprehensive records management program
 - Think strategically about the case and the evidence from the beginning of the case
 - Memorialize each step of the collection and production process to bolster reliability
 - Use every opportunity during discovery to authenticate potential evidence
- Examples:**
- For pretrial disclosures under FRCP 26(a)(3), you have 14 days to file objections or possible waiver
 - Document produced by opposing party are presumed to be authentic under Rule 801(d)(2) – burden shifts
 - FRCP 36 Requests for Admissions
 - Request stipulation of authenticity from opposing counsel
- Be prepared to provide the court with enough information to understand the technology issues as they relate to the reliability of the evidence at hand
 - Be creative and consider whether there are case management tools that might assist the court and the other parties in addressing evidentiary problems concerning some of the more complex issues (such as "dynamic" data in a database or what is a "true and accurate copy" of ESI)
 - Keep your audience in mind. Will this be an issue for the judge or the jury? (e.g. Rule 104(a) or (b))



Electronic copies of the chart are available at no charge. To obtain a copy or for more information, contact Kevin Brady at Kevin.Brady@vw.com.

THE SEDONA CONFERENCE COMMENTARY ON THE
PROPER IDENTIFICATION OF ASSERTED TRADE SECRETS IN
MISAPPROPRIATION CASES

*A Project of The Sedona Conference Working Group on Trade
Secrets (WG12)*

Author:

The Sedona Conference

Editors-in-Chief:

James Pooley

Victoria Cundiff

Managing Editor:

Jim W. Ko

Senior Editors:

David Almeling

Charles Tait Graves

Contributing Editors:

Demarron Berkley

Thomas A. Brown

Steven M. Kayman

Mark Klapow

Sid Leach

Patrick J. O'Toole, Jr.

Dean A. Pelletier

Michael Risch

WG12 Judicial Advisors:

Hon. Denise Cote

Hon. Gail J. Standish

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference's Working Group 12. They do not necessarily

Copyright 2020, The Sedona Conference.
All Rights Reserved.

represent the views of any of the individual participants or their employers, clients, or any organizations to which they may belong, nor do they necessarily represent official positions of The Sedona Conference.

We thank all of our Working Group Series Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors, click on the “Sponsors” navigation bar on the homepage of our website.

This publication may be cited as follows:

The Sedona Conference, *Commentary on the Proper Identification of Asserted Trade Secrets in Misappropriation Cases*, 22 SEDONA CONF. J. 223 (2021).

PREFACE

Welcome to the final October 2020 version of *The Sedona Conference Commentary on the Proper Identification of Asserted Trade Secrets in Misappropriation Cases*, a project of The Sedona Conference Working Group on Trade Secret Law (WG12). This is one of a series of Working Group commentaries published by The Sedona Conference, a 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of anti-trust law, complex litigation, intellectual property rights, and data security and privacy law. The mission of The Sedona Conference is to move the law forward in a reasoned and just way.

The mission of WG12, formed in February 2018, is “to develop consensus and nonpartisan principles for managing trade secret litigation and well-vetted guidelines for consideration in protecting trade secrets, recognizing that every organization has and uses trade secrets, that trade secret disputes frequently intersect with other important public policies such as employee mobility and international trade, and that trade secret disputes are litigated in both state and federal courts.” The Working Group consists of members representing all stakeholders in trade secret law and litigation.

The WG12 *Commentary* drafting team was launched in May 2018. Earlier drafts of this publication were a focus of dialogue at the WG12 Annual Meeting in Charlotte, North Carolina, in November 2019, the WG12 Inaugural Meeting in Los Angeles, California, in November 2018, and the Inaugural Sedona Conference on Developing Best Practices for Trade Secret Issues in Scottsdale, Arizona, in December 2017. The *Commentary* was published for public comment in April 2020. The editors have carefully considered the comments received through the Working Group Series review and comment process and, where appropriate, incorporated them into this final version.

This *Commentary* represents the collective efforts of many individual contributors. On behalf of The Sedona Conference, I thank

in particular James Pooley, the Chair of WG12, and Victoria Cundiff, the Vice-Chair of WG12, who serve as the Editors-in-Chief of this publication, and David Almeling and Charles Tait Graves, both WG12 Steering Members, who serve as the Senior Editors of this publication. I also thank everyone else involved for their time and attention during this extensive drafting and editing process, including our Contributing Editors Demarron Berkley, Thomas A. Brown, Steven M. Kayman, Mark Klapow, Sid Leach, Patrick J. O'Toole, Jr., Dean A. Pelletier, and Michael Risch.

The Working Group had the benefit of candid comments by the Judicial Advisors designated to this Commentary drafting team effort—Hon. Denise Cote and Hon. Gail Standish. The statements in this *Commentary* are solely those of the nonjudicial members of the Working Group; they do not represent any judicial endorsement of any recommended practices.

The drafting process for this *Commentary* has also been supported by the Working Group 12 Steering Committee and Judicial Advisors.

We encourage your active engagement in the dialogue. Membership in The Sedona Conference Working Group Series is open to all. The Series includes WG12 and several other Working Groups in the areas of electronic document management and discovery, cross-border discovery and data protection laws, international data transfers, data security and privacy liability, patent remedies and damages, and patent litigation best practices. The Sedona Conference hopes and anticipates that the output of its Working Groups will evolve into authoritative statements of law, both as it is and as it should be.

Craig W. Weinlein
Executive Director
The Sedona Conference
October 2020

FOREWORD

A fundamental question in every case involving a claim of trade secret misappropriation is: what are the alleged trade secrets that are the subject of the claim? This question sets apart trade secret law from other major categories of intellectual property, such as patents and copyrights, in which the alleged intellectual property is defined and registered with a regulatory body before litigation begins.

The burden is on the party asserting trade secret misappropriation to answer this question by “identifying” the alleged trade secrets. While this requirement for “identification” is ubiquitous, the rules for doing so are not clear or consistent. At the federal level, neither the criminal statute (Economic Espionage Act) nor the civil statute (Defend Trade Secrets Act) explicitly addresses identification. At the state level, California and Massachusetts define certain, but not all, aspects of identification by statute, while other states set out certain rules in case law, and a few states appear yet to have addressed the issue.

The Sedona Conference’s Working Group 12 (WG12) resolved that its first commentary on trade secret law would address the identification question. This commentary represents WG12’s views about certain aspects of identification, including when an identification must be provided, what an identification must contain, and how an identification can be amended. The proposals below and across our WG12 commentary drafting team efforts are not intended to displace current law, which is diverse with respect to numerous substantive and procedural issues in trade secret law and litigation, and thus often does not lend itself to the development of more authoritative Best Practice recommendations. Rather, they are intended to summarize WG12’s consensus Principles and Guidelines regarding the identification of alleged trade secrets in trade secret litigation,

which if adopted in whole or in part would advance The Sedona Conference's mission, "moving the law forward in a reasoned and just way."

The editors would like to express their appreciation to the members of the drafting team and the judicial advisors for their valuable input and thoughtful commentary.

James Pooley

Victoria Cundiff

Editors-in-Chief and Working Group 12 Steering Committee

Chair and Vice-Chair

David S. Almeling

Charles T. Graves

Senior Editors and Working Group 12 Steering Committee

Members

TABLE OF CONTENTS

THE PROPER IDENTIFICATION OF ASSERTED TRADE SECRETS IN MISAPPROPRIATION CASES PRINCIPLES AT A GLANCE	231
THE PROPER IDENTIFICATION OF ASSERTED TRADE SECRETS IN MISAPPROPRIATION CASES GUIDELINES AT A GLANCE	232
I. INTRODUCTION.....	234
A. Why the Identification of Asserted Trade Secrets Poses Special Challenges in Trade Secret Cases	234
B. Benefits of Identifying Asserted Trade Secrets with Reasonable Particularity.....	237
C. Stages of Identifying Asserted Trade Secrets.....	239
D. Common Areas of Dispute over Trade Secret Identifications	240
E. The Structure and Purpose of this Commentary	241
II. IDENTIFICATION IS NOT AN ADJUDICATION AND IS NOT A SUBSTITUTE FOR DISCOVERY.....	243
III. TIMING OF THE IDENTIFICATION.....	245
IV. REQUIRED LEVEL OF PARTICULARITY: CASE LAW EXAMPLES	247
A. Examples of Identifications Courts Have Deemed Sufficient	249
B. Examples of Identifications Courts Have Deemed Insufficient.....	253
C. A Proposed Format for Identification	257
D. Guidelines for a Trade Secret Identification.....	260
V. AMENDMENT OF THE IDENTIFICATION.....	263

APPENDIX A: Proposed Model Local Rule For Trade Secret Misappropriation Cases	265
APPENDIX B: Examples of Case Law Regarding Identification of Asserted Trade Secrets During Litigation.....	270

**THE PROPER IDENTIFICATION OF ASSERTED TRADE SECRETS IN
MISAPPROPRIATION CASES PRINCIPLES AT A GLANCE**

- Principle No. 1:** The identification of an asserted trade secret during a lawsuit is not an adjudication of the merits and is not a substitute for discovery.
- Principle No. 2:** The party claiming misappropriation of a trade secret should identify in writing the asserted trade secret at an early stage of the case.
- Principle No. 3:** The party claiming the existence of a trade secret must identify the asserted trade secret at a level of particularity that is reasonable under the circumstances.
- Principle No. 4:** The identification of an asserted trade secret may be amended as the case proceeds.

THE PROPER IDENTIFICATION OF ASSERTED TRADE SECRETS IN MISAPPROPRIATION CASES GUIDELINES AT A GLANCE

Guideline 1: Where the trade secret plaintiff requests preliminary relief, the scope of an identification will depend on the relief sought.

Guideline 2: Where the plaintiff does not request preliminary relief, the plaintiff should identify an asserted trade secret with reasonable particularity by the outset of merits discovery.

Guideline 3: The description of an asserted trade secret in a publicly filed pleading, or other publicly filed document, may be general if necessary to avoid destroying the status of information asserted to be a trade secret.

Guideline 4: The identification of an asserted trade secret under a protective order or equivalent agreement between the parties should be made with sufficient particularity to allow the defendant to meaningfully compare an asserted trade secret to information that is generally known or readily ascertainable and to permit the parties and the court to understand what information is claimed to be a trade secret.

Guideline 5: While an asserted trade secret should be identified at a level of particularity that is reasonable under the circumstances, a defendant should not use this standard as a tool to delay litigation by demanding particularity beyond that reasonably necessary for the defendant to develop its defenses and for the court to evaluate the claims and defenses.

Guideline 6: The plaintiff should not identify an asserted trade secret exclusively by reference to a document or other item, or exclusively by cross-reference to another asserted trade secret, unless such document, other item, or cross-reference sets forth the asserted trade secret.

Guideline 7: The plaintiff should verify its identification of an asserted trade secret under oath or affirmation.

Guideline 8: If the plaintiff claims that the defendant has taken files or other materials, the court may allow motion practice and/or discovery relating to the return or inspection of such files or materials prior to requiring identification of an asserted trade secret contained within such files or materials.

I. INTRODUCTION

A. *Why the Identification of Asserted Trade Secrets Poses Special Challenges in Trade Secret Cases*

Unlike other major categories of intellectual property (patents and copyrights), trade secrets are not registered with, or examined by, any regulatory body before a litigation commences. Unless and until a court or a jury decides whether that information constitutes a trade secret, the status of an asserted trade secret is generally a matter for private contractual protection and self-help security measures.

Specific trade secrets are thus often expressly identified as such in writing for the first time during a litigation in which they are asserted. While the procedure and logistics vary widely across the country, once litigation is underway, it is common for attorneys, whether in-house or outside counsel, to work with expert witnesses and/or company employees to develop documents identifying and otherwise describing an asserted trade secret that are then served on the opposing party.

Identification of asserted trade secrets poses special challenges in litigation because the expression of what a plaintiff¹ believes to constitute its intellectual property (in the form of trade secrets) is typically only a subset of the plaintiff's technology or business information. All businesses employ a mix of secret and nonsecret information, with different types of protection (patent, copyright, trademark, and trade secret) in play. Especially in technology cases, a body of engineering work (in a product or in research) may contain trade secrets, public

1. This *Commentary* uses the term "plaintiff" throughout to refer to the trade secret claimant.

information, patented subject matter, and copyrighted expression all at the same time.

Thus, when a plaintiff alleges that a defendant has misappropriated a trade secret, it is not always evident at the outset of a lawsuit what information the plaintiff claims as a trade secret in the case. Further, it may be that some individual elements of what the plaintiff contends to be a trade secret are known to or readily ascertainable by the relevant public, while the plaintiff's particular combination of public information that the plaintiff claims has value may not be readily ascertainable. Thus, it can be challenging to distinguish information that qualifies as a proprietary secret from information that does not.

Early identification of trade secrets in a lawsuit is important for several reasons. It avoids situations where an asserted trade secret becomes a moving target, leading to significant inefficiencies for the parties and the court. It defines the contours of discovery, leading to a more streamlined process of evidence exchange. And it allows the court to focus its attention on the relevant issues early on, allowing more effective judicial oversight of the litigation as a whole.

At the same time, parties asserting a trade secret cause of action are not, and should not be, required to specifically describe an asserted trade secret in a public filing. To do so would destroy the trade secret itself. As a result, lawsuits begin without a detailed identification of the asserted trade secrets, unless the plaintiff is able to file its complaint under seal or identifies the asserted trade secrets in a separate document available to opposing counsel pursuant to a protective order or confidentiality agreement. Courts have observed, however, that pleading requirements such as those set forth in *Bell Atlantic Corp. v.*

*Twombly*² and *Ashcroft v. Iqbal*³ require something more than a statement that the defendant has allegedly misappropriated “trade secrets” without some explanation of the nature of the alleged trade secrets (technical information relating to particular topics, compilations of customer data, or the like), how they have been subject to reasonable measures to protect them, and their actual or potential economic value.⁴ Otherwise, virtually every trade secret complaint would survive a motion to dismiss.

Once a protective order is in place, however, disagreements frequently arise regarding when and how a plaintiff must deliver a written identification of an asserted trade secret to opposing counsel and what happens if that identification is insufficient. Courts began ruling on these issues regularly in the 1970s, and the number of such rulings has particularly increased since the 2000s. One reason for this is that the volume of trade secret litigation has steadily increased in state and federal courts in recent years. Nevertheless, neither the federal Defend Trade Secrets Act nor the states’ and territories’ respective versions of the Uniform Trade Secrets Act and related procedural statutes (with three exceptions, California, Puerto Rico, and Massachusetts⁵) provide a procedure for identification of an asserted trade secret during litigation.

2. 550 U.S. 544 (2007).

3. 556 U.S. 662 (2009).

4. *See, e.g., Elsevier Inc. v. Doctor Evidence, LLC*, No. 17-CV-5540 (KBF), 2018 WL 557906 (S.D.N.Y. Jan. 23, 2018).

5. *See* CAL. CIV. PROC. CODE § 2019.210 (enacted in 1985), P.R. LAWS title 10, § 4139(a) (2011), and MASS. GEN. LAWS ANN. ch. 93 § 42D(b) (2018). We do not propose, nor do we advise against, the adoption, in any jurisdiction or case, of the statutory approach employed in California and Massachusetts, in which discovery regarding the trade secret claims is effectively stayed pending determination that the plaintiff has provided an adequate—*i.e.*, “reasonable”—identification. We instead present the specific proposals

To assist parties and courts in solving identification problems, this Commentary provides a set of principles, guidelines for identifying different types of asserted trade secrets, and a reference library of case law from around the country. The proposals below are not intended to displace existing statutory law, case law, or local rules and practices. Rather, these proposals are intended to supplement such authority and practices and to provide guidance for all courts and litigants with a consensus baseline.

Early identification is but one step in the parties' litigation of the elements of a trade secret cause of action. A plaintiff, for example, might argue that the defendant took an asserted trade secret and modified it, so that the defendant's end product differs but still reflects use of the asserted trade secret. Or, the defendant might argue that the end product is the result of independent development by persons whose efforts predated the hiring of an employee from the plaintiff, and thus is not the result of access to the information the plaintiff asserts as a trade secret. Early identification of an asserted trade secret assists the parties and the court in adjudicating the substantive claims and defenses, but it does not alter the substantive law that applies to determine whether there is liability for misappropriation of those trade secrets.

B. Benefits of Identifying Asserted Trade Secrets with Reasonable Particularity

We recommend that courts adopt a local rule or proposed order governing the identification of trade secrets. This would benefit: judges who must adjudicate different kinds of trade

defined below in this document. In this section, we merely explain why having the claimed trade secrets identified with reasonable particularity at some point and in some way is advantageous for all involved in the litigation.

secret claims (and issues relating to the appropriate scope of discovery and the proper reach of any remedies); attorneys who represent trade secret owners or those accused of trade secret misappropriation (whether a business or individual employees); third parties subject to discovery in trade secret lawsuits; and academics whose interests may include consistency in the application of state and federal intellectual property laws.

A local rule or proposed order should provide clarity sufficient to reduce disputes in trade secret litigation. A clear approach to identification should also help parties understand the scope of their preservation obligations. To that end, we use the phrase most commonly adopted by courts around the country—"reasonable particularity"—while recognizing implementation of that standard can differ when, for example, the types of information in dispute differ.

Requiring the identification of asserted trade secrets makes litigation more efficient by, among other things, providing notice of what trade secrets the defendant is alleged to have misappropriated, reducing disputes about when and how to identify trade secrets, helping define the scope of discovery, and streamlining substantive motion practice, pretrial proceedings, and trial. At the same time, rules for identification should not be disproportionately burdensome and strict when compared with the efficiency and other benefits they provide. Moreover, accommodation should be provided for cases in which there is clear evidence of improper acquisition coupled with an urgent need for temporary relief, where the plaintiff may not be fully aware of the scope of the information that has been taken. In many such cases, the emphasis at the outset will understandably be on obtaining an order directing the quarantine, appropriate inspection, and, as further adjudicated, removal or remediation of more specifically identified information from the devices or computer systems of the accused party. Indeed, in

cases where a party merely seeks an order to preserve or return specific documents that have (for example) been downloaded, identification of asserted trade secrets would not be required in that narrow context.

Trial judges, whether in state or federal court, sometimes adjudicate cases where the information in dispute is highly technical or scientific. A local rule or proposed order for the identification of trade secrets should improve the understanding of judges who are not specialists in the relevant technical or scientific field, allowing them to guide the parties during discovery and assess the allegations and defenses being asserted.

The proposals advanced in this *Commentary* come after a number of state and federal courts have adopted approaches to trade secret identification. Especially where states have adopted a process via statute or published case law, WG12 does not intend to “reinvent the wheel,” and we have taken those approaches into account. At the same time, WG12 recognizes that a clear local rule or proposed order may assist courts that have not yet considered this problem, and may inform a process of reconsideration or evaluation, in individual cases or overall, for those that have.

C. Stages of Identifying Asserted Trade Secrets

Courts have examined the identification of an asserted trade secret in different ways, and at different times, during litigation. By way of example, a plaintiff may seek a temporary restraining order to preserve evidence or to require the segregation or return of downloaded files. Such cases may present different issues, at least at an early stage, from those in which a plaintiff seeks a preliminary injunction restricting ongoing or future activity by the defendant, where the degree of detail required in federal courts may be subject to the requirements of Rule 65(d).

Or, a court may address identification on a motion for summary judgment, in order to assess whether there remains a triable issue of fact. Disputes over written discovery, depositions, or even third-party subpoenas might also involve identification in some manner. Finally, courts overseeing jury trials may decide the means by which jury members will be exposed to an asserted trade secret.

Decisions made about identification early in a lawsuit may affect other aspects of the litigation over time—such as the scope of discovery, responses to written discovery, deposition testimony, document production, expert reports, summary judgment, and how the jury is instructed during trial.

D. Common Areas of Dispute over Trade Secret Identifications

Disputes over identification can be costly, time consuming, and burdensome for the court and parties. In the worst instances, parties trade accusations of delay and impropriety.

In jurisdictions or federal district courts that have not yet settled on a process, litigants sometimes dispute whether early identification is required at all, and, if so, to what degree, which side has the burden in a challenge, and what are the consequences for noncompliance.

Most disagreements center not on the wisdom of early identification, but rather on the degree of particularity that should be required. For example, disputes may center on: (1) whether information should be broken down and described in a numbered list; (2) whether claimed secrets can be identified in whole or in part by reference to documents; (3) whether and how “combination” or “compilation” trade secrets should be described; and—most commonly—(4) the degree of detail required for identification.

Parties sometimes dispute the standard for subsequent amendments to an identification of an asserted trade secret. On the one hand, defendants contend they should not have to prepare a defense to a moving target, while on the other hand, plaintiffs are reluctant to too narrowly define the trade secrets at issue before they discover the extent of the alleged misappropriation.

Parties also may argue over whether discovery should proceed while the sufficiency of a trade secret identification remains disputed, including discovery not only on the trade secret claim but also on other claims that are based on the same general set of allegations and directed to the same body of information as the trade secret claim.

While many litigants approach these issues reasonably, the intensity of disagreements can increase at the extremes. For example, a plaintiff may seek to avoid providing meaningful information that a defendant could use to adequately defend itself, or a defendant may file repetitive motions challenging identification of an asserted trade secret and seek delay, even where the plaintiff has already provided an identification reasonable minds would find sufficiently particular.

E. The Structure and Purpose of this Commentary

The adoption of the proposals below is intended to reduce disputes and make trade secret litigation fairer and more efficient. To that end, this *Commentary* contains four guiding principles; additional commentary regarding these principles—including sample identifications of particular types of asserted trade secrets; and a model local rule and provision for a case management order (also known as a scheduling order). While by no means comprehensive, it also includes a library of relevant case law from state and federal courts for reference.

In addition to the principles, commentary, model local rule, and case law library, WG12 presents in Appendix A a Proposed Local Rule or Case Management Order that courts can use in trade secret cases around the country, taking into account the four principles enumerated below. Throughout this *Commentary*, WG12 cites various case decisions and other authorities that illustrate elements of these principles. WG12 cautions, however, that these are only examples, and WG12 does not express an opinion on the correctness or incorrectness of any particular authority or its reasoning. Many opinions and orders in trade secret cases are deliberately crafted to be opaque so that trade secrets are not exposed to non-parties. The cited cases provide color and can be instructive but should not be taken as fully representative of judicial decision-making in this field.

II. IDENTIFICATION IS NOT AN ADJUDICATION AND IS NOT A SUBSTITUTE FOR DISCOVERY

Principle No. 1: The identification of an asserted trade secret during a lawsuit is not an adjudication of the merits and is not a substitute for discovery.

Any court ruling on the sufficiency of identification of an asserted trade secret during a lawsuit is not a determination of whether the information is a trade secret or has been misappropriated. Identifying an asserted trade secret is a procedural notice issue—a drafting step to provide clarity so that merits issues can separately and later be determined in a facilitated manner.⁶

Courts have repeatedly held that the acceptance by a court or party of a trade secret identification as adequate is not a ruling or an admission as to whether the information at issue is in fact a trade secret.⁷ The identification is instead intended to put the court and parties on notice of the plaintiff’s misappropriation claim and to facilitate evaluation and resolution of issues such as a request for preliminary relief, case management (including the scope of discovery), appropriate protection of all parties’ claimed proprietary information, and relevance.

Just as a sufficient identification says nothing about whether the identified subject matter actually constitutes a trade secret, it also does not address the question whether misappropriation

6. See *Brescia v. Angelin*, 172 Cal. App. 4th 133, 144 (Cal. Ct. App. 2009) (emphasizing that the step of identifying an asserted trade secret during a lawsuit is not a mini-trial on the merits, and instead is a preliminary step before reaching the merits).

7. See, e.g., *Advanced Modular Sputtering, Inc. v. Superior Court*, 132 Cal. App. 4th 826, 835–36 (2005) (identification does not call for a “miniature trial on the merits”).

has occurred. The plaintiff must have a sufficient basis to assert that a trade secret was misappropriated. But in most instances, the identification need only describe the information reasonably believed to have been misappropriated. It does not and need not specify how that misappropriation took place and whether the kind of misappropriation alleged (improper acquisition, use, or disclosure) varies as to each specific trade secret.

Nor is the identification of a trade secret the end of the parties' ability to inquire through discovery into how the trade secret works or is used, how it has been protected or not, and whether the alleged trade secret constitutes a trade secret. Identification is an important step in trade secret litigation, but it is not an end point or a substitute for discovery.

III. TIMING OF THE IDENTIFICATION

Principle No. 2: The party claiming misappropriation of a trade secret should identify in writing the asserted trade secret at an early stage of the case.

A party claiming the misappropriation of a trade secret should describe the asserted trade secret in written form. The plaintiff should not be allowed to refuse to provide any written identification or be allowed to identify only documents and state that its asserted trade secret may be found in those documents.⁸

Guideline 1: Where the trade secret plaintiff requests preliminary relief, the scope of an identification will depend on the relief sought.

If there is a request for preliminary relief, then the asserted trade secret should be identified as part of that process. The extent and scope of an identification may vary, however, depending on the nature of the relief sought, the urgency of the claimed need for relief, and the timing of the request. On this latter point, where a party seeks a temporary restraining order at the outset of a lawsuit or seeks a preliminary injunction, whether the same principles for identification of an asserted trade secret detailed below should apply depends on the nature of the relief sought.

As an exception, one circumstance where a plaintiff seeking early injunctive relief would not be required to serve an identification that complies with the rule or order proposed in Appendix A below is when there is evidence that a defendant downloaded or otherwise took documents or information, and the

8. *But see* Guideline 6, *infra*.

plaintiff seeks an order to (1) preserve evidence relating to that cause of action; and/or (2) return the documents or information alleged to have been taken.⁹

In other situations in which a plaintiff seeks early injunctive relief, the plaintiff would be required to serve on the defendant an identification of its asserted trade secret that complies with the rule or order proposed below.¹⁰ In cases where the plaintiff seeks early injunctive relief regarding some, but not all, of its asserted trade secret as to some, but not all, named defendants, and where such asserted trade secrets are reasonably separable by subject matter or otherwise, the plaintiff need only identify the asserted trade secret at issue in the motion and need only serve such identification on the defendants against or from whom such relief is being sought.¹¹

Guideline 2: Where the plaintiff does not request preliminary relief, the plaintiff should identify an asserted trade secret with reasonable particularity by the outset of merits discovery.

If there is no request for preliminary relief, the asserted trade secret should be identified with reasonable particularity by the outset of merits discovery. Courts may implement this principle with attention to the needs of the particular case and applicable statutes, and case law rules.

9. See Appendix A, Paragraph 3-1, *infra*.

10. See Appendix A, Paragraph 3-2, *infra*.

11. *Id.*

IV. REQUIRED LEVEL OF PARTICULARITY: CASE LAW EXAMPLES

Principle No. 3: The party claiming the existence of a trade secret must identify the asserted trade secret at a level of particularity that is reasonable under the circumstances.

Any type of information is eligible to be a trade secret. In other words, the subject matter of an asserted trade secret varies widely, from operational information like customer lists to financial information like prerelease prices to technical information like formulas and inventions. A rule or order requiring trade secret identification should therefore be flexible enough for litigants and courts to use in cases involving small businesses or large corporations and in relatively simple or complex cases, regardless of the type of information at issue. Given the wide variety of information and technology that can be at issue in trade secret cases, there is no single format by which a plaintiff can properly identify its asserted trade secret. However, a proper identification must reach a level of particularity that is reasonable under the circumstances, including taking into account the alleged urgency of the need for relief and the nature of the relief sought. For example, a request for relief directing the segregation and inspection of particular information alleged to have been wrongfully acquired, retained, or transferred in bulk may require a different level of particularity than a claim that a defendant has used a particular chemical formula to advance its research, development, or manufacturing processes.

Guideline 3: The description of an asserted trade secret in a publicly filed pleading, or other publicly filed document, may be general if necessary to avoid destroying the status of information asserted to be a trade secret.

Because most complaints are public documents, a plaintiff is not expected to provide in a complaint details that publicly disclose an asserted trade secret. Ordinarily, it is sufficient for a plaintiff to provide descriptions of the categories of the asserted trade secrets in a complaint. We address here the situation after the plaintiff has filed the complaint and the parties or the court have arrived at procedures to protect the trade secret from public disclosure. Issues surrounding the use of protective orders to shield parties' confidential information during litigation are addressed in the forthcoming Working Group 12 *Commentary on Protecting Trade Secrets in Litigation About Them*.

Guideline 4: The identification of an asserted trade secret under a protective order or equivalent agreement between the parties should be made with sufficient particularity to allow the defendant to meaningfully compare an asserted trade secret to information that is generally known or readily ascertainable and to permit the parties and the court to understand what information is claimed to be a trade secret.

Because there is no “one-size-fits-all” format for the identification of trade secrets, WG12 provides here instead: (1) examples of identifications that courts have deemed sufficient, (2) examples of identifications that courts have deemed insufficient, (3) a suggested format for an identification, and (4) additional guidelines that should be helpful.

Guideline 5: While an asserted trade secret should be identified at a level of particularity that is reasonable under the circumstances, a defendant should not use this standard as a tool to delay litigation by demanding particularity beyond that reasonably necessary for the defendant to develop its defenses and for the court to

evaluate the claims and defenses.

The Principles and Guidelines in this Commentary are designed to assist in the efficient administration of trade secret cases, not to prolong or expand the scope of litigation. For example, in a particular case, the trade secret may be that a specific ingredient—say, aluminum—is used for a specific purpose in a certain way as part of a formula claimed to be at issue. There may be several different types of aluminum, each sourced from a different vendor. If the plaintiff does not claim that the efficacy of the secret formula is affected by the source of the aluminum, however, or the defendant is not accused of having misappropriated information regarding the source of the aluminum, it would not be efficient to require the plaintiff to identify the particular source of aluminum it uses.

Absent a statute, rule, or existing order that requires trade secrets to be identified with particularity before the plaintiff may engage in discovery, the progress of the case, including such discovery, should be delayed only if the defendant successfully moves for a protective order or other ruling to that effect. This Commentary does not express an opinion on whether or under what circumstances such an application should be granted. Where the court declines to enter such an order, it should consider alternative measures to ensure that a reasonable identification is provided at an appropriate stage of the litigation.

A. Examples of Identifications Courts Have Deemed Sufficient

Courts have found that the identification of a trade secret is sufficiently particular where it enables the court to manage the

scope of discovery and the defendant to prepare a defense.¹² Importantly, consistent with Principle No. 1 above, courts have distinguished the sufficiency of the identification from whether the identified information is in fact a trade secret.¹³ The format of the identification and the scope and amount of information necessary to identify a trade secret may vary depending on the nature of the secret that is being claimed. In some cases, the identification can be straightforward and likely uncontroversial. For example, if the claim is that a physical prototype embodies the trade secret, the plaintiff should so state and provide a copy of the prototype, provide access to the prototype, or provide photographs that sufficiently display the prototype. If the claim is that a particular algorithm is the trade secret, the algorithm itself should be disclosed. If the claim is that a customer list is the trade secret, the list should be provided, with a statement that the compilation of all the information in the list (or a specified subset of that information) is the trade secret or, instead, a statement that specified information about each customer, such as its name, is a separate trade secret. And if immediate relief is sought for the preservation or return of specified files alleged to have been improperly downloaded or taken, there may be no need at that stage of the proceeding to address the specific trade secret information contained in those files.

12. *See, e.g.,* M.A. Mobile Ltd. v. Indian Inst. of Tech. Kharagpur, No. C08-02658 RMW (HRL), 2011 WL 92734 (N.D. Cal. Jan. 10, 2011).

13. *See, e.g.,* Prolifiq Software Inc. v. Veeva Sys. Inc., No. C 13-03644 SI, 2014 WL 2527148, at *3 & n.4 (N.D. Cal. June 4, 2014) (decided on motion to dismiss Third Amended Complaint; holding that California's procedural rule concerning identification of trade secrets "does not create a procedural device to litigate the ultimate merits of the case—that is, to determine as a matter of law on the basis of evidence presented whether the trade secret actually exists") (citing *Brescia v. Angelin*, 172 Cal. App. 4th 133, 149 (2009)).

There is a limited body of reported cases providing guidance on the adequacy of early trade secret identifications.¹⁴ Some examples of efforts the courts have found to be adequate include:

- a listing of specific computer files with reference to specific pages of documents;¹⁵
- a flow chart identifying the structural aspects of a computer program;¹⁶
- compilations of data specifically identifying related text files allegedly at issue (as opposed to an earlier identification of “Plaintiff’s Data-Source Database,” which was held to be too broad and not sufficiently particular);¹⁷
- a “schematic depicting a [billing] database’s structure”;¹⁸
- a formula setting forth 15 specific ingredients and their percentages and a manufacturing process for combining and processing them;¹⁹

14. Detailed discussions about the adequacy of trade secret identifications are frequently filed only under seal, so the publicly available case law on this topic is limited. California has produced a disproportionate amount of these published cases due to the fact that it enacted its version of the United Trade Secrets Act (UTSA) in 1985 with the then novel “reasonable particularity” requirement for pre-discovery identification of trade secrets, resulting in a wave of disputes over interpretation and application of the statute.

15. *CBS Interactive, Inc. v. Etilize, Inc.*, No. BC410579, 2009 WL 8514005 (Cal. Super. Ct. Nov. 20, 2009).

16. *Id.*

17. *Id.*

18. *TelSwitch, Inc. v. Billing Sols. Inc.*, No. C 12-00172 EMC LB, 2012 WL 3877645, at *2 (N.D. Cal. Sept. 6, 2012) (unreported).

19. *Brescia v. Angelin*, 172 Cal. App. 4th 133, 141 (2009).

- specification of pricing of products sold to specifically identified customers, profit margins and production costs on those products, and promotional discounts, pricing concessions, advertising allowances, volume rebates and marketing concessions, rebate incentives, trade discounts, and payment terms offered to particular customers;²⁰
- identification of eight elements that, in combination with one another, were alleged to form the trade secrets at issue;²¹
- identification of a specific combination of flow charts, even though some of the individual charts contained public information;²²
- identification of claimed trade secrets that the defendant had identified as warranting protection as trade secrets while employed by the plaintiff;²³
- identification of seven discrete aspects of an adjustable, weighted golf club design, including details of design elements and degrees employed in the product.²⁴

A description in a trade secret identification that uses terms like “comprising” or “including” may be appropriate where it

20. *Whyte v. Schlage Lock Co.*, 101 Cal. App. 4th 1443, 1455 (2002).

21. *Advanced Modular Sputtering, Inc. v. Superior Court*, 132 Cal. App. 4th 826, 836 (2005).

22. *Air Facts, Inc. v. deAmezaga*, 909 F. 3d 84, 97 (4th Cir. 2018).

23. *In the Matter of Certain Crawler Cranes and Components Thereof*, No. 337-TA-887, USITC Pub. 556530 (May 6, 2015).

24. *Triple Tee Golf, Inc. v. Nike, Inc.*, 485 F.3d 253, 258–59 (5th Cir. 2007).

provides guidance on how the parties and the court are to determine whether particular information falls within the scope of what is claimed as a trade secret.²⁵ However, a description saying that the trade secret “includes but is not limited to” particular information may sometimes be misused as a way of preparing to “spring” an entirely new claimed secret on a defendant at a later stage of the litigation without providing fair notice to either the court or the parties. Therefore, courts that are asked to assess the adequacy of an identification should be attentive to whether plaintiff’s use of such terms is an attempt at gamesmanship rather than a thoughtful, although nonexhaustive, identification.

Sometimes an asserted trade secret such as a formula, a computer program, or a process may have many elements. If the plaintiff does not intend to place all of those elements at issue, the plaintiff likely will not have to identify all elements of the formula or program. The time of the court and the litigants should not be spent on achieving granular identification of information plaintiff does not contend is at issue.

B. Examples of Identifications Courts Have Deemed Insufficient

Courts and parties should focus on whether a proposed identification can be reasonably understood by the court and parties to identify the information at issue. Generally, a “data dump” without further particularity will not satisfy that objective.

25. As discussed in detail in Sect. IV.C. (A Proposed Format for Identification), *infra*, Working Group 12 does not intend to adopt or require the use of patent-drafting concepts or terms in the identification of asserted trade secrets in misappropriation cases.

In some cases, including those set forth below, courts have determined after the motion-to-dismiss stage that a proposed trade secret identification was insufficient:

- submission of 20 pages of formulas and machine operating settings²⁶
- a claim to the dimensions and tolerances of particular components, without stating what were the specific dimensions and tolerances²⁷
- referring to an entire computer program containing thousands of lines of code as a trade secret where the plaintiff does not contend that the entirety of the program has been misappropriated²⁸
- providing only hundreds of file names, or a 43-page description of “the methods and processes

26. *Loparex, LLC v. MPI Release Techs., LLC*, No. 09 Civ. 1411, 2011 WL 1135906, at *5 (S.D. Ind. Mar. 25, 2011).

27. *Imax Corp. v. Cinema Techs., Inc.*, 152 F.3d 1161 (9th Cir. 1998). Note that if the plaintiff had not claimed these dimensions and tolerances as trade secrets, however, their identification would not have been necessary.

28. *Canter v. W. Publ'g Co., Inc.*, 1999 U.S. Dist. LEXIS 3815 (N.D. Cal. Jan. 6, 1999) (granting summary judgment to defendant where plaintiff made a blanket reference to an entire computer program as a trade secret, leaving defendant and the court to speculate as to what plaintiff claimed to be its trade secrets). *See also* *Staffbridge, Inc. v. Gary D. Nelson Assocs., Inc.*, No. 0124912BLS, 2004 WL 1429935, at *2 (Mass. Super. June 11, 2004) (unpublished opinion), (requiring plaintiff to separate its alleged trade secrets from the “vast body” of its own source code, but not to go further and separate it from information in the public domain, as the latter was a merits issue, not an identification issue).

underlying and the inter-relationships among various features” of plaintiff’s software²⁹

- merely referring to voluminous documents without identifying the particular information within them that is the claimed trade secret³⁰

29. *IDX Sys. Corp. v. Epic Sys. Corp.*, 285 F.3d 581, 583 (7th Cir. 2002) (holding the proposed identification was “both too vague and too inclusive, effectively asserting that all information in or about its software is a trade secret.”); *Integral Dev. Corp. v. Tolat*, No. C 12-06575 JSW, 2014 WL 721844, at *3 (N.D. Cal. Feb. 24, 2014) (granting summary judgment to defendant on trade secret claim, finding a “description of the category, or even of the sub-categories of information within a category, does not comply with the requirement to identify the actual matter that is claimed to be a trade secret . . . Listing hundreds of file names without identifying the trade secret information contained within the files, is insufficient.” (internal citations omitted)).

30. *See, e.g., Safety Today, Inc. v. Roy*, 2014 WL 12749231, at *2–3. (S.D. Ohio Feb. 11, 2014) (rejecting effort to produce documents in lieu of identifying claimed trade secret, finding that “[o]nly the employer will know what portion of that myriad of information known to its employees can legitimately be claimed as a trade secret, and no amount of record production . . . can provide the appropriate answer to that question,” quoting *Xerox Corp. v. IBM*, 64 F.R.D. 367, 371 (S.D.N.Y. 1974)); *U.S. Gypsum Co. v. LaFarge N. Am., Inc.*, 508 F. Supp. 2d 601, 636 (N.D. Ill. 2007) (holding that plaintiff could not “simply point to an 11,000-page document covering many diverse topics and assert that the entire document constitutes a trade secret that defendants must refute, page-by-page.”); *Utah Med. Prods., Inc. v. Clinical Innovations Assocs., Inc.*, 79 F. Supp. 2d 1290 (D. Utah 1999), *aff’d* 251 F. 3d 171 (Fed. Cir. 2000) (dismissing trade secret claim on motion for summary judgment where plaintiff alleged only that much information in 17,000 pages of documents was a trade secret).

- a combination of “[p]erhaps five” and “[p]erhaps four, probably not fewer” than any four of thirteen elements in a claimed combination³¹

Determining whether a proposed trade secret identification is sufficient is often contextual; indeed, in other cases, courts have found that identifications that appear similar in some respects to those described above were actually sufficient. One distinguishing factor in some instances may have been that the defendant seemed to have some prior business knowledge of what trade secrets were at issue.³²

31. *Maxtech Consumer Prods., Ltd. v. Robert Bosch Tool Corp.*, 255 F. Supp. 3d 833, 854 (N.D. Ill. 2017); *see also, e.g., Tesla Wall Systems, LLC v. Related Companies, L.P.*, No. 17-cv-5966, 2018 WL 2225002, at *5 n.4 (S.D.N.Y. May 14, 2018) (granting summary judgment to defendant based on claim that defendant had misappropriated an unspecified “unique compilation of data”); *Vesta Corp. v. Amdocs Mgmt. Ltd.*, 147 F. Supp. 3d 1147, 1155–56 (D. Ore. 2015), *Switch Commc’ns Grp. v. Ballard*, No. 2:11-CV-00285-KJD, 2012 WL 2342929 (D. Nev. June 19, 2012), and *Hill v. Best Med. Int’l Inc.*, No. 2:2007-cv- 01709, 2010 WL 2546023, at *3–4 (W.D. Pa. June 24, 2010) (notably, awarding attorney’s fees to defendant in a subsequent decision for the failure to identify trade secrets with particularity), all citing *Struthers Sci. & Int’l Corp. v. Gen. Foods Corp.*, 51 F.R.D. 149, 153 (D. Del. 1970), for the proposition that where the trade secret is a combination, “[plaintiff] should . . . specifically describe what particular combination of components it has in mind, how these components are combined, and how they operate in unique combination”; *Sit-Up Ltd. v. IAC Interactive Corp.*, No. 05-cv-9292, 2008 WL 463884, at *10 (S.D.N.Y. Feb. 20, 2008) (rejecting generalized description of compilation, phrasing the standard as requiring the plaintiff to “describe the secret with sufficient specificity that its protectability can be assessed and to show that its compilation is unique”; granting summary judgment for defendants on trade secret claim).

32. *See, e.g., Alta Devices, Inc. v. LG Elecs., Inc.*, 343 F. Supp. 3d 868, 881 (N.D. Cal. 2018) (finding that schedules to the parties’ pre-dispute contracts had defined trade secrets and that plaintiff’s trade secret identification in its

C. *A Proposed Format for Identification*

Patent claims are drafted in accordance with specific requirements set forth in statutes and case law.³³ For a host of legal and practical reasons, courts should not import those requirements into trade secret litigation and require trade secret plaintiffs to compose their identifications as if they were patent claims. For example, the broader scope of trade secrets (applying to “information” and not just “inventions”), together with the special requirements of the patent statute, make it inappropriate to equate the two.

With that said, one way to identify trade secrets that is consistent with the principles and guidelines in this *Commentary* is provided below and includes three parts:

1. A short introductory description for the general subject matter, i.e., general subject matter or description, such as:
 - a. A formula for a
 - b. A program, or code, to do b
 - c. A process for making c
 - d. A design of d
 - e. A combination or compilation of a type of information
 - f. A use of e for doing f
 - g. A fact
 - h. A piece of information
 - i. A teaching from research and development of g

complaint enabled the defendant to ascertain at least the boundaries in which the trade secret lies).

33. See, e.g., 35 U.S.C. § 112.

- j. An image, diagram, drawing or other rendering of h
With the a, b, c, d, type of information, e, f, fact, piece of information, g or h specifically identified as noted in part (3) below,
2. A transitional term, such as “comprising,” “consisting of,” “for,” “of,” and
3. Specific, identifying information, such as: the trade secret elements, components, ingredients, steps, algorithms, and other specific details the plaintiff contends constitute the trade secret at issue.

To provide applications of this proposed format of identification of trade secrets that would presumptively meet the trade secret plaintiffs’ burden at an early stage of a trade secrets case, WG12 provides the following examples:

If the plaintiff contends its trade secret is an entire formula, then it should identify the entire formula, including, where the plaintiff contends such information is at issue, the ingredients and the respective amounts, proportions, and ranges of ingredients. If the plaintiff contends individual ingredients or amounts, proportions, or ranges of individual ingredients are additional trade secrets, then it should separately identify those ingredients, amounts, proportions, and ranges as additional trade secrets.

If the plaintiff contends the sources or other specific aspects of particular ingredients are additional trade secrets, then it should separately identify each such source or aspect as an additional trade secret.

If the plaintiff contends its trade secret is an entire computer program, then the plaintiff should identify the entire program. If the plaintiff contends a trade secret is only a portion of a

program, such as certain lines of code, module, or an underlying algorithm, then the plaintiff should identify that portion, i.e., those lines of code, the module, or the algorithm, as a trade secret.

If the plaintiff contends its entire manufacturing process is a trade secret, then it should describe the entire process, including relevant details such as sequencing rates and methods of adding ingredients, temperatures, pressures, cure or preparation times, and the like. If the plaintiff contends its trade secret consists of only one or several aspects of a process, then it should identify those specific aspects that it claims to be a trade secret.

If the plaintiff contends its trade secret is an entire design, then it should identify the entire design, including all aspects of the design, such as dimensions, features, and materials. If the plaintiff contends individual aspects, such as individual dimensions, features, and materials, are additional trade secrets, then it should separately identify each such aspect as an additional trade secret.

If the plaintiff contends its trade secret is a combination or compilation of information, including information that is generally known, such as a combination of elements or a compilation of data (whether any, some, or all of the individual elements or data are generally or publicly known or not), then it should identify the entire combination or compilation. If the plaintiff contends its trade secret consists of only one or several elements of the claimed combination or only a subset of data in the compilation, then it should identify that element(s) or subset(s) as a trade secret(s).

If the plaintiff contends its trade secret is a use of an item, including a generally or publicly known item, for a certain purpose or function, then it should identify the use of the item and the purpose or function.

A plaintiff can choose to identify a trade secret in a format other than the three-part format described above. However, a differently formatted identification likewise must meet the reasonable particularity standard, i.e., a level of particularity that is reasonable under the circumstances, taking into account the nature of the information at issue, the nature of the parties and the industry (e.g., the extent to which there is knowledge in the industry or by the defendant that is not claimed to be a trade secret), the relief sought, and the urgency of the relief requested. One example of a differently formatted identification might concern details in a customer list that a plaintiff contends is its trade secret. In that instance, given the relative simplicity of expressing such information, the plaintiff might simply provide the database at issue, or list for each customer or group of customers the specific associated details alleged to be at issue.

The same guidelines apply if a plaintiff seeks to identify negative information, such as knowledge of something that did not work well or at all. In making such an identification, it may be appropriate, depending on context, to identify the trade secret as, for example, "the fact that adding ingredient x to the formula does not create a stronger material, as evidenced in detail in the lab results set forth in Exhibit 1 pertaining to material strength." Or, in the case of extensive research and development leading to a preferred outcome, the information developed through experimentation may be described by reference to the records of it.

D. Guidelines for a Trade Secret Identification

Guideline 6: The plaintiff should not identify an asserted trade secret exclusively by reference to a document or other item, or exclusively by cross-reference to another asserted trade secret, unless such document,

other item, or cross-reference sets forth the asserted trade secret.

A trade secret should not be identified exclusively by reference to a document or other item, unless such document (e.g., specification, drawing, schematic diagram or formula/formulation sheet, computer program (source code), or customer list) or other item (e.g., prototype or model) constitutes the alleged trade secret. A trade secret should not be identified exclusively by reference to another asserted trade secret, unless such reference provides an identification that meets the reasonable particularity standard. If the plaintiff references a document or other item as setting forth more than one trade secret, then where possible the plaintiff should specify which portion of the document or other item identifies each such trade secret. The plaintiff may choose to add document or other-item references to its written identification, but such references do not excuse the plaintiff's obligation to provide a proper identification that meets the reasonable particularity standard.

Guideline 7: The plaintiff should verify its identification of an asserted trade secret under oath or affirmation.

The identification, regardless of format, should be verified,³⁴ i.e., the plaintiff should have the identification signed by at least one witness (other than counsel) knowledgeable about what the plaintiff contends are the trade secrets. The witness should certify that, to the best of that person's knowledge, information, and belief, formed after a reasonable inquiry, the identification

34. See, e.g., *Decision Insights, Inc. v. Sentia Grp., Inc.*, 311 F. App'x 586, 589–90 (4th Cir. 2009) (directing plaintiff to produce a “clear and express verified statement containing only those items which Plaintiff considers to be actual trade secrets and which Plaintiff has reasonable grounds to believe were misappropriated by Defendant.”).

is accurate and complete as to the asserted trade secrets. The same should be true of any permitted amendments to the identification. Verification by a party's employee or officer, rather than by counsel, will aid in accuracy by encouraging a party's qualified employee or officer, such as a qualified engineer or other employee, to play an active role in identification efforts. It also will facilitate discovery by identifying at least one person knowledgeable about each asserted trade secret.

Guideline 8: If the plaintiff claims that the defendant has taken files or other materials, the court may allow motion practice and/or discovery relating to the return or inspection of such files or materials prior to requiring identification of asserted trade secret contained within such files or materials.

There may be instances in which the plaintiff knows, at the time of filing its trade secret misappropriation claim, that files or other materials have been taken, but does not know the contents and/or the extent of what was taken. For example, the party taking the files or materials may have removed the only existing version. In such instances, it may be appropriate to allow motion practice and/or discovery relating to the return or inspection of such files or materials prior to requiring particularized identification of an asserted trade secret.

V. AMENDMENT OF THE IDENTIFICATION

Principle No. 4: The identification of an asserted trade secret may be amended as the case proceeds.

The identification of an asserted trade secret may be amended in appropriate circumstances.

If a party wishes to amend its prior identification, it should first confer with the opposing party. If the parties are unable to reach an agreement, the party proposing the amendment may then file a motion for leave to amend, with the motion subject to the court's discretion based on all the circumstances, including all the factors set forth below.

Important and often dispositive factors in deciding such a motion are whether the party seeking leave to amend was diligent and whether the opposing party would be unduly prejudiced by amendment. These factors are not specific to trade secret law but reflect traditional concerns where amendment of pleadings or disclosures is sought.

There are other factors that may be considered. Although these factors exist in other areas of law as well, they may have special relevance in trade secret cases because of information asymmetry—for example, a plaintiff may learn details about the claimed misappropriation, including a defendant's internal technology or processes or dissemination of the information at issue, only during discovery. Thus, these factors include whether the proposed amendments are based on facts that were newly learned (such as learning through discovery that a defendant has misappropriated a trade secret the plaintiff previously did not believe was at issue), in which case an amendment is more likely to be allowed, depending on overall case management considerations. They also include the stage of the litigation, with amendments being more likely to be allowed if they

occur at an earlier stage; whether the amendment will delay the trial date, with amendments being more likely to be allowed if they do not necessitate a delay; and whether this is the first attempt or a later attempt to amend, with the first attempt more likely to be allowed.

Another factor that may be considered is the nature of the amendment: whether it simply clarifies or explains a prior identification or whether it materially alters or expands a prior identification. In the latter two cases, it will generally be appropriate for the party seeking leave to amend the identification to explain the circumstances and rationale for the proposed change. If the proposed amendment materially narrows (i.e., reduces) the number of asserted trade secrets, the party seeking to amend the identification should explain why the amendment was not sought earlier. In some cases, discovery may have revealed that contrary to initial reasonable belief, fewer trade secrets are at issue than initially suspected, or certain information initially at issue may have ceased being a trade secret at a relevant point in time. In many such cases, a narrowing amendment could likely be assented to as appropriate without material court intervention. In other cases, however, if a prior identification of trade secrets is found to have been overly expansive by design in order to unduly expand the scope of discovery, protract litigation, or drive up expenses, it may be that sanctions directed against the plaintiff for the initial overclaiming are appropriate. Such cases will likely be rare, and, in general, narrowing amendments should be viewed favorably as streamlining the case.

APPENDIX A: Proposed Model Local Rule For Trade Secret Misappropriation Cases

The goal of Appendix A is to propose a rule appropriate for the full range of trade secret misappropriation cases—whether a relatively simple customer list case or a highly complex technology case. WG12’s proposed rule is thus presented here in two formats—a model local rule, and a provision for use in a case management order—so that courts and parties can select the format more appropriate for them.

The identification of asserted trade secrets must be in writing and it must be kept confidential using appropriate procedures as determined by the court.

Courts may adjudicate cases concerning the rule or provision, and address violations of the rule or provision, in the same way that they would address other issues concerning discovery sequencing and conduct.

This format is for use in cases in which a plaintiff alleges a trade secret misappropriation claim. Whether a rule or provision requiring identification should be extended to a breach of contract claim or other claim arising out of the defendant’s access to the plaintiff’s confidential information may depend on the degree of overlap with the trade secret misappropriation claim. Whether a rule or provision requiring identification of asserted trade secrets should also apply to tort claims that the plaintiff alleges may depend on factors such as whether the particular jurisdiction treats its version of the Uniform Trade Secrets Act as preempting such tort claims, along with case-specific factors.

1. Scope of Rules

1-1. Title. These are the Local Rules of Practice for Trade Secret Cases before the [insert name of federal district court or state court].

1-2. Application. These rules apply to all civil actions filed in or transferred to this Court in which a trade secret misappropriation cause of action is asserted, whether the cause of action arises out of state or federal law or both. They are intended to supplement, and not to replace, the rules that otherwise apply in this Court, in order to provide for better management of issues likely to arise in trade secret litigation.

1-3. Modification. Based on the circumstances of the particular case, the Court may modify any aspect of these Local Rules in the interests of justice and efficient case management. Before seeking any modification, the parties shall confer and attempt to reach agreement. The burden for obtaining any such modification rests on the party proposing the modification.

1-4. Effective date. These rules apply to any case filed in or transferred to this Court after [insert date], and may be applied in previously pending cases as the Court deems appropriate.

2. Identification of Asserted Trade Secrets

2-1. Obligation to identify trade secrets. Subject to applicable law regarding the timing and content of trade secret identification and following the entry of a protective order, a party claiming the existence of a trade secret

must, by the outset of merits discovery (or, subject to Rule 3-1, with a motion for preliminary relief) identify in writing and serve on the parties, with a level of particularity that is reasonable under the circumstances, each asserted trade secret. The required particularity of this identification differs from what may be adequate in a publicly filed pleading under applicable pleading rules such as Federal Rule of Civil Procedure 8 or analogous state procedural rules, which may allow more generalized identification to avoid disclosing the substance of the asserted secrets and destroying the trade secret status of the information. The identification is not intended to and should not require the Court to make a threshold finding as to whether it is a trade secret or was misappropriated. These issues are ultimate issues to be decided by the Court or jury.

2-2. Initial identification. The identification required by Paragraph 2-1 must be sufficiently particularized to allow the other party to meaningfully compare the asserted trade secret to information that is generally known or readily ascertainable and to permit the parties and the Court to understand what information is claimed to be the trade secret. The identification does not need to specify the differences between the alleged trade secret and publicly available information. The identification should separate, to the extent practical, different asserted trade secrets into numbered paragraphs. Documents may be appended as a supplement to the identification but may not be used as a substitute for the identification unless the document itself is claimed to be the trade secret. In cases where an entire document or portions thereof constitutes the trade secret, the written identification must

identify the content in such documents or portions of such documents in language sufficient to meet the standards in this Paragraph 2-2.

2-3. Amendments. A party that has provided an initial identification under Paragraph 2-1 may amend that identification if the parties so agree or if the Court permits such amendment. The parties must confer regarding the timing and terms of the proposed amendment. If the parties are unable to reach an agreement, the party proposing the amendment may apply to the Court for an order allowing the proposed amendment. In determining whether to grant leave to amend the identification, the Court shall consider whether the party seeking amendment was diligent and whether the party opposing amendment would be unduly prejudiced by the amendment. Other factors that may be considered include, but are not limited to: whether the proposed amendment is based on facts that were newly learned in discovery; the stage of the litigation; whether the amendment will expand discovery and/or delay the trial date; and whether the amendment adds, removes, or materially modifies asserted trade secrets, or merely clarifies an existing identification.

2-4. Verification. The identification of each asserted trade secret shall be verified under oath or affirmation by one or more employees or officers of the party asserting trade secret misappropriation.

2-5. Purpose of the identification. The purpose of the identification under Paragraph 2-1 is to facilitate the resolution of trade secret cases and to inform the Court and parties of the information at issue. The process of

identifying trade secrets should not become a protracted and repetitive exercise in evaluating proffered identifications that satisfy this purpose.

3. Applications for Preliminary Relief

3-1. Orders to preserve evidence and/or return documents or information. Where a party has evidence that an opposing party improperly downloaded or otherwise took documents, things, or information from the party, and the party files a lawsuit that includes a trade secret misappropriation cause of action, and then, by motion, seeks an early court order requiring only that the defendant (1) preserve evidence; and/or (2) return the specific documents, things or information that were allegedly taken, the moving party is not required to prepare or serve an identification of its asserted trade secret that complies with Paragraph 2.

3-2. Identification of asserted trade secrets in requests for other early injunctive relief. In all other situations in which a party asserting trade secret misappropriation seeks such relief, the moving party must comply with Paragraph 2 as to the trade secrets for which it seeks early injunctive relief to the extent it has not already done so.

3-3. This Paragraph 3 is subject to Federal Rule 65(d) or state law equivalents and other applicable statutory requirements.

APPENDIX B: Examples of Case Law Regarding Identification of Asserted Trade Secrets During Litigation

This list, although not comprehensive, identifies most of the best-known cases regarding trade secret identification and many other examples from jurisdictions around the country through late 2018. It separates cases by the applicable stage of the litigation and also by jurisdiction. As with other cited authority, WG12 does not necessarily embrace any of these decisions as representing consensus views or controlling law on the issues they address.

Early Discovery and/or Discovery Stay Case

State & Territorial Statutes:

CAL. CIV. PROC. CODE § 2019.210 (enacted in 1985).

PUERTO RICO LAWS title 10, § 4139(a) (2011).

MASS. GEN. LAWS ANN. ch. 93 § 42D(b) (2018).

State Cases:

California: *Perlan Therapeutics, Inc. v. Superior Court*, 178 Cal. App. 4th 1333, 1339 (2009); *Brescia v. Angelin*, 172 Cal. App. 4th 133, 144 (2009); *Advanced Modular Sputtering, Inc. v. Superior Court*, 132 Cal. App. 4th 826, 834–35 (2005).

Delaware: *Engelhard Corp. v. Savin Corp.*, 505 A.2d 30 (Del. 1986).

Florida: *AAR Mfg., Inc. v. Matrix Composites, Inc.*, 98 So. 3d 186, 187 (Fla. Dist. Ct. App. 2012).

New Hampshire: *Vention Med. Advanced Components, Inc. v. Pappas*, 2015 N.H. Super. LEXIS 7 (July 15, 2015).

North Carolina: *DSM Dyneema, LLC v. Thagard*, 2014 WL 5317770 (N.C. Super. Ct. Oct. 17, 2014).

Federal Cases:

Arizona: *BioD, LLC v. Amnio Tech., LLC*, 2014 WL 3864658 (D. Ariz. Aug. 6, 2014).

California: *VIA Techs., Inc. v. Asus Computer Int'l*, 2016 U.S. Dist. LEXIS 141581 (N.D. Cal. Oct. 12, 2016) & 2016 U.S. Dist. LEXIS 63676 (N.D. Cal. May 13, 2016), & 2016 WL 1056139 (N.D. Cal. Mar. 17, 2016); *Lilith Games (Shanghai) Co. v. uCool, Inc.*, 2015 WL 4149066 (N.D. Cal. July 9, 2015); *Loop AI Labs Inc. v. Gatti*, 195 F. Supp. 3d 1107 (N.D. Cal. July 6, 2016) & 2015 WL 9269758 (N.D. Cal. Dec. 21, 2015); *Prolifiq Software Inc. v. Veeva Sys., Inc.*, 2014 WL 2527148 (N.D. Cal. June 4, 2014); *Phoenix Techs., Ltd. v. DeviceVM, Inc.*, 2010 WL 8590525 (N.D. Cal. March 17, 2010).

Connecticut: *Powerweb Energy, Inc. v. Hubbell Lighting, Inc.*, 2012 WL 3113162, at *1–2 (D. Conn. July 31, 2012).

Colorado: *L-3 Commc'ns Corp. v. Jaxon Eng'g & Maint., Inc.*, 2011 WL 10858409, at *3–4 (D. Col. Oct. 12, 2011).

Georgia: *DeRubeis v. Witten Techs., Inc.*, 244 F.R.D. 676, 682 (N.D. Ga. 2007).

Illinois: *AutoMed Techs., Inc. v. Eller*, 160 F. Supp. 2d 915, 925–26 (N.D. Ill. 2001).

Michigan: *Giasson Aerospace Sci., Inc. v. RCO Eng'g, Inc.*, 2009 WL 1384179, at *2 (E.D. Mich. May 14, 2009).

Minnesota: *Porous Media Corp. v. Midland Brake Inc.*, 187 F.R.D. 598, 600 (D. Minn. 1999).

Nevada: *Switch Commc'ns Grp. v. Ballard*, 2012 WL 2342929, at *4–5 (D. Nev. June 19, 2012).

New Jersey: *Osteotech, Inc. v. Biologic, LLC*, 2008 WL 686318 (D.N.J. Mar. 7, 2008); *Reckitt Benckiser Inc. v. Tris Pharma, Inc.*, 2011 WL 773034 (D.N.J. Feb. 28, 2011).

North Carolina: *Ikon Office Sols., Inc. v. Konica Minolta Bus. Sols. USA, Inc.*, 2009 WL 4429156, at *4 (W.D.N.C. Nov. 25, 2009).

Ohio: *A&P Tech., Inc. v. Lariviere*, 2017 WL 6606961 (S.D. Ohio Dec. 27, 2017).

Oregon: *Vesta Corp. v. Amdocs Mgmt. Ltd.*, 147 F. Supp. 3d 1147, 1156 (D. Or. 2015); *Nike, Inc. v. Enter Play Sports, Inc.*, 305 F.R.D. 642, 646 (D. Or. 2015); *St. Jude Medical S.C., Inc. v. Janssen-Coulotte*, 305 F.R.D. 630, 632 (D. Or. 2015).

Texas: *Huawei Techs. Co. v. Huang*, 2018 U.S. Dist. LEXIS 136929 (E.D. Tex. Aug. 14, 2018); *Zenimax Media, Inc. v. Oculus Vr, Inc.*, 2015 WL 11120582 (N.D. Tex. Feb. 13, 2015); *StoneEagle Servs. v. Valentine*, 2013 WL 9554563 (N.D. Tex. June 5, 2013); *United Serv. Auto Ass'n v. Mitek Systems, Inc.*, 289 F.R.D. 244, 248 (W.D. Tex. Feb. 15, 2013); *Polydyne Software, Inc. v. Celestica Int'l, Inc.*, 2014 WL 12479201 (W.D. Tex. Dec. 31, 2014).

Utah: *Storagecraft Tech. Corp. v. Symantec Corp.*, 2009 WL 361282, at *2 (D. Utah Feb. 11, 2009).

Interrogatory Dispute

State courts:

Massachusetts: *Alnylam Pharms. v. Discerna Pharms. Inc.*, 2016 Mass. Super. LEXIS 140 (Apr. 6, 2016).

Tennessee: *Cryosurgery, Inc. v. Rains*, 2016 Tenn. Bus. LEXIS 11 (Chancery Ct. Tenn. May 25, 2016).

Federal courts:

California: *Attia v. Google LLC*, 2018 U.S. Dist. LEXIS 84196 (N.D. Cal. May 10, 2018); *Excelligence Learning Corp. v. Oriental Trading Co., Inc.*, 2004 WL 2452834, at *3–4 (N.D. Cal. June 14, 2004); *Phoenix Techs., Ltd. v. DeviceVM, Inc.*, No. 09-cv-4697, 2010 WL 8590525 (N.D. Cal. Mar. 17, 2010).

Florida: *Knights Armament Co. v. Optical Sys. Tech., Inc.*, 254 F.R.D. 463, 467 (M.D. Fla. 2008).

Georgia: *DeRubeis v. Witten Techs., Inc.*, 244 F.R.D. 676, 680 (N.D. Ga. 2007).

Illinois: *Compuware Corp. v. Health Care Serv. Corp.*, 2002 WL 485710, at *7 (N.D. Ill. Apr. 1, 2002).

Kentucky: *Caudill Seed & Warehouse Co. v. Jarrow Formulas, Inc.*, 2017 WL 4799815 (W.D. Ky. Oct. 24, 2017); *Babcock Power, Inc. v. Kapsalis*, 2015 WL 9244487 (W.D. Ky. Dec. 17, 2015).

Michigan: *Dow Corning Corp. v. Jie Xiao*, 2011 WL 6739403 (E.D. Mich. Dec. 22, 2011); *Dura Global Tech., Inc. v. Magna Donnelly Corp.*, 2007 WL 4303294, at *4 (E.D. Mich. Dec. 6, 2007).

Minnesota: *Luminara Worldwide, LLC v. Liown Elecs. Co.*, 2015 WL 9861106 (D. Minn. Oct. 5, 2015).

Nevada: *Switch Communs. Grp. v. Ballard*, 2012 WL 2342929 (D. Nev. June 19, 2012).

New Jersey: *Givaudan Fragrances Corp. v. Krivda*, 639 F. App'x 840 (3d Cir. 2016); *Givaudan Fragrances Corp. v. Krivda*, 2013 WL 5781183 (D.N.J. Oct. 25, 2013); *Vital State Canada, Ltd. v. Dream-Pak, LLC*, 303 F. Supp. 2d 516 (D.N.J. 2003).

New York: *Uni-Sys., LLC v. U.S. Tennis Ass'n*, 2017 WL 4081904 (E.D.N.Y. Sept. 13, 2017); *Norbrook Labs. Ltd. v. G.C. Hanford Mfg. Co.*, 297 F. Supp. 2d 463 (N.D.N.Y. 2003).

North Carolina: *Ikon Office Sols., Inc. v. Konica Minolta Bus. Sols. U.S.A., Inc.*, 2009 WL 4429156 (W.D.N.C. Nov. 25, 2009).

Ohio: *Safety Today, Inc. v. Roy*, 2014 WL 12749231 (S.D. Ohio Feb. 11, 2014).

Oregon: *Vesta Corp. v. Amdocs Mgmt., Ltd.*, 2016 U.S. Dist. LEXIS 45741 (D. Or. Apr. 1, 2016).

Pennsylvania: *Syngy, Inc. v. ZS Assocs.*, 2013 WL 3716518, at *15 (E.D. Pa. July 15, 2013); *Hill v. Best Med. Int'l, Inc.*, 2010 WL 2546023, at *1–3 & n.4 (W.D. Pa. June 24, 2010).

Texas: *Triple Tee Golf, Inc. v. Nike, Inc.*, 485 F.3d 253, 258–59 (5th Cir. 2007) (description of interrogatory response in case summary); *Vianet Grp. PLC v. Tap Acquisition, Inc.*, 2016 WL 9559913 (N.D. Tex. Mar. 8, 2016).

Utah: *StorageCraft Tech. Corp. v. Symantec Corp.*, 2009 WL 112434 (D. Utah Jan. 16, 2009).

Washington: *StonCor Grp., Inc. v. Campton*, 2006 WL 314336 (W.D. Wash. Feb. 7, 2006).

Court of Federal Claims: *Demodulation, Inc. v. United States*, 122 Fed. Cl. 652 (2015).

Motion for Summary Judgment, Pretrial, or Post-Trial Stage

Freeman Inv. Mgmt. Co., LLC v. Frank Russell Co., 2016 WL 5719819 (S.D. Cal. Sept. 30, 2016).

Fortinet, Inc. v. Sophos, Inc., 2015 WL 5971585 (N.D. Cal. Oct. 14, 2015).

PTT, LLC v. Gimme Games, 2014 WL 5798148 (D.N.J. Nov. 6, 2014).

Waymo LLC v. Uber Techs., Inc., 2017 U.S. Dist. LEXIS 182197 (N.D. Cal. Nov. 2, 2017) (MSJ stage).

Loparex, LLC v. MPI Release Techs., LLC, 2012 WL 6094141 (S.D. Ind. Dec. 7, 2012) (post-trial stage).

Sit-Up Ltd. v. IAC/InterActiveCorp., 2008 WL 463884 (S.D.N.Y. Feb. 20, 2008) (MSJ stage).

Amendment

Swarmify, Inc. v. Cloudflare, Inc., 2018 WL 2445515 (N.D. Cal. May 31, 2018).

Neothermia Corp. v. Rubicor Med., Inc., 345 F. Supp. 2d 1042 (N.D. Cal. 2004).

LifeCell Corp. v. Tela Bio, Inc., No. SOM-C-12013-15 (N.J. Ch. Apr. 30, 2015).

Fast Food Gourmet, Inc. v. Little Lady Foods, Inc., 2007 WL 3052944 (N.D. Ill. Oct. 18, 2007).

Dura Global Techs., Inc. v. Magna Donnelly Corp., 2011 WL 4527576 (E.D. Mich. Sept. 29, 2011).

Montgomery v. eTreppid Techs., LLC, 2008 WL 2277118 (D. Nev. May 29, 2008).

Morgardshammar, Inc. v. Dynamic Mill Servs. Corp., 2009 WL 10685154 (W.D.N.C. Nov. 19, 2009).

THE SEDONA CONFERENCE COMMENTARY
ON THE ENFORCEABILITY IN U.S. COURTS
OF ORDERS AND JUDGMENTS ENTERED UNDER GDPR

*A Project of The Sedona Conference Working Group
on Data Security and Privacy Liability (WG11)*

Author:

The Sedona Conference

Editors-in-Chief:

Alex M. Pearce

Contributing Editors:

Joseph A. Dickinson

Eric P. Mandel

Starr Turner Drum

Shoshana E. Rosenberg

Marcel Duhamel

Meredith L. Schultz

Ronald J. Hedges

David Shonka

Steering Committee Liaison:

Bob Cattanach

Staff Editors:

David Lumia

Michael Pomarico

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 11. They do not necessarily

Copyright 2021, The Sedona Conference.
All Rights Reserved.

represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

We thank all of our Working Group Series Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors, just click on the “Sponsors” navigation bar on the homepage of our website.

This publication may be cited as follows:

The Sedona Conference, *Commentary on the Enforceability in U.S. Courts of Orders and Judgments Entered under GDPR*, 22 SEDONA CONF. J. 277 (2021).

PREFACE

Welcome to the January 2021 final version of *The Sedona Conference Commentary on the Enforceability of Orders and Judgments Entered under GDPR* (“*Commentary*”), a project of The Sedona Conference Working Group 11 on Data Security and Privacy Liability (WG11). This is one of a series of Working Group commentaries published by The Sedona Conference, a 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, intellectual property rights, and data security and privacy law. The mission of The Sedona Conference is to move the law forward in a reasoned and just way.

The mission of WG11 is to identify and comment on trends in data security and privacy law, in an effort to help organizations prepare for and respond to data breaches, and to assist attorneys and judicial officers in resolving questions of legal liability and damages.

The Sedona Conference acknowledges Editor-in-Chief Alex Pearce for his leadership and commitment to the project. We also thank contributing editors Joseph Dickinson, Starr Drum, Marcel Duhamel, Ron Hedges, Eric Mandel, Shoshana Rosenberg, Meredith Schultz, and David Shonka for their efforts. We also thank Bob Cattanach for his contributions as Steering Committee liaison to the project. We thank Claire Spencer for her contributions.

In addition to the drafters, this nonpartisan, consensus-based publication represents the collective effort of other members of WG11 who reviewed, commented on, and proposed edits to early drafts of the *Commentary* that were circulated for feedback from the Working Group membership. Other members provided feedback at WG11 annual and midyear meetings where drafts of the *Commentary* were the subject of the dialogue.

The publication was also subject to a period of public comment. On behalf of The Sedona Conference, I thank all of them for their contributions.

We encourage your active engagement in the dialogue. Membership in The Sedona Conference Working Group Series is open to all. The Series includes WG11 and several other Working Groups in the areas of electronic document management and discovery, cross-border discovery and data protection laws, international data transfers, patent litigation, patent remedies and damages, and trade secrets. The Sedona Conference hopes and anticipates that the output of its Working Groups will evolve into authoritative statements of law, both as it is and as it should be. Information on membership and a description of current Working Group activities is available at <https://thesedonaconference.org/wgs>.

Craig Weinlein
Executive Director
The Sedona Conference
January 2021

TABLE OF CONTENTS

INTRODUCTION.....	284
I. AN OVERVIEW OF GDPR’S EXTRATERRITORIAL SCOPE.....	288
A. GDPR’s Territorial Scope under Article 3.....	288
B. Enforcement Activity Directed at Non-EU Organizations.....	291
II. RECOGNITION AND ENFORCEMENT OF FOREIGN JUDGMENTS IN U.S. COURTS: OVERVIEW OF CURRENT LAW	295
A. Origins of the law of recognition and enforcement of foreign judgments.....	295
B. Foundational requirements for recognition and enforcement of foreign judgments	298
C. The rule against recognition of foreign fines and penal judgments	300
D. Other grounds for nonrecognition of foreign judgments	302
E. Recognition of foreign administrative orders ..	303
F. Procedural considerations and burdens of proof.....	304
III. RECOGNITION AND ENFORCEMENT OF GDPR ORDERS AND JUDGMENTS IN U.S. COURTS: PRIVATE ACTIONS BY DATA SUBJECTS AND REPRESENTATIVE ORGANIZATIONS	306
A. General considerations for private causes of action	306
B. Data subject compensation claims under GDPR Article 82.....	310
1. Overview and general considerations	310

2.	Enforceability under U.S. law	312
C.	Injunctions and nonmonetary orders issued under GDPR Article 79.....	313
1.	Overview and general considerations	313
2.	Enforceability under U.S. law	314
IV.	RECOGNITION AND ENFORCEMENT OF GDPR ORDERS AND JUDGMENTS IN U.S. COURTS: CORRECTIVE ORDERS ENTERED BY EU SUPERVISORY AUTHORITIES.....	315
A.	Overview and general considerations.....	315
B.	Nonmonetary orders issued under Article 58: enforceability under U.S. law	317
C.	Administrative fines issued under Articles 58.2(i) and 83: enforceability under U.S. law ...	318
V.	POTENTIAL DEFENSES UNDER U.S. LAW TO AN ACTION SEEKING RECOGNITION AND ENFORCEMENT OF A GDPR ORDER OR JUDGMENT	321
A.	Lack of personal jurisdiction over the defendant in the EU	321
1.	Personal jurisdiction under GDPR Article 3.1	326
2.	Personal jurisdiction under Article 3.2	328
3.	Data Protection Officers and Article 27 representatives: impact on personal jurisdiction in the EU.....	330
4.	Execution of data processing and data transfer agreements: impact on personal jurisdiction in the EU	332
B.	Repugnancy to federal or state public policy...	334

VI. ALTERNATIVE ROUTES TO GDPR ENFORCEMENT IN U.S. COURTS: THE FEDERAL TRADE COMMISSION AND CONTRACT CLAIMS..... 337

 A. The Federal Trade Commission: Section 5 of the FTC Act and Privacy Shield remedies ... 337

 B. Contract actions associated with data protection..... 340

 1. Contracts between data subjects and data controllers..... 340

 2. Contracts between data controllers and data processors under GDPR Article 28 341

 3. Data transfer contracts based on Standard Contractual Clauses 342

VII. CONCLUSION 343

INTRODUCTION

This *Commentary* evaluates the enforceability in a United States court of an order or judgment entered under the European Union (EU) General Data Protection Regulation (GDPR)¹ by an EU court, or by an EU Member State supervisory authority, against a U.S.-based controller or processor. The goal of the *Commentary* is to provide guidance to stakeholders in the EU² and in the U.S. on the factors—both legal and practical—that speak to the enforcement of GDPR mandates through U.S. legal proceedings.

The question how and under what circumstances GDPR mandates can be enforced through U.S. legal proceedings arises as a result of the GDPR's broad territorial scope. To that end, GDPR constitutes a "significant evolution" of the territorial scope of EU data protection law compared to its predecessor and reflects an intention "to ensure comprehensive protection of the rights of data subjects in the EU and to establish . . . a level playing field for companies active on the EU markets, in a

1. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Re-pealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119/1), *available at* <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679#PP3Contents> [hereinafter GDPR].

2. GDPR has been incorporated into the European Economic Area (EEA) Agreement by the EEA Joint Committee and thus applies to all Member States of the EEA, i.e., Member States of the EU plus Iceland, Liechtenstein and Norway (note: Switzerland has not ratified the EEA Agreement, and GDPR has no direct application in that country). *See General Data Protection Regulation incorporated into the EEA Agreement*, EUROPEAN FREE TRADE ASSOCIATION, July 6, 2018, <https://www.efta.int/EEA/news/General-Data-Protection-Regulation-incorporated-EEA-Agreement-509291>. Thus, for simplicity's sake, this *Commentary* will use the term "EU" to refer to all Member States of the EEA.

context of worldwide data flows.”³ Because of this evolution in territorial scope, organizations based outside the EU—including in the U.S.—that previously were not subject to EU data protection rules, or the consequences of violating them, can now be subject to both. But as a recent report from the Internet & Jurisdiction Policy Network explains, “a state’s ability to enforce its laws is often more limited than the claims it makes regarding the reach of its laws.”⁴ Questions will thus inevitably arise about how supervisory authorities and data subjects can enforce the GDPR against these non-EU organizations.

In some cases, the answer will be straightforward. When an organization maintains a branch, subsidiary, or other assets in the EU, European supervisory authorities and data subjects can enforce GDPR mandates against the organization within the EU’s borders.

The answer is less clear, however, if an organization violates the GDPR but does not maintain a physical presence or other assets in the EU. In that case, EU supervisory authorities and data subjects could issue an order or obtain a judgment against the organization. But unless the organization is willing to comply voluntarily with that order or judgment, the supervisory authority or data subject may require foreign assistance to enforce it.

3. European Data Protection Board, Guidelines 3/2018 on the territorial scope of the GDPR (Article 3) Version 2.1, at 4 (Nov. 12, 2019), https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_3_2018_territorial_scope_after_public_consultation_en_1.pdf [hereinafter Territorial Scope Guidelines].

4. DAN JERKER B. SVANTESSON, INTERNET & JURISDICTION POLICY NETWORK, INTERNET & JURISDICTION GLOBAL STATUS REPORT 2019 59 (2019), https://www.internetjurisdiction.net/uploads/pdfs/GSR2019/Internet-Jurisdiction-Global-Status-Report-2019_web.pdf.

When the violator is a U.S.-based organization, one potential source of assistance is the U.S. court system. There is an established body of U.S. law concerning the recognition and enforcement by U.S. courts of foreign judgments in other contexts.

This *Commentary* addresses the application of that body of law to GDPR orders and judgments. It explores the options for a party in the EU—whether a supervisory authority, individual data subject, or a not-for-profit body acting on behalf of data subjects—to obtain a U.S.-based organization’s compliance through resort to a proceeding in a U.S. court.

Part I of the *Commentary* provides an overview of GDPR’s extraterritorial scope under GDPR Article 3 and briefly examines how EU supervisory authorities have interpreted that provision since GDPR entered into force in May 2018.

Part II addresses the state of the law in the U.S. regarding the recognition and enforcement of foreign country orders and judgments. As we explain, some states have addressed the issue by adopting statutes, and others have relied on the common law. Each approach, however, relies on a set of common principles. Part II describes those principles, touching on questions about enforcement of private money judgments and injunctions as well as public orders prohibiting or mandating certain conduct or levying fines or other penalties for violations of foreign laws.

Building on that discussion of general principles, Parts III, IV, and V address how those general principles apply to claims by private plaintiffs (Part III) and claims by EU supervisory authorities (Part IV), and the potential defenses they create for U.S. defendants (Part V).

Finally, Part VI briefly addresses the ways that GDPR’s requirements might be enforced other than through the direct enforcement of an existing EU order or judgment entered under

GDPR. These could include contract-based claims arising from GDPR-mandated data processing agreements, and claims brought against U.S. organizations by the U.S. Federal Trade Commission (FTC) using and individual data subjects under the EU-U.S. Privacy Shield and using its authority under Section 5 of the FTC Act.

I. AN OVERVIEW OF GDPR'S EXTRATERRITORIAL SCOPE

A. GDPR's Territorial Scope under Article 3

GDPR Article 3 defines GDPR's territorial scope according to two key criteria: the "establishment" criterion under Article 3.1 and the "targeting" criterion under Article 3.2.⁵

Under GDPR Article 3.1, GDPR applies to "the processing of personal data in the context of the activities of an establishment of a controller or a processor in the [EU], regardless of whether the processing takes place in the Union or not."⁶ Although GDPR does not specifically define "establishment" for this purpose, its recitals explain that the term implies "the effective and real exercise of activities through stable arrangements" in the EU.⁷ "The legal form of such arrangements, whether through a branch or a subsidiary with a legal personality, is not the determining factor in that respect."⁸

GDPR Article 3.2 extends the law to a controller or processor with no establishment in the EU, when the controller or processor processes the personal data of data subjects in the EU in connection with (a) the offering of goods or services to data subjects in the EU (irrespective of whether payment is required),⁹ or (b) the monitoring of those data subjects' behavior when they are in the EU.¹⁰ Both conditions imply the purposeful "targeting" of data subjects located within the EU by an organization outside

5. Territorial Scope Guidelines, *supra* note 3, at 4.

6. GDPR, *supra* note 1, art. 3.1.

7. *Id.*, Recital 22.

8. *Id.*

9. *Id.*, art. 3.2(a).

10. *Id.*, art. 3.2(b).

the EU, and focus on processing activities related to that targeting.¹¹

Since GDPR entered into force in May 2018, the European Data Protection Board (EDPB)—an independent European body composed of representatives of member state supervisory authorities established under GDPR Article 68¹²—has issued Guidelines that interpret Article 3.¹³ Those Guidelines confirm an organization outside the EU can trigger GDPR’s extraterritorial application without engaging in extensive or significant activities—physical or virtual—within the EU’s borders.

With respect to the “establishment” criterion under GDPR Article 3.1, the EDPB Guidelines explain that the threshold “can actually be quite low” and can be satisfied if a non-EU entity has “one single employee or agent” in the EU, “if that employee or agent acts with a sufficient degree of stability.”¹⁴ Put another way, “[t]he fact that the non-EU entity responsible for the data processing does not have a branch or subsidiary in a[n EU] Member State does not preclude it from having an establishment there within the meaning of EU data protection law.”¹⁵

The EDPB’s interpretation of the limits of the “targeting” criterion is similarly expansive. The Guidelines explain that the application of GDPR Article 3.2(a) depends on the controller or processor’s “intention to offer goods or services” to data subjects in the EU, which can be shown through factors such as “the mention of an international clientele composed of customers domiciled in various EU member states,” and offering delivery

11. Territorial Scope Guidelines, *supra* note 3, at 14.

12. GDPR, *supra* note 1, art. 68.1

13. Territorial Scope Guidelines, *supra* note 3.

14. *Id.* at 6.

15. *Id.* at 6–7.

of goods to EU member states.¹⁶ The Guidelines also explain that “monitoring” sufficient to trigger application of GDPR Article 3.2(b) can include activities commonly performed through commercial websites, including behavioral advertisements and “online tracking through the use of cookies.”¹⁷

Of particular note, the Guidelines also explain that a non-EU processor who would not otherwise fall within GDPR’s scope can become subject to GDPR under Article 3.2(b) when a non-EU controller for which the processor provides processing services engages in targeting activities.¹⁸ The Guidelines acknowledge that the decision to target individuals in the EU “can only be made by an entity acting as a controller.”¹⁹ They conclude, however, that a non-EU processor can fall within GDPR’s scope under Article 3.2(b) when its processing activities on the controller’s behalf are “related to carrying out the [controller’s] targeting,” even when those processing activities are limited to providing data storage to the controller.²⁰

When an organization falls within GDPR’s territorial scope under GDPR Article 3.2, GDPR Article 27 requires the organization to appoint a representative in the EU, subject to certain narrow exceptions.²¹ The representative must be mandated to receive—on behalf of the non-EU controller or processor—requests and inquiries from EU supervisory authorities and data subjects on all issues related to processing that falls within GDPR’s scope. In practical terms, this often means that the

16. *Id.* at 17.

17. *Id.* at 20.

18. *Id.* at 21.

19. *Id.*

20. *Id.*

21. GDPR, *supra* note 1, arts. 27.1, 27.2.

representative will pass those requests and inquiries on to the controller or processor to formulate a response that the representative will then pass to the inquirer. To be clear, the representative is not merely a receiver of legal process. In fact, GDPR provides that the representative “should be subject to enforcement proceedings in the event of non-compliance by the controller or processor.”²² Guidelines in turn explain that supervisory authorities can launch enforcement proceedings “through the representative” against the controller or processor, including by “address[ing] corrective measures or administrative fines and penalties imposed on the controller or processor . . . to the representative.”²³

The Guidelines also conclude that the representative’s direct liability under GDPR is “limited to its direct obligations referred to in articles [sic] 30 [record keeping] and article 58(1)a [responding to orders of a supervisory authority] of the GDPR.”²⁴ As the EDPB explains, the representative cannot itself be held directly liable for the controller or processor’s GDPR violations because “the GDPR does not establish a substitutive liability of the representative in place of the controller or processor it represents.”²⁵

B. Enforcement Activity Directed at Non-EU Organizations

To date, there have been two notable instances of GDPR enforcement activity directed toward non-EU controllers with no discernible physical presence or assets in the EU. They offer

22. *Id.*, Recital 80.

23. Territorial Scope Guidelines, *supra* note 3, at 28.

24. *Id.*

25. *Id.* at 27–28.

contrasting views on the limitations on the reach of EU supervisory authorities' enforcement power under those circumstances.

First, according to reporting by *The Register* in November 2018, a United Kingdom (UK) data subject made a complaint to the UK Information Commissioner's Office (ICO) regarding the cookie consent practices on the website of *The Washington Post*.²⁶ According to the complaint, the *Post*'s website impermissibly tied readers' consent to the use of cookies, tracking, and advertising to access to the website's content.²⁷ The ICO, according to the *The Register*'s reporting, agreed that the practice violated Article 7 of GDPR (which requires that consent be "freely given") and issued a written warning that directed the newspaper to change its practices.²⁸ The ICO concluded, however, that it had no ability to compel *The Washington Post*'s compliance with that direction, explaining in a statement to *The Register* that "[w]e hope that the *Washington Post* will heed our advice, but if they choose not to, there is nothing more we can do in relation to this matter."²⁹

Second, in July 2018, the ICO served an enforcement notice on a Canadian company called Aggregate IQ Data Services Ltd. ("AIQ"), which contracted with various UK political organizations to target political advertising messages to UK data subjects

26. Rebecca Hill, *Washington Post offers invalid cookie consent under EU Rules—ICO*, THE REGISTER (Nov. 19, 2018), https://www.theregister.co.uk/2018/11/19/ico_washington_post/.

27. *Id.*

28. *Id.*

29. *Id.*

on social media.³⁰ That enforcement notice claimed that AIQ was subject to GPDR under Article 3.2(b),³¹ and that the company's data collection and advertising activities violated various provisions of GDPR, including GDPR Articles 5, 6, and 14.³² The enforcement notice demanded that AIQ cease processing any personal data of UK or EU citizens for the purposes of data analytics, political campaigning, or any other advertising purposes.³³

As a report issued earlier by the ICO explained, however, AIQ initially contended that the company was "not subject to the jurisdiction of the ICO."³⁴ As a result, the ICO notified the Canadian government that AIQ refused to participate in the ICO's investigation, and Canadian privacy authorities subsequently announced investigations into the company's practices.³⁵

Ultimately, the ICO issued a new enforcement notice against AIQ in October 2018 that "varie[d] and replace[d]" the July 2018 notice.³⁶ Notably, that new notice said nothing about the ICO's

30. United Kingdom Information Commissioner's Office, Enforcement Notice to AggregateIQ Data Services Ltd, (July 6, 2018), <https://ico.org.uk/media/2259362/r-letter-ico-to-aiq-060718.pdf>.

31. *Id.* at ¶ 2.

32. *Id.* at ¶¶ 9–12.

33. *Id.* at ¶ 14; Annex 1.

34. United Kingdom Information Commissioner's Office, Investigation into the use of data analytics in political campaigns: investigation update (July 11, 2018), <https://ico.org.uk/media/action-weve-taken/2259371/investigation-into-data-analytics-for-political-purposes-update.pdf>, at 37.

35. *Id.*

36. United Kingdom Information Commissioner's Office, Enforcement Notice to AggregateIQ Data Services Ltd, (Oct. 24, 2018), <https://ico.org.uk/>

jurisdiction. The notice also imposed far narrower sanctions on AIQ: rather than a complete ban on the relevant processing, the company would simply have to erase the personal data of individuals in the UK that was maintained on the company's servers.³⁷

The contrast between *The Washington Post* and AIQ cases suggest that EU supervisory authorities' willingness to pursue enforcement actions against non-EU organizations may depend on various factors. Those may include the seriousness of the alleged violation, the willingness of a local regulator to cooperate in enforcement efforts, and the defendant's willingness to engage with EU and local authorities.

media/action-weve-taken/enforcement-notice/2260123/aggregate-iq-en-20181024.pdf.

37. *Id.* at ¶ 14; Annex 1.

II. RECOGNITION AND ENFORCEMENT OF FOREIGN JUDGMENTS IN U.S. COURTS: OVERVIEW OF CURRENT LAW

This part of the *Commentary* summarizes the general principles under existing U.S. law that govern the recognition and enforcement of foreign country orders and judgments. It is not intended to be a comprehensive primer on the law in this area. Rather, its purpose is to identify and summarize those principles that are most relevant to the enforceability of a judgment or order entered by a court or other enforcement authority.

A. *Origins of the law of recognition and enforcement of foreign judgments*

The question of recognition and enforcement of foreign judgments and orders arises from the foundational principle that under U.S. law, any judgment from a country or U.S. state outside a given forum is considered “foreign” and cannot be directly enforced in that forum without a basis to “recognize” the judgment domestically.³⁸ The Full Faith and Credit Clause in Article IV of the Constitution provides that basis for judgments rendered in any other court—state or federal—in the United States.³⁹

The Full Faith and Credit Clause does not apply, however, to judgments rendered by courts in foreign countries. Nor is there any U.S. federal statute or treaty dealing generally with foreign country judgment recognition. Instead, recognition of foreign country judgments is primarily a matter of state law,

38. Yuliya Zeynalova, *The Law on Recognition and Enforcement of Foreign Judgments: Is It Broken and How Do We Fix It?*, 31 BERKELEY J. INT’L L. 150, 154 (2013).

39. See U.S. CONST. art. IV § 1.

and its historical roots can be traced back to the U.S. Supreme Court's 1895 decision in *Hilton v. Guyot*.⁴⁰

In *Hilton*, the U.S. Supreme Court concluded that absent a treaty, U.S. courts asked to recognize a foreign judgment should turn to the principle of comity, which the court explained is "neither a matter of absolute obligation . . . nor a mere courtesy and good will," but rather "the recognition which one nation allows within its territory to the legislative, executive or judicial acts of another nation, having due regard both to international duty and convenience, and to the rights of its own citizens or of other persons who are under the protection of its laws."⁴¹ After reviewing the leading authorities on the subject at the time, the *Hilton* court set forth the following considerations that would justify recognizing the judgment of a foreign court:

[W]here there has been opportunity for a full and fair trial abroad before a court of competent jurisdiction, conducting the trial upon regular proceedings, after due citation or voluntary appearance of the defendant, and under a system of jurisprudence likely to secure an impartial administration of justice between the citizens of its own country and those of other countries, and there is nothing to show either prejudice in the court, or in the system of laws under which it was sitting, or fraud in procuring the judgment, or any other special reason why the comity of this nation should not allow it full effect, the merits of the case should

40. Ronald A. Brand, *Federal Judicial Center International Litigation Guide: Recognition and Enforcement of Foreign Judgments*, 74 U. PITT. L. REV. 491, 496 (2013) (citing *Hilton v. Guyot*, 159 U.S. 113 (1895)) [hereinafter Brand, *FJC Guide*].

41. *Hilton*, 159 U.S. at 163–64.

not, in an action brought in this country upon the judgment, be tried afresh.⁴²

Using *Hilton* as a “conceptual backdrop,” U.S. states generally follow one of two approaches to recognizing foreign country judgments: (1) recognition at common law as a matter of comity; or (2) recognition under state statutes that are based on one of two model acts promulgated by the Uniform Law Commission.⁴³

Courts in a minority of U.S. states—sixteen—follow the first approach.⁴⁴ They generally rely on *Hilton*, the Restatement (Third) of Foreign Relations Law⁴⁵ (recently succeeded by the Restatement (Fourth) of Foreign Relations Law⁴⁶), and the Restatement (Second) of Conflict of Laws.⁴⁷

The other thirty-four U.S. states and the District of Columbia have adopted one of two model recognition acts:⁴⁸ (1) the 1962 Uniform Foreign Money Judgments Recognition Act (the “1962

42. *Id.* at 123.

43. Tanya J. Monestier, *Whose Law of Personal Jurisdiction? The Choice of Law Problem in the Recognition of Foreign Judgments*, 96 B.U. L. REV. 1729, 1736 (2016).

44. Ronald A. Brand, *The Continuing Evolution of U.S. Judgments Recognition Law*, 55 COLUM. J. TRANSNAT'L L. 277, 295 (2017) [hereinafter Brand, *The Continuing Evolution*].

45. RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW (Am. Law Inst. 1987) [hereinafter RESTATEMENT (THIRD)].

46. RESTATEMENT (FOURTH) OF FOREIGN RELATIONS LAW (Am. Law Inst. 2018) [hereinafter RESTATEMENT (FOURTH)].

47. RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 98 (Am. Law Inst. 1971) [hereinafter RESTATEMENT (SECOND)].

48. Brand, *The Continuing Evolution*, *supra* note 44, at 295.

Recognition Act”),⁴⁹ or (2) the 2005 Uniform Foreign-Country Money Judgments Recognition Act (the “2005 Recognition Act”)⁵⁰ (collectively, the “Recognition Acts”).

While U.S. law regarding foreign judgment recognition may thus seem to be a disparate patchwork,⁵¹ the common law and both Recognition Acts are largely consistent as to both the foundational requirements to recognize a foreign judgment and the primary grounds for nonrecognition.

B. Foundational requirements for recognition and enforcement of foreign judgments

Under the common law and both Recognition Acts, to be recognizable by a U.S. court a foreign judgment must be (1) final, (2) conclusive, and (3) enforceable in the rendering country.⁵² A judgment is “final” for this purpose when it “is not subject to additional proceedings in the rendering court other than execution.”⁵³ Both contested and default judgments can meet these criteria.⁵⁴ While being subject to an appeal “does not deprive it

49. Uniform Foreign Money-Judgments Recognition Act (Unif. Law Comm’n 1962) [hereinafter 1962 Recognition Act].

50. Uniform Foreign-Country Money Judgments Recognition Act (Unif. Law Comm’n 2005) [hereinafter 2005 Recognition Act].

51. Monestier, *supra* note 43, at 1735.

52. 1962 Recognition Act, *supra* note 49, § 2; 2005 Recognition Act, *supra* note 50, § 3(a)(2); RESTATEMENT (THIRD), *supra* note 45, § 481; RESTATEMENT (FOURTH), *supra* note 46, § 481.

53. RESTATEMENT (THIRD), *supra* note 45, § 481 cmt. e. *See also* RESTATEMENT (FOURTH), *supra* note 46, § 481 cmt. d.

54. *See* Brand, *FJC Guide*, *supra* note 40, at 524 (explaining that “any decision on the merits that could have been litigated in the originating court will have preclusive effect in the recognizing court,” but noting that “this does not prevent challenges based on lack of personal jurisdiction or lack of

of its character as a final judgment,”⁵⁵ a U.S. court may—but need not—stay the recognition of a foreign judgment until the appeal has run its course in the rendering country.⁵⁶

Notably, the 1962 Recognition Act and the 2005 Recognition Act are limited by their own terms to judgments that grant or deny recovery of a sum of money.⁵⁷ The common-law approach, however, also allows for a U.S. court to *recognize* foreign judgments that grant injunctions, declare parties’ rights, or determine parties’ legal status.⁵⁸

Whether and under what circumstances a U.S. court will *enforce* these nonmonetary judgments, however, is less clear. The Restatement (Third) of Foreign Relations Law and the Restatement (Fourth) of Foreign Relations Law suggest that U.S. courts are *not* required to enforce these judgments by granting the

proper notice in the originating court, or other grounds for non-recognition otherwise available under the applicable statute or common law”).

55. RESTATEMENT (THIRD), *supra* note 45, § 481 cmt. e.

56. 1962 Recognition Act, *supra* note 49, § 6; 2005 Recognition Act, *supra* note 50, § 8; RESTATEMENT (THIRD), *supra* note 45, § 481 cmt. e.; RESTATEMENT (FOURTH), *supra* note 46, § 481 cmt. e.

57. 1962 Recognition Act, *supra* note 49, §§ 1(2), 3; 2005 Recognition Act, *supra* note 50, § 3(a)(1).

58. RESTATEMENT (THIRD), *supra* note 45, § 481 cmt. b (“Judgments granting injunctions, declaring rights or determining status . . . may be entitled to recognition under this and the following sections.”); RESTATEMENT (FOURTH), *supra* note 46, § 488 (“[A] final and conclusive judgment of a court in a foreign state in an action seeking an injunction or a comparable nonmonetary remedy is entitled to recognition by courts in the United States.”); RESTATEMENT (SECOND), *supra* note 47, § 102 cmt. g (“A valid decree rendered in a foreign nation that orders or enjoins the doing of an act will usually be recognized in the United States.”).

relief ordered by the rendering court.⁵⁹ The Restatement (Second) of Conflict of Laws, by contrast, concludes that foreign injunctive decrees *can* be enforced as long as such enforcement is “necessary to effectuate the [foreign court’s] decree and will not impose an undue burden upon the American court and provided further that in the view of the American court the decree is consistent with fundamental principles of justice and of good morals.”⁶⁰ At least two federal courts have relied on that statement to conclude that they could enforce injunctions entered by foreign courts under the principle of comity.⁶¹

C. The rule against recognition of foreign fines and penal judgments

The general rule in favor of recognizing foreign country judgments that meet the foundational requirements above is subject to a key exception: under both the Recognition Acts and the common law, U.S. courts generally do not recognize or enforce foreign judgments for the collection of taxes, fines, or penalties.⁶²

A judgment is “penal” under this rule when it is “in favor of a foreign state or one of its subdivisions, and primarily punitive

59. RESTATEMENT (THIRD), *supra* note 45, § 481 cmt. b (“Judgments granting injunctions, declaring rights or determining status . . . are not generally entitled to enforcement.”); RESTATEMENT (FOURTH), *supra* note 46, § 488 (“[T]he question of what remedies to grant as a result of recognition of the foreign judgment, including whether to provide injunctive relief, does not depend on the remedies provided by the rendering court.”).

60. RESTATEMENT (SECOND), *supra* note 47, § 102 cmt. g.

61. See *Siko Ventures Ltd. v. Argyll Equities, LLC*, No. SA-05-CA-100-OG, 2005 WL 2233205, at *3 (W.D. Tex. Aug. 5, 2005); *Pilkington Bros. P.L.C. v. AFG Indus. Inc.*, 581 F. Supp. 1039, 1043 (D. Del. 1984).

62. See RESTATEMENT (THIRD), *supra* note 45, § 483; RESTATEMENT (FOURTH), *supra* note 46, § 489; 1962 Recognition Act, *supra* note 49, § 1(2); 2005 Recognition Act, *supra* note 50, § 3(b).

rather than compensatory in character.”⁶³ The rule against recognizing such judgments reflects “a reluctance of courts to subject foreign public law to judicial scrutiny . . . combined with a reluctance to enforce law that may conflict with the public policy of the forum state.”⁶⁴

The Recognition Acts both expressly exclude foreign fines and penal judgments from their provisions for recognition.⁶⁵ The 2005 Recognition Act, however, includes a savings clause that leaves room to recognize these judgments on other grounds, such as comity under the common-law approach.⁶⁶

Under the Restatement (Third) of Foreign Relations Law, the common-law rule against recognizing fines and penal judgments is phrased as being permissive, rather than mandatory.⁶⁷ As a comment explains, nonrecognition is permitted on this basis, but not required, as “no rule of United States law or of international law would be violated if a court in the United States enforced a judgment of a foreign court for payment of taxes or

63. RESTATEMENT (THIRD), *supra* note 45, § 483 cmt. b. *See also* RESTATEMENT (FOURTH), *supra* note 46, § 489 cmt. b.

64. RESTATEMENT (THIRD), *supra* note 45, § 483 reporter’s note 2.

65. 1962 Recognition Act, *supra* note 49, § 1(2) (defining “foreign judgment” that is subject to recognition as excluding “a judgment for taxes, a fine, or other penalty”); 2005 Recognition Act, *supra* note 50, § 3(b) (providing that the act does not apply “to the extent that the judgment is . . . a fine or other penalty”).

66. *Id.* § 11 (“This act does not prevent the recognition under principles of comity or otherwise of a foreign-country judgment not within the scope of this act.”).

67. RESTATEMENT (THIRD), *supra* note 45, § 483 (“Courts in the United States are not required to enforce [penal judgments].”).

comparable assessments that was otherwise consistent” with the standards for recognition.⁶⁸

The Restatement (Fourth) of Foreign Relations Law, by contrast, simply states that courts “do not” recognize or enforce foreign judgments “to the extent such judgments are for taxes, fines, or other penalties, unless authorized by a statute or an international agreement.”⁶⁹

D. Other grounds for nonrecognition of foreign judgments

Assuming a foreign judgment meets the foundational requirements above and is not subject to nonrecognition as a fine or penalty, both the common-law approach and the Recognition Acts provide several other grounds for nonrecognition.

Some of these grounds are mandatory. A U.S. court cannot enforce a foreign judgment, for example, if the rendering court lacked personal jurisdiction over the defendant.⁷⁰ There is some question as to whose law governs the U.S. court’s determination of that issue: the law of the rendering country, the law of the U.S. forum, or some combination thereof.⁷¹ Setting aside that choice-of-law issue, however, both the common-law approach and the Recognition Acts provide several criteria that can preclude a U.S. court from refusing to recognize a foreign judgment for lack of personal jurisdiction over the defendant.⁷² These criteria identify activities by a defendant that make an assertion of

68. *Id.* § 483 cmt. a.

69. RESTATEMENT (FOURTH), *supra* note 46, § 489.

70. *See* RESTATEMENT (THIRD), *supra* note 45, § 482(1)(b); RESTATEMENT (FOURTH), *supra* note 46, § 483(b); 1962 Recognition Act, *supra* note 49, § 4(a)(2); 2005 Recognition Act, *supra* note 50, § 4(b)(2).

71. *See* Monestier, *supra* note 43, at 1739–44.

72. *See* RESTATEMENT (THIRD), *supra* note 45, §§ 482(1)(b), 421(2); 1962 Recognition Act, *supra* note 49, § 5; 2005 Recognition Act, *supra* note 50, § 5.

personal jurisdiction by the rendering court presumptively reasonable.⁷³

The common-law approach and the Recognition Acts also provide several discretionary grounds for nonrecognition, meaning the U.S. court may—but need not—treat them as precluding recognition of a foreign judgment.⁷⁴ Of particular relevance here, a U.S. court may decline to recognize a foreign country judgment if the judgment is “repugnant to the public policy” of the United States or of the U.S. state in which recognition is sought.⁷⁵

E. Recognition of foreign administrative orders

The Recognition Acts apply by their own terms to “judgments,” and thus cannot be used to recognize foreign administrative acts that have not been the subject of a final, conclusive, and enforceable judgment between the defendant and the party seeking recognition. As a result, in the absence of a treaty, the only basis for recognizing a foreign administrative act that has not been reduced to a “judgment” in a U.S. court is the common law.⁷⁶

As the Restatement (Third) of Foreign Relations Law and the Restatement (Fourth) of Foreign Relations Law acknowledge, however, the common law is unclear as to whether foreign

73. See Part V.A, *infra*.

74. See RESTATEMENT (THIRD), *supra* note 45, § 482(2); RESTATEMENT (FOURTH), *supra* note 46, § 484; 1962 Recognition Act, *supra* note 49, § 4(b); 2005 Recognition Act, *supra* note 50, § 4(c).

75. RESTATEMENT (THIRD), *supra* note 45, § 482(2)(d); RESTATEMENT (FOURTH), *supra* note 46, § 484(c); 1962 Recognition Act, *supra* note 49, § 4(b)(3); 2005 Recognition Act, *supra* note 50, § 4(c)(3).

76. John C. Reitz, *Recognition of Foreign Administrative Acts*, 62 AM. J. COMP. L. 589, 602 (Supp. 2014).

administrative acts can be recognized in a U.S. court.⁷⁷ The reporter's notes to the Restatement (Fourth) explain that "[a] handful of State-court decisions have indicated that a final, conclusive and enforceable administrative determination can be eligible for recognition if the administrative body employed proceedings generally consistent with due process, at least if the person opposing recognition had an opportunity to obtain judicial review."⁷⁸

The Restatement (Third) of Foreign Relations Law, however, confirms that the rule against recognizing foreign penal judgments applies equally to foreign administrative orders that impose fines or penalties, explaining that "[a]ctions may be penal in character . . . even if they do not result from judicial process, for example when a government agency is authorized to impose fines or penalties for violation of its regulations."⁷⁹

F. Procedural considerations and burdens of proof

Under both the common law and the 2005 Recognition Act, the procedure for seeking recognition of a foreign country

77. RESTATEMENT (THIRD), *supra* note 45, § 481 cmt. f ("The rule [in favor of recognizing foreign court judgments] is less clear with regard to decisions of administrative tribunals, industrial compensation boards, and similar bodies."); RESTATEMENT (FOURTH), *supra* note 46, § 481 cmt. f (explaining that the rule's application to the decisions of administrative tribunals is "less clear").

78. *Id.* § 481 Reporter's Note 6 (citing *Alberta Sec. Comm'n v. Ryckman*, 30 P.3d 121, 126–27 (Ariz. Ct. App. 2001) and *Regierungspraesident Land Nordrhein-Westfalen v. Rosenthal*, 232 N.Y.S.2d 963 (N.Y. 1st Dep't 1962)); *see also* *Petition of Breau*, 565 A.2d 1044, 1050 (N.H. 1989) (recognizing determination of Canadian administrative body regarding teacher's lack of good moral character by giving preclusive effect to body's findings in New Hampshire credential revocation proceedings).

79. RESTATEMENT (THIRD), *supra* note 45, § 483 cmt. b.

judgment is to initiate a civil action in a U.S. court.⁸⁰ A party to an already existing proceeding in a U.S. court can also seek recognition by raising the issue in that proceeding, for instance through a counterclaim or cross-claim, or as an affirmative defense.⁸¹

Once the issue is before the U.S. court, the party seeking recognition bears the initial burden of establishing that the foreign judgment meets the foundational requirements for recognition under the common law and the Recognition Acts: the judgment is final, conclusive, and enforceable in the rendering jurisdiction, and is not a judgment for taxes, fines, or penalties.⁸²

Once a party seeking recognition makes that showing, the burden shifts to the party resisting recognition to establish that the foreign judgment is subject to one or more of the mandatory or discretionary grounds for nonrecognition, such as lack of personal jurisdiction in the rendering forum or that the judgment is repugnant to U.S. public policy.⁸³

80. RESTATEMENT (FOURTH), *supra* note 46, § 482; 2005 Recognition Act, *supra* note 50, § 6(a).

81. RESTATEMENT (FOURTH), *supra* note 46, § 482; 2005 Recognition Act, *supra* note 50, § 6(b).

82. *See* RESTATEMENT (FOURTH), *supra* note 46, § 485(1); 2005 Recognition Act, *supra* note 50, § 3(c). While the 1962 Recognition Act does not contain any specific provisions on the burden of proof, courts deciding cases under that Act also typically place the initial burden of establishing that a judgment is within the Act's scope on the party seeking recognition. *See* Brand, *FJC Guide*, *supra* note 40, at 524 (citing *Bridgeway Corp. v. Citibank*, 45 F. Supp. 2d 276, 285 (S.D.N.Y. 1999); *S.C. Chimexim S.A. v. Velco Enters. Ltd.*, 36 F. Supp. 2d 206, 212 (S.D.N.Y. 1999)).

83. *See* RESTATEMENT (FOURTH), *supra* note 46, § 485(3); 2005 Recognition Act, *supra* note 50, § 4(d).

III. RECOGNITION AND ENFORCEMENT OF GDPR ORDERS AND JUDGMENTS IN U.S. COURTS: PRIVATE ACTIONS BY DATA SUBJECTS AND REPRESENTATIVE ORGANIZATIONS

This part of the *Commentary* explores the different kinds of GDPR orders and judgments that a private plaintiff—whether an individual EU data subject or a representative organization acting on behalf of a group of EU data subjects—might seek to enforce through a U.S. court and how U.S. law would apply to those orders and judgments.

A. *General considerations for private causes of action*

If a U.S.-based data controller or data processor lacks a physical presence, assets, or other financial ties to the EU and is unwilling to comply voluntarily with a judgment or order issued by an EU court or supervisory authority, an aggrieved EU plaintiff could file a civil action in a U.S. court seeking recognition and enforcement of that judgment or order within the United States. To succeed, that plaintiff will first need to clear the jurisdictional hurdles that confront all would-be litigants in the U.S. court system. First, the plaintiff will need to identify and commence the action in a forum in which the defendant is subject to personal jurisdiction.⁸⁴ While a detailed discussion of personal jurisdiction is beyond the scope of this *Commentary*, in general, personal jurisdiction in both federal and state courts will be governed by the law on personal jurisdiction that is in force in the

84. See, e.g., RESTATEMENT (FOURTH), *supra* note 46, § 482 Reporter's Note 3 ("A court entertaining a separate action to obtain recognition of a foreign judgment must obtain jurisdiction over every person on whom its decision will have conclusive effect.").

state where the court is located,⁸⁵ and by the Due Process Clause of the U.S. Constitution.⁸⁶

Second, the plaintiff will need to establish that the court has subject-matter jurisdiction over the action. As with personal jurisdiction, a detailed discussion of subject-matter jurisdiction is beyond the scope of this *Commentary*. But one important threshold requirement to establish subject-matter jurisdiction is standing to sue.

In federal court, Article III of the U.S. Constitution requires that a plaintiff establish standing to sue by demonstrating that she “(1) suffered an injury in fact, (2) that is fairly traceable to the challenged conduct of the defendant, and (3) that is likely to be redressed by a favorable judicial decision.”⁸⁷ It appears no federal court has squarely addressed the question whether a party seeking to enforce a foreign judgment has standing to do so. It is nonetheless highly likely that a party seeking to do so would be able to establish standing when: (1) the judgment awards money damages to the plaintiff; and (2) the defendant is the party against whom the foreign judgment was issued. Under these circumstances, the plaintiff can convincingly argue that she has suffered an injury in fact, insofar as she was awarded a money judgment that has not been satisfied, and the defendant’s failure to satisfy that judgment would be “fairly traceable” to that defendant.⁸⁸ Finally, the party seeking

85. See, e.g., FED. R. CIV. P. 4(k)(1)(a).

86. See, e.g., *Daimler AG v. Bauman*, 571 U.S. 117 (2014); *Goodyear Dunlop Tires Operations, S.A. v. Brown*, 564 U.S. 915 (2011).

87. *Spokeo, Inc. v. Robins*, 136 S. Ct. 1540, 1547 (2016).

88. Cf. *ACLI Gov’t Secs., Inc. v. Rhoades*, 653 F. Supp. 1388, 1390 (S.D.N.Y. 1987), *aff’d sub nom. ACLI Gov. v. Rhoades*, 842 F.2d 1287 (2d Cir. 1988) (providing that owner of the judgment against defendant had standing in action to pursue collection).

damages could likely also show that recognition and enforcement of the judgment by the federal court would redress the injury caused by the defendant's failure to satisfy it.

Although not governed by Article III, a substantial majority of U.S. state courts apply analogous standing requirements.⁸⁹ To that end, many of these courts also require that a plaintiff show she has suffered an injury that is attributable to the defendant's conduct.⁹⁰ As in federal court, a plaintiff's possession of a judgment issued in her favor by an EU court against the defendant should be sufficient to satisfy these state court standing requirements

The standing analysis can be more complicated, however, in cases that involve judgments obtained by representative bodies on individual data subjects' behalf. GDPR Article 80 expressly allows for one or more data subjects to be represented in a private GDPR enforcement action in EU courts by "a not-for-profit body, organisation or association."⁹¹ The organization must have been "properly constituted in accordance with the law of a Member State, ha[ve] statutory objectives which are in the public interest, and [be] active in the field of the protection of data subjects' rights and freedoms with regard to the protection of their personal data."⁹²

Such a body, organization, or association can either be requested by a data subject to lodge a complaint and obtain compensation under Article 82 on that individual's behalf,⁹³ or may

89. See generally Wyatt Sassman, *A Survey of Constitutional Standing in State Courts*, 8 KY. J. EQUINE, AGRIC. & NAT. RESOURCES L. 349 (2016).

90. *Id.*

91. GDPR, *supra* note 1, art. 80.1.

92. *Id.*

93. See *id.*

act independently on the behalf of individual or multiple data subjects to submit matters to a supervisory authority under Article 77, or to a court under Articles 78 and 79, as provided by the law of their local Member State.⁹⁴

If an organization that has obtained a judgment on behalf of data subjects in the EU seeks to obtain recognition and enforcement of that judgment in a U.S. court, its claims could be analyzed under the doctrine of “representational standing.” To that end, the United States has long recognized that groups or organizations can maintain actions on behalf of their members in federal court when certain conditions are met. In *Hunt v. Washington State Apple Advertising Commission*,⁹⁵ the U.S. Supreme Court held that “an association has standing to bring suit on behalf of its members when: (a) its members would otherwise have standing to sue in their own right; (b) the interests it seeks to protect are germane to the organization’s purpose; and (c) neither the claim asserted nor the relief requested requires the participation of individual members in the lawsuit.” When a foreign organization seeks to maintain representational standing, U.S. courts often make an additional inquiry into the law of the organization’s place of incorporation to determine whether the organization is permitted to pursue claims on behalf of its members.⁹⁶ Significantly, when an organization satisfies all of these requirements, the organization itself does not have to suffer an injury to maintain standing; it merely has to show that its members have suffered an injury.

94. *Id.*, art. 80.2.

95. 432 U.S. 333, 343 (1977).

96. *Cf.* *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014) (associations authorized by foreign law to administer their foreign members’ copyrights had standing to bring action); *Matter of Oil Spill by Amoco Cadiz Off Coast of France* on March 16, 1978, 954 F.2d 1279, 1319–20 (7th Cir. 1992).

Given that the GDPR requires that the representative organization be “in the public interest” and “active in the field of the protection of data subjects’ rights,” and assuming an EU Court or Supervisory Authority has already found an organization to satisfy those requirements, that organization could convincingly argue that it meets the requirements for representational standing under *Hunt*.

B. Data subject compensation claims under GDPR Article 82

Under GDPR Article 82, individuals can receive compensation for damages suffered as a result of a controller’s or processor’s GDPR violation.⁹⁷ This part provides an overview of this aspect of GDPR and evaluates the enforceability in U.S. courts of money judgments issued by EU courts in favor of data subjects, or not-for-profit bodies who bring suit on their behalf, under GDPR Article 82.

1. Overview and general considerations

Prior to GDPR’s implementation, claims for damages by EU data subjects for privacy breaches were limited to claims against data controllers and did not apply universally across all EU Member States. This right was not widely exercised. GDPR Article 82 expanded the rights of individuals to seek compensation directly from both data controllers and data processors for “any material or non-material damage as a result of an infringement”⁹⁸ of GDPR, thereby increasing the scope of compensatory claims and the parties against whom they can be brought.

97. GDPR, *supra* note 1, art. 82.1.

98. U.S. readers should be mindful that “material and immaterial” may not mean the same thing to those in the U.S. that they do to those in the EU. Perhaps a better way for a U.S. reader to consider these terms is “tangible”

Under GDPR, individuals or not-for-profit entities are permitted to file a direct legal claim for compensation in the courts of the Member State where the controller or processor is established or in the courts where the data subject(s) maintain a “habitual residence.”⁹⁹ Claims for compensation need not be preceded by a determination of fault by a supervisory authority, or any other administrative or nonjudicial finding or remedy.

GDPR provides that “[d]ata subjects should receive full and effective compensation for the damage they have suffered.”¹⁰⁰ Compensation may be recovered for both pecuniary and non-pecuniary losses that might include, but are not limited to, claims for distress, anxiety, or reputational damage.¹⁰¹ GDPR does not impose any caps or limits on the amount of damages recoverable by a data subject harmed by a controller’s or processor’s violation.

As discussed below, an EU party that is able to present a U.S. court with a compensatory monetary judgment issued by an EU court of competent jurisdiction does have a reasonable probability of securing recognition and enforcement of that order in the United States.

and “intangible.” An immaterial injury, like an intangible one, can be substantial.

99. GDPR, *supra* note 1, art. 79.2. If the controller or processor is a public authority of a Member State exercising its public powers, an action must be brought in that Member State. *Id.*

100. *Id.*, Recital 146 (“The concept of damage should be broadly interpreted in the light of the case-law of the Court of Justice in a manner which fully reflects the objectives of this Regulation.”).

101. European Commission, Can my company/my organisation be liable for damages?, *available at* https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/enforcement-and-sanctions/sanctions/can-my-company-my-organisation-be-liable-damages_en (last visited Dec. 15, 2020).

2. Enforceability under U.S. law

Of the various types of orders and judgments that can be issued under GDPR, EU-based plaintiffs are most likely to be able to establish a *prima facie* case in U.S. courts for recognition of money judgments obtained through EU court proceedings under GDPR Article 82.

First, assuming they are final and conclusive between the parties, these judgments should qualify as judgments that grant recovery of a sum of money and therefore fall comfortably within the scope of both Recognition Acts and the common-law approach.¹⁰² Examples abound of U.S. courts recognizing and enforcing foreign judgments from EU Member States by applying analyses that would likely be applied to Article 82 recognition and enforcement actions.¹⁰³

Second, these judgments are unlikely to violate the rule against enforcing “penal” judgments because their primary purpose is to compensate data subjects—rather than punish the U.S.-based defendant—and they do not serve to benefit public authorities.¹⁰⁴

102. See Parts II.A & I.B, *supra*.

103. See, e.g., *de Fontbrune v. Wofsy*, 838 F.3d 992, 1005 (9th Cir. 2016) (finding that a French judgment awarding damages under the French concept of *astreinte* could be recognized under Californian law because it could “be seen as fulfilling a function akin to statutory damages in American copyright law”); *Societe d’Aménagement et de Gestion de l’Abri Nautique v. Marine Travelift Inc.*, 324 F. Supp. 3d 1004, 1005 (E.D. Wis. 2018) (recognizing French products liability judgment); *ABC Arbitrage S.A. v. Caen*, No. CV 16-07014 SJO (Ex), 2017 WL 7803784, at *3 (C.D. Cal. Feb. 28, 2017) (finding compensatory damages for fraud and breach of contractual monetary awards enforceable).

104. See RESTATEMENT (THIRD), *supra* note 45, § 483 cmt. b; see also *de Fontbrune*, 838 F.3d at 1005 (“[T]he purpose of the award was not to punish a harm against the public, but to vindicate [the judgment creditor’s] personal

C. *Injunctions and nonmonetary orders issued under GDPR Article 79*

In addition to compensation claims that would require a U.S. defendant to pay damages to EU data subjects, an EU-based plaintiff might also seek and obtain an injunction, or an order for specific performance, against a U.S.-based defendant under GDPR Article 79. This part of the *Commentary* discusses these types of orders and evaluates their enforceability in U.S. courts.

1. Overview and general considerations

GDPR Article 79 guarantees each EU data subject the nonexclusive right to “an effective judicial remedy where he or she considers that his or her [GDPR] rights under have been infringed as a result of the processing of his or her personal data in non-compliance with [GDPR].”¹⁰⁵

While GDPR Article 82 provides for compensatory damages to data subjects for noncompliance, monetary payments may not, by themselves, provide a sufficient judicial remedy. In such cases, an EU court can issue injunctive orders to prevent ongoing violations, or orders for specific relief or performance that require a data controller or data processor to either take or cease taking specific actions.

interest in having his copyright respected and to deter further future infringements by [the judgment debtor.]”); *Plata v. Darbun Enters., Inc.*, Case No. D062517, 2014 WL 341667, at *5 (Cal. Ct. App. Jan. 31, 2014) (“[T]he issue whether a monetary award is a penalty within the meaning of the [Recognition Act] requires a court to focus on the legislative purpose of the law underlying the foreign judgment. A judgment is a penalty even if it awards monetary damages to a private individual if the judgment seeks to redress a public wrong and vindicate the public justice, as opposed to affording a private remedy to a person injured by the wrongful act.”).

105. GDPR, *supra* note 1, art. 79.1.

2. Enforceability under U.S. law

As noted in Part II.B, the Recognition Acts apply only to judgments that grant or deny recovery of a sum of money. Even under the relatively permissive view of the Restatement (Second) of Conflict of Laws, enforcing injunctions dealing with the processing of personal data might arguably run afoul of its mandate that to be enforced, an injunction must “not impose an undue burden upon the American court.”¹⁰⁶

Thus, while a private plaintiff may be able to make a prima facie case for recognition of a foreign judgment imposing an injunction on a U.S. defendant, or ordering specific performance, the circumstances under which a U.S. court could actually provide that relief are limited.

106. RESTATEMENT (SECOND), *supra* note 47, § 102 cmt. g.

IV. RECOGNITION AND ENFORCEMENT OF GDPR ORDERS AND JUDGMENTS IN U.S. COURTS: CORRECTIVE ORDERS ENTERED BY EU SUPERVISORY AUTHORITIES

This part of the *Commentary* discusses the types of corrective orders that an EU supervisory authority might seek to enforce against a U.S. defendant through U.S. courts, and how U.S. law would apply to those orders.

A. *Overview and general considerations*

GDPR grants supervisory authorities broad authority to exercise “corrective powers” for violations of GDPR’s requirements. Specifically, GDPR Article 58.2(c)-(j) enumerates “corrective powers”:

- (c) to order the controller or the processor to comply with the data subject’s requests to exercise his or her rights pursuant to this Regulation;
- (d) to order the controller or processor to bring processing operations into compliance with the provisions of this Regulation, where appropriate, in a specified manner and within a specified period;
- (e) to order the controller to communicate a personal data breach to the data subject;
- (f) to impose a temporary or definitive limitation including a ban on processing;
- (g) to order a rectification or erasure of personal data or restriction of processing pursuant to Articles 16, 17 and 18 and the notification of such actions to recipients to whom the personal data have been disclosed pursuant to Article 17.2 and Article 19;

- (h) to withdraw a certification or to order the certification body to withdraw a certification issued pursuant to Articles 42 and 43, or to order the certification body not to issue certification if the requirements for the certification are not or are no longer met;
- (i) to impose an administrative fine pursuant to Article 83, in addition to, or instead of measures referred to in this paragraph, depending on the circumstances of each individual case; [and]
- (j) to order the suspension of data flows to a recipient in a third country or to an international organisation.¹⁰⁷

These corrective powers are discretionary in nature and consist both of affirmative (clauses c-e, i) and prohibitive actions (clauses f-h, j). The former require affirmative acts of compliance by controllers or processors, while the latter impose restrictions on their activities. These powers are not plenary, but rather are expressly subject to “appropriate safeguards, including effective judicial remedy and due process.”¹⁰⁸ Further, GDPR Article 78 provides “each natural or legal person” with “the right to an effective judicial remedy against a legally binding decision of a supervisory authority concerning them.”¹⁰⁹

107. GDPR, *supra* note 1, art. 58.2.

108. *Id.*, art. 58.3.

109. *Id.*, art. 78.1.

B. Nonmonetary orders issued under Article 58: enforceability under U.S. law

Would or could a U.S. court enforce a supervisory authority's nonmonetary order under GDPR Article 58.2? There is currently little, if any, basis for U.S. judicial enforcement of these types of orders, for at least three reasons.

First, to the extent a supervisory authority's nonmonetary order has not been reduced to a final judgment through proceedings in a court of competent jurisdiction in the EU, there is very little precedent for the recognition of that order in a U.S. court. As noted in Part II.B, the Recognition Acts are generally limited to recognizing "judgments" that are final, conclusive, and enforceable in the rendering jurisdiction. And as discussed in Part II.E, there is very little precedent under the common law for recognizing administrative orders that have not been reduced to judgments.

Second, as also noted in Part II.B, the Recognition Acts apply only to judgments that grant or deny recovery of a sum of money. Nonmonetary orders issued under GDPR Article 58.2 therefore cannot be recognized or enforced under the Recognition Acts. And while the common law may allow for these orders to be *recognized*—given legal effect for purposes such as *res judicata* or collateral estoppel—there is little authority for invoking the authority of a U.S. court to lend its power to *enforcing* them against a U.S. defendant, especially when the order has not been reduced to a judgment in an EU court.¹¹⁰ Even under the relatively permissive view of the Restatement (Second) of Conflict of Laws regarding the enforcement of foreign injunctions, some of the corrective powers—including, for example, an order to "bring processing operations into compliance" with

110. See Part II.B, *supra*.

GDPR,¹¹¹ or imposing a “ban on processing”¹¹²— would seem to require a level of involvement by a U.S. court that would run afoul of its mandate that to be enforced, an injunction must “not impose an undue burden upon the American court.”¹¹³

Third, an order issued by a supervisory authority using its corrective powers could run afoul of the rule against the recognition of penal judgments outlined in Part II.C. Orders to “bring processing operations into compliance” with GDPR under Article 58.2(d), or that impose a ban on processing under Article 58.2(g), for instance, would arguably be “penal” insofar as they are “in favor of a foreign state . . . and primarily punitive rather than compensatory in character,” and would require a U.S. court to scrutinize and enforce foreign public law.¹¹⁴

In sum, a plaintiff seeking to enforce a nonmonetary order issued by a supervisory authority under GDPR Article 58.2 would face several challenges.

C. Administrative fines issued under Articles 58.2(i) and 83: enforceability under U.S. law

GDPR Article 58.2(i) gives supervisory authorities the authority to issue an administrative fine “in addition to, or instead of” the nonmonetary orders listed in the preceding Section, depending on the circumstances of each individual case. GDPR Article 83.1 provides that these fines should be “effective, proportionate and dissuasive.”¹¹⁵ To that end, GDPR Article 83.2 lists the criteria to be considered in determining whether to

111. GDPR, *supra* note 1, art. 58.2(d).

112. *Id.*, art. 58.2(f).

113. RESTATEMENT (SECOND), *supra* note 47, § 102 cmt. g.

114. RESTATEMENT (THIRD), *supra* note 45, § 483 cmt. b.

115. GDPR, *supra* note 1, art. 83.1.

impose a fine and the amount. These include, *inter alia*, “the nature, gravity, and duration of the infringement,”¹¹⁶ “any relevant previous infringements by the controller or processor,”¹¹⁷ the controller or processor’s “degree of cooperation with the supervisory authority,”¹¹⁸ and “any other aggravating or mitigating factor applicable to the circumstances of the case.”¹¹⁹ Taken together, these provisions confirm that administrative fines issued under GDPR are punitive—rather than compensatory—in character.

Accordingly, administrative fines are in most circumstances subject to nonrecognition under the Recognition Acts, both of which expressly exclude foreign fines and penal judgments from their provisions for recognition.¹²⁰ They can also be subject to nonrecognition under the common law.¹²¹ These conclusions likely apply whether or not an administrative fine is incorporated into a court judgment.

There is, however, a potential exception to the rule against recognizing foreign fines and penal judgments. As noted in Part II.C above, the 2005 Recognition Act’s savings clause might still allow for a foreign penal judgment to be recognized under the

116. *Id.*, art. 83.2(a).

117. *Id.*, art. 83.2(e).

118. *Id.*, art. 83.2(f).

119. *Id.*, art. 83.2(k).

120. 1962 Recognition Act, *supra* note 49, § 1(2) (defining “foreign judgment” that is subject to recognition as excluding “a judgment for taxes, a fine, or other penalty”); 2005 Recognition Act, *supra* note 50, § 3(b) (providing that the act does not apply “to the extent that the judgment is . . . a fine or other penalty”).

121. See RESTATEMENT (THIRD), *supra* note 45, § 483; RESTATEMENT (FOURTH), *supra* note 46, § 489.

common law.¹²² And under the Restatement (Third) of Foreign Relations Law, the common-law rule against recognition of foreign penal judgments is permissive, rather than mandatory, insofar as it provides that courts in the United States “are not required” to recognize or enforce penalties rendered by courts of other states.¹²³ Thus, it is conceivable that a U.S. court could recognize and enforce an administrative fine under GDPR that had been reduced to a judgment in an EU court, provided that the judgment was not subject to nonrecognition on another mandatory or discretionary basis.

But enforcement of such a judgment would seem unprecedented. Although U.S. courts sometimes *recognize* foreign penal judgments in the context of criminal prosecutions and sentencing,¹²⁴ no U.S. court appears to have ever *enforced* a foreign judgment or order that called for the payment of a fine to a foreign government body in the absence of a treaty that required it.

122. See 2005 Recognition Act, *supra* note 50, § 11 (“This Act does not prevent the recognition under principles of comity or otherwise of a foreign-country judgment not within the scope of this act.”).

123. RESTATEMENT (THIRD), *supra* note 45, § 483 cmt. a (“No rule of United States law or of international law would be violated if a court in the United States enforced a judgment of a foreign court for payment of taxes or comparable assessments that was otherwise consistent with the standards of §§ 481 and 482.”).

124. *Id.* at Reporter’s Note 3.

V. POTENTIAL DEFENSES UNDER U.S. LAW TO AN ACTION SEEKING RECOGNITION AND ENFORCEMENT OF A GDPR ORDER OR JUDGMENT

The party seeking to enforce the order or judgment bears the initial burden of establishing a *prima facie* case for recognition.¹²⁵ The issues the plaintiff might face—and that a defendant might exploit—in that regard are discussed in Parts II and III.

Assuming the plaintiff establishes a *prima facie* case for recognition, the burden switches to the U.S. defendant to establish that the judgment or order is subject to one of the mandatory or discretionary grounds for nonrecognition.¹²⁶ U.S. defendants might be especially likely to raise two grounds for nonrecognition, one mandatory and one discretionary: (1) that the rendering forum in the EU lacked personal jurisdiction over the defendant, and (2) that the order sought to be enforced is repugnant to U.S. public policy.

This part of the *Commentary* provides an overview of those defenses.

A. *Lack of personal jurisdiction over the defendant in the EU*

Under the common law and the Recognition Acts, lack of personal jurisdiction over the defendant in the rendering court is a mandatory ground for nonrecognition of a foreign judgment in a U.S. court.¹²⁷ Thus, a U.S. court will recognize a foreign judgment only if the foreign court had personal jurisdiction over the party against whom the judgment is to be enforced. A key issue in that regard is what law controls that question: the law of the

125. See Part II.F, *supra*.

126. *Id.*

127. See Part I.D, *supra*.

country in which the judgment was rendered, or U.S. law.¹²⁸ The common law and the Recognition Acts diverge somewhat on this point.

The Restatement (Third) of Foreign Relations Law takes the view that under the common law, a U.S. court should look to both the law of the rendering country and U.S. law. Specifically, Section 482 of the Restatement declares that a court in the United States “may not” recognize a foreign judgment if “the court that rendered the judgment did not have jurisdiction over the defendant in accordance with the law of the rendering state *and* with the rule set forth in § 421.”¹²⁹ Section 421 of the Restatement (Third), in turn, lists several grounds that make an exercise of personal jurisdiction over a defendant presumptively reasonable:

- (2) In general, a state’s exercise of jurisdiction to adjudicate with respect to a person or thing is reasonable if, at the time jurisdiction is asserted:
 - (a) the person or thing is present in the territory of the state, other than transitorily;
 - (b) the person, if a natural person, is domiciled in the state;
 - (c) the person, if a natural person, is resident in the state;
 - (d) the person, if a natural person, is a national of the state;

128. For a comprehensive discussion of this question, *see* Monestier, *supra* note 43.

129. RESTATEMENT (THIRD), *supra* note 45, § 482(1)(b) (emphasis added).

- (e) the person, if a corporation or comparable juridical person, is organized pursuant to the law of the state;
- (f) a ship, aircraft or other vehicle to which the adjudication relates is registered under the laws of the state;
- (g) the person, whether natural or juridical, has consented to the exercise of jurisdiction;
- (h) the person, whether natural or juridical, regularly carries on business in the state;
- (i) the person, whether natural or juridical, had carried on activity in the state, but only in respect of such activity;
- (j) the person, whether natural or juridical, has carried on outside the state an activity having a substantial, direct, and foreseeable effect within the state, but only in respect of such activity; or
- (k) the thing that is the subject of adjudication is owned, possessed, or used in the state, but only in respect of a claim reasonably connected with that thing.¹³⁰

In addition, Section 421 of the Restatement provides that a defense of lack of jurisdiction is generally considered to be waived “by any appearance by or on behalf of a person . . . if the appearance is for a purpose that does not include a challenge to the exercise of jurisdiction.”¹³¹

130. *Id.* § 421(2).

131. *Id.* § 421(3).

Thus, under the Restatement (Third)'s construction, a U.S. court first inquires whether the foreign court had personal jurisdiction under its own law, and then whether the exercise of that jurisdiction is "reasonable" in accordance with standards provided by U.S. common law and as set out in the Restatement.

The Restatement (Fourth), by contrast, suggests that only U.S. law governs the question of personal jurisdiction. Its rule makes no mention of the rendering country's law regarding personal jurisdiction, and its comments provide that "[c]ourts in the United States will not recognize a foreign judgment if the court rendering the judgment would have lacked personal jurisdiction under the minimum requirements of due process imposed by the U.S. Constitution."¹³²

Both the 1962 and 2005 Recognition Acts also prohibit a U.S. court from recognizing a judgment rendered by a foreign court that lacked personal jurisdiction over the defendant.¹³³ Neither Recognition Act identifies the source of law that should govern that question in the U.S. court. Like the Restatement (Third), however, the Recognition Acts identify several factors that, once established, prohibit nonrecognition for lack of personal jurisdiction. Under the 2005 Recognition Act, for instance, a U.S. court "may not" refuse to recognize a foreign judgment for lack of personal jurisdiction if:

- (1) the defendant was served with process personally in the foreign country;
- (2) the defendant voluntarily appeared in the proceeding, other than for the purpose of protecting property seized or threatened with seizure in

132. RESTATEMENT (FOURTH), *supra* note 46, § 483(b) and cmt. e.

133. 1962 Recognition Act, *supra* note 49, § 4(a)(2); 2005 Recognition Act, *supra* note 50 § 4(b)(2).

the proceeding or of contesting the jurisdiction of the court over the defendant;

(3) the defendant, before the commencement of the proceeding, had agreed to submit to the jurisdiction of the foreign court with respect to the subject matter involved;

(4) the defendant was domiciled in the foreign country when the proceeding was instituted or was a corporation or other form of business organization that had its principal place of business in, or was organized under the laws of, the foreign country; [or]

(5) the defendant had a business office in the foreign country and the proceeding in the foreign court involved a [cause of action] [claim for relief] arising out of business done by the defendant through that office in the foreign country[.]¹³⁴

As to this choice-of-law question, at least one commentator has argued—with some force—that a U.S. court generally should not attempt to resolve the question of whether the foreign court actually had jurisdiction over the defendant under its own laws.¹³⁵ Perhaps more importantly for purposes of this *Commentary*, that same commentator has also argued that even when U.S. courts purport to look to foreign law, the end result is the same: they rarely end their analysis at the question of the application of foreign law, and their decisions most often turn on the application of U.S. law to the question of whether the foreign court's assertion of personal jurisdiction was "reasonable,"

134. *Id.* § 5(a).

135. Monestier, *supra* note 43, at 1743–63.

“permitted,” or consistent with a “minimum contacts” analysis.¹³⁶

Without opining on the usefulness of an inquiry into the foreign country’s law, this *Commentary* focuses on the question whether a U.S. court will consider an EU Member State’s assertion of personal jurisdiction under Article 3 of GDPR to be reasonable or permitted under U.S. legal standards. In other words, the *Commentary* assumes that the assertion of personal jurisdiction by the hypothetical EU court is consistent with GDPR and the law of personal jurisdiction within the relevant EU Member State.

GDPR Articles 3.1 and 3.2 provide the most likely starting point for an EU court or Data Protection Authority’s exercise of personal jurisdiction over a U.S. defendant.

1. Personal jurisdiction under GDPR Article 3.1

In the case of GDPR Article 3.1, the question appears fairly straightforward insofar as that provision relies on the existence of an “establishment” in the EU:

This Regulation applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not.¹³⁷

An assertion of personal jurisdiction on this basis would likely be held to be reasonable under the both the common law and the Recognition Acts:

136. *Id.* at 1759–60.

137. GDPR, *supra* note 1, art. 3.1.

- If the conduct at issue was “in the context of an establishment of a controller or a processor in the Union,” the existence of an “establishment” in the EU would likely support a finding that the defendant was “present in the territory of the state” for purposes of Section 421 of the Restatement.
- Similarly, the 2005 Recognition Act’s view that the exercise of jurisdiction is permitted where the defendant “had a business office in the foreign country and the proceeding in the foreign court involved a [cause of action] [claim for relief] arising out of business done by the defendant through that office in the foreign country” would appear to be satisfied whenever Article 3.1 is triggered by the existence of an “establishment” in the EU.

Granted, GDPR Article 3 purports to apply “regardless of whether the processing takes place in the Union or not,” while the Recognition Act requires that the cause of action arise out of business “done by the defendant through that office in the foreign country.” However, the Recognition Act’s use of the word “through,” rather than “in,” would likely apply to a showing that the processing was “in the context of the activities of an establishment” of the defendant. The fact that the processing itself did not take place “in” that establishment would seem to be of little help to a defendant if that processing was “in the context of the activities of” that establishment.¹³⁸

138. Precisely what it might mean for processing that does not take place “in” a particular business establishment to nonetheless be “in the context of the activities of” that establishment is a question of the substantive application of GDPR that is beyond the scope of this *Commentary*.

2. Personal jurisdiction under Article 3.2

GDPR Article 3.2, by contrast, may prove more difficult as a ground for personal jurisdiction over a U.S.-based defendant, because neither of its grounds for application of GDPR depends on the physical presence of that defendant within the EU. That provision provides:

This Regulation applies to the processing of personal data of data subjects who are in the Union by a controller or processor not established in the Union, where the processing activities are related to:

- a. the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or
- b. the monitoring of their behaviour as far as their behaviour takes place within the Union.¹³⁹

As an illustrative scenario in which the issue of personal jurisdiction could be especially relevant, consider a U.S.-based retailer operating a website clearly and unambiguously marketing the sales of goods or services to EU residents, but otherwise having no physical presence or stable relationships in the EU. Under the Recognition Acts, the retailer could argue that none of the criteria for the permissible exercise of jurisdiction are present absent some showing of personal service within the EU or some appearance in the EU proceedings other than for the purpose of contesting jurisdiction.

139. GDPR, *supra* note 1, art. 3.2.

The situation under the common law may, perhaps, be slightly more favorable for the party seeking to enforce the judgment or order if that party could show that the defendant's "offering of goods or services" to data subjects in the EU constituted "regularly carr[ying] on business" within the EU for purposes of Section 421(h) of the Restatement (Third) of Foreign Relations Law. Application of GDPR Article 3.2(a), however, is not restricted to situations in which the controller or processor "regularly" offers goods or services, and it is therefore likely that GDPR at least in some instances facially purports to extend its effect to U.S. businesses in a manner in which most U.S. courts would be unlikely to recognize.

A defendant might have an even better chance at mounting a successful challenge to personal jurisdiction where the EU's assertion of jurisdiction over that defendant arose under GDPR Article 3.2(b) based solely on the "monitoring of behavior" of data subjects within the EU. Take, for example, a scenario contemplated by the EDPB in its Guidelines on the Territorial Scope of the GDPR, in which a U.S. company (acting as a controller) develops a health and lifestyle app that allows users to record detailed health and fitness information, and monitors the behavior of individuals in the EU who use that app.¹⁴⁰ For purposes of data storage, that company engages a processor—a cloud service provider—established in the U.S.¹⁴¹ The EDPB concludes that in this scenario, the controller is subject to GDPR under Article 3.2, but also that the *cloud service provider* is subject to GDPR under Article 3.2 because it is engaging in

140. Territorial Scope Guidelines, *supra* note 3, at 21.

141. *Id.*

processing—data storage—that is “related to” the targeting of individuals in the EU by the controller.¹⁴²

In the hypothetical, the cloud provider would not likely satisfy any of the criterion required for a “reasonable” assertion of personal jurisdiction under the Restatement (Third) of Foreign Relations Law or a “permitted” one under the Recognition Acts. The cloud provider could thus argue convincingly that any judgment or order obtained against it in the EU related to the processing performed on behalf of the health and lifestyle app company is subject to mandatory nonrecognition under the Recognition Acts and the common law.

3. Data Protection Officers and Article 27 representatives: impact on personal jurisdiction in the EU

A U.S. entity that does not trigger any of the standards that make an assertion of jurisdiction presumptively reasonable through its day-to-day operations might nonetheless submit itself to jurisdiction of an EU court or regulator through the appointment of an agent in the EU. The comments and reporters’ notes to the Restatement (Third) of Foreign Relations Law suggest that conducting activity in a foreign state through an “agent” could be a basis to find a waiver of lack of personal jurisdiction as a ground for nonrecognition.¹⁴³

Two potential grounds for this “agency” theory of waiver are the defendant’s appointment of a “representative” in the EU pursuant to GDPR Article 27 or the designation of a Data Protection Officer (DPO) under GDPR Article 37.

142. *Id.*

143. See RESTATEMENT (THIRD), *supra* note 45, § 481, Reporter’s Note 3; § 482, cmt. c.

Under GDPR Article 27, an EU representative appointed by a controller or processor not established in the EU “shall be mandated by the controller or processor to be addressed in addition to or instead of the controller or the processor by, in particular, supervisory authorities and data subjects, on all issues related to processing, for the purposes of ensuring compliance with this Regulation.”¹⁴⁴ Arguably, this could be seen as either the explicit or implied expression of consent to submit to personal jurisdiction within the Member State where the representative is appointed, particularly because the appointment is “without prejudice to legal actions which could be initiated against the controller or the processor themselves.”¹⁴⁵ The mandate that the representative is “to be addressed” by data subjects and supervisory authorities “for the purposes of ensuring compliance with this Regulation” is likely to be seen as a voluntary designation of an agent for the purpose of securing personal jurisdiction over the appointing entity.

It thus seems likely that a U.S. business that has appointed a representative under Article 27 will be found to have consented to the personal jurisdiction of the EU courts and regulators. More difficult situations would involve U.S. businesses that fail to appoint an EU representative, whether because they do not know they are obligated to do so, incorrectly determine they are not obligated to do so, or deliberately refuse to appoint an EU representative in a purposeful attempt to avoid enforcement. Under such circumstances, the U.S. court would need to determine if the U.S. entity was subject to the EU’s long-arm jurisdiction despite the failure to appoint. In any case, the court’s

144. GDPR, *supra* note 1, art. 27.4.

145. *Id.*, art. 27.5.

decision would likely turn on the particular facts and circumstances presented in the evidence.

A U.S. organization's appointment of a DPO under GDPR Article 37 might also lead a U.S. court to conclude that the organization consents to jurisdiction in the EU. Among the responsibilities of a DPO designated under GDPR Article 37 is that she "cooperate with the supervisory authority," "act as the contact point with the supervisory authority on issues relating to processing," and be available for contact by data subjects "with regard to all issues related to processing of their personal data and to the exercise of their rights under this Regulation."¹⁴⁶ This, too, may be sufficient to imply consent to jurisdiction. Even if not, when a DPO is physically present in the EU, that presence may allow for personal service on the organization through an agent or at a place of business, a sufficient basis for personal jurisdiction under both the Restatement and the Recognition Acts.

4. Execution of data processing and data transfer agreements: impact on personal jurisdiction in the EU

A U.S. organization could also make itself subject to a presumptively reasonable exercise of personal jurisdiction in the EU by entering into data processing or data transfer agreements with EU-based organizations in which the U.S. organization consents to such jurisdiction. To that end, both the Restatement (Third) of Foreign Relations Law¹⁴⁷ and the 2005 Recognition Act¹⁴⁸ provide grounds for a U.S. court to find that an EU court validly exercised jurisdiction over a defendant when that

146. *Id.*, arts. 38.4, 39.1(d)-(e).

147. *See* RESTATEMENT (THIRD), *supra* note 45, § 482(2)(g).

148. 2005 Recognition Act, *supra* note 50, § 5(a)(3).

defendant previously agreed to submit to the jurisdiction of the foreign court.

U.S. organizations that sign data processing and data transfer agreements that include EU Commission-approved standard contractual clauses to facilitate transatlantic data transfers may waive—at least in part—any defense based on lack of personal jurisdiction in the EU on this basis. In that regard, both the controller-to-controller and controller-to-processor versions of the standard contractual clauses give data subjects the right to enforce certain clauses against the data importer as third-party beneficiaries.¹⁴⁹ The clauses provide in turn that the data importer agrees to accept jurisdiction in the data exporter’s country of establishment with respect to claims by data subjects in that capacity.¹⁵⁰

A proposed new set of standard contractual clauses released by the European Commission in November 2020 go even further: in this new proposed set of clauses, the data importer “agrees to submit itself to the jurisdiction of the competent supervisory authority in any procedures aimed at ensuring compliance with these clauses,” including inquiries and audits.¹⁵¹

149. See Commission Decision 2010/87 on standard contractual clauses for the transfer of personal data to processors established in third countries under directive 95/46/EC of the European Parliament and of the Council, 2010 O.J. L (39/5), Annex Standard Contractual Clauses, cl. 3.1; Commission Decision 2004/915/EC amending Decision 2001/497/EC as regards the introduction of an alternative set of standard contractual clauses for the transfer of personal data to third countries, 2004 O.J. L (385/74), Annex Set II, cl. III(b).

150. *Id.*

151. See 12 November 2020 Draft Annex to the Commission Implementing Decision on Standard contractual clauses for the transfer of personal data to third countries pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council, cl. 9(b), available at <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12741-Commission->

These draft clauses also provide that the parties “agree to submit themselves to the jurisdiction of [the] courts” of an EU member state specified by the parties.¹⁵² If and when these clauses are approved by the EU Commission, any U.S. organization that signs them may have difficulty successfully asserting lack of personal jurisdiction in the EU as a basis for nonrecognition of a GDPR order or judgment.

B. Repugnancy to federal or state public policy

Under both the common law and the Recognition Acts, a U.S. court may decline to recognize a foreign country judgment if the judgment is “repugnant to the public policy” of the United States or of the U.S. state in which recognition is sought.¹⁵³ “Repugnancy,” however, is a stringent standard.¹⁵⁴ Courts have held that simple “inconsistency” between state or federal law and the foreign law does not render a foreign judgment unenforceable because of “repugnancy.”¹⁵⁵ But although repugnancy

Implementing-Decision-on-standard-contractual-clauses-for-the-transfer-of-personal-data-to-third-countries.

152. *Id.* cl. 3(a).

153. RESTATEMENT (THIRD), *supra* note 45, § 482(2)(d); RESTATEMENT (FOURTH), *supra* note 46, § 484(c); 1962 Recognition Act, *supra* note 49, § 4(b)(3); 2005 Recognition Act, *supra* note 50, § 4(c)(3).

154. RESTATEMENT (FOURTH), *supra* note 46, § 484 cmt. e (“The test for public policy is therefore a stringent one A foreign judgment violates local public policy only if its recognition would tend clearly to injure public health, public morals, or public confidence in the administration of law, or would undermine settled expectations concerning individual rights.”).

155. *See, e.g.,* Ohno v. Yasuma, 723 F.3d 984, 1002 (9th Cir. 2013) (“California courts have set a high bar for repugnancy under the Uniform Act. The standard . . . measures not simply whether the foreign judgment or cause of action is contrary to our public policy, but whether either is ‘so offensive to our public policy as to be ‘prejudicial to recognized standards of morality and to the general interests of the citizens.’”); Loucks v. Standard Oil Co. of N.Y., 120

presents a high bar, there are several examples of cases in which courts have repugnancy as the basis for nonrecognition of foreign judgments.¹⁵⁶

As one obvious potential area of repugnancy, enforcement of foreign judgments or administrative orders issued under GDPR may raise serious questions under the First Amendment of the U.S. Constitution. One such example arises from the “right to be forgotten” under GDPR Article 58.2(g). Any such order would likely be repugnant to public policy because it might violate the First Amendment as an impermissible prior restraint on publication.¹⁵⁷

N.E. 198, 201 (N.Y. 1918) (Cardozo, J.) (“We are not so provincial as to say that every solution of a problem is wrong because we deal with it otherwise at home.”).

156. See, e.g., *Telnikoff v. Matusevitch*, 702 A.2d 230 (Md. 1997) (declining to enforce an English libel judgment under principles of comity because English defamation law is “totally different” from Maryland’s law “in virtually every significant respect” and “so contrary . . . to the policy of freedom of the press underlying Maryland law.”); *Pentz v. Kuppinger*, 107 Cal. Rptr. 540 (Cal. Ct. App. 1973) (concluding that a Mexican decree of divorce was repugnant to California law when it required husband to continue to pay alimony even after remarriage of wife).

157. See *Near v. Minnesota ex rel. Olson*, 283 U.S. 697, 716 (1931) (“[L]iberty of the press, historically considered and taken up by the Federal Constitution, has meant, principally although not exclusively, immunity from previous restraints or censorship”); *Bantam Books, Inc. v. Sullivan*, 372 U.S. 58, 70 (1963) (“Any system of prior restraints of expression comes to this Court bearing a heavy presumption against its constitutional validity.”); See also Kurt Wimmer, *Free Expression and EU Privacy Regulation: Can the GDPR Reach U.S. Publishers?*, 68 SYR. L.R. 547, 574 (2018) (“When a foreign judgment is one that would violate the First Amendment, courts have found that it violates public policy and is thus unenforceable. . . . Because an order or fine under GDPR related to the right to be forgotten would almost certainly violate the First Amendment, a U.S. court would likely refuse to enforce such an order from an EU court.”). Note also the Securing the Protection of Our

Repugnancy to public policy may also be reflected in the Securing the Protection of Our Enduring and Established Constitutional Heritage (SPEECH) Act, 22 U.S.C. §§ 4101-05. Interpreted broadly, the SPEECH Act suggests that all foreign judgments that would violate the First Amendment or chill free speech could be unenforceable through the U.S. court system if those cases are deliberately brought in jurisdictions whose laws are less protective of free speech—as would likely be the case with right-to-be-forgotten actions brought against U.S. companies abroad.¹⁵⁸

In addition to First Amendment issues, GDPR orders and judgments could also raise procedural due-process concerns under the Fifth and Fourteenth Amendments of the U.S. Constitution, depending on the procedures used in the EU to issue or obtain them.¹⁵⁹

Enduring and Established Constitutional Heritage (SPEECH) Act, 28 U.S.C. §§ 4101–05, which interpreted broadly suggests that all foreign judgments that would violate the First Amendment or chill free speech should be unenforceable through the U.S. court system if those cases are deliberately brought in jurisdictions whose laws are less protective of free speech—as would likely be the case with right-to-be-forgotten actions brought against U.S. companies abroad. *See* Wimmer at 574–75.

158. *See id.*

159. *See, e.g.,* *Koster v. Automark Indus., Inc.*, 640 F.2d 77 (7th Cir. 1981) (Dutch statute governing service of process on defendants who reside in foreign countries provided insufficient assurances of actual notice to comport with American due-process requirements, and thus Dutch default judgment could not be enforced in U.S. courts).

VI. ALTERNATIVE ROUTES TO GDPR ENFORCEMENT IN U.S. COURTS: THE FEDERAL TRADE COMMISSION AND CONTRACT CLAIMS

Where a U.S.-based organization has violated GDPR, there may be mechanisms for obtaining relief against that organization that do not, strictly speaking, arise under GDPR or involve the recognition or enforcement of GDPR judgments or orders. This part of the *Commentary* discusses two significant possibilities in that regard: (1) enforcement by the U.S. Federal Trade Commission of GDPR-related promises under its authority to police unfair and deceptive acts and practices under Section 5 of the FTC Act; and (2) contract-based actions arising out of agreements that U.S.-based organizations enter for GDPR-related purposes.

A. *The Federal Trade Commission: Section 5 of the FTC Act and Privacy Shield remedies*

The FTC enforces several privacy-related U.S. laws (e.g., the Fair Credit Reporting Act, and the Children’s Online Privacy Protection Act, to name just two); but its primary enforcement authority in privacy and data security cases is based on Section 5 of the FTC Act, which prohibits unfair or deceptive acts or practices.¹⁶⁰ FTC has used that authority repeatedly to bring enforcement actions against companies that fail to abide by the commitments they make in privacy policies and other public statements about their privacy practices.¹⁶¹ The FTC does not, however, have any power either to enforce non-U.S. laws or to

160. 15 U.S.C. § 45(a)(1).

161. See generally Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 628–30 (2014).

bring actions on behalf of individual private persons who may have suffered a privacy or data-breach-related injury.

The FTC's Section 5 authority could nevertheless be used to police a U.S. company's compliance with GDPR, to the extent that company makes broad promises about GDPR compliance that extend to U.S. customers. To that end, in 2018 a representative from the agency explained that "[i]f a company chooses to implement some or all of GDPR across their entire operations, and as a result makes promises to U.S. consumers about their specific practices," then the company must live up to those commitments, as "the FTC could initiate an enforcement action if the company does not comply with" its GDPR-related promises with respect to U.S. consumers.¹⁶²

Section 5 of the FTC Act therefore offers a potential alternative route to enforce GDPR against U.S. companies, albeit only with respect to failures to comply with GDPR-related promises made to U.S. consumers.

Notably, the FTC's authority under Section 5 of the FTC Act also includes the authority to enforce commitments made by U.S. companies that have certified to the EU-U.S. Privacy Shield program. Notwithstanding the Court of Justice of the European Union's judgment in the so-called "Schrems II" decision, which invalidated the European Commission's decision on the adequacy of the Privacy Shield program,¹⁶³ the FTC's enforcement

162. Daniel R. Stoller, *FTC Could Police U.S. Companies' Promises on EU Data Privacy Law*, BLOOMBERG LAW (June 20, 2018), <https://bna.news.bna.com/privacy-and-data-security/ftc-could-police-us-companies-promises-on-eu-data-privacy-law>.

163. See Case C-311/18, *Data Protection Commissioner v Facebook Ireland Limited and Maximillian Schrems*, July 16, 2020 E.C.J., available at <http://curia.europa.eu/juris/document/document.jsf?text=&docid=228677&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=9710189>.

authority remains in place over Privacy Shield program participants who previously received personal data of EU data subjects through the program. To that end, in the wake of the *Schrems II* decision, U.S. Secretary of Commerce Wilbur Ross issued a statement indicating that the Department of Commerce would continue to administer the Privacy Shield program, and that the *Schrems II* decision “does not relieve participating organizations of their Privacy Shield obligations.”¹⁶⁴ Thus, any U.S. organization that remains certified to the EU-US Privacy Shield program and continues to process data received under the program faces a risk of FTC enforcement if it fails to adhere to its commitments.

Such an organization could also face claims by data subjects in the EU. Specifically, Annex I to the EU-U.S. Privacy Shield provides that data subjects have a right to binding arbitration if they have first complained to the relevant company, given it an opportunity to correct its actions, resorted to the (free) independent recourse mechanisms set up in Principle 7, then complained to the relevant supervisory authority and given the U.S. Department of Commerce an opportunity to resolve the matter.¹⁶⁵ The arbitrators in each instance are selected by the parties from a list of at least 20 arbitrators developed by the U.S. Department of Commerce and the European Commission, and the

164. Press Release, U.S. Secretary of Commerce Wilbur Ross Statement on Schrems II Ruling and the Importance of EU-U.S. Data Flows (July 16, 2020), available at <https://www.commerce.gov/news/press-releases/2020/07/us-secretary-commerce-wilbur-ross-statement-schrems-ii-ruling-and>.

165. See Commission Implementing Decision of July 12, 2016 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the EU-US Privacy Shield, 2016 O.J. (L 207/1), Annex 2 Arbitral Model, Annex I.

ensuing arbitration may be conducted over the telephone.¹⁶⁶ Although the arbitration panel lacks any authority to grant monetary remedies to data subjects, it has the authority to impose nonmonetary relief such as granting access, correction, deletion, or return of the personal data in question.¹⁶⁷ EU-U.S. Privacy Shield companies are required to advise data subjects of their rights to binding arbitration and the procedures they need to follow to invoke those rights.¹⁶⁸ At least with respect to EU-U.S. Privacy Shield companies that violate GDPR, these mechanisms can provide individual data subjects with a viable alternative to seeking enforcement of a judgment by a U.S. court.

B. Contract actions associated with data protection

1. Contracts between data subjects and data controllers

There are myriad contractual arrangements entered between EU data subjects and data controllers on a daily basis that expressly involve the collection and retention of personal data. Some may be related to long-term, essential relationships, such as contracts for employment, housing, or financial arrangements. Others may be highly transactional in nature, such as the use of internet browser tracking mechanisms and one-off online transactions. Still others occupy a middle space: ongoing relationships of a nonessential nature. Many of these are represented by the omnipresent “I Agree” button that must be clicked to use some new software for a computer or mobile device, or to sign up for an online Software as a Service (SaaS). These agreements often contain a hyperlink to a “privacy policy” that has been “incorporated by reference” to the agreement and sets

166. *Id.*

167. *Id.*

168. *Id.*, Annex II § II.1.a.xi.

out the nonnegotiable terms for processing personal data that are difficult to understand by even the most experienced attorneys.

An EU data subject who has established a contractual relationship with a U.S.-based controller that includes data protection provisions, depending on the contract's choice-of-forum provisions, might thus be able to seek enforcement of her rights in a breach-of contract action filed directly against the controller in a U.S. court.

2. Contracts between data controllers and data processors under GDPR Article 28

Under GDPR Article 28, when a controller engages a processor, the parties are obligated to enter into a contract that governs the processing, is "binding on the processor with regard to the controller," and includes various mandatory terms relating to the processing.¹⁶⁹ The processor is in turn obligated to impose the same obligations on any other processors it engages to carry out that processing.¹⁷⁰

When a U.S.-based processor signs a contract pursuant to Article 28 with an EU-based controller or processor, that contract provides another means through which the GDPR's requirements might be enforced against the U.S.-based processor. If the U.S.-based processor violates the requirements of the processing agreement, the EU-based controller or processor can—depending on the contract's choice of forum—enforce that contract directly against the U.S.-based processor in a U.S. court.

169. GDPR, *supra* note 1, art. 28.3.

170. *Id.* at art. 28.4.

3. Data transfer contracts based on Standard Contractual Clauses

As noted in Part V.A.4 above, U.S. organizations acting as controllers or processors may also enter into data processing and data transfer agreements that incorporate standard contractual clauses approved by the EU Commission to address GDPR's restrictions on cross-border data transfers.¹⁷¹

An EU data exporter with whom the U.S. organization has executed the standard contractual clauses, or a data subject with status as a third-party beneficiary under those clauses,¹⁷² could bring an action to enforce the clauses against the U.S. organization in a U.S. court.

171. *See id.* art. 46.2(c).

172. *See* Commission Decision 2010/87 on standard contractual clauses for the transfer of personal data to processors established in third countries under Directive 95/46/EC of the European Parliament and of the Council, 2010 O.J. L (39/5), Annex Standard Contractual Clauses, cl. 3; Commission Decision 2004/915/EC amending Decision 2001/497/EC as regards the introduction of an alternative set of standard contractual clauses for the transfer of personal data to third countries, 2004 O.J. L (385/74), Annex Set II, cl. III(b).

VII. CONCLUSION

Since its entry into force, GDPR has given rise to important questions about the reach of European data protection rules, and the ability of individuals and supervisory authorities to enforce those rules against foreign defendants. The answers to those questions can be especially complex for U.S.-based organizations that do not maintain a physical presence or other assets in the EU, but still fall within GDPR's extraterritorial reach. The *Commentary* discusses whether and how a party in the EU—whether a supervisory authority, individual data subject, or not-for-profit body acting on behalf of data subjects—can obtain such an organization's compliance with GDPR through resort to a U.S. court proceeding.

The *Commentary* outlines the considerations, both legal and practical, that U.S.-based organizations and parties in the EU should consider when faced with the question of how a U.S. court might address a request to enforce a GDPR order or judgment. As the *Commentary* shows, the enforceability of GDPR orders and judgments in a U.S. court will depend on several factors, including the nature of the relief sought through the order or judgment, the nature of the underlying violation and the process through which the order or judgment was initially obtained in the EU, and the U.S. organization's contacts with the EU. By exploring how those factors might influence a court's application of the existing body of U.S. law regarding the recognition and enforcement of foreign judgments, the *Commentary* provides a framework that parties on both sides of the Atlantic can use to evaluate whether, in a given case, the long arm of the GDPR might reach a U.S. defendant.

THE SEDONA CONFERENCE COMMENTARY
ON A REASONABLE SECURITY TEST

*A Project of The Sedona Conference Working Group
on Data Security and Privacy Liability (WG11)*

Author:

The Sedona Conference

Editor-in-Chief:

William R. Sampson

Contributing Editors:

David Thomas Cohen

Chris Cronin

Hon. Joseph C. Iannazzone

James Pizzirusso

Ruth Promislow

Samuel S. Rubin

Joseph W. Swanson

James Trilling

Hon. Thomas I. Vanaskie (ret.)

Steering Committee Liaison:

Douglas H. Meal

Staff Editors:

David Lumia

Michael Pomarico

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 11. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of

Copyright 2021, The Sedona Conference.
All Rights Reserved.

the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

We thank all of our Working Group Series Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors, just click on the “Sponsors” navigation bar on the homepage of our website.

This publication may be cited as follows:

The Sedona Conference, *Commentary on a Reasonable Security Test*, 22 SEDONA CONF. J. 345 (2021).

PREFACE

Welcome to the February 2021 final version of *The Sedona Conference Commentary on a Reasonable Security Test* (“*Commentary*”), a project of The Sedona Conference Working Group 11 on Data Security and Privacy Liability (WG11). This is one of a series of Working Group commentaries published by The Sedona Conference, a 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, intellectual property rights, and data security and privacy law. The mission of The Sedona Conference is to move the law forward in a reasoned and just way.

The mission of WG11 is to identify and comment on trends in data security and privacy law, in an effort to help organizations prepare for and respond to data breaches, and to assist attorneys and judicial officers in resolving questions of legal liability and damages.

The Sedona Conference acknowledges Editor-in-Chief Bill Sampson for his leadership and commitment to the project. We also thank Contributing Editors David Cohen, Chris Cronin, Judge Joe Iannazzone, James Pizzirusso, Ruth Promislow, Sam Rubin, Joe Swanson, Jim Trilling, and Judge Tom Vanaskie for their efforts, and Doug Meal for his contributions as Steering Committee liaison to the project. We thank Alyssa Coon, Colman McCarthy and Jim Shook for their contributions.

In addition to the drafters, this nonpartisan, consensus-based publication represents the collective effort of other members of WG11 who reviewed, commented on, and proposed edits to early drafts that were circulated for feedback from the Working Group membership. Other members provided feedback at WG11 annual and midyear meetings where drafts of the *Commentary* were the subject of dialogue. The publication was also subject to a period of public comment.

On behalf of The Sedona Conference, I thank all of them for their contributions.

We encourage your active engagement in the dialogue. Membership in The Sedona Conference Working Group Series is open to all. The Series includes WG11 and several other Working Groups in the areas of electronic document management and discovery, cross-border discovery and data protection laws, international data transfers, patent litigation, patent remedies and damages, and trade secrets. The Sedona Conference hopes and anticipates that the output of its Working Groups will evolve into authoritative statements of law, both as it is and as it should be. Information on membership and a description of current Working Group activities is available at <https://thesedonaconference.org/wgs>.

Craig Weinlein
Executive Director
The Sedona Conference
February 2021

TABLE OF CONTENTS

INTRODUCTION.....	351
I. THE TEST.....	358
A. Articulation of the Test.....	360
B. Explanation of the Test.....	362
1. Controls	362
2. Foresight Versus Hindsight.....	362
3. Burden to the Information Steward and Others (Costs).....	363
4. Benefit to the Claimant and Others (Benefits) 364	
5. Industry Custom	366
6. Effect of Violating a Statute, Regulation, or Ordinance.....	368
7. Determining Likelihood of Burden and Benefit.....	370
8. When to Apply the Test	371
9. Availability of Resources	372
10. Poor Implementation of a Control.....	373
11. Illustrations of the Application of the Test..	373
II. DISCUSSION	375
A. The Work That Led to the Test.....	375
1. Judicial Opinions.....	375
2. Statutes and Regulations.....	385
3. Marketplace	388
(a) Mandated Minimum Controls	388
(b) Prescriptive But Flexible Controls	389
(c) Standards/ Frameworks	389
(d) Open Requirements.....	390

B.	All the Things “Ruled Out”	392
1.	Specific Controls.....	392
2.	Definition of Personal Information	392
3.	Breach Requirement	392
4.	Causation in Fact.....	393
5.	Proximate Cause	394
6.	Damages.....	394
7.	Existence of Obligation to Have “Reasonable” Security	394
8.	Fault/Liability	395
C.	The Importance of Flexibility	396
1.	The Data to Be Protected.....	396
2.	Threats and Risks	396
III.	CONCLUSION	398
	APPENDIX A—EXEMPLAR CASES	400

INTRODUCTION

Objective

This *Commentary* addresses what “legal test” a court or other adjudicative body should apply in a situation where a party has, or is alleged to have, a legal obligation to provide “reasonable security” for personal information, and the issue is whether the party in question has met that legal obligation.

Roadmap

The *Commentary* begins with a brief summary of the importance of having a test, the reasoning behind a cost/benefit approach for the test, and what issues the test does not address. Part I sets out the proposed test and the explanation of how it is applied. Part II provides review and analysis of existing resources that offer guidance on how reasonable security has been defined and applied to date and explains how they bear upon the test. It includes a summary review of statutes and regulations that require organizations to provide reasonable security with respect to personal information, decisions of courts and other administrative tribunals with respect to the same, applicable industry standards, and marketplace information. Following this discussion, the *Commentary* identifies those items that are not included in the proposed test (also referenced in the Introduction section) and concludes with a discussion regarding the importance of flexibility.

The Importance of Having a Test

This *Commentary* proposes a reasonable security test. In the course of developing it, the drafters considered whether a “reasonable security” test is even needed.

The reasons are there, and they are important. First, there is no one-size-fits-all cybersecurity program. Different

organizations face different data security risks and have different levels of resources available to address those risks.

While approaches such as the National Institute of Standards and Technology (NIST) Cybersecurity Framework provide a helpful structure for identifying protections an organization may need to counter risks particular to its business, few frameworks set out a structure for determining what is “reasonable” in the circumstances—a necessary consideration when adapting such a framework to an organization.

Statutes and regulations require subject organizations to implement reasonable security with respect to the protection of personal information. But here, as well, most of these statutes and regulations require the organization to determine what is reasonable in the circumstances. Review of existing laws and regulations¹ found different requirements. Because fewer than

1. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L119/1) *available at* <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679#PP3Contents>; Health Insurance Portability and Accountability Act (HIPAA), 45 C.F.R. Parts 160 and 164, Subpart C (2002); Federal Trade Commission Gramm-Leach-Bliley Act Safeguards Rule, 16 C.F.R. Part 314 (2002); Children’s Online Privacy Protection Act (COPPA) 15 U.S.C. §§ 6501–6506 (1998); Standards for the Protection of Personal information of Residents of the Commonwealth (Massachusetts data breach notification law), 201 MASS. CODE REGS. 17.00 (2010); California Consumer Privacy Act, CAL. CIV. CODE § 1798.150(a)(1) (2020); California Customer Records Act, CAL. CIV. CODE § 1798.81.5 (2000); New York SHIELD Act, N.Y. GEN. BUS. LAW § 899-bb; New York Department of Financial Services Regulation, N.Y. COMP. CODES R. & REGS. tit. 23 § 500 (2017); Personal Information Protection and Electronic Documents Act (PIPEDA), S.C. 2000 c. 5, (Can.); Ontario Personal Health Information Protection Act, SO 2004, c. 3; Alberta Personal Information Protection Act, SA 2003, c. P-6.5; British Columbia Personal Information Protection Act, SBC 2003, c. 63; New

half explicitly required a common component, the question of how to determine what is reasonable continues unanswered.

Certain regulators have tried to address this situation by offering “guidance” to organizations on how to implement reasonable security.² Such guidance, however, is not legally binding.³ Accordingly, organizations that follow (or fail to follow) the guidance would not necessarily be found to have implemented (or to have failed to implement) reasonable security.

Even if it were legally binding, the guidance provides limited instruction on the question of “reasonableness.” The Federal Trade Commission’s (FTC) guidance, for example, provides high-level descriptions of security management programs and specific controls. These controls are by no means comprehensive and cannot account for the many factors that might be pertinent for any given organization.

California’s guidance describes the measures specified in the Center for Internet Security’s Critical Security Controls as furnishing the minimum security measures that the California Attorney General believed to be necessary ingredients of reasonable security.⁴ Yet, because it is keyed to an identified set

Brunswick Personal Health Information Privacy and Access Act, SNB 2009, c. P-7.05; Newfoundland Personal Health Information Act, SNL 2008 c. P-7.01; Nova Scotia Personal Health Information Act, SNS 2010 c.41.

2. *See, e.g.*, FED. TRADE COMM’N, PROTECTING PERSONAL INFORMATION: A GUIDE FOR BUSINESS (2016), <https://www.ftc.gov/tips-advice/business-center/guidance/protecting-personal-information-guide-business>; FED. TRADE COMM’N, START WITH SECURITY: A GUIDE FOR BUSINESS (2015), <https://www.ftc.gov/system/files/documents/plain-language/pdf/0205-startwithsecurity.pdf> ; KAMALA D. HARRIS, CAL. DEP’T OF JUSTICE, CALIFORNIA DATA BREACH REPORT (2016), <https://www.oag.ca.gov/sites/all/files/agweb/pdfs/dbr/2016-data-breach-report.pdf>.

3. *See, e.g.*, *Wos v. E.M.A. ex rel. Johnson*, 568 U.S. 627, 643 (2013).

4. HARRIS, *supra* note 2, at 30–32.

of 20 controls, the guidance is both cumbersome and static. In sum, regulatory guidance has not provided a test for determining reasonableness.

The importance of a reasonable security test is further underscored by the reported legal cases. Taken together, they indicate “unreasonable” security may be a necessary element of a data security claim; but they do not clearly define reasonableness. This point is highlighted by the Pennsylvania Supreme Court decision in *Dittman v. University of Pittsburgh Medical Center*,⁵ where the Court affirmed the preexisting, negligence-based duty to safeguard personal information where an employer had required employees to provide personal information and then stored it in a manner that permitted an undetected breach of that information. The imposition of a negligence-based duty to safeguard personal information highlights the utility of a test to assess whether an organization has implemented reasonable security.

Cybersecurity reasonableness crosses both legal and technology issues. Reasonable security is a standard that both legal and technology professionals seek to apply. It can be difficult for information technology (IT) organizations to understand how to apply legal concepts to their organizations; it is similarly difficult for lawyers, compliance/risk professionals, and even judges to understand IT well enough to apply it to the legal concepts they know. A reasonableness test would help to bridge that divide.

Having said all of this, the proposed test is not intended to impose on organizations an affirmative legal duty to make one or another information security decision. Instead, the test determines the reasonableness of an organization’s security measures based on the outcomes—measured by costs and

5. 196 A.3d 1036 (Pa. 2018).

benefits—that reasonably would be expected to flow from whatever information security measures the organization did or did not provide.

A Cost/Benefit Approach: How and Why

The statutes and regulations summarized in this *Commentary* commonly identify the following themes with respect to reasonable security:

- Sensitivity of information: Personal information should be protected by safeguards appropriate to the sensitivity of the information. More sensitive information should be safeguarded by a higher level of protection.
- Cost/benefit analysis: The analysis should include a consideration of the sensitivity of the information, the associated risk of harm arising either from unauthorized access to it or from the deprivation, loss or destruction of the information, the available controls to protect the information, and the cost of those measures to the organization.

The cost/benefit analysis that is embedded in some statutes and regulations weaves together the concept that reasonable security is informed by the sensitivity of information with a second concept: it is important to count the cost of implementing security to the organization relative to the cost of the potential harm of failing to do so. While a cost/benefit approach provides a useful, overall structure, further guidance is important when determining how the themes underlying the cost/benefit analysis should work together in defining reasonableness.

Costs and benefits may come in many forms, relevant both to organizations that are required to implement reasonable security measures and to others that are not. Organizations consider these costs and benefits as they design security controls; only later are adjudicators asked to consider the balance between them. The test should accommodate the variety of costs and benefits that should be considered in a cost/benefit analysis, including the utility or benefit of the risk, organizational costs (including financial and operational costs), and harm (including harms that alternative controls may cause). The test should take these costs and benefits into account not only as to the organization and the claimant in question, but also as to all persons who would incur such costs and benefits, such as the data subjects whose information the organization elects to place at risk.

What the Test Does Not Address or Require

The test does not address other issues that may arise in cybersecurity litigation and regulatory enforcement proceedings, nor does the test require the presence of certain events or items. Those issues, events, and items include the following, all of which are outside the scope of this paper:⁶

1. The test does not mandate particular controls as part of a “reasonable” cybersecurity program.
2. The test does not define “personal information.”
3. The test does not require a breach or similar incident to have occurred.
4. Causation in fact is not part of the test.

6. These issues are discussed in more detail in Section II.B, *infra*.

5. Similarly, proximate cause is irrelevant to application of the test.
6. Although the test addresses the issue of “harm,” it does not address the issue of “damages.”
7. The test does not address whether any particular information steward has an obligation to maintain “reasonable” security for personal information.
8. Legal fault and liability are not part of the test.

I. THE TEST

The proposed test for reasonable security is designed to be consistent with models for determining reasonableness that have been used in various other contexts by courts, in legislative and regulatory oversight, and in information security control frameworks. All of these regimes use a form of risk analysis to balance cost and benefit. The proposed test provides a practical method for expressing cost/benefit analysis that can be applied in data security regulatory actions, to litigation, and to information security practitioners using their current evaluation techniques. The *Commentary* also explains how the analysis should apply in the data security context. Because the test is rooted in commonly held principles, the drafters believe it offers methods for deriving reasonableness that are familiar to all interested parties. But it should be noted that depending on their text, individual laws or rules that require reasonable security might require use of a different analysis.

Since the organizations addressed in this paper are, by definition, those that have or are alleged to have an obligation to maintain reasonable security for personal information, this *Commentary* refers to them below as “information stewards.”

As described below, two particular points warrant acknowledgement: (1) courts have often looked to industry customs to inform a reasonableness analysis;⁷ and (2) in some instances, legislatures and regulatory agencies have already identified particular security measures or “controls” to be worth the cost of implementation and have required them.⁸ In

7. See, e.g., *McDermott v. Connecticut*, 113 A.3d 419, 428 (Conn. 2015); *Brooks v. Beech Aircraft Corp.*, 902 P.2d 54, 63 (N.M. 1995); *Schultz v. Consumers Power Co.*, 506 N.W.2d 175, 180 (Mich. 1993); *Pierce v. Platte Clay Elec. Coop., Inc.*, 769 S.W.2d 769, 772 (Mo. 1989); and *D.L. ex rel. Friederichs v. Huebner*, 329 N.W.2d 890, 907 (Wis. 1983).

8. E.g., 201 MASS. CODE REGS. 17.04; NEV. REV. STAT. ANN. § 603A.215.

connection with these two points, this *Commentary's* position is that evidence of an information steward's noncompliance with an "industry custom" that required a specific security control as to the personal information in question, in a way that increased the risk of a security breach, will be sufficient to establish that the information steward's information-security controls for that personal information were not reasonable. Unreasonableness would remain the conclusion unless the information steward adequately counters the effect of this evidence (1) by questioning the intelligence of the custom, (2) by showing that its operation poses different or less serious risks than those occasioned by others engaging in seemingly similar activities, (3) by showing that it has adopted an alternative method for reducing or controlling risks that is at least as effective as the customary method, or (4) by establishing, through application of the cost/benefit test, that implementation of the industry-custom-required controls in question would have burdened the information steward and others by at least as much as the implementation of the controls would have benefitted the claimant and others.

Evidence of an information steward's noncompliance with a statute, regulation, or ordinance that required implementation of the specific controls for the personal information in question will be sufficient to establish a presumption that the information steward's information security for that personal information was not reasonable. The force of such a presumption will depend on the application of the governing substantive law, which might include the doctrine of negligence per se that many states in the United States have adopted in one form or another. If permitted by applicable law, such presumption could be rebutted if the information steward establishes, by applying the cost/benefit test, that implementation of the legally required controls would have burdened the information steward and

others by at least as much as the implementation of the controls would have benefited the claimant and others.

Further, the test addresses the fact that information-security risks stemming from the absence of a control may affect more than just the claimant. The public may have its own risks; even the information steward may have some. The corollary also applies: controls that place burdens on information stewards can place the same or different burdens on the claimant and the public. To deal with this, the test compares the risks and burdens for all parties while protected by the control to the risks and burdens for all parties without the control.

A. *Articulation of the Test*

An information steward's information security controls for personal information are not reasonable when implementation of one or more additional or different controls would burden the information steward and others by less than the implementation of such controls would benefit the claimant and others.

The test may be expressed as a formula similar to the rule that Judge Learned Hand famously summarized in *United States v. Carroll Towing Co.*:⁹

$$B_2 - B_1 < (P \times H)_1 - (P \times H)_2$$

Where B represents the burden, P represents the probability of harm, H represents the magnitude of harm, subscript 1 represents the controls (or lack thereof) at the time the information steward allegedly had unreasonable security in place, and subscript 2 represents the alternative or supplementary control.

9. 159 F.2d 169, 173 (2nd Cir. 1947). The *Commentary* provides a detailed consideration of *Carroll Towing* at p. 379, *infra*.

“Burden” to the information steward and others from implementation of one or more additional controls is the net burden on the information steward and others that likely would result from such implementation. The calculation is the product of the cost or value of such burden and the likelihood of such burden resulting from the implementation of the controls. The burden would include (1) the incremental cost to the information steward and others of implementing the controls in question, (2) the cost or value to the information steward and others of any other lost or diminished, or any gained or increased, utility by reason of the implementation of such controls, and (3) the cost of new threats that may be introduced by the controls.

“Benefit” to the claimant and others from implementation of one or more additional controls means the net benefit to the claimant and others that likely would result from such implementation. The calculation is the product of the cost or value of such benefit and the likelihood of such benefit resulting from the implementation of the controls. The benefit would include (1) the incremental value to the claimant and others resulting from the implementation of the controls in question as measured by the magnitude of the harm they would likely incur from unauthorized access to or disclosure or use of the information in question in the absence of the controls, and (2) the cost or value to the claimant and others of any lost or diminished, or any other gained or increased, utility by reason of the implementation of such controls.

An information steward is not responsible for failing to address risks that were neither known nor reasonably knowable at the time of the alleged violation of the duty to provide reasonable security.

B. *Explanation of the Test*

1. Controls

The controls being evaluated include the known or reasonably knowable technical, physical, or administrative measures that secure or could secure the personal information in question.

2. Foresight Versus Hindsight

An information steward should not be responsible for failing to address risks that were neither known nor reasonably knowable at the time of the alleged violation of the duty to have in place reasonable security measures. In the analogous product liability context, for instance, courts frequently determine whether a defectively designed product was unreasonably dangerous by applying a risk/benefit analysis based on what was known or reasonably knowable at the time the product left the defendant's control, rather than what is known or reasonably knowable at the time of trial.¹⁰ A similar approach should apply in the data security context.¹¹

10. DAVID G. OWEN & MARY J. DAVIS, OWEN AND DAVIS ON PRODUCTS LIABILITY § 5:33 (4th ed. 2019) ("Almost all courts focusing on the issue in recent years have agreed, rejecting the hindsight test and limiting a manufacturer's responsibility to risks that are foreseeable."); Aaron Twerski & James A. Henderson Jr., *Manufacturer's Liability for Defective Product Designs: The Triumph of Risk-Utility*, 74 BROOK. L. REV. 1061, 1065 (2009) ("most American courts do not hold product sellers responsible for information not available at time of sale"). For an examination of the policy rationales for and against a "time of trial" approach, see Guido Calabresi & Alvin K. Klevorick, *Four Tests for Liability in Torts*, 14 J. LEGAL STUD. 585 (1985). In the United States, the courts applying that approach are in the minority. OWEN & DAVIS, *supra*, § 5:33; Twerski & Henderson, *supra*, at 1065.

11. See, e.g., Remarks Before the Congressional Bipartisan Privacy Caucus (statement of Fed. Trade Comm'r Maureen K. Ohlhausen), 2014 WL 585465, at *2 (Feb. 3, 2014) (noting that the FTC, in assessing whether a

Accordingly, in assessing the costs and benefits under the proposed test, an adjudicator should look to what was known or reasonably knowable at whatever points in time the information steward allegedly failed to have reasonable security in place.¹² In a case stemming from a data breach, this would normally be the time of the breach.

3. Burden to the Information Steward and Others (Costs)

The “incremental cost to the information steward and others of implementing the controls in question” would include any of the following: the out-of-pocket costs to acquire or create such controls; the labor costs to identify, implement, maintain, and monitor such controls; and the interruption of normal business operations by reason of the foregoing actions. The “cost or value to the information steward and others of any other lost or diminished, or of any gained or increased, utility” would include but not necessarily be limited to the cost or value to the information steward and others of any loss or improvement of quality of service or products by reason of the implementation of the controls in question, the cost or value of any increased or decreased risk to the information steward and others by reason of such implementation, and the harm from unauthorized access to or disclosure or use of the information in question—all to the extent such costs and values have not separately been

company’s security was “reasonable,” “examines factors such as whether the risks at issue were well known or reasonably foreseeable . . .”).

12. If the information steward previously conducted an assessment of its own data security risks, the product of that assessment may include evidence of whether a particular threat or harm was foreseeable. This *Commentary* provides three scenarios, each that use a different risk assessment approach, to illustrate how risk assessments may be used to exercise the test.

taken into account in applying the other components of the test.¹³

4. Benefit to the Claimant and Others (Benefits)

The decrease, by reason of the implementation of the controls in question, in the likelihood and/or in the magnitude of the harm the claimant and others¹⁴ would likely incur from unauthorized access to or disclosure or use of the information in question would be determined by taking into account any security risks that would have been reduced by the implementation or maintenance of the additional security controls in question as well as security risks that would have been introduced or increased by implementation of the same controls.¹⁵ The task would be to develop a “net” change in the

13. “Cost” for the purposes of this test is any interference with self-interested business goals. Business costs interfere with profitability, growth, returns on investment, and meeting strategic objectives. Costs to nonprofit organizations interfere with maintaining a balanced budget or reaching fundraising goals. “Utility” is a good provided to others. Grocery stores and restaurants provide food to their customers. A research university hospital provides many utilities, including health to patients, education to medical students, and advances in medical knowledge. A hotel provides safe, comfortable, and quiet lodging near places of interest.

14. One might ask why benefits to “others” than the claimant should come into the unreasonableness equation, as doing so might enable a claimant to predicate an unreasonable security claim entirely on the harm that the information steward’s information security practices caused or threatened to cause to persons *other than* the claimant. This concern, to the extent it is valid, can be addressed through the legal principles that govern a claimant’s standing to make the claim in question and/or the required showing of injury and causation of injury in order to prevail on that claim (all of which are issues beyond the scope of this paper).

15. A security control may reduce some risks while increasing others. For example, encrypting communications between two computers may safeguard sensitive data. But it may also obscure cyberattacks that are occurring between those computers. Controls that delay authorized users’

probability and/or magnitude of harm by reason of the implementation of such controls.

The “harm” to be taken into account here is the harm that is legally actionable under the law of the jurisdiction where the harm was incurred. The law on what constitutes legally actionable harm in the data security context is evolving. Whether and when intangible harms such as emotional distress or invasion of privacy are actionable and how such harms would be quantified are critical questions that are receiving different answers in different courts. The *Commentary* takes no position on them here. It simply notes that whatever harm is recognized as legally actionable under the law of the jurisdiction where the harm was incurred should be considered in the reasonableness analysis, as those are the harms the jurisdiction has identified as warranting a legal remedy.

The “cost or value to the claimant and others of any lost or diminished, or of any other gained or increased, utility” would include the cost or value to the claimant and others of any loss or increase of quality of service and/or products by reason of the implementation of the controls in question, the cost or value of any increased or decreased risk to the claimant and others by reason of such implementation, and the harm from unauthorized access to or disclosure or use of the information in question.¹⁶

access to sensitive data may encourage users to share data among themselves. Organizations commonly avoid implementing common safeguards because of other risks they may increase. Such technical and business considerations should be considered in the test.

16. While evaluating the risk of a breach—either at the time of the breach or in consideration of alternative or additive controls—an information steward may articulate the utility of its conduct so it may be included in its risk assessment, or presented to an adjudicator for its consideration as it exercises the test.

5. Industry Custom

“Industry custom” refers to a practice that is both generally followed within the relevant industry and sufficiently well known that the information steward may fairly be charged with knowledge of it.¹⁷

“Utility” may be understood as a benefit to the public or to an individual that results from the conduct that creates risk. Organizations presumably use personal information to provide a benefit other than their sole enrichment. For example, banks use their customer’s personal information to provide beneficial services to their individual customers. These services, and the customer’s financial goals, could not plausibly be met without the bank’s processing customer personal information. Some personal information can be analyzed, aggregated, or otherwise used to provide a broader public good, such as by schools who educate children, epidemiologists who track and control pandemics, or health-application vendors who provide individual coaching to subscribers based on the outcomes of their large user base.

When an adjudicator applies the test, parties may present an estimation of risk to the utility along with other factors such as costs of controls and harm to others. Adjudicators may evaluate the applicability and use of utility factors based on several criteria, such as whether a plaintiff or the public directly benefited from the conduct that put them at risk, and whether equally available and affordable alternatives presented a lower risk to the plaintiff or the public and therefore reduced the necessity of the information steward’s risky conduct.

An adjudicator may properly refuse to credit any forms of utility from the handling of personal information that society does not regard as appropriate, just as an adjudicator hearing an ordinary negligence action may refuse to recognize the feeling of excitement a motorist feels from racing a train towards a highway crossing. *See* RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL & EMOTIONAL HARM § 3 cmt. h (AM. LAW INST. 2010).

17. Further discussion on how courts have defined industry custom in situations where it is used to decide what was “reasonable” is discussed at length at Section II.A.1, p. 375 and pp. 382–85, *infra*. (*See, e.g., Silverpop Sys., Inc. v. Leading Market Techs., Inc.*, 641 F. App’x 849, 852 (11th Cir. 2016); *McDermott v. Connecticut*, 113 A.3d 419, 428 (Conn. 2015); *Brooks v. Beech Aircraft Corp.*, 902 P.2d 54, 63 (N.M. 1995); *Schultz v. Consumers Power Co.*,

Courts have historically seen industry custom not as conclusive, but as a relevant factor in the reasonableness inquiry.¹⁸ Industry custom is not merely an indication of whether a practice is cost-efficient; it is also evidence of acceptable, reasonable behavior. And this *Commentary* maintains evidence of it may be offered by either the claimant or the information steward in the reasonableness analysis. Courts often give industry custom significant weight. But a defendant may counter this evidence by questioning the intelligence of the custom, by showing its own operation poses risks that are less serious or altogether different than those posed by others in the same industry, or by showing it has adopted an alternative method for addressing risks that is at least as effective as the customary method.¹⁹

Evidence of industry custom should be relevant whether offered by the claimant or the information steward. Evidence offered by the claimant that the custom existed, that the custom called for implementation of the control, and that the information steward failed to adhere to the custom should be sufficient to shift the burden to the information steward to justify the lack of the control. Evidence offered by the

506 N.W.2d 175, 180 (Mich. 1993); *Pierce v. Platte-Clay Elec. Coop., Inc.*, 769 S.W.2d 769, 772 (Mo. 1989); ; D.L. *ex rel.* *Friederichs v. Huebner*, 329 N.W.2d 890, 907 (Wis. 1983); *In re City of New York*, 522 F.3d 279, 285 (2d Cir. 2008); *Sours v. Gen. Motors Corp.*, 717 F.2d 1511, 1517 (6th Cir. 1983); *cf.* *U.S. Fid. & Guar. Co. v. Plovidba*, 683 F.2d 1022, 1028–29 (7th Cir. 1982); *Hoffman v. Enter. Leasing Co. of Minn., LLC*, No. A16-869, 2017 WL 1210123, at *4 (Minn. Ct. App. June 20, 2017); *cf.* *Friendship Heights Assocs. v. Koubek*, 785 F.2d 1154, 1162 (4th Cir. 1986); and *Beard v. Goodyear Tire & Rubber Co.*, 587 A.2d 195, 199 (D.C. 1991)).

18. *See, e.g., McDermott*, 113 A.3d at 428; *Brooks*, 902 P.2d at 63; *Schultz*, 506 N.W.2d at 180; *Pierce*, 769 S.W.2d at 772; D.L. *ex rel.* *Friederichs*, 329 N.W.2d at 907.

19. RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM, *supra* note 16, at § 13 cmt. c.

information steward that an industry custom existed and that the steward adhered to it is likewise relevant. But, as discussed in Comment b to Section 13 of the Restatement (Third) of Torts: Liability for Physical & Emotional Harm, such evidence is not entitled to the same weight. As set out therein, it is conceivable the entire industry has lagged in the implementation of reasonable standards.

Private contractual requirements, such as the Payment Card Industry Data Security Standard or other private contractual standards, to the extent they meet the standard for “industry custom” set forth above, may create an industry custom.²⁰

6. Effect of Violating a Statute, Regulation, or Ordinance

Evidence of an information steward’s noncompliance with a statute, regulation, or ordinance that required the implementation of a specific control as to the personal information in question will be sufficient to establish a

20. Because risk analysis is a common practice in cybersecurity management and is often required by regulations, statutes, and information security frameworks, organizations may have conducted a risk assessment prior to a breach. The results of such *ex ante* risk analysis may be used by those organizations to counter a *prima facie* claim by Complainant, or an expert risk analysis presented by Complainant (although the adjudicator of course will be free to question the accuracy of either party’s analysis). In this regard, the cybersecurity community offers many risk-assessment methods that an organization may consider using to evaluate their risks and controls. As of this writing, methods such as the International Organization for Standardization’s ISO/IEC 27005, NIST Special Publications 800-30, Center for Internet Security Risk Assessment Method (CIS RAM), Operationally Critical Threat, Asset, and Vulnerability EvaluationSM (OCTAVE), Factor Analysis for Information Risk (FAIR), RISK IT, and Applied Information Economics (AIE) all estimate the likelihood and magnitude of harm and may be used to conduct analysis that is similar to the proposed test.

presumption that the information steward's security for that personal information was not reasonable.

This position finds support in various sources related to the doctrine of negligence per se, which has been adopted in one form or another by many states in the United States.²¹ Under this doctrine, statutes, regulations, or ordinances applicable to the conduct at issue set the applicable standard of care, and failure to comply is presumptively unreasonable.²²

Applicable law may make this presumption irrebuttable; and in those situations the adjudicator must follow the law. Where applicable law does not impose that requirement, a rebuttable presumption is better suited to the data security context. Technology and business practices change rapidly.²³ A rebuttable presumption strikes a useful balance. It allows the information steward charged with a violation the opportunity

21. *E.g.*, RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM, *supra* note 16, at §§ 14–15; RESTATEMENT (SECOND) OF TORTS, §§ 286, 288A, 288B (AM. LAW INST. 1965). The California Evidence Code explicitly creates a presumption that may be rebutted with proof that “[t]he person violating the statute, ordinance, or regulation did what might reasonably be expected of a person of ordinary prudence, acting under similar circumstances, who desired to comply with the law[.]” CAL. EVID. CODE § 669.

22. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS, *supra* note 21, at § 288B(1) (“The unexcused violation of a legislative enactment or an administrative regulation which is adopted by the court as defining the standard of conduct of a reasonable man, is negligence in itself.”); *Pratico v. Portland Terminal Co.*, 783 F.2d 255, 265 (1st Cir. 1985) (negligence per se “allows the presence of a statutory regulation to serve as irrefutable evidence that particular conduct is unreasonable.”).

23. *E.g.*, *In re LabMD*, slip op. at 14 (F.T.C. Jan. 16, 2014) (“The Commission has long recognized that information security is an ongoing process of assessing risks and vulnerabilities: no one static standard can assure appropriate security, as security threats and technology constantly evolve.”).

to demonstrate that falling out of technical compliance was reasonable because the costs of achieving such technical compliance would have matched or exceeded the benefits of doing so.²⁴ If the law allows it, the presumption that arises here should be found rebutted if the information steward establishes, by applying the cost/benefit test, that implementation of the legally required controls would have burdened the information steward and others by at least as much as the implementation of the controls would have benefitted the claimant and others.

We include statutes, regulations, and ordinances alike as potential sources for the presumption. All carry the force of law,²⁵ and the doctrine of negligence per se has recognized all three.²⁶

7. Determining Likelihood of Burden and Benefit

A cynic would say that because there is no usable framework for determining probability, the fact finder applying the proposed test will achieve the desired result by plugging in the degree of likelihood necessary to achieve it. In fact, the

24. Negligence is a question ordinarily resolved by the trier of fact, and the strict liability concept of negligence per se is an exception. There is a difference among jurisdictions as to whether the presumption created by the violation of a statute or regulation is rebuttable or not. Some larger jurisdictions, such as California, New York, and Georgia, use a rebuttable presumption standard. Some recent literature suggests that negligence per se should be abandoned. Barry L. Johnson, *Why Negligence Per Se Should Be Abandoned*, 20 N.Y.U J. LEGIS. & PUB., 247 (2017). Based on these factors, a rebuttable presumption is favored.

25. It is worth noting here that regulatory guidance, policy statements, opinion letters, and the like do not have the force of law. *See, e.g.*, *Wos v. E.M.A. ex rel. Johnson*, 568 U.S. 627, 643 (2013). As a result, violation of such regulatory pronouncements would not trigger the presumption.

26. *E.g.*, RESTATEMENT (SECOND) OF TORTS, *supra* note 21, at § 288B (“legislative enactment or an administrative regulation”); CAL. EVID. CODE § 669 (“statute, ordinance, or regulation of a public entity”).

information security community has broad experience with this. Likelihood of harm can be estimated, for example, using one of several techniques that are provided by the information security community. NIST Special Publications 800-30,²⁷ ISO 27005,²⁸ Center for Internet Security Risk Assessment Method,²⁹ Applied Information Economics,³⁰ and Factor Analysis for Information Risk³¹ all provide guidance for estimation of likelihood or probability.

8. When to Apply the Test

The cost/benefit analysis should be applied as of the time the information steward is or was allegedly violating its obligation to maintain reasonable security, and not as of the time the adjudicator is conducting the cost/benefit analysis. In a breach case, that would typically be at the time of the breach. In a case involving an agency accusation of unreasonableness not tethered to a breach, it would be as of the time of the events on which the agency's accusation is based.³²

27. NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, NIST SPECIAL PUBLICATION 800-30 REVISION 1, GUIDE FOR CONDUCTING RISK ASSESSMENTS (2012).

28. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, ISO/IEC 27005:2018, INFORMATION TECHNOLOGY—SECURITY TECHNIQUES—INFORMATION SECURITY RISK MANAGEMENT (2018).

29. CENTER FOR INTERNET SECURITY, RISK ASSESSMENT METHOD (CIS RAM) ver. 1.0 (2018).

30. DOUGLAS W. HUBBARD & RICHARD SEIERSEN, HOW TO MEASURE ANYTHING IN CYBERSECURITY RISK (1st ed. 2016).

31. JACK FREUND & JACK JONES, MEASURING AND MANAGING INFORMATION RISK: A FAIR APPROACH (2014).

32. In order to apply the test, the parties and the adjudicator will need to consider the question of over what period of time the burden and the benefit are to be measured. To the extent either the burdens or the benefits of the added security measure(s) in question would reasonably be incurred

9. Availability of Resources

While the test does not directly consider whether the information steward in question had the resources necessary to implement the additional controls that application of the test would require in order for that information steward's data security measures to be found reasonable, the availability of those resources may affect the results of the test indirectly. To explain, the "burdens" included in the test take into account the lost utility that would result from the additional controls. That being the case, if an information steward had insufficient resources to manage a high likelihood and high magnitude of harm, and if the benefit of engaging the risk were low, then the test could result in the additional controls being deemed necessary even where the information steward lacked the resources to implement them . . . and would go out of business trying to do so.

On the other hand, if a similarly under-resourced information steward provided a highly beneficial utility, then the test might demonstrate a commensurately high loss of benefit with the additional controls in place. And this could result in a finding that the information steward was not required to implement the additional controls in order to maintain reasonable security. In other words, it was reasonable to proceed without implementing the controls.

beyond the initial period (*e.g.*, "year one") into a subsequent period (*e.g.*, "out-year(s)"), those reasonably expected benefits and burdens would need to be included in the analysis. Having said that, the methodology by which the "out-year" burdens and benefits are to be factored into the analysis would be determined not by application of a pre-set formula but rather on a case-by-case basis, which would depend on the evidence presented as to the amount and duration of those burdens and benefits, the appropriateness of discounting them to present value, and, if appropriate, the manner of accomplishing such discounting.

10. Poor Implementation of a Control

There may be instances where the information steward has determined a security control is necessary but has implemented it poorly, or not at all. Indeed, such a fact pattern may occur frequently. The question presented is how the adjudicator applying the proposed test should account for the poor implementation of the control. As an example, this could occur if the information steward had determined enhanced training was required for all individuals handling certain types of data but failed to identify everyone who handled it, leaving out individuals in a given location or business unit. As another example, the information steward may have assigned responsibility appropriately, but the individual charged with implementing the control failed to do it. Under the test, the adequacy of the design is not determinative. Even an excellent design will not protect the information steward where a consideration of the costs and benefits of the failed control shows its proper implementation would have been “worth it.” The test satisfactorily addresses this issue.

11. Illustrations of the Application of the Test

To demonstrate the practical utility of the proposed test, three hypothetical illustrations are included in an Appendix. The exemplars do not represent any one case and do not name actual organizations. However, the facts, issues, and causes in each exemplar are common components of breaches in which members of Working Group 11 have been professionally involved.

The reader will note the third exemplar uses quantitative scoring based on the nonquantitative assessments of such things as potential utility, cost, and harm. An adjudicator should first look for quantitative information on both sides of the cost/benefit analysis and should endeavor to apply the

reasonable security test using quantitative information. However, information stewards do sometimes resort to nonquantitative inputs in order to conduct a cost/benefit analysis. The *Commentary* takes no position on how an adjudicator should apply the test in a situation where it does not have quantitative information available, or on whether the adjudicator should do so at all. The third exemplar is included only to illustrate how an adjudicator might apply the test where quantitative information was not available.

II. DISCUSSION

A. *The Work That Led to the Test*

Extensive, separate reviews of the treatment of reasonable security were conducted in three distinct areas: (1) judicial opinions; (2) statutes and regulations; and (3) the marketplace. A summary of that work follows.

1. Judicial Opinions

A review of judicial opinions in which courts considered the issue of reasonable security highlights the benefits of articulating a test to determine what it is.

In *LabMD v. Federal Trade Commission*, the Eleventh Circuit overturned a cease-and-desist order the FTC had entered requiring LabMD to implement “reasonable” data security.³³ The court held that the reasonableness requirement in the FTC’s order, which did not specify what measures are “reasonable” or set forth a standard for “reasonableness,” was so vague that being subject to penalties for violating it could violate due process: it subjected LabMD to the prospect of being found in violation of the order without having been given fair notice of what conduct is prohibited. The court added it would also be impossible for the FTC to enforce the order as a practical matter. Without a governing standard for reasonableness, a court would have no way to determine whether LabMD violated the order.

Several earlier data security cases suggested a standard for reasonable data security, but only in discrete contexts. In *Federal Trade Commission v. Wyndham Worldwide Corp.*, the FTC asserted what Wyndham did was “unreasonable” and thus “unfair” under Section 5 of the FTC Act. Wyndham responded it lacked

33. No. 16-16270, 2018 WL 3056794, at *7–12 (11th Cir. June 6, 2018).

fair notice of what data security measures the FTC claimed were reasonable. Here, the Third Circuit concluded the “unfairness” provision of Section 5 at issue in *Wyndham* provided sufficient notice as to what conduct would comply with its requirements for purposes of Wyndham’s motion to dismiss: the text of Section 5 expressly cabins the FTC’s authority to declare an act unfair to situations where the act or practice in question “causes or is likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition.”³⁴ Finding the statute “informs parties that the relevant inquiry here is a cost-benefit analysis that considers a number of relevant factors, including the probability and expected size of reasonably unavoidable harms to consumers given a certain level of cybersecurity and the costs to consumers that would arise from investment in stronger cybersecurity,”³⁵ the Third Circuit rejected Wyndham’s position.

In *Silverpop Systems, Inc. v. Leading Market Technologies, Inc.*, the Eleventh Circuit held in an unpublished opinion that the plaintiff’s negligence claim against its service provider, which suffered a cybersecurity breach, failed at the summary judgment stage because the plaintiff “failed to present evidence to establish the applicable standard of care.”³⁶ Observing that “evidence of custom within a particular industry, group, or organization is admissible as bearing on the standard of care in determining negligence,” the court noted the plaintiff failed to identify any “standards that are ordinarily employed in [the defendant’s] industry.”³⁷ Accordingly, as the plaintiff “failed to

34. 799 F.3d 236, 255–59 (3d Cir. 2015) (quoting the statute).

35. *Id.* at 255.

36. 641 F. App’x 849, 852 (11th Cir. 2016).

37. *Id.*

present evidence establishing the standard of care,” it could not “establish a breach of the standard of care.”³⁸

In *Razuki v. Caliber Home Loans, Inc.*, the court held “Razuki could have identified what made Caliber’s security measures unreasonable by comparison to what other companies are doing.”³⁹

Additional decisions have likewise pointed to industry custom or standards as a potentially relevant consideration.⁴⁰

In the context of an order that could subject a party to contempt sanctions for failing to have reasonable cybersecurity, *LabMD* suggests “reasonableness” currently has no enforceable meaning. *Wyndham* clarifies that “reasonableness” has meaning to the extent it is the standard for unfairness liability under Section 5 of the FTC Act, since Section 5 itself expressly sets forth a cost/benefit test. *Silverpop* and other private data-security litigation cases show industry standards and/or industry custom play a role in an analysis of “reasonable data security.”

In *Dittman v. University of Pittsburgh Medical Center (UPMC)*,⁴¹ the Pennsylvania Supreme Court recognized a negligence-based duty to safeguard personally identifiable information (PII) where the plaintiffs alleged the employer

38. *Id.*

39. No. 17CV1718-LAB (WVG), 2018 WL 6018361, at *1 (S.D. Cal. Nov. 15, 2018).

40. *See, e.g.*, *Dugas v. Starwood Hotels & Resorts Worldwide, Inc.*, No. 316CV00014GPCBLM, 2016 WL 6523428, at *10 (S.D. Cal. Nov. 3, 2016) (denying dismissal where plaintiffs alleged that defendants “failed to employ reasonable security measures to protect . . . PII, such as the utilization of industry-standard encryption”); *In re Hannaford Bros. Co. Customer Data Sec. Breach Litig.*, 613 F. Supp. 2d 108, 119 (D. Me. 2009) (“reasonable” security “might include meeting industry standards”), *aff’d in part, rev’d in part on other grounds sub nom.* *Anderson v. Hannaford Bros. Co.*, 659 F.3d 151 (1st Cir. 2011).

41. 196 A.3d 1036 (Pa. 2018).

(UPMC) required its employees “to provide certain personal and financial information, which UPMC collected and stored on its internet-accessible computer system without use of adequate security measures, including proper encryption, adequate firewalls, and an adequate authentication protocol.”⁴²

A trio of recent cases from the Northern District of Georgia embraced the view that federal statutes and regulations can provide an ascertainable standard of conduct for state-law claims of negligence per se.⁴³ These cases looked to both Section 5 of the FTC Act and, in one case, to the Safeguards Rule of the Gramm-Leach-Bliley Act as providing an applicable standard of conduct.⁴⁴ These cases also found a negligence-based duty under Georgia law to provide reasonable security.⁴⁵ An intervening Georgia Supreme Court case appears to negate such a duty but does not affect the Northern District’s findings regarding negligence per se.⁴⁶

A review of case law where a standard of reasonableness was applied outside the data security context showed two

42. *Id.* at 1047; *but see* Cooney v. Chicago Pub. Sch., 943 N.E.2d 23, 28-29 (Ill. App. Ct. 2010) (no duty to safeguard personal information under Illinois law).

43. *In re* Equifax, Inc., Customer Data Sec. Breach Litig., 362 F. Supp. 3d 1295, 1325 (N.D. Ga. 2019); *In re* Equifax, Inc., Customer Data Sec. Breach Litig., 371 F.Supp.3d 1150 (N.D. Ga. 2019); *In re* Arby’s Rest. Grp. Inc. Litig., No. 1:17-CV-0514-AT, 2018 WL 2128441, at *5 (N.D. Ga. Mar. 5, 2018).

44. *In re* Equifax, 362 F. Supp. 3d at 1327; *In re* Equifax, 371 F. Supp. 3d, 1173-76; *In re* Arby’s, 2018 WL 2128441, at *8; *but see, e.g., In re* Supervalu, Inc., 925 F.3d 955, 963 (8th Cir. 2019) (violation of Federal Trade Commission Act (FTCA) could not establish breach of duty for negligence claim in data breach case part because “Congress empowered the Commission—and the Commission alone—to enforce the FTCA. Implying a cause of action would be inconsistent with Congress’s anticipated enforcement scheme.”).

45. *E.g., In re* Arby’s, 2018 WL 2128441, at *5.

46. *Dep’t of Labor v. McConnell*, 828 S.E.2d 352, 358 (Ga. 2019).

approaches: a cost/benefit test and a consideration of industry custom.

Judge Learned Hand famously summarized the test for reasonableness with his algebraic expression in *United States v. Carroll Towing Co.*⁴⁷ *Carroll Towing* considered whether the owner of a barge should be held liable when the barge broke away from its moorings while the bargee was absent. Recognizing there would be occasions when a barge breaks away from its moorings, the potential liability of the barge owner involved the assessment of (1) the probability of the barge breaking away, referred to as “P,” (2) the gravity of the loss if the barge did break away, referred to as “L,” and (3) the burden of adequate precautions, referred to as “B.” Liability would seem to be warranted when B (the cost of adequate precautions) is less than the product of P multiplied by L.

The test in *Carroll Towing* is keyed to applying safeguards that are no more burdensome than the risks they protect against. Thus, the burden of the safeguards must not be greater than the probability and liability of the harmful event. And while the harm from a barge that escapes its moorings is almost always more determinable than the harm from sensitive, personal information that escapes its server, there is nevertheless good reason to believe that the Learned Hand Formula can be usefully applied to both.⁴⁸

47. 159 F.2d 169, 173 (2nd Cir. 1947).

48. In other cases, Judge Hand questioned, or even rejected, the quantitative test outlined in *Carroll Towing* as being unworkable. See, e.g., *Conway v. O'Brien*, 111 F.2d 611, 612 (2d Cir. 1940), *rev'd on other grounds*, 312 U.S. 492 (1941). In *Moisan v. Loftus*, 178 F.2d 148, 149–50 (2d Cir. 1949), for example, authored by Judge Hand after *Carroll Towing*, he recognized the “inherent uncertainties . . . in applying such a formula” to an “incommensurable subject matter.” But even then, in *Moisan*, he supported the *Carroll Towing* test and observed that, if nothing else, the test is helpful in

Product liability cases were examined as well. At least one court has rejected the adoption of a strict liability test in the data breach context.⁴⁹ Nonetheless, the case law and scholarship associated with product liability cases is useful in supporting a reasonable security test resting on a cost/benefit analysis.⁵⁰ For example, Section 2 of the *Restatement (Third) of Torts: Products Liability* provides that a product is defective in design where the foreseeable risks of harm could have been reduced or avoided with a reasonable alternative design.⁵¹ That section, the *Restatement* continues,

adopts a reasonableness (“risk-utility balancing”) test as the standard for judging the defectiveness of product designs. More specifically, the test is whether a reasonable alternative design would, at reasonable cost, have reduced the foreseeable risks of harm posed by the product and, if so, whether the omission of the alternative design by the seller or a predecessor in the

identifying which of those factor(s) will be determinative in any given case. *Id.* at 149.

This *Commentary* and its proposed test draw inspiration from *Carroll Towing* while noting the difficulties that may arise in a strictly quantitative application of the test. In that regard, the *Restatement (Third) of Torts* section 3(e) and the accompanying Reporters’ Note for section 3(d) use *Carroll Towing*, *Moisan*, and *Conway* as examples of courts’ applying the *Restatement’s* proposed cost/benefit approach to negligence determinations. Such authority provides further support for the approach proposed in this *Commentary*.

49. See *In re Hannaford Bros. Co. Customer Data Sec. Breach Litig.*, 613 F. Supp. 2d 108, 119, 125 (D. Me. 2009).

50. See Mark A. Geistfeld, *Protecting Confidential Information Entrusted to Others in Business Transactions: Data Breaches, Identity Theft, and Tort Liability*, 66 DEPAUL L. REV. 385, 399–401 (2016).

51. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 2(b) (AM. LAW INST. 2012).

distributive chain rendered the product not reasonably safe.⁵²

The case law outside the data security context also recognizes a defendant's compliance with or departure from industry custom is evidence either of due care or negligence but is not dispositive.⁵³

This view of industry custom has been adopted by the leading commentators.⁵⁴

52. *Id.* at cmt. d.

53. *See, e.g.,* *McDermott v. Connecticut*, 113 A.3d 419, 428 (Conn. 2015) ("The trier of fact is not bound by the industry standard, but rather should consider it in light of the totality of the evidence presented in the case."); *Brooks v. Beech Aircraft Corp.*, 902 P.2d 54, 63 (N.M. 1995) ("We adhere to the principle that evidence of industry custom or usage, and evidence of compliance with applicable regulations, is relevant to whether the manufacturer was negligent or whether the product poses an unreasonable risk of injury, but that such evidence should not conclusively demonstrate whether the manufacturer was negligent or the product was defective."); *Schultz v. Consumers Power Co.*, 506 N.W.2d 175, 180 (Mich. 1993) ("While it may be evidence of due care, conformity with industry standards is not conclusive on the question of negligence where a reasonable person engaged in the industry would have taken additional precautions under the circumstances."); *Pierce v. Platte-Clay Elec. Coop., Inc.*, 769 S.W.2d 769, 772 (Mo. 1989) ("[E]vidence of industry standards is generally admissible as proof of whether or not a duty of care was breached. However, compliance with an industry's own safety standards is never a complete defense in a case of negligence."); *D.L. ex rel Friederichs v. Huebner*, 329 N.W.2d 890, 907 (Wis. 1983) ("Customary practice is not ordinary care but is evidence of ordinary care.").

54. *See* RESTATEMENT (THIRD) OF TORTS: PHYSICAL & EMOTIONAL HARM, *supra* note 16, at § 13 ("An actor's compliance with the custom of the community . . . is evidence that the actor's conduct is not negligent but does not preclude a finding of negligence. An actor's departure from the custom of the community . . . in a way that increases risk is evidence of the actor's negligence but does not require a finding of negligence."); 57A AMERICAN JURISPRUDENCE 2d, Negligence § 165 (2019) ("[C]ompliance or noncompliance with customary or industry practices is not dispositive of the

The *Restatement (Third) of Torts: Physical & Emotional Harm* states “there is no minority rule,” and modern decisions frequently cite Justice Holmes’s opinion in *Texas & Pacific Railway Co. v. Behymer*, and Judge Hand’s opinion in *The T.J. Hooper*.⁵⁵

While industry custom is not conclusive on the issue of reasonableness, it often has “significant weight.”⁵⁶ However, a “party who has departed from custom can counter the effect of this evidence by questioning the intelligence of the custom, by showing that its operation poses different or less serious risks than those occasioned by others engaging in seemingly similar activities, or by showing that it has adopted an alternative method for reducing or controlling risks that is at least as effective as the customary method.”⁵⁷

issue of due care, but constitutes only some evidence thereof.”); WILLIAM LLOYD PROSSER & W. PAGE KEETON, PROSSER & KEETON ON TORTS § 33 (5th ed. 1984) (“Much the better view, therefore, is that of the great majority of cases, that every custom is not conclusive merely because it is a custom, that must meet the challenge of learned reason, and be given only the evidentiary weight which the situation deserves. . . . But, as a general rule, the fact that a thing is done in an unusual manner is merely evidence to be considered in determining negligence and is not in itself conclusive.”).

55. RESTATEMENT (THIRD) OF TORTS: PHYSICAL & EMOTIONAL HARM, *supra* note 16, at § 13 Reporter’s Note cmt. b; *see also* *Texas & Pacific Ry. Co. vs. Behymer*, 189 U.S. 468, 470 (“What usually is done may be evidence of what ought to be done, but what ought to be done is fixed by a standard of reasonable prudence, whether it usually is complied with or not.”); *The T.J. Hooper*, 60 F.2d 737, 740 (“Indeed in most cases reasonable prudence is in fact common prudence; but strictly it is never its measure; a whole calling may have unduly lagged in the adoption of new and available devices. It never may set its own tests, however persuasive be its usages.”).

56. RESTATEMENT (THIRD) OF TORTS: PHYSICAL & EMOTIONAL HARM, *supra* note 16, at § 13 Reporter’s Note cmt. c.

57. *Id.*

In general, industry custom relates to the feasibility and acceptance of alternative measures and whether the defendant was, or should have been, aware of those measures.⁵⁸ In addition, if the defendant complied with industry custom, this fact cautions the jury that its ruling on the particular actor's negligence has implications for large numbers of other parties. A companion caution is that the industry "may have been pursuing self-interest in a way that has encouraged the neglect of a reasonable precaution."⁵⁹

Industry custom is an important factor the adjudicator would take into account in determining whether the defendant exercised reasonable care. But industry standards are not dispositive.⁶⁰

In the context of contractual relationships, merchants bargain against the backdrop of industry custom, and those customs will be implied in a contract unless the agreement indicates a specific intent to depart from them.⁶¹ Even in tort cases, the existence of a contractual or other special relationship between the plaintiff and the defendant can affect the weight

58. *Id.*, cmts. b & c.

59. *Id.*, cmt. b.

60. See *In re City of New York*, 522 F.3d 279, 285 (2d Cir. 2008) ("Fortunately, we need not reason from a blank slate in applying the Hand formula; we can look to guideposts like industry custom and government regulations in determining the standard of care . . ."); *Sours v. Gen. Motors Corp.*, 717 F.2d 1511, 1517 (6th Cir. 1983) ("GM's alleged compliance with FMVSS 216, along with its other evidence of adherence to industry customs and standards, was properly left to the jurors to factor into the calculus that comprises reasonable design in a case of strict products liability."); cf. *U.S. Fid. & Guar. Co. v. Plovidba*, 683 F.2d 1022, 1028–29 (7th Cir. 1982) (Posner, J.) (observing that, at least under a no-fault liability regime, industry practice should reflect efficient risk allocation).

61. See RESTATEMENT (SECOND) OF CONTRACTS § 220 cmt. f & § 222 (AM. LAW INST. 1981).

given to industry standards: “The prospect of unreasonable conduct by all potential defendants who engage in a line of activity is especially great when potential victims do not enter into contractual or other consensual relationships with those defendants. By contrast, when potential victims are the patrons of defendants who engage in a particular line of commercial activity, the customs that those defendants accept might be expected to give considerable weight to their patrons’ desires.”⁶² Likewise, in professional malpractice cases, the standard of care is largely defined by professional standards and customs, although industry custom would be given less weight in a products liability case.⁶³

The case law also explains what industry practices constitute an “industry custom” for this purpose. William Lloyd Prosser and W. Page Keeton state in *Prosser & Keeton on Torts*: “A custom, to be relevant, must be reasonably brought home to the actor’s locality, and must be so general, or so well known, that the actor may be charged with knowledge of it or with negligent ignorance.”⁶⁴ That a few members of the industry may use a particular safety measure is not sufficient to show a custom.⁶⁵ An industry standard that is not generally followed or that is

62. RESTATEMENT (THIRD) OF TORTS: PHYSICAL & EMOTIONAL HARM, *supra* note 16, at § 13 cmt. b.

63. *See id.*

64. PROSSER & KEETON, *supra* note 54, at § 33.

65. *See In re City of N.Y.*, 522 F.3d at 285 (“And while the precautions taken by the one ferry operator with ships comparable to the Staten Island Ferry may be prudent, those practices have not become universal enough to suggest an industry custom.”); *The T.J. Hooper*, 60 F.2d 737, 740 (2d Cir. 1932) (Hand, J.) (“[H]ere there was no custom at all as to receiving sets; some had them, some did not; the most that can be urged is that they had not yet become general.”).

merely aspirational will not establish industry custom.⁶⁶ But neither do industry standards require 100 percent adherence by the industry members in order to become recognized as industry custom.⁶⁷

2. Statutes and Regulations

A broad set of U.S., Canadian, Australian, and European privacy legislation was reviewed to identify themes employed there. The review focused in particular on requirements for the protection of personal information that were common across the several statutory regimes.

66. See *Hoffman v. Enter. Leasing Co. of Minn., LLC*, No. A16-869, 2017 WL 1210123, at *4 (Minn. Ct. App. June 20, 2017) (unpublished) (expert failed to demonstrate that industry recommendations rose “any higher than best practices” or were “relied on or followed in the rental-car or tire-repair industry”).

67. Cf. *Friendship Heights Assocs. v. Koubek*, 785 F.2d 1154, 1162 (4th Cir. 1986) (the standard of care could be shown “through testimony describing steps ordinarily taken” by members of the profession); *Beard v. Goodyear Tire & Rubber Co.*, 587 A.2d 195, 199 (D.C. 1991) (evidence that the merchants’ own “procedures conform to those generally used by members of their industry, or at least by many of them” was relevant to the standard of care).

Courts appear to use terms like “industry custom,” “industry standard,” and the like interchangeably, or as equivalents. See, e.g., *In re City of N.Y.*, 522 F.3d at 285 (referring to “[c]ustom or standard practice in the industry”); *Tzilianos v. N.Y. City Transit Auth.*, 936 N.Y.S.2d 159, 161 (N.Y. App. Div. 2012) (referring to “an industry standard or a generally accepted safety practice”). For the purposes of this *Commentary*, the term “industry custom” is preferable because it tracks the language used by the *Restatement (Third) of Torts: Physical & Emotional Harm*. Terms like “industry standard” may imply a formal standard, which is not necessary to establish industry custom, and may not be sufficient to do so. See *RESTATEMENT (THIRD) OF TORTS: PHYSICAL & EMOTIONAL HARM*, *supra* note 16, § 13; see also *Hoffman*, 2017 WL 1210123, at *4. Terms like “common practice” are vague and could cover situations in which the practice has been adopted by only a minority of industry members. See *The T.J. Hooper*, 60 F.2d at 740.

Here are key findings:

- (a) Sensitivity of information: Personal information should be protected by safeguards appropriate to the sensitivity of the information. More sensitive information is expected to be safeguarded by a higher level of protection.
- (b) Availability of resources: The size, sophistication, and availability of resources of an information steward can be relevant to what is required in given circumstances.
- (c) Cost/benefit analysis: Reasonable security entails consideration of the sensitivity of the information, the associated risk of harm arising from unauthorized access to it or from the deprivation, loss or destruction of the information, the available measures to protect the information, and the cost of those measures to the information steward.
- (d) Industry standards: Industry standards may be considered to determine what is reasonable in a particular context.

Examples of legislative requirements that appear throughout the sources include the following:

- Comprehensive, written, information-security program/policies;
- Commitment to protect information through “reasonable” security measures;
- Designation of responsible person(s);
- Performance of risk assessment;

- Restrictions on physical access to personal information;
- Encryption of sensitive personal information;
- Incident response planning;
- Limiting access privileges to those with a need to know;
- Employee training and compliance monitoring;
- Evaluating and improving the means for detecting and preventing security system failures;
- Disciplinary measures for violations;
- Oversight of the data security practices of third parties, subcontractors, vendors, and the like; and
- Secure user-authentication protocols.

Even where the statutes/regulations set out specific requirements for the protection of personal information, a determination of what is reasonable in a particular circumstance is always required. A “check-here-and-you’re-done” form does not exist.

The Ohio Data Protection Act is of great interest. In Ohio, an information steward can claim a “safe harbor” against tort claims if it has “reasonably conformed” with a specified, industry-recognized cybersecurity framework. However, the Ohio Data Protection Act relies on the same factors found in other statutes/regulations. In particular, the Act provides that the scale and scope of a covered information steward’s cybersecurity program is appropriate if it is based on all of the following factors:

- Size and complexity of the covered information steward;

- Nature and scope of the activities of the covered information steward;
- Sensitivity of the information to be protected;
- Cost and availability of tools to improve information security and reduce vulnerabilities; and
- The resources available to the covered information steward.

It's important to note that the Ohio Data Protection Act does not specify how these factors are to be prioritized when determining whether the information steward has "reasonably conformed" to the industry-recognized cybersecurity framework. For example, if an information steward has highly sensitive personal information but limited resources, will it be afforded a safe harbor if it does not implement the entire industry-recognized framework?

Overall, the themes embedded in the statutes and regulations provided useful guidance for assessing reasonable security, but they did not make clear how the several principles should be weighed against each other.

3. Marketplace

Marketplace standards of reasonable conduct in cybersecurity preparedness included the following approaches: (a) mandated minimum controls; (b) prescriptive but flexible controls; (c) standards/frameworks and; (d) open requirements. Here are examples of each:

(a) Mandated Minimum Controls

- The Payment Card Industry Data Security Standard requires specific technical controls for

information stewards that handle payment card information.

- The National Institute of Standards and Technology Special Publication 800-171 is a list of required controls that federal contractors must apply when safeguarding “Sensitive but Unclassified” data. These controls are a subset of NIST SP 800-53 and apply to what NIST believes are the most common causes of security concerns federal agencies encounter with their contractors.

(b) Prescriptive But Flexible Controls

- A familiar example is the Health Insurance Portability and Accountability Act (HIPAA) Security Rule, which requires covered information stewards to ensure the confidentiality, integrity, and availability of electronic Protected Health Information. But the Rule allows “flexibility of approach” in how that data protection is achieved, based on the information steward’s size, complexity, and other factors such as risk.
- The Center for Internet Security Controls (CIS Controls) lays out no fewer than 20 high-level controls, each of which contains subordinate implementation guidance.

(c) Standards/ Frameworks

- Some information security standards provide listings and descriptions of controls. For example, the NIST Cybersecurity Framework (NIST CSF) includes high-level control groupings (Identify, Protect, Detect, Respond, Recover) but does not

require specific, technical controls. Instead, NIST CSF subcategories reference specific controls from other standards, such as the CIS Controls, ISO 27001, and NIST SP-800-53.

- Other information security standards describe how to analyze information security risks so that controls can be implemented in a way that is reasonable or acceptable for each environment. NIST SP 800-30 and ISO 27005 provide guidance for evaluating controls for their risk acceptability, while the CIS Risk Assessment Method provides guidance for evaluating controls for their reasonableness. Some methods such as Factor Analysis for Information Risk and Applied Information Economics help quantify information security risks.
- The Federal Financial Institutions Examination Council Cybersecurity Assessment Tool (CAT) identifies controls that should be found in commercial and retail banks and organizes them in five different maturity levels. The CAT classifies banks by size, complexity, and volume of business, then indicates the maturity of controls that banks in those classifications should achieve.

(d) Open Requirements

- An excellent example is the European Union's General Data Protection Regulation (GDPR), whose language notes that, "considering the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and

the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk.”

- The Australian Essential Eight Maturity Model was developed by the Australian Signals Directorate. The idea of the Essential Eight is to implement a “baseline” of eight cyber-threat mitigation strategies that can be deployed against common threats in a cost-effective way. These include application whitelisting, patching, and restricting administrative privileges. The Essential Eight is a one-size-fits-all approach, which has the benefits of simplicity and broad applicability.
- The U.S. Department of Health and Human Services has published guidance documents (“HHS Guidance”) on best practices for health-care information stewards to reduce cybersecurity risks. The HHS Guidance outlines prevalent threats to the health-care industry and identifies best practices to mitigate these threats. The HHS Guidance identifies the following ten specific practices to be considered by an information steward according to its size, complexity, and type:
 - (a) email protection systems
 - (b) endpoint protection systems
 - (c) access management
 - (d) data protection and loss protection
 - (e) asset management
 - (f) network management
 - (g) vulnerability management

- (h) incident response
- (i) medical device security
- (j) cybersecurity policies

The HHS Guidance includes a discussion of specific examples of steps a health-care information steward may take in the context of each of the ten practices but does not identify any framework for assessing what comprises reasonable security.

B. All the Things “Ruled Out”

As the drafters reviewed and discussed their sources and moved on to developing the proposed test, some things were ruled out. Included were:

1. Specific Controls

This project began and ended with the belief there is not and should not be a one-size-fits-all cybersecurity program. Because application of the proposed test will necessarily depend on the particular circumstances faced by the information steward, mandating particular controls would be inconsistent with a cost/benefit approach.

2. Definition of Personal Information

The proposed test does not seek to define personal information but is intended to be flexible enough to apply to any definition of “personal information.”

3. Breach Requirement

Consideration was given to whether a reasonableness test should apply only when a breach or incident has actually occurred. But there are many instances where a determination of reasonableness is important regardless of whether a breach

has occurred. In addition, the proposed cost/benefit test does not focus on, nor is it limited to, the occurrence of a breach. Rather, the test focuses on the probability and magnitude of the costs and benefits that would reasonably have been expected to flow from the adoption and implementation of the additional or different security measures under consideration.

4. Causation in Fact

Just as the proposed test does not depend on the presence of a breach, the operation of the cost/benefit analysis is separate from the issue of causation in fact. The cost/benefit analysis addresses reasonably expected costs and benefits with an eye toward the potential for a breach, rather than looking for and focusing on what caused the breach. Indeed, since the test can be applied whether or not there is a breach, it can be applied whether or not causation in fact is an issue that needs resolution.

That causation in fact is not a necessary part of the test becomes concrete where the presence or absence of a particular control is blamed for an incident. Post-incident analyses invariably conclude that implementation of one or more controls could have prevented the incident. But a breach can take any one of many paths. That a brilliant attacker found a new door to walk through should not in and of itself mean the information steward failed to implement reasonable security. Under the test, then, the question is never whether absence of a particular control is to blame for an incident. Instead, the test is always whether, at the time of the incident, the reasonably anticipatable benefits of the control in question outweighed its reasonably expected costs.

Because it was concluded that causation in fact is not necessary to an inquiry into whether the security was reasonable, it was not incorporated as part of the proposed test. Still, in saying that, it is recognized that in many cases the

claimant will need to prove the information steward's unreasonable security controls were a but-for cause of the injury on which the claimant's claim is based.

5. Proximate Cause

Consideration of proximate cause was excluded because, like causation in fact, it is irrelevant to application of the cost/benefit test. Again, it is acknowledged that in many cases the claimant will need to prove the information steward's unreasonable security controls were the proximate cause of the injury on which the claimant's claim is based.

6. Damages

The issue of "damages" is not addressed as a component of the test, but "harm" is included. The concepts are related, but different. While proof of actual damages (or for that matter actual harm or injury) is not necessary to application of the cost/benefit test, in many cases the claimant may be able to use such proof. It could establish the magnitude of reasonably foreseeable harm to the claimant and others that was potentially avoidable by implementation of the additional controls in question; and it could establish that the information steward's unreasonable security caused the injury and damages to the claimant.

7. Existence of Obligation to Have "Reasonable" Security

The *Commentary* takes no position as to whether any particular information steward is, in fact, under an obligation to maintain reasonable security for personal information. While it is indisputable that some are under such an obligation, that is not clear for all information stewards.

8. Fault/Liability

If the application of the test results in a finding that the information steward did not maintain reasonable security, it does not necessarily follow that the information steward is “at fault” and liable to the claimant, or subject to some adverse finding and penalty by a regulator or court. Legal fault, and any liability that may flow from it, will be determined according to the law applicable to the claim in question. In order for there to be liability under the applicable law, a claimant may need to show fault or other culpability on the part of the information steward in addition to a showing that the information steward’s security for personal information was unreasonable. For example, if the information steward acted in response to advice from experienced third-party consultants and attorneys, that “advice of counsel” might provide a complete defense. The *Commentary* takes no position on whether a showing of fault or other culpability on the part of the information steward is required to impose liability on an information steward for failure to have reasonable security for personal information.⁶⁸

68. On a related but different note, just as the test would not require an information steward to implement a particular control where the burden of doing so is greater than or equal to the benefit to be derived from it, one could argue the steward should still have responsibility in this setting. Under this line of thinking, where the costs of employing a control are \$100,000 and the probability-adjusted costs to others from not employing it are \$100,000, and the information steward who declines the control is found to have reasonable security . . . but will also have saved \$100,000, individuals who are impacted by the absence of the control should be compensated up to the limit of the savings. In response, another could argue that such position would make the information steward a guarantor against some degree of loss, no matter how reasonable its security. While it is not the position of this *Commentary* that the information steward should always have responsibility to a claimant, irrespective of the reasonableness of its security, the *Commentary* acknowledges that such an argument exists.

C. *The Importance of Flexibility*

If one accepts there is no one-size-fits-all cybersecurity program, it follows that a reasonableness test must be flexible.

Some of the flexibility factors that were identified include:

1. The Data to Be Protected

As the loss or compromise of different types of data presents different kinds of harm, different levels of protection are appropriate. The source or owner of the data should also be considered. An information steward holding data about others, particularly personal data, must consider the value of that data to the owners and to itself. Maybe the information steward should not hold the information in the first place. If it does hold the information, and if the information is sensitive enough, the information steward may not only be obligated to employ the very highest level of protection but may also have to pay damages no matter how or why the information was compromised—the so-called “plutonium covenant.” Conversely, if the data held belongs to the information steward—such as intellectual property—then absent law, regulation, industry standard, or fiduciary obligation to shareholders, the information steward should have considerable flexibility in how to protect it.

2. Threats and Risks

Bad actors have varying levels of sophistication and resources. Protecting against a sophisticated team operating at the nation-state level may well be impossible. Still, as nation states do not threaten the majority of information stewards, threat identification can be an important component of evaluating reasonableness, as it will inform the analysis of what threats were reasonably knowable at the time of the claimed

violation, and what threats were not. Such an analysis is important to the application of the proposed cost/benefit test.

III. CONCLUSION

In the data security space, the reasonableness of a protection has a kind of half-life, and probably a short one. Even regulators concede the point. As set out in the Cybersecurity Requirements for Financial Services Companies:⁶⁹

Given the seriousness of the issue and the risk to all regulated entities, certain regulatory minimum standards are warranted, while not being overly prescriptive so that cybersecurity programs can match the relevant risks and keep pace with technological advances. Accordingly, this regulation is designed to promote the protection of customer information as well as the information technology systems of regulated entities. This regulation requires each company to assess its specific risk profile and design a program that addresses its risk in a robust fashion. [Emphasis added.]

While there is some guidance for assessing reasonable security in the existing judicial opinions, in statutes and regulations, and in the marketplace, the guidance is not uniform and is not always helpful. Further clarity will help custodians of personal information determine whether they have complied with their obligations, and it will assist the courts when they are asked to rule on the efforts to do so. The clarity can be achieved in the form of a test; one keyed to a rigorous analysis of risk.

Risk analysis is particularly appropriate to a consideration of the threats to and the responsibility for data security. Here, the expectations of protection are high and are increasingly endorsed by statute; here, the threats to privacy are real, constant, serious, and rapidly changing; here, the cost of providing the protections can be daunting. Just as how to

69. N.Y. COMP. CODES R. & REGS. tit. 23 § 500 (2017).

identify and assess these considerations is important, the actual assessment can be difficult.

For questions of reasonableness concerning the handling of personal information, a test keyed to risk analysis is most likely to yield the right answers, and it is in that context that this test is offered for consideration.

APPENDIX A—Exemplar Cases

The test was applied using the three exemplars below. The scenarios they present do not reflect any one case, and they do not name actual organizations. However, the facts, issues, and causes in each scenario are common components of breaches that the drafters have been professionally involved in.

These scenarios were developed with the intention that they fit the following criteria:

- The breach scenarios should involve facts (controls and causes) that are commonly found in breaches of personal information.
- Each hypothetical information steward's identifying features should not match any organization that any of the drafters worked with relevant to a breach.
- Each scenario should present facts that call for application of each factor of the test.
- Cybersecurity attacks, analysis, vulnerabilities, and alternatives are complex and would be difficult to treat in their full complexity in these exemplar cases. Each exemplar uses facts, descriptions of controls, and descriptions of control alternatives that have been simplified in order to demonstrate the application of the test within a limited space. For the same reason, the exemplars simplify the analysis of harm from a breach. As well, the exemplar cases simplify the litigation process, such as by treating as undisputed inputs that, in a real lawsuit, would be vigorously disputed.
- The estimations, values, and decisions presented in the exemplar cases are not intended to represent actual or normative evaluations or expected

outcomes. They are presented for illustrative purposes only.

- The exemplar cases evaluate only the cost/benefit analysis that would be considered in a data breach case and do not address violations of industry custom, regulations, statutes, or ordinances.
- The test is stated as a formula, and there are many approaches an adjudicator could use to arrive at the inputs to be inserted into that formula. That the first two of these exemplars use quantitative information while the third uses nonquantitative information is an acknowledgement only that organizations use both. As the *Commentary* states, an adjudicator should endeavor to use quantitative information if it is available.
- Solely in an effort to illustrate how these varying analysis methods can operate, the test was applied using Applied Information Economics (AIE), Factor Analysis for Information Risk (FAIR), and Center for Information Security's Risk Assessment Method (CIS RAM). Whether any of these methods, or a different method, should be used is beyond the scope of this *Commentary*.
- Some of the exemplars include risk assessments that were performed ex ante by information stewards to determine, as part of their normal security management program, whether their controls were suitable for the risks posed by their information technology environment.
- Although the exemplars provide both quantitative and nonquantitative risk analysis to demonstrate the varieties of risk assessment that may be presented by

parties as they advocate for how the test should be applied, the exemplars were carefully created based on risk assessment actually used in the field. The exemplars also assume that the risk assessment was conducted by qualified professionals, with appropriate and available evidence to substantiate their estimates.

- Depictions of *ex ante* risk assessments are not intended to be approved applications of the test. Instead, they illustrate how industry has, in practice, assessed what risk may be prior to or at the time of a breach. Adjudicators in the exemplars reference the *ex ante* risk assessments as evidence of what was reasonably foreseeable at the time of the breach.
- The risk assessment methods employed in the exemplars all have strengths and limitations, and all require an element of subjective estimation and modeling by experts. Yet they have attained legitimacy among practitioners by producing measurable and predictable results. Stated simply, the methods shown in the exemplars have been adopted by government agencies and information security practice organizations because they have proved useful.
- Quantitative risk analysis is useful when all factors are expressible numerically and can be compared to each other in a numerical form. As stated previously, quantitative risk analysis should be used whenever quantitative inputs are available.
- Qualitative (nonquantitative) risk analysis should not come into play unless some factors cannot readily be expressed numerically or cannot be compared to each

other when in numerical form. Qualitative risk analysis will be impractical to apply to the test when qualitative inputs are expressed in terms that are not measurable, are arbitrary, are vague, or are not comparable to other inputs.

- These limitations and capabilities are explored by the adjudicators in these exemplars.
- The *Commentary* does not take a position on whether one risk assessment method is better than others. The goal of the exemplars is to illustrate how adjudicators may use different methods as they apply the test.
- Quantitative risk analysis is generally expressed as ranges of probable values, rather than as distinct, singular values.⁷⁰ However, to demonstrate the test in the simplest quantitative form, and in recognition that in certain cases application of the test to a range of possible values would result in one outcome as to one point in the range and a differing outcome as to another point in the range and hence will require the adjudicator to use the most likely point in the range in applying the test, the examples using AIE and FAIR will express distinct dollar values, as is possible in some variations of AIE and FAIR methods.

70. HUBBARD & SEIERSEN, *supra* note 30; FREUND & JONES, *supra* note 31.

SCENARIO I: A Vulnerable API at STS

Company: Small Tech Startup, Inc. (STS)

Number of employees: 22.

Revenue: \$0 (Venture Capital funded).

Industry: Tech/real estate.

Products/Services: Aggregation of consumer home-loan mortgage data.

Sensitive customer information: Residential home-loan mortgage packages, including detailed and extensive personally identifiable information (PII) for approximately 20 million U.S.-based borrowers.

Network environment: Google Cloud and Amazon Web Services (AWS).

Background:

STS has been in business for 1.5 years. The company collects and aggregates residential mortgage loan data from the major national lenders, which includes full loan packages for tens of millions of borrowers. It sells analysis and a feed of this data, which is purchased by large financial services firms and hedge funds, across a subscription-based Application Programming Interface (API).

Security Posture:

- While tech-savvy and generally aware of information security best practices, the company has no formal information security program. Because of its distributed workforce and heavy reliance on cloud-based services, STS's security posture is loosely based on the "zero trust" model, where all access from

inside or outside is untrusted until properly authenticated.

- The Chief Technology Officer is responsible for security. She delegates various security operational responsibilities, including patch management and network security, to engineers who are otherwise overworked building the company's products.
- The company has taken steps to secure its AWS environment, including adding controls to prevent unauthorized remote access to the AWS infrastructure.
- STS also encrypts sensitive loan data at rest in its AWS databases; that data is decrypted as needed in response to authenticated requests from the API. API responses are also TLS (Transport Layer Security) encrypted (i.e., encrypted in transit).
- In Q2 2019, the company retained a third-party cybersecurity company to perform a network and application vulnerability scan and penetration test. The findings of that engagement identified a web application vulnerability in the company's main API product.
- The company considered the web application vulnerability to be low priority because it was thought to be exploitable only in rare circumstances that would not occur in a normal production environment. Nonetheless, STS created a plan to remediate the web vulnerabilities after six months' time, when the API was scheduled to be overhauled.
- Besides the technical vulnerability assessment, the company has not otherwise conducted an information security risk assessment.

- Because of pressure from clients and data suppliers, STS has a longer-term plan to formalize its security program to earn an independently verified certification, but that process isn't scheduled to start until next quarter.

The Incident:

On December 3, 2019, an STS employee received an email message from a purported "security researcher" who, in broken English, claimed to have identified and exploited an SQL⁷¹ injection vulnerability in the STS API. The researcher/attacker included a link to a Twitter account that contained screenshots of database tables containing STS's data with the sensitive customer loan information. For a "consulting fee" of 72 Bitcoin (~\$250,000) the researcher offered to reveal the API vulnerability to STS and to delete the obtained copy of data.

STS decided to ignore the threat.

Two weeks later, the attacker uploaded a 10-gigabyte dump of STS's sensitive customer information to the website "Pastebin." Security bloggers found the data, and within weeks the breach was making national headlines.

A plaintiffs' class action lawsuit followed. Included in the claims was the allegation that STS failed to properly protect plaintiffs' sensitive personal information that was contained in the loan documents.

The Dispute:

Did STS employ "reasonable security" to protect the personal information it maintained?

71. Structured Query Language. See *The Sedona Conference Glossary: eDiscovery & Digital Information Management, Fifth Edition*, 21 SEDONA CONF. J. 263 (2020).

The plaintiff argued that although the data was encrypted in storage and in transit, it was presented to the web application as clear text through the API that had a known vulnerability for many months. Further, STS knew that the data was in the hands of hackers who expressed an intent to breach that data if certain demands were not met. STS should have repaired the API immediately when it discovered it was vulnerable during its vulnerability scan.

STS argued it allowed the API vulnerability to continue because it understood the vulnerability was difficult to exploit. As a start-up it had too few resources to address all vulnerabilities, so it scheduled a fix for July 2020 during the API's normal maintenance routine. As for the timing of its response to the breach, STS argued that it was investigating the breach to verify that the data came from STS and was discussing with its attorneys whether to pay the ransom, which could have increased incentives for hackers to continue hacking the company and others. The hackers did not give STS deadlines, and STS did not have any means of knowing when or whether any exposure would happen.

The Risk Assessment Method:

In this exemplar, the plaintiff's expert conducts a risk assessment using quantitative, probabilistic methods known as Applied Information Economics (AIE). AIE helps organizations estimate probabilities of harm stated in quantities such as financials, time, populations, degrees of harm, or as binaries. AIE offers multiple, evidence-based methods for experts to express probability using subjective judgment and the data available to them.

As a probabilistic method, AIE may provide results as ranges of probable outcomes, or as distinct values. AIE results are presented as distinct values in this scenario.

Application of the Test:

The adjudicator applies the test by evaluating the claims and evidence.

1. The adjudicator is unable to determine that a time period within which to resolve known vulnerabilities has been established as an industry custom. The adjudicator therefore allows STS and the plaintiff each to present a cost/benefit analysis of the utility of repairing the vulnerability on an emergency basis.
2. The plaintiff's expert conducts a risk assessment of the breach, estimating the probability of harm with and without the API repair in place.
 - (a) The expert uses information provided by STS to estimate the likelihood of the API being breached and the likelihood and impact of personal information being abused, both when the API's vulnerability is present and when it is not present.
 - (b) The plaintiff's expert estimates a probability of a breach and misuse of personal information (meaning that information will be breached and will be used in a manner that harms others) through exploitation of the vulnerable API at 12.2 percent during the period between discovery of and the scheduled fix to the vulnerability (the "at-risk period"). Using publicly available data about consumers' out-of-pocket expenses that result from a data breach, the expert estimates the financial impact to the plaintiff and others from such a breach to be

\$740,000.⁷² Applying the probability of a breach and misuse occurring (12.2 percent) to the estimated costs to the plaintiff and others from a breach, the expert estimates a probability-adjusted benefit of repairing the vulnerability immediately, rather than in July 2020, as \$90,280.

- (c) The expert next estimates that a probability of a breach and misuse of personal information through the repaired API to be at 1.7 percent during the at-risk period. The estimated financial impact to consumers from such misuse remained at \$740,000 in the repaired API scenario, so the probability-adjusted benefit of a repaired API was \$12,580.
- (d) The expert then applies the 10.5 percent (i.e., 12.2 percent minus 1.7 percent) net reduced probability of a breach and misuse occurring if the vulnerability had been repaired immediately to the likely costs to others of such a breach to estimate a probability adjusted benefit to consumers of immediately repairing the vulnerability at \$77,700.
- (e) The expert then applies the burden side of the test by concluding that STS would have incurred \$5,000 of one-time additional burden to repair

72. Such publicly available sources of breach cost data may be found in Verizon's Data Breach Investigations Report, NetDiligence's Claims Study, and the Ponemon Institute's Cyber Crime Study. Each provides an annual analysis on the costs of cyber breaches—generally costs to breached organizations and their insurers. Their estimates vary between \$0.58/record to \$200/record cost, or higher. Regardless of the sources or estimation methods used, experts would provide a rationale for their estimates.

the API's vulnerability immediately rather than during its normal maintenance period.

- (f) Because the \$5,000 of burden would thereby have generated a benefit to others of \$77,700, the plaintiff's expert concludes that STS's security for personal information was rendered unreasonable by API's failure to repair the vulnerability immediately.
3. STS rejects the plaintiff's expert's analysis. STS asserts that had it stopped developing its application long enough to repair the vulnerability out-of-cycle (as an emergency change), it would have risked disruption of functionality of the API for days, which would have reduced the value of STS's clients' use of the API during that time. The utility of the application to STS's financial service provider customers would have suffered.
- (a) STS produces information about daily dollar value of use of the application's features. It estimates \$35,000 per day of value enjoyed by its clients.
 - (b) STS's expert estimates the likelihood of the API being unavailable to STS's clients if it were to have repaired the API as an emergency change. The experts estimate a 59 percent likelihood that the API would have gone down had the repair been attempted; and had that happened, the API would have been down for two days. Because the API creates \$35,000 per day in value to STS's clients, the STS expert calculates \$41,300 in expected loss of utility (i.e., $\$35K \times .59 \times 2 = \$41.3K$) because of an emergency repair.

- (c) STS agrees that the repair itself would have involved a one-time incremental cost of \$5,000 if done on an emergency basis, without consideration of the potential loss in utility of the API if the upgrade failed. STS also does not dispute the plaintiff's estimate that the repaired vulnerability would have produced a 10.5 percent net reduced probability of a breach and misuse occurring during the at-risk period.
 - (d) STS states that the reduced utility of the API if it failed during an emergency repair (\$41,300) should be added to the burden associated with repairing the vulnerability on an emergency basis in applying the test.
4. The adjudicator employs STS's and the plaintiff's experts' analyses in applying the test.
- (a) The adjudicator determines that a benefit of \$77,700 would have been realized from repairing the API on an emergency basis, with a \$5,000 additional cost burden.
 - (b) The adjudicator also adds to STS's cost burden of repairing the vulnerability immediately the reduced utility of \$41,300 associated with doing the repair on an emergency basis, bringing STS's total calculated burden of immediately repairing the vulnerability to \$46,300 (adding the \$41,300 reduced utility to the \$5,000 repair).
 - (c) *The adjudicator determines that the one-time added cost burden of \$46,300 would have provided a benefit of \$77,700, and therefore that STS's failure to repair the API vulnerability*

*immediately rendered its security for personal information unreasonable.*⁷³

73. While utility, cost, and benefit can often be quantified, organizations may find it difficult *ex ante*, and adjudicators may find it difficult *ex post* to evaluate some factors using quantitative information. Factors such as the education of students, the care of patients, development of health science, and the promotion of safety are not as obviously associated with quantitative information as are budgets. Further, court cases such as *Grimshaw v Ford Motor Co.*, 174 Cal. Rptr. 348 (Cal. Ct. App. 1981) and social science research (W. Kip Viscusi, "Jurors, Judges, and the Mistreatment of Risk by the Courts, 30 J. OF LEGAL STUD. 107 (2001))" demonstrate negative juror and jurist reactions to a purely quantitative risk analysis. In circumstances where organizations or adjudicators consider quantitative methods to be impracticable, they may feel inclined to consider the possibility of opting for other methods. The third exemplar sets forth a methodology by which a nonquantitative analysis might be done. Alternatively, in circumstances where a quantitative analysis is considered to be impracticable, organizations and adjudicators may conclude the test is unworkable and look instead to industry custom and/or statutory requirements to evaluate the reasonableness of the information steward's data security. The question of whether and to what extent a nonquantitative analysis may be used in such circumstances is beyond the scope of this paper.

SCENARIO II: Advanced Persistent Threat Attack at MMT Labs, Inc.

Company: Medium Medical Testing Labs, Inc. (MMT)

Number of employees: 500.

Revenue: \$120 million.

Industry: Health care.

Products/Services: Medical testing services for hospitals and doctors' offices.

Sensitive customer information: Drug testing and other patient health test results going back five years.

Background:

Founded in 2003, MMT operates clinical labs in five states, offering health testing services to hospitals and other health-care providers. It maintains its test results in databases in its secure network environment.

Security Posture:

- MMT has worked hard to improve its security posture over time, formalizing its policies and procedures and implementing controls to achieve and stay in compliance with professional and regulatory requirements.
- MMT maintains a risk-based cybersecurity program that includes regular audits, penetration tests, and corrective actions where noncompliant controls or vulnerable controls are identified.
- Lab test results are maintained in encrypted form in Microsoft SQL Server databases. After five years, records are purged from the databases.

- MMT has a team of eight full time IT personnel, with one person in charge of network security.
- MMT has conducted annual risk assessments using Factor Analysis for Information Risk (FAIR) analysis methods. While many risks have been identified, not all information assets or threats have been analyzed. As is common practice, MMT's risk analysis evaluates harm to themselves, including loss of reputation, incurred costs, and regulatory fines that could result from breaches.

The Risk Assessment Method:

MMT has conducted annual risk assessments using Factor Analysis for Information Risk (FAIR) analysis methods. FAIR uses subjective estimates by experts to estimate component factors that comprise risk, such as the commonality and strength of attacks, the robustness of controls, the diligence of attackers, and multiple factors that contribute to post-incident costs. As is common practice, its FAIR risk assessment estimated the likelihood of cost that MMT would incur as the result of a security incident but did not estimate the likelihood of harm that others (their customers or patients) would incur.

The Incident:

On December 24, 2019, a system administrator noticed a large compressed file on the database server called "EXPORT.RAR." The administrator opened the file and found a dump of the database tables in decrypted format.

Further investigation revealed that similar export files had been created on the other database servers, and company firewall logs established that the data had been exfiltrated from the network and was therefore stolen.

Forensic investigators found a database administrator's account had been used to log into the database servers and export the data. They did not discover how the attackers obtained the database administrator's credentials.

The investigators determined that the attackers got into the network via a phishing attack that occurred seven months prior. A billing manager opened an email attachment with a weaponized Excel file that installed hybrid trojan malware on his computer. The malware opened up a port-forwarding back door using PowerShell, allowing the attackers to remotely control his computer, even through the firewall.

From there, the attackers found and cracked the credentials for a domain administrator who had previously logged onto the billing manager's computer; and they used that account to move laterally across the network environment.

Based on the tactics, tools, and procedures used by the attacker, the forensic team believed that MMT had been victimized by a sophisticated Advanced Persistent Threat actor.

MMT reported the breach to the State Attorney General (AG) in each of the five states the identified patients resided in and notified each of the 2.5 million affected patients.

The Dispute:

State Attorneys General allege MMT did not employ "reasonable security" in protecting the patient medical data in its care.

State AGs argued that the unsafe configuration of the billing manager's computer and the cached domain administrator's credentials on that machine permitted the hack to occur. Additionally, MMT's technical audits and penetration tests found these vulnerabilities, yet MMT accepted the risk and left the vulnerabilities in place.

MMT argued that its security program that includes phishing tests, encryption, continuously improved policies, Microsoft Advanced Threat Protection, patch management, etc. demonstrated due diligence. It further argued the billing manager's computer was configured in a weaker state than the others because the billing manager did not access sensitive data, and the computer needed to run a client health network's billing software application, which required a less-secure configuration in order to operate. MMT presented its *ex ante* FAIR risk analysis as evidence during discovery.

Application of the Test:

The adjudicator applies the test by evaluating the claims and evidence.

1. The adjudicator reviews MMT's *ex ante* risk assessment and sees that MMT evaluated the risk posed by the billing manager's less-secure computer. The adjudicator also sees that MMT only considered the risk of costs to themselves (not the risk of harm to their customers), and that MMT accepted the risk. While not providing explicit criteria for accepting the risk, MMT explains the computer needed to be in the less-secure state in order to operate a critical billing application.
2. State AGs submit their own risk analysis (also using FAIR) to the adjudicator to estimate the probability of harm to residents of their state. State AGs' risk analysis compares the risk to residents with the standard protections on the billing manager's computer to the risks without those protections.
 - (a) State AGs' expert estimates the likelihood of the hacker's successful attack and subsequent harm

to states' residents as it would have been estimated at the time of the breach and with the billing manager's computer configured in its nonsecured state. Given the known vulnerabilities introduced by the billing software, the State AGs' expert estimates the Loss Event Frequency⁷⁴ at 29.2 percent per annum and the Loss Event Magnitude⁷⁵ to the states' residents as \$10,000,000⁷⁶ at the time of the breach.

- (b) State AGs' expert then estimates the likelihood of the hacker's successful attack and consequent harm had the billing manager's computer been configured as securely as his colleagues' computers. The State AGs' expert estimates the Loss Event Frequency at 1 percent per annum (meaning that enhancing the security on the billing manager's computer would have decreased the probability of harmful abuse of personal information from 29.2 percent to 1

74. "Loss Event Frequency" is FAIR's term for per-annum probability of a loss event.

75. "Loss Event Magnitude" is FAIR's term for the sum of losses experienced during a loss event when paired with a "Loss Event Frequency." In essence, it is developed by considering the probability of a breach and a range of the financial impact of possible breach outcomes in order to create the probability-weighted range of losses anticipated if the event in fact occurs.

76. Loss event magnitudes are most often expressed as a range of potential values (minimum, most likely, and highest values). If in a given case application of the test to a range of possible values would result in one outcome as to one point in the range and a differing outcome as to another point in the range, the adjudicator should use the most likely point in the range in applying the test.

percent, or 28.2 percent, per annum).

- (c) State AGs assert that had MMT made a one-time \$1,000 investment to secure the billing manager's computer, that investment would have generated a year-one benefit to their states' residents of 28.2 percent x \$10,000,000 (i.e., \$2,820,000).
3. The adjudicator solicits MMT's evaluation of risk at the time of the breach, and the risk had the billing manager's computer been configured similarly to MMT's other systems.
- (a) MMT argues that the test should be applied by including the burden that would have resulted had the billing manager not run the health network client's invoicing software. MMT's largest client would only have done business with MMT had MMT used the client's billing software, which could only operate on a computer configured with moderate security controls. Because the billing manager's computer was the only one that was atypically unsecured, MMT agrees with AGs' assumption that \$1,000 is an appropriate estimate for the one-time cost of applying controls to that one system, including the added labor for doing so.
 - (b) Additionally, MMT believes that the evidence supporting the risk assessment it conducted prior to the breach should be considered for purposes of determining what the reasonably foreseeable likelihood and magnitude of a breach was at the time of the breach. MMT provides the annual billings from its largest

client that it would not have earned had it secured the billing manager's computer and not used the client's billing software. The net profits from these billings average \$1,800,000 per year.

- (c) MMT produces analysis from its *ex ante* risk assessment showing that (i) enhancement of the billing manager's computer security would reasonably be expected to result in only a 2.7 percent decrease in the per annum probability of a breach that resulted in harm (from 13.2 percent to 10.5 percent) and (ii) such a breach would lead to \$5,000,000 in damages to MMT alone (without consideration of harm to others).
- (d) MMT presents analysis of risk at the time of the breach by multiplying (i) its 2.7 percent estimate of the per annum decreased likelihood of a breach resulting in harm after enhancing the billing manager's computer security by (ii) the State AGs' estimate that a breach would have resulted in costs to their residents of \$10,000,000, to show that the public would have received a year-one benefit of \$270,000.
- (e) MMT then compares its estimated \$1.8 million of year-one lost profits from implementing the enhancement to the \$270,000 year-one benefit to the states' residents. It argues that the reasonably expected burden of the enhancement outweighed its reasonably expected benefit, as both would reasonably have been estimated prior to the breach.
- (f) MMT acknowledges that its prior risk assessment estimated only costs to MMT from a

breach that resulted in harm and did not separately estimate the potential costs of a breach to others, including the patients whose personal health information (PHI) was stored by MMT, even though such costs to others were reasonably foreseeable at the time of the estimate and the breach.

- (g) MMT also attempts to introduce its utility of producing accurate and fast test results but fails to produce a coherent financial model for that utility. State AGs respond that MMT has competitors who also provide fast and accurate results, so its customers could have used safer alternatives while enjoying the same benefits, rendering the utility claim moot.
4. The adjudicator applies MMT's and State AGs' analysis to the test.
- (a) The adjudicator notes that MMT and State AGs agree that at the time of the breach, the likely harm to the states' residents from such a breach was \$10,000,000 without consideration of the likelihood of such a breach occurring.
 - (b) The adjudicator notes that the State AGs and MMT agree that MMT's burden in securing the billing manager's computer in the manner advocated by the State AGs would have been the loss of \$1.8 million in net profits in year one, plus the \$1,000 to secure the billing manager's computer.
 - (c) The adjudicator notes that MMT and States' AGs disagree on the net decreased likelihood of a breach of this sort occurring had the billing

manager's computer been secured in the manner advocated by the State AGs.

- (d) If the adjudicator finds State AGs' expert's likelihood-of-breach estimates persuasive (perhaps the billing application's vulnerabilities being widely known to the hacking community is a deciding factor) it would apply the test as follows:
- (i) The adjudicator would calculate the net year-one benefit of applying additional security to the billing manager's computer by multiplying the expected harm from such a breach by the State AG's estimates of the per annum probability of such a breach occurring with, and without, the billing manager's computer secured in the manner advocated by the State AGs. The adjudicator would therefore conduct multiple calculations.
- 1) Risk to states' residents at the time of the breach: 29.2 percent annual likelihood x \$10,000,000 = \$2,920,000.
 - 2) Risk to states' residents if the billing manager's computer was secured: 1 percent likelihood x \$10,000,000 = \$100,000.
 - 3) Net year-one benefit from the additional security measures advocated by State AGs = \$2,820,000 (i.e., \$2,920,000 minus \$100,000).

- (ii) The adjudicator would compare the year-one benefit of \$2,820,000 to the year-one burden of \$1,800,000. *Based on that comparison, and in the absence of any evidence that the benefit would not exceed the burden after year one, the adjudicator would find that the failure to secure the billing manager's computer in the manner advocated by the State AGs rendered MMT's security for personal information unreasonable.*
- (e) If the adjudicator instead finds MMT's likelihood-of-breach estimates persuasive (perhaps MMT's history of penetration tests and audits make a convincing case of MMT's estimate), the adjudicator would apply the test as follows:
 - (i) The adjudicator would calculate the net year-one benefit of applying additional security to the billing manager's computer by multiplying the range of expected harm from such a breach by MMT's estimates of the per annum probability of such a breach occurring with, and without, the billing manager's computer secured in the manner advocated by the State AGs. The adjudicator would therefore conduct multiple calculations.
 - 1) Risk to states' residents at the time of the breach: 13.2 percent annual likelihood x \$10,000,000 = \$1,320,000.
 - 2) Risk to states' residents if the billing manager's computer was secured: 10.5

percent annual likelihood x \$10,000,000 =
\$1,050,000.

- 3) Net year-one benefit from the additional security measure advocated by State AGs = \$270,000 (i.e., \$1,320,000 minus \$1,050,000).
- (ii) The adjudicator would compare the year-one benefit of \$270,000 to the year-one burden of \$1,800,000. ***Based on that comparison, and in the absence of any evidence that the burden would not exceed the benefit after year one, the adjudicator would find that the failure to secure the billing manager's computer in the manner advocated by the State AGs did not render MMT's security for personal information unreasonable.***

SCENARIO III: Lost Mobile Device at a Research University Hospital

Company: Research University Hospital (RUH)

Number of employees: 4,000.

Revenue: \$3.2 billion patient revenue; \$350 million research grants.

Industry: Academic medical center.

Products/Services: Patient care, clinical education, medical science research, clinical studies.

Sensitive customer information: Patients' protected health information (PHI).

Background: Founded in 1957, RUH serves its community through direct patient care, supports its affiliated university through clinical education of its medical students, and advances medical knowledge through hard science research and clinical trials.

Security Posture:

- RUH funds its security program comparably to other research universities of similar size. It collaborates with other hospitals, security experts, and regulators to determine, communicate, and improve best practices for securing PHI.
- RUH operates an information security program that conforms to the HIPAA Security Rule. The hospital's risk management program has defined when controls are "reasonable and appropriate" in alignment with the rule. RUH tests and improves its controls and maintains a record of their risks, vulnerabilities, and needs for improvement.

- RUH operates a set of secured mobile devices (tablets) to be used in its public outreach programs. Clinicians regularly visit underserved, remote communities to provide free checkups, examinations, and primary care to patients who cannot afford them. To prepare for these remote visits, clinicians download patient records from the Electronic Health Record (EHR) to a set of tablets, enabling fast, easy access to local patients' records. Access to these records does not require multifactor authentication, but a four-digit password is required to access the tablet's interface. RUH accepted the risks involved in this configuration because access to the EHR and multifactor authentication systems from remote locations is unreliable, and timely access to accurate patient charts is critical for providing safe, effective care.

The Risk Assessment Method:

In compliance with the HIPAA Security Rule, RUH evaluated its risk of potential breaches using a risk assessment. RUH used the Center for Information Security's Risk Assessment Method (CIS RAM), a nonquantitative risk assessment method, to model and prioritize its risks. CIS RAM evaluates risk in terms of an organization's duty of care by evaluating the likelihood and impact of harm to themselves and others, by delineating between acceptable and unacceptable harm, and by determining whether additional controls are more burdensome than the risks they reduce.

RUH evaluated risk to five factors: its three utilities of patient health outcomes, educating clinicians, and advancing medical science; its objectives to balance its budget; and its obligation to protect the privacy of patients. As it evaluated

these risks, it estimated the likelihood of harm in plain-language terms using associated integers (1 through 5) for the likelihood scores and impact scores.

Likelihood scores 1 through 5 used plain-language terms to describe the estimated plausibility and commonality of breaches. Impact scores 1 through 5 indicated degrees of measurable harm that would result from the breaches. Scores of 1 and 2 indicated harms that in its judgment would not require correction or remedy by any party. Score 3 indicated harms that would require some correction or remedy. Score 4 indicated harms that would be potentially severe but recoverable. Score 5 indicated unrecoverable harms such as death, or the hospital's inability to provide the care in question.

RUH's risk assessment determined that multifactor authentication on the tablets would have created a greater risk to the delivery of patient care than any potential harm to patient privacy if the tablets were lost or stolen.

The Incident:

In August 2019, a physician left his tablet behind at a school building where he and a medical outreach team were running a three-day remote clinic. Records for approximately 20,000 patients were stored on the tablet in case any patients from the region attended the clinic. The records were encrypted while stored in the EHR application but were viewable on a one-record-at-a-time basis.

The four-digit passcodes used to access the tablets could have been viewed by patients due to the clinic's public setting.

Once the tablets were accessed, no further credentials were required to access patient records on the local EHR application. This was meant to avoid clinicians being delayed while accessing patient records in critical situations or preventing

lockouts when multiple attempts at tapping in complex passwords failed.

Once the physician realized he left his tablet behind, he immediately alerted his IT team and requested that they remotely wipe the device while a member of his staff drove back to the school to retrieve the tablet. But the staff member was not able to locate the tablet when he arrived at the school, the IT team could not confirm the tablet was remotely wiped, and the tablet did not contain a cellular network chip to assist in the recovery or wipe. Rather, it required only a local wi-fi network to connect to the internet.

RUH appears to have complied with the HIPAA Breach Notification Rule. It notified the Department of Health and Human Services Office for Civil Rights (OCR) the day after the tablet was left behind. Further, RUH informed the patients whose data was on the tablet and provided them with identity theft protection services.

The Center for Medicare and Medicaid Services noted several recent and apparently fraudulent Medicare claims had been made in the names of patients whose PHI was on the outreach clinic tablets.

The Dispute:

The OCR claimed RUH should have used multifactor authentication as a “reasonable and appropriate safeguard” to protect the patients’ PHI.

OCR argued multifactor authentication was used to provide access to patient records in all other uses of the EHR, and the tablets were at higher risk of breach due to their mobility.

RUH argued the risk to patients would have been higher had the tablets used multifactor authentication and enforced passwords. The remote location of the clinics and the user

interface provided by tablets made it difficult to assure access to PHI if normal access controls were used.

The Test:

The adjudicator applies the test by evaluating the claims and evidence.

1. The adjudicator agrees with OCR's position that multifactor authentication to access patient health records on the tablets would have been industry custom, and that in fact RUH had used multifactor authentication to access the records on other systems, so the technology was known and available to them. The adjudicator therefore determines that OCR has presented evidence sufficient to support a finding that RUH's failure to use multifactor authentication on the tablets rendered its security for personal information unreasonable.
2. The adjudicator then allows RUH to seek to rebut OCR's evidence case by demonstrating that failing to use multifactor authentication on the tablets was reasonable under the test.
3. In an effort to describe the risk of breach from the nonuse of multifactor authentication for the tablets as it was known at the time of the breach, RUH presents its pre-breach risk assessment.
4. The adjudicator determines that RUH's pre-breach risk assessment was not quantitative and asks why RUH used a nonquantitative method to determine the risk of breach.
 - (a) RUH states that their multiple utilities—patient health outcomes, educating clinicians, and advancing medical science—were very difficult

to quantify in financial terms, and to do so in a way that retained the meaning of those utilities. Moreover, hospital management was concerned that comparing budgetary impacts to financial representations of the benefits resulting from educated clinicians and advanced medical knowledge would have sent the wrong message to its staff and the community about its multiple missions.

5. RUH offers the adjudicator results from RUH's risk assessment that RUH argues the adjudicator can use to apply the test.
 - (a) RUH first presents its analysis of the budgetary costs at the time of the breach of using multifactor authentication on the tablets.
 - (i) RUH's experts calculated in their risk assessment that without the multifactor authentication control the risk to RUH's budget was 5 out of 25. This calculation reflected RUH's assessment that not adopting multifactor authentication would certainly (likelihood 5) have had a negligible impact (impact 1) to its budget; the 5 score was the result of multiplying the likelihood score by the impact score (i.e., $5 \times 1 = 5$).
 - (ii) RUH's experts then estimated that with multifactor authentication controls in place, the risk to RUH's budget would have been the same: 5 out of 25 and for the same reasons.
 - (iii) With the two scores being the same, RUH's experts concluded that the incremental cost to

RUH's budget of employing multifactor authentication on the tablets would have been zero (i.e., $5 \text{ minus } 5 = 0$) at the time of the breach and thus should have no impact on application of the test.

- (b) RUH next presents analysis of the risk of patient privacy harm at the time of the breach, first without and then with multifactor authentication being employed on the tablets.
 - (i) RUH's experts calculated in their risk assessment that without the multifactor authentication control, the risk of privacy harm to patients was 8 out of 25. This calculation reflected RUH's assessment that not adopting such authentication would plausibly (likelihood 2) have had redressable privacy impact to thousands of patients (impact 4) whose information may have been exposed one record at a time in the encrypted application. The 8 score was the result of multiplying the likelihood score by the impact score (i.e., $2 \times 4 = 8$).
 - (ii) RUH's experts then estimated that with multifactor authentication controls in place, the risk of privacy harm to patients would have been 4 out of 25. This calculation reflected RUH's assessment that even though upon adopting multifactor authentication a lost or stolen tablet would not be accessible, the patients would still plausibly (likelihood 2) be concerned about their unexploitable privacy, although they would not suffer a

particularized or concrete harm (impact 2). The 4 score was the result of multiplying the likelihood score by the impact score (i.e., $2 \times 2 = 4$).

- (iii) Based on this analysis, RUH argues that the net benefit of employing multifactor authentication on the tablets was 4 (i.e., $8 \text{ minus } 4 = 4$) at the time of the breach.
- (c) RUH next presents analysis of the risk to patient health outcomes at the time of the breach, first without and then with multifactor authentication being employed on the tablets.
- (i) RUH's experts calculated in their risk assessment that without the multifactor authentication control in place on the tablets, the risk to patient health outcomes would have been 5 out of 25. This calculation reflects RUH's assessment that not adopting multifactor authentication would certainly (likelihood 5) have had a negligible impact (impact 1) to patient health outcomes; the 5 score was the result of multiplying the likelihood score by the impact score (i.e., $5 \times 1 = 5$).
 - (ii) RUH's experts then estimated that with multifactor authentication controls in place, the risk to patient care outcomes would have been 12 out of 25. This calculation reflects RUH's assessment that in rural environments where internet connectivity is not reliable, multifactor authentication communications

also would not be reliable, with the result being that physicians would occasionally (likelihood 3) not gain access to patient records and would misdiagnose or erroneously provide harmful treatments to patients that worsen health outcomes short of permanent damage or death (impact 4). The 12 score was the result of multiplying the likelihood score by the impact score (i.e., $3 \times 4 = 12$).

- (iii) Based on this analysis, RUH argues that at the time of the breach the net burden of employing multifactor authentication on the tablets was '7' (i.e., $12 \text{ minus } 5 = 7$) in terms of patient health-care outcomes and 0 in terms of its impact on RUH's budget, for a total burden of 7.
- (d) RUH then argues that its failure to employ multifactor authentication on the tablets did not render its security for personal information unreasonable under the test because 7 is greater than 4.
- (e) OCR challenges RUH's proposed application of the test on the following grounds: (i) RUH's methodology for calculating the net benefit of employing multifactor authentication on the tablets does not provide a reliable estimate of that net benefit, as it is entirely the product of RUH's own subjective qualitative value judgments and RUH's arbitrary formulas for quantifying and comparing those judgments; (ii) RUH's methodology for calculating the net burden of employing multifactor authentication

on the tablets does not provide a reliable estimate of that net burden, as it too is entirely the product of RUH's own subjective qualitative value judgments and RUH's arbitrary formulas for quantifying and comparing those judgments; and (iii) even if they did yield reliable estimates of net benefit and net burden, RUH's methodologies for determining net benefit and net burden differ from one another so fundamentally in regard to the subjective qualitative value judgments and the formulas on which they are based that the output of one methodology (here '4') cannot be compared to the output of the other methodology (here '7') for purposes of comparing the benefits and the burdens of an additional security measure.

6. If the adjudicator rejects OCR's challenges to RUH's application of the test, and instead concludes that RUH's methodologies for calculating the net benefit and net burden of employing multifactor authentication on the tablets provide a reliable estimate of that net benefit and net burden that are themselves reliable and may reliably be compared to one another for purposes of applying the test, the adjudicator would apply the test as follows:
 - (i) The adjudicator would first determine the incremental benefit of employing multifactor authentication on the tablets from the reduction of privacy harm to patients achieved by the use of such authentication. The risk score for harm without multifactor authentication was 8, while the risk score for harm with multifactor authentication was 4.

The incremental benefit, therefore, would be $8 - 4 = 4$.

- (ii) The adjudicator then would determine the incremental burden of employing multifactor authentication on the tablets from the impact on RUH's budget of adopting such authentication and increased risk to patient care outcomes caused by its use. The budgetary impact score was 5 both with and without multifactor authentication, so the incremental budgetary impact of adopting it for the tablets would be $5 - 5 = 0$. The risk score for patient care outcomes with multifactor authentication was 12, whereas the risk score for patient care outcomes without it was 5, so the incremental burden to patient care outcomes caused by the use of multifactor authentication, therefore, was $12 - 5 = 7$.
- (iii) *Because the use of multifactor authentication on the tablets has greater incremental burden (7) than it has incremental benefit (4), the adjudicator therefore concludes that RUH's failure to use multifactor authentication on the tablets did not render its security for personal information unreasonable.*

THE SEDONA CONFERENCE COMMENTARY
ON EPHEMERAL MESSAGING

*A Project of The Sedona Conference Working Group on International
Electronic Information Management, Discovery, and Disclosure
(WG6)*

Author:

The Sedona Conference

Editor-in-Chief:

Philip J. Favro

Contributing Editors:

Bennett Arthur

Starr Turner Drum

Stacey Blaustein

David K. Gaston

Oliver Brupbacher

Alan Geolot

Guillermo Santiago Christensen

Jennifer L. Joyce

Andrea D'Ambra

Agnieszka McPeak

Robert DeCicco

Hon. Anthony E. Porcelli

Steering Committee Liaisons:

Denise E. Backhouse

Wayne Matus

Taylor Hoffman

Staff Editors:

David Lumia

Michael Pomarico

Copyright 2021, The Sedona Conference.
All Rights Reserved.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 6. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

We thank all of our Working Group Series Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors, just click on the “Sponsors” navigation bar on the homepage of our website.

This publication may be cited as follows:

The Sedona Conference, *Commentary on Ephemeral Messaging*, 22 SEDONA CONF. J. 435 (2021).

PREFACE

Welcome to the July 2021 final version of The Sedona Conference *Commentary on Ephemeral Messaging* (“*Commentary*”), a project of The Sedona Conference Working Group 6 on International Electronic Information Management, Discovery, and Disclosure (WG6). This is one of a series of Working Group commentaries published by The Sedona Conference, a 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, intellectual property rights, and data security and privacy law. The mission of The Sedona Conference is to move the law forward in a reasoned and just way.

The mission of WG6 is to develop principles, guidance and best practice recommendations for information governance, discovery and disclosure involving cross-border data transfers related to civil litigation, dispute resolution and internal and civil regulatory investigations.

The Sedona Conference acknowledges Editor-in-Chief Phil Favro for his leadership and commitment to the project. We also thank Contributing Editors Bennett Arthur, Stacey Blaustein, Oliver Brupbacher, Guillermo Christensen, Andrea D’Ambra, Robert DeCicco, Starr Drum, David Gaston, Alan Geolot, Jennifer Joyce, Professor Agnieszka McPeak, and Judge Anthony Porcelli for their efforts, and Denise Backhouse, Taylor Hoffman, and Wayne Matus for their guidance and input as Steering Committee liaisons to the drafting team. We also thank Natascha Gerlach for her contributions.

In addition to the drafters, this nonpartisan, consensus-based publication represents the collective effort of other members of WG6 who reviewed, commented on, and proposed edits to early drafts of the *Commentary* that were

circulated for feedback from the Working Group membership. Other members provided feedback at a WG6 meeting where drafts of this *Commentary* were the subject of the dialogue. The publication was also subject to a period of public comment. On behalf of The Sedona Conference, I thank both the membership and the public for all of their contributions to the *Commentary*.

We encourage your active engagement in the dialogue. Membership in The Sedona Conference Working Group Series is open to all. The Series includes WG6 and several other Working Groups in the areas of electronic document management and discovery, data security and privacy liability, international data transfers, patent litigation, patent remedies and damages, and trade secrets. The Sedona Conference hopes and anticipates that the output of its Working Groups will evolve into authoritative statements of law, both as it is and as it should be. Information on membership and a description of current Working Group activities is available at <https://thesedonaconference.org/wgs>.

Craig Weinlein
Executive Director
The Sedona Conference
July 2021

TABLE OF CONTENTS

I.	INTRODUCTION.....	441
II.	EPHEMERAL MESSAGING—NATURE AND SCOPE	446
	A. Automated Disposition of Message Content	446
	B. E2E Encryption	447
	C. Other Characteristics of Ephemeral Messaging.....	449
	1. Purely Ephemeral Messaging	449
	2. Quasi-Ephemeral Messaging	451
	3. Non-Ephemeral Messaging.....	452
III.	TENSIONS ASSOCIATED WITH THE USE OF EPHEMERAL MESSAGING APPLICATIONS	454
	A. Benefits of Ephemeral Messaging.....	454
	1. Organizational Benefits.....	454
	(a) Information Governance	454
	(b) Legal Compliance Support	456
	(c) Privacy by Design	459
	(d) Data Security.....	460
	(e) Productivity	461
	2. Benefits to Individual Users	462
	B. Risks of Ephemeral Messaging	464
	1. Regulatory Risks	465
	2. Legal Risks	468
	3. Operational Risks.....	469
IV.	GUIDELINES	471
	A. Guideline One: Regulators and Courts Should Recognize that Ephemeral Messaging May Advance Key Business Objectives	471

- B. Guideline Two: Organizations Should Take Affirmative Steps to Manage Ephemeral Messaging Risks474
- C. Guideline Three: Organizations Should Make Informed Choices and Develop Comprehensive Use Policies for Ephemeral Messaging Applications476
- D. Guideline Four: Regulators, Courts, and Organizations Should Consider Practical Approaches, Including Comity and Interest Balancing, to Resolve Cross-Jurisdictional Conflicts over Ephemeral Messaging480
- E. Guideline Five: Reasonableness and Proportionality Should Govern Discovery Obligations Relating to Ephemeral Messaging Data in U.S. Litigation482

I. INTRODUCTION

Ephemeral messaging is increasingly used around the globe. With its ability to automate the deletion of content shared with others, ephemeral messaging offers organizations a robust option to strengthen aspects of their corporate information governance programs. This feature, combined with end-to-end encryption (“E2E encryption”) that enables secure communications, may also facilitate compliance with data protection and privacy laws. Indeed, these laws—including the European Union (EU) General Data Protection Regulation (GDPR)¹—are among the considerations driving organizations toward the use of ephemeral messaging.

Beyond these considerations are issues such as convenience and ease of use. Users find that by keeping discussions confidential, ephemeral messaging enhances their ability to collaborate and exchange information without significant information technology (IT) infrastructure. These collective factors make ephemeral messaging a potentially attractive communication option for organizations and their employees.

Despite the growing use of ephemeral messaging, there are concerns about its widespread adoption.² Government

1. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119/1), *available at* <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679#PP3Contents> [hereinafter GDPR]. GDPR is a single, binding, EU-wide regulatory framework that became effective on May 25, 2018.

2. The Council of the European Union recently renewed its consideration of a resolution regarding the use of encrypted messaging applications that attempts to balance the needs of data subjects for strong encryption against government security interests seeking access to

regulators at the U.S. Department of Justice (U.S. DOJ) and the U.S. Securities & Exchange Commission (U.S. SEC) worry that ephemeral messaging can lead to increased criminal activity such as bribery, fraud, and money laundering. The U.S. DOJ and the U.S. SEC have implemented policies that discourage organizational adoption of ephemeral messaging without careful consideration of their compliance obligations. While the U.S. DOJ recently modified its policy toward a potentially more accommodating view in the context of corporate compliance programs,³ the fact remains that certain government regulators around the world disfavor the use of ephemeral messaging absent strong corporate governance.⁴

Other complications related to the use of ephemeral messaging include the legal obligation in common law countries that parties preserve evidence for litigation. For example, civil litigation in U.S. federal and state courts generally requires that litigants (at a minimum) keep information relevant to the claims and defenses in a particular action. Once the common law duty to preserve attaches, use of

encrypted data. See Natasha Lomas, *What's all this about Europe wanting crypto backdoors?*, TECH CRUNCH (Nov. 9, 2020), <https://techcrunch.com/2020/11/09/whats-all-this-about-europe-wanting-crypto-backdoors/>.

3. See Section III.B.1, *infra*.

4. See, e.g., Financial Conduct Authority, *Newsletter on market conduct and transaction reporting issues* (Jan. 2021), <https://www.fca.org.uk/publications/newsletters/market-watch-66> (warning that encrypted messaging applications may enable regulated companies to circumvent mandatory recordkeeping obligations); Sarah Basford Canales, *Australia's Controversial Encrypted Messaging Laws, Explained*, GIZMODO (Aug. 7, 2020), <https://www.gizmodo.com.au/2020/08/assistance-and-access-law-encrypted-messaging-explained/> (discussing the status and impact of Australia's new encryption cracking law, which impacts the use of encrypted messaging applications).

ephemeral messaging may cause relevant data to be discarded, which could violate that duty.⁵

These and similar competing demands spotlight a clear tension that has created a quandary for organizations wishing to implement ephemeral messaging. In the face of that tension, organizations need direction on how they should address these competing demands. This is particularly the case for organizations seeking to use ephemeral messaging to comply with cross-border data protection directives without violating other legal requirements.

This tension is also apparent for government regulators and judges who have been tasked with evaluating an organization's efforts at compliance with a particular law or regulation. These decision-makers may be inclined to presume that ephemeral messaging is being used to prevent regulators, courts, litigation adversaries, or the public from obtaining critical information about the inside workings of an organization. A closer, more thorough inspection could provide a more balanced perspective, revealing that a corporate ephemeral messaging program is meritorious and designed to advance business objectives, including compliance with cross-border data protection regimes. Just as organizations could profit from guidance on the issues, regulators and courts may also benefit from direction on how

5. *See* WeRide Corp. v. Kun Huang, No. 5:18-cv-07233, 2020 WL 1967209 (N.D. Cal. Apr. 24, 2020) (criticizing defendants and imposing terminating sanctions for, among other things, implementing an enterprise grade ephemeral messaging application to conceal relevant communications from discovery); *Herzig v. Arkansas Found. for Med. Care, Inc.*, No. 2:18-cv-02101, 2019 WL 2870106 (W.D. Ark. July 3, 2019) (holding that plaintiffs' use of Signal during litigation was designed to prevent discovery of relevant communications, was "intentional, bad-faith spoliation of evidence," and justified the imposition of sanctions).

to address ephemeral messaging. In particular, regulators and courts should understand how to identify and distinguish a legitimate ephemeral messaging program from uses of this technology that may be inappropriate.

All of which has led The Sedona Conference Working Group 6 to prepare The Sedona Conference *Commentary on Ephemeral Messaging* (“*Commentary*”). Section II of the *Commentary* defines the nature and scope of ephemeral messaging, while Section III provides a detailed sketch of the tension and competing demands facing organizations that wish to use these tools.⁶ Section IV encompasses a series of guidelines that provide direction to organizations on how to navigate the landscape of uncertainty surrounding the use of ephemeral messaging.⁷ The guidelines also offer recommendations to regulators and judges for evaluating good-faith uses of corporate ephemeral messaging.

In particular, Guideline One provides that regulators and courts should recognize that ephemeral messaging may advance key business objectives. Guideline Two proposes that organizations recognize—and take affirmative steps to manage—ephemeral messaging risks. Guideline Three states that organizations should make informed choices and develop comprehensive use policies for ephemeral messaging applications. Guideline Four recommends that regulators, courts, and organizations consider practical approaches, including comity and interest balancing, to resolve cross-jurisdictional conflicts over corporate uses of ephemeral messaging. Guideline Five emphasizes how reasonableness and proportionality should govern discovery obligations relating to ephemeral messaging data in U.S. litigation.

6. See Sections II & III, *infra*.

7. See Section IV, *infra*.

These guidelines are designed to help organizations and their counsel, in addition to regulators and courts, as they evaluate and address conflicting obligations for organizations regarding their use of ephemeral messaging.

II. EPHEMERAL MESSAGING—NATURE AND SCOPE

Ephemeral messaging refers to secure written communications between one or more parties that are generally considered dynamic, nonstatic,⁸ and “lasting a very short time.”⁹ The two central components of ephemeral messaging that distinguish this technology from other electronic communication media are: (1) automated disposition of message content on the sender’s application *and* that of the recipient; and (2) E2E encryption functionality.

A. Automated Disposition of Message Content

As ephemeral messages are intended to be short-lived, the applications used to generate these communications are designed to enable automatic disposition or expiration of the messages. The specialized functionality of ephemeral messaging applications to delete these messages automatically or after a predefined duration (most often a very short time) also eliminates the message and (in some cases) the underlying metadata residing on the user’s application *and* on the applications of those who either sent or received the messages in question.¹⁰

8. See The Sedona Conference, *Primer on Social Media, Second Edition*, 20 SEDONA CONF. J. 1, 10 (2019) (discussing the dynamic characteristics of social media and messaging application content including that such information “may be easily modified or destroyed by the user, the recipient, the application provider, or by the technology itself.”).

9. <https://www.merriam-webster.com/dictionary/ephemeral>.

10. Wickr’s ephemeral messaging offering is one such example. *How private are my Wickr messages?*, WICKR, <https://support.wickr.com/hc/en-us/articles/115005145108-How-private-are-my-Wickr-messages> (“Wickr then deletes all metadata from its communications and our Secure File Shredder cleans the RAM after each message or picture is opened.”).

For some technologies, the deletion of such content is instantaneous upon closing the message.¹¹ For others, users can set a period of time—from moments to days or even months—before such information is discarded.¹² They can also modify retention and deletion periods by sender or recipient.¹³

B. E2E Encryption

Another significant point of distinction between ephemeral messaging and certain electronic communication tools is that of E2E encryption.¹⁴ Encryption involves the use of cryptography to take a plain text and, through use of keys and algorithms, transforms that plain text into coded text that cannot be read. At the other end, the process is reversed to

11. See *Your Confidential Messenger*, CONFIDE, <https://getconfide.com> (“Confide messages self-destruct. After they are read once, they are gone.”).

12. See *Disappearing messages for Signal*, SIGNAL, <https://signal.org/blog/disappearing-messages/> (“... any conversation can be configured to delete sent and received messages after a specified interval. The configuration applies to all parties of a conversation, and the clock starts ticking for each recipient once they’ve read their copy of the message.”).

13. See *What makes Wickr different from other productivity tools?*, WICKR, <https://support.wickr.com/hc/en-us/articles/115002632813-What-makes-Wickr-different-from-other-productivity-tools-> (“In Wickr, administrators can enforce policies for message retention similar to email retention policies. Retention can be customized for different groups of users or teams depending upon internal policies and compliance requirements.”).

14. Non-ephemeral messaging applications like iMessage may also offer users E2E encryption. *Privacy*, APPLE, <https://www.apple.com/privacy/features/> (“Your Messages and FaceTime conversations are encrypted end-to-end, so they can’t be read while they’re sent between devices.”). In contrast, workplace collaboration tools may not have the most robust forms of encryption necessary to safeguard user confidentiality. See Gennie Gebhart, *What if All Your Slack Chats Were Leaked?*, NEW YORK TIMES (July 1, 2019), <https://www.nytimes.com/2019/07/01/opinion/slack-chat-hackers-encryption.html>.

decrypt a message sent to an intended recipient. Encryption enhances privacy by making it more difficult for hackers and other unintended data recipients to read encrypted data. Encryption can take many forms and provides varying degrees of protection depending on the sophistication of the keys and algorithm.

E2E encryption provides the user with enhanced control over the disposition of messages and enables ephemeral messaging technology to support the objective of transient message content.¹⁵ This type of encryption safeguards communicated data by making it unintelligible in the absence of the algorithm and keys before the data is scheduled for expiration. By so doing, E2E encryption ensures there are no other points in the transmission chain where the data would be accessible to a third party (barring a technical flaw in the implementation of the encryption). This, in turn, typically prevents third parties from obtaining or viewing message content and other transmission details. To further enhance security of the communications and notions of user control, many ephemeral messaging technologies implement endpoint encryption schemes that typically provide no external key management or escrowing capability. This, in effect, shields message content from third parties, including the ephemeral messaging provider, its data stores, and its employees.¹⁶

15. See *Primer on Social Media, Second Edition*, *supra* note 8, at 15 (“Different applications offer competing features, including the ability to control distribution of messages (to a small group versus a community of users), message encryption, private messaging capability, prevention of screenshots, untraceable messages, and removal of messages from others’ devices.”).

16. See, e.g., *Viber Encryption Overview*, RAKUTEN VIBER, <https://www.viber.com/app/uploads/viber-encryption-overview.pdf> (“... all of Viber’s core features are secured with end-to-end encryption . . . This

C. *Other Characteristics of Ephemeral Messaging*

Beyond automated disposition and E2E encryption, ephemeral messaging applications have a variety of characteristics and features. To better understand the nature of their functionality and the corresponding impact they have on senders and recipients, this *Commentary* categorizes ephemeral messaging applications as follows: purely ephemeral, quasi-ephemeral, and non-ephemeral. These categories provide additional understanding for determining whether a messaging application is actually ephemeral and what other features might distinguish an ephemeral messaging application from one that is non-ephemeral. These categories are not mutually exclusive. Some applications may have features from more than one category. Nor are the factors delineated under the respective categories exhaustive. Certain applications may have additional features not discussed in this *Commentary*.

1. Purely Ephemeral Messaging

The following features generally characterize purely ephemeral messaging applications.

- *Deliberate, Permanent, and Automated Message Deletion Built into the Application.* This is one of the core components of an ephemeral messaging application for both the sender and the recipient of a message.

means that the encryption keys are stored only on the clients themselves and no one, not even Viber itself, has access to them.”); *Telegram FAQ*, TELEGRAM, <https://telegram.org/faq#secret-chats> (“All messages in secret chats use end-to-end encryption. This means only you and the recipient can read those messages—nobody else can decipher them, including us here at Telegram.”).

- *Unchangeable Deletion Trigger.* Once a time frame (e.g., 24 hours) or trigger (e.g., once viewed by recipient) is established for deletion, it cannot be changed after a message is sent. The time frame may be shortened or lengthened for future messages, typically with a corresponding notification to a recipient through that conversation or channel. For some applications, these triggers are built into the application's functionality as a "read and burn" function and cannot be modified.
- *No Archiving or Storage Capability.* Purely ephemeral messaging applications disable archiving and storage capacity to better ensure that content and metadata are permanently deleted. They also have mechanisms such as forwarding protection and message overwriting to safeguard message deletion. Nevertheless, indirect means of archiving, such as screen shots, are always possible. While some applications provide a warning when a screen shot is made on the same device, this is easily bypassed with a second device.¹⁷
- *Deletion Consistent within the Application for Senders and Recipients.* Senders cannot retain messages that are removed from a recipient's application and vice-versa.

17. See *United States v. Engstrom*, No. 2:15-cr-00255-JAD-PAL, 2016 WL 2904776 (D. Nev. May 16, 2016) (observing that Wickr's screen protection feature could be circumvented by taking "pictures of texts with a camera to document them.").

- *E2E encryption.* Third parties, including the application provider, cannot access message content without encryption keys.

2. Quasi-Ephemeral Messaging

The following features may characterize quasi-ephemeral messaging applications.

- *Preservation Possible in Certain Circumstances.* Applications that are quasi-ephemeral provide senders, recipients, or administrators with the ability to set deletion as a default while also configuring the application to preserve certain message content. In like manner, senders, recipients, or administrators also have the ability to override preservation as a default and implement ephemeral deletion mechanisms for certain messages, senders, recipients, or components of the application.
- *Deletion May be Impeded by External Mechanisms.* Quasi-ephemeral applications do not disable external mechanisms such as message forwarding or screenshots that prevent total deletion.
- *Content is Deleted, But Metadata is Preserved.* Quasi-ephemeral messages are completely deleted and their content is not preserved, but certain metadata—including the time a message was sent or received or the identity of the sender or recipients—is retained.
- *Combination of Other Features.* Messaging applications may be quasi-ephemeral if they

combine a series of features from both purely ephemeral and non-ephemeral applications.

3. Non-Ephemeral Messaging

The following features often characterize non-ephemeral messaging applications and are included to distinguish ephemeral messaging technologies from those that are non-ephemeral.

- *Deliberate and Permanent Message Deletion not Built into the Application.* The intentional, irrevocable deletion of messages is a key component of an ephemeral messaging application. Applications that do not provide this feature in some form cannot be considered ephemeral.
- *Deletion is Not Consistent across Senders and Recipients.* If a sender cannot automate deletion of the message from both the sender's device *and* the recipient's device, the application from which the message was sent is not ephemeral.
- *Deletion from the Application Does Not Delete Content from Other Sources.* If a message can be deleted from an application but is still kept in some format on a server, backups, or other storage mediums, the application from which the message was sent is not ephemeral.
- *Deletion Time Frame is Variable.* Where the time frame for message deletion is indefinite, can be determined or modified after the message is sent, or can be based on nontemporal factors that could accelerate deletion (such as size

limitations), the application from which the message was sent is not ephemeral.

- *Lack of E2E Encryption.* Encryption is either entirely lacking or is limited to data that is in transit and at rest. Under either scenario, third parties—including the provider—have the ability to access messages, making the application from which the message was sent non-ephemeral.

The aforementioned descriptions provide important context on how the *Commentary* views ephemeral messaging, both in terms of understanding the tensions associated with its operation and delineating guidelines regarding the use of this technology.

III. TENSIONS ASSOCIATED WITH THE USE OF EPHEMERAL MESSAGING APPLICATIONS

The widespread use of ephemeral messaging applications reduces a number of privacy and data protection risks but also creates new challenges for governments and private sector organizations. Organizations and their counsel must consider how to balance these opposing interests, taking into account the views of government regulators and courts. Section III of this *Commentary* explores the underlying nature of these considerations by examining the laws, practices, and perspectives that support and oppose the use of ephemeral messaging.

A. *Benefits of Ephemeral Messaging*

1. Organizational Benefits

There are a number of benefits of ephemeral messaging—both for organizations and for individual users. For organizations, in particular, ephemeral messaging supports information governance best practices by reducing unnecessary data. It also facilitates, among other things, compliance with legal requirements to protect personal data, privacy by design, and data security objectives.

(a) Information Governance

The massive growth in data volumes has driven organizations to adopt policies that seek to manage the life cycle of data. The focus of those policies is on retention of data with ongoing business value and early identification and action to discard data without such value. Responsible usage of ephemeral messaging tools can offer significant economies in data storage and records management. Established record retention policies naturally weigh the business value of a data

asset against the costs of retention and remove data assets that have aged beyond their use in an organization.

In practice, enforcing the deletion of obsolete data is difficult and generally not prioritized by organizations. Stale data is often challenging to destroy because its value is hard to ascertain later in time. It may require laborious review long after the reason for its creation or retention has been forgotten.

Effective governance of messaging and emails is more likely when the method is built on a “read then delete/action/store” process versus the more common accumulation without limit or until the mailbox exceeds its quota. The consequences for adopting the latter, laissez-faire approach include enforcement actions and fines against organizations that fail to remediate “data graveyards” with “years-old private data.”¹⁸ The €14.5 million fine that the Berlin Commissioner for Data Protection and Freedom of Information imposed on Deutsche Wohnen in 2019 for failing to implement an effective information management system exemplifies the folly of this approach.¹⁹

Ephemeral messaging can assist with implementation of the life-cycle process by eliminating data with no ongoing business value, particularly since a sizeable portion of the data growth involves this type of information (e.g., routine communications, meeting requests, duplicative email chains to large groups, etc.). Such a practice removes large volumes of low-value data, offering significant benefits to the organization. Likewise, information governance policies that

18. European Data Protection Board, *Berlin Commissioner for Data Protection Imposes Fine on Real Estate Company* (Nov. 5, 2019), https://edpb.europa.eu/news/national-news/2019/berlin-commissioner-data-protection-imposes-fine-real-estate-company_en.

19. *Id.*

prioritize data assets with business value, rather than controlling all information equally, enhance the usefulness of retained information and are more responsive to changing end-user preferences.

(b) Legal Compliance Support

The 1981 Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (“Convention 108”) was the first binding international law addressing privacy and data protection.²⁰ Convention 108 mandates a number of personal data protection and privacy requirements that are facilitated by ephemeral messaging, including the implementation of security measures to protect personal data, data minimization, storage limitation, and the right of individuals to have their personal data deleted. In the decades since, Convention 108 has been ratified by fifty-five countries. Numerous additional data protection laws have been adopted across the globe with similar requirements.

One of the most significant pieces of recent data protection legislation is the GDPR, which establishes data protection and privacy requirements for personal data of individuals within the European Economic Area (EEA) and governs the export of personal data outside the EEA. Like Convention 108 and the EU Data Protection Directive, the GDPR requires the implementation of security measures to protect personal data, including by imposing the principles of data minimization and storage limitation on all personal data processing operations. The GDPR also provides individuals with the right to have their personal data deleted.

20. Council of Europe, European Treaty Series No. 108 (Jan. 28, 1981), <https://rm.coe.int/1680078b37>.

The use of ephemeral messaging can facilitate GDPR compliance. The automated deletion features of ephemeral messaging applications can help meet GDPR data minimization and storage limitation requirements. Ephemeral messaging can also minimize the effort required to respond to data subject deletion or access requests, since certain data will be subject to automatic erasure. Finally, the encryption protections and automatic deletion of personal data through ephemeral messaging platforms reduces exposure in the event of a breach. Notification to data subjects is not required where the breach is not likely to result in a “high risk” to their rights and freedoms, and regulatory notification is not required where the breach is “unlikely to result in a risk to the rights and freedoms of natural persons.”²¹ The protections afforded by ephemeral messaging can reduce or eliminate these risk factors, regardless of the sensitive nature of any information communicated through ephemeral messaging.

The GDPR is particularly significant given its broad reach. It applies to organizations established within the EEA. It also applies to organizations located outside the EEA that offer goods or services in the EEA, monitor behavior of data subjects within the EEA, or to which EU law applies due to public international law. Violations of the GDPR can carry severe consequences, including regulatory penalties of up to €20,000,000 or 4 percent of global revenues, whichever is greater.²²

21. GDPR arts. 33, 34.

22. GDPR art. 82. *See* Adam Satariano, *Google Is Fined \$57 Million Under Europe’s Data Privacy Law*, NEW YORK TIMES (Jan. 21, 2019) <https://www.nytimes.com/2019/01/21/technology/google-europe-gdpr-fine.html> (discussing €50 million fine imposed by French data protection authority on Google for not disclosing how user’s data is collected across its services).

The GDPR is not the only significant data protection law that has taken effect recently. Various countries have enacted or updated data protection laws to enhance privacy safeguards in the digital age, including Australia,²³ Bermuda,²⁴ Brazil,²⁵ and Israel.²⁶ In the U.S., the Federal Trade Commission enforces data protection pursuant to Section 5 of the FTC Act,²⁷ though much of the movement on data protection has originated with state governments. For example, some state data breach statutes impose proactive storage limitation requirements.²⁸ In 2016, New York State promulgated cybersecurity regulations requiring financial institutions to develop and implement cybersecurity policies, including “policies and procedures for the secure disposal on a periodic basis of [certain] Nonpublic Information.”²⁹ The California Consumer Privacy Act (CCPA) incentivizes organizations to reduce their data footprint and enhance security protections in

23. Privacy Amendment (Notifiable Data Breaches) Act 2017.

24. Personal Information Protection Act 2016.

25. Lei Geral de Proteção de Dados Pessoais [Brazilian General Data Protection Act], Law No. 13,709/2018.

26. Protection of Privacy Regulations (Data Security) 5777-2017.

27. *See* In the Matter of Snapchat, Inc., FTC Docket No. C-4501, FTC File No. 132-3078 (December 23, 2014) (consent order) (approving final order settling charges that Snapchat misrepresented the ephemeral nature of messages sent through the service); FEDERAL TRADE COMM’N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESS AND POLICYMAKERS (March 2012), <https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf> (calling for enhanced focus on privacy, data security, and data minimization of consumer personal data).

28. *See, e.g.*, ALA. CODE 1975 § 8-38-10; COLO. REV. STAT. § 6-1-713.

29. Cybersecurity Requirements for Financial Services Companies, N.Y. DEPT. OF FIN. SERV., 23 NYCRR 500.13 (2016).

the face of statutory penalties for breaches of personal information.³⁰ The CCPA also provides individuals with new rights to access and delete their personal information.³¹

As a result of increased data protection legislative activity, ephemeral messaging may gain even more traction as a beneficial tool for legal risk mitigation.

(c) Privacy by Design

Privacy by design is an increasingly popular information management approach. It includes privacy and security protection as fundamental goals, embedding privacy into the design of the information technology system and business practices as a core functionality. This policy is designed to be proactive rather than reactive. It requires end-to-end security for the data at issue and directs operators to keep privacy as the default mode to ensure a user's privacy is protected without the user having to take any action. Operators are

30. The CCPA allows California residents the right to know the personal data collected about them, to access such data, to know whether their data has been sold or disclosed to another organization, and to refuse to allow the sale of their personal data. *See* CAL. CIV. CODE § 1798.100 *et seq.* (West 2020). Companies that suffer a security breach of personal information can be subject to a civil lawsuit and be ordered to pay California residents statutory damages of \$100-\$750 “per consumer per incident or actual damages, whichever is greater.” CAL. CIV. CODE § 1798.150(a)-(b) (West 2020).

31. Effective Jan. 1, 2023, the California Privacy Rights Act (CPRA) will replace the CCPA. The CPRA will generally augment the duties of regulated businesses toward California consumers and impose new limitations on their use of consumers' personal information. *See* Cynthia Cole, Matthew R. Baker, & Katherine Burgess, *Move Over, CCPA: The California Privacy Rights Act Gets the Spotlight Now*, BLOOMBERG LAW (Nov. 16, 2020), <https://news.bloomberglaw.com/us-law-week/move-over-ccpa-the-california-privacy-rights-act-gets-the-spotlight-now>.

accountable for the collection of data, maintaining data security, making data available to the user upon request, and protecting data with appropriate security measures.³² This emphasis on privacy encourages corporate adoption of ephemeral messaging technologies to address privacy issues.

(d) Data Security

Organizations may actively seek to use ephemeral messaging in situations where data security is paramount. For example, organizations bringing a new product to market or otherwise handling sensitive information relating to intellectual property may rely on ephemeral messaging to better ensure communications are secure and reduce the likelihood they are subject to interception.

Ephemeral messaging tools minimize the amount of data vulnerable to compromise.³³ This is one of the most effective means of ensuring data security and may prevent hackers from gaining access to important information. Even if a mobile device is lost or otherwise compromised, for example, the automatic deletion of data provides protection against loss.

Another advantage that flows indirectly from the use of ephemeral messaging is derived from E2E encryption that is integral to these platforms.³⁴ The use of reliable and easy to implement E2E encryption allows for more effective authentication of each user, something that is more difficult to do at scale with email or text messaging. This helps to secure an organization's networks by mitigating the risk of spoofed

32. See Ann Cavoukian, *Privacy by Design: The Seven Foundational Principles*, IAPP (2011) https://iapp.org/media/pdf/resource_center/pbd_implement_7found_principles.pdf.

33. See Section III.A.1.a, *supra*.

34. See Section II.B, *supra*.

senders and by ensuring the integrity and confidence in the identity of a particular sender or a group.

Even when well implemented, encryption is not foolproof. For example, it is possible to take a screenshot of an ephemeral message that has been decrypted and appears on an intended recipient's screen.³⁵ Depending on the level of security required, it may be necessary to use encryption in conjunction with other ephemeral data management methods.

(e) Productivity

Large organizations are also taking advantage of ephemeral messaging to facilitate collaboration among employees in different locales. Certain messaging applications allow personnel to work on a collaborative basis. Those applications establish data minimization processes that govern data retention on a platform so that information is not retained unnecessarily and provide E2E encryption of data, which limits access to authorized users. These features allow users to work together across the globe while reducing unnecessary retention of incidental communications and prioritizing the retention of those critical to the organization's mission. This has the further benefit of providing customers in certain circumstances with greater security regarding a corporate relationship, product or other intellectual property

35. See Section II.C, *supra*. Indeed, many encryption systems typically contain flaws of various kinds that enable decryption or allow discovery of a shortcut to the clear text. The field of cryptography is full of examples of cryptographic systems that have failed to protect the communications involved because of flaws or other design features in some part of the device or software. See Greg Miller, *How the CIA Used Crypto AG Encryption Devices to Spy on Countries for Decades*, THE WASHINGTON POST (Feb. 11, 2020), <https://www.washingtonpost.com/graphics/2020/world/national-security/cia-crypto-encryption-machines-espionage/>.

development, or joint business venture. These benefits stand in contrast to non-ephemeral workplace collaboration tools, which may have comparatively relaxed limitations on authorized users of the platform and may also lack E2E encryption to safeguard confidentiality.³⁶

2. Benefits to Individual Users

Concern over data privacy and user control of data has grown in importance in recent years. Given the raft of business and government data breaches and news stories that service providers are more focused on monetizing the value of customer data than protecting it, users have become aware that their online data may not be secure.³⁷ As a result, interest has grown in tools that give users more protection and control over their data and allow them to reduce their individual data footprints. As concepts such as data minimization and erasure gain further traction globally, ephemeral messaging offers individual users a check against unknown retention schemes and objectives.

36. See Gennie Gebhart, *What if All Your Slack Chats Were Leaked?*, NEW YORK TIMES (July 1, 2019), <https://www.nytimes.com/2019/07/01/opinion/slack-chat-hackers-encryption.html>. (“Right now, Slack stores everything you do on its platform by default—your username and password, every message you’ve sent, every lunch you’ve planned and every confidential decision you’ve made. That data is not end-to-end encrypted, which means Slack can read it, law enforcement can request it, and hackers—including the nation-state actors highlighted in Slack’s S-1—can break in and steal it.”).

37. See Christopher Mele, *Data Breaches Keep Happening. So Why Don’t You Do Something*, NEW YORK TIMES (Aug. 1, 2018), <https://www.nytimes.com/2018/08/01/technology/data-breaches.html>; Kevin Granville, *Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens*, NEW YORK TIMES (Mar. 19, 2018), <https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html>.

Certain ephemeral messaging platforms have been widely adopted on a worldwide basis. For example, WhatsApp, a messaging application offering E2E encryption and a limited automated deletion feature, is estimated to have over 2 billion users in 180 countries.³⁸ Another popular messenger service, Snapchat, which features deletion of messages after review, reports that it has approximately 238 million daily active users, with approximately 90 million North American active users and 71 million European active users.³⁹ Ephemeral messaging has become popular in part due to the enhanced control it provides to users in disseminating and deleting data as they choose.⁴⁰ Wide-scale acceptance of these applications suggests that ephemeral messaging may continue to be popular into the foreseeable future.

38. See Mansoor Iqbal, *WhatsApp Revenue and Usage Statistics (2020)*, BUSINESS OF APPS (Jan. 12, 2021), <https://www.businessofapps.com/data/whatsapp-statistics/>. The largest WhatsApp country markets are India (340 million users) and Brazil (99 million users). In some markets, including the Netherlands, Spain and Italy, WhatsApp has achieved penetration of over 80 percent.

39. Snapchat's services appear to be particularly popular with the young, reaching over 80 percent of those between the ages of 18-24 in the U.S. See Mansoor Iqbal, *Snap Inc. Revenue and Usage Statistics (2020)*, BUSINESS OF APPS (Nov. 27, 2020), <https://www.businessofapps.com/data/snapchat-statistics/>.

40. Ephemeral messaging that provides secure encryption or deletes messages after review can also have an important political role in authoritarian countries. Applications that provide users control over dissemination of data allow dissidents to engage in more secure communications, with less fear that their data and messages will be subject to interception by government officials. See Ron Synovitz, *Encrypted messaging apps struggle against authoritarian regimes*, RADIO FREE EUROPE/RADIO LIBERTY, https://internetfreedom.io/rferl_encrypted-messaging-apps.html.

B. Risks of Ephemeral Messaging

Longstanding government regulatory policies and litigation practices in the U.S. and elsewhere discourage the use of ephemeral messaging, sometimes directly but more often informally or indirectly. Organizations typically face legal and regulatory risks from the improper or sometimes unintended deletion of data. The focus in some regulated settings and in litigation contexts is often on the importance of long-term access to relevant data, and as a consequence, negative consequences can arise when such access is denied or diminished due to a failure to preserve. Because ephemeral messaging might be misused, those charged with risk management in organizations may be reluctant to adopt these technologies if they perceive a likelihood that the organization will be seen as uncooperative with law enforcement, regulators, or in litigation.

Ephemeral messaging can also disrupt traditional approaches to information governance. When data may be destroyed immediately after creation, use, or consumption, organizations will have to adjust their retention policies to either redirect certain communications to a different channel or adopt software that disables or otherwise controls data deletion in certain situations. Additionally, ephemeral messaging applications are dynamic platforms, i.e., features may be removed, changed, or added without the knowledge or consent of the organization. This aspect of ephemeral messaging injects unpredictability to data resources that are volatile by design. Accordingly, the risks and consequences of improper data deletion may be amplified and should be considered before an ephemeral messaging application is deployed. Specific regulatory and legal risks are considered in turn below.

1. Regulatory Risks

As noted above, the focus in some regulated settings is often on the importance of long-term access to relevant data, which conversely can lead to serious negative consequences when such access is denied or diminished due to a failure to preserve. Complying with regulatory controls that require strict retention protocols, including various reporting and audit requirements, is often seen as a key inhibitor to adopting ephemeral messaging. In addition, certain organizations must securely retain particular classes of information or risk robust penalties for noncompliance.

For example, the U.S. SEC's National Office of Compliance Inspections and Examinations advises regulated entities to specifically prohibit "business use of apps and other technologies that can be readily misused by allowing an employee to send messages or otherwise communicate anonymously, allowing for automatic destruction of messages, or prohibiting third-party viewing or back-up."⁴¹ This guidance, coupled with the requirement that brokers, dealers, and traders keep all communications "relating to its business as such" for three years, could limit the ability of organizations in the financial services industry to use ephemeral messaging.⁴²

Organizations seeking to demonstrate cooperation in Foreign Corrupt Practices Act (FCPA) investigations must

41. See *National Exam Program Risk Alert*, OFFICE OF COMPLIANCE INSPECTIONS AND EXAMINATIONS (Dec. 14, 2018), <https://www.sec.gov/files/OICIE%20Risk%20Alert%20-%20Electronic%20Messaging.pdf>.

42. See 17 C.F.R. § 240.17a-4(b)(4). See also Sridhar Natarajan, Michelle Davis & Dan Wilchins, *JPMorgan Puts Senior Credit Trader on Leave Over WhatsApp Use*, BLOOMBERG (Jan. 13, 2020), <https://www.bloomberg.com/news/articles/2020-01-13/jpmorgan-puts-senior-credit-trader-on-leave-over-whatsapp-use>.

satisfy standards that the U.S. DOJ has promulgated regarding the use of ephemeral messaging. The most recent FCPA guidance states that cooperation can be shown by “appropriate retention of business records . . . including implementing appropriate guidance and controls on the use of personal communications and ephemeral messaging platforms.”⁴³ Nevertheless, this guidance should be construed in the context of the U.S. DOJ’s historical antipathy toward the use of ephemeral messaging.⁴⁴

Similarly, the U.S. DOJ’s Antitrust Division recently promulgated guidance regarding the role of information governance as it relates to developing antitrust compliance programs that require regulatory approval. According to the U.S. DOJ, a key aspect of such information governance should include controls for evaluating “new methods of electronic communication” and addressing “the antitrust risk associated with these new forms of communication.”⁴⁵ While the DOJ guidance does not specifically mention ephemeral messaging,

43. See FCPA Corporate Enforcement Policy (2018), United States Department of Justice, Justice Manual, 9-47.120(3)(c), <https://www.justice.gov/jm/jm-9-47000-foreign-corrupt-practices-act-1977>.

44. The U.S. DOJ previously published FCPA guidance on November 29, 2017, generally disapproving the use of ephemeral messaging. See Philip Favro, *Ephemeral Messaging: Balancing the Benefits and Risks*, at *6, PRACTICAL LAW (2020). That guidance declared as follows: “The following items will be required for a company to receive full credit for timely and appropriate remediation . . . Appropriate retention of business records, and prohibiting the improper destruction or deletion of business records, including prohibiting employees from using software that generates but does not appropriately retain business records or communications.”

45. Press Release, U.S. Dept. of Justice, Antitrust Division Announces New Policy to Incentivize Corporate Compliance, (July 11, 2019), <https://www.justice.gov/opa/pr/antitrust-division-announces-new-policy-incentivize-corporate-compliance>.

an organization may want to consider developing a written policy that sets out its business needs for use of an ephemeral messaging application and provides guidance for using that application. As detailed in Guideline Two and Guideline Three of this *Commentary*, the policy could also discuss the benefits and risks of the application and identify appropriate risk mitigation strategies that the organization has implemented.

Beyond the U.S., regulators in other countries and regions have expressed concerns with encrypted messaging applications, which encompass ephemeral messaging. These concerns focus on both the lack of information that encrypted messages retain for investigative purposes and how they may prevent organizations from monitoring message content. These concerns have resulted in enforcement actions in the United Kingdom (U.K.) and in Europe against organizations and individuals using encrypted messaging.⁴⁶ In particular, the U.K.'s Financial Conduct Authority has taken action against firms and individuals that have used WhatsApp to transmit sensitive information and conduct deal and investment-related activities.⁴⁷ To address concerns, organizations may consider selecting ephemeral messaging applications that have features and functionality that allow for retention of message content.⁴⁸ Organizations may also consider memorializing their

46. See, e.g., Council of the European Union, *Council Resolution on Encryption—Security through encryption and security despite encryption*, (Nov. 24, 2020), <https://data.consilium.europa.eu/doc/document/ST-13084-2020-REV-1/en/pdf>.

47. See Financial Conduct Authority, *Newsletter on market conduct and transaction reporting issues* (Jan. 2021), <https://www.fca.org.uk/publications/newsletters/market-watch-66>.

48. See Section IV.C, *infra*.

technology selection process into an overall ephemeral messaging use policy.⁴⁹

2. Legal Risks

The use of ephemeral messaging poses risks to any party that must retain information for a legal matter. Noncompliance with common law or court-imposed retention requirements may impact the organization's ability to assert or defend its claims in legal actions, run afoul of discovery obligations in litigation, or invite further scrutiny into its affairs.

A primary consideration for organizations that are subject to U.S. jurisdiction is the duty to preserve information relevant to reasonably anticipated or pending litigation in the U.S.⁵⁰ Failure to comply with this duty may expose an organization to legal consequences that can significantly add to the time and costs required to litigate a matter, regardless of the merits of the underlying lawsuit. As a result, the duty to preserve creates a separate and distinct set of risks that may involve records beyond those normally retained for operational utility. Once a duty to preserve has been triggered, a company must take steps to preserve data as required by a particular jurisdiction. Organizations may need to have policies and procedures to allow for the suspension of the use of ephemeral messaging for affected custodians or alternatively disable the ephemerality function as to affected custodians until a preservation obligation has been satisfied.⁵¹

49. *See id.*

50. *See* DR Distributions v. 21 Century Smoking, Inc., --- F. Supp. 3d ---, 2021 WL 185082, at *54 (N.D. Ill. Jan. 19, 2021) (“Once a party reasonably anticipates litigation, it is duty-bound to take good faith steps to preserve documents and data that may be relevant to the litigation.”).

51. *See* Section IV.E, *infra*.

3. Operational Risks

The massive increase in data volumes in most organizations is primarily due to the daily flow of operations information. The challenge for most organizations is managing this information—including communications—in a way that does not overwhelm their systems. Organizations need to ensure business records are both accessible and properly retained to safeguard their integrity. At the same time, they should also develop protocols and procedures to dispose of nonessential materials. Depending on the industry, various corporate communications may fall outside the ambit of business records and may not require long-term retention. In that event, organizations may use ephemeral messaging in the same way as email, or they may choose to limit the scope of use to text messages.

Adoption of ephemeral messaging systems may pose operational risks to organizations regarding the governance of its information. Information governance is premised on notions of transparency regarding the information an organization generates, receives, and maintains. It generally requires the implementation of corporate policies and procedures both to enforce these principles and to accomplish corporate information objectives. Policies and procedures can define and implement controls regarding the types of business records that a company requires to be stored for certain periods of time or in certain locations. The policies and procedures can also be designed to disallow the use of ephemeral messaging with respect to certain categories of records, to provide guidance on the types of records that require retention, and to identify those that may be appropriate for ephemeral systems, such as those with no ongoing business value. Without such policies or procedures, organizations may risk not retaining essential records,

communications, or other information required for business purposes and legal and regulatory needs.

IV. GUIDELINES

Against the backdrop of these conflicting considerations, this *Commentary* has promulgated five guidelines regarding use of ephemeral messaging. These guidelines provide recommendations for organizations and their counsel, along with government regulators and courts, that spotlight how they can best implement, evaluate, or address organizational use of ephemeral messaging.⁵²

A. Guideline One: Regulators and Courts Should Recognize that Ephemeral Messaging May Advance Key Business Objectives

Regulators and courts should acknowledge that ephemeral messaging applications may be a valuable aspect of an organization's information governance program. Ephemeral messaging offers automated message deletion and E2E encryption, which can confer significant business benefits. Those benefits include confidentiality and security for sensitive electronic information in the face of increasing threats of inadvertent disclosure of such information.⁵³

Other benefits include data minimization, which ephemeral messaging facilitates by reducing data volumes and safeguarding personal information. Limiting the retention of corporate data that has no ongoing business value and decreasing the risk of exposing personal data to third parties are recognized as proper information governance practices and

52. Guideline One and Guideline Two, which respectively address the benefits and risks of ephemeral messaging, should be considered holistically.

53. *Cf.* Health Insurance Portability and Accountability Act of 1996, 45 CFR § 164.306 (requiring covered entities and business associates to implement security policies and procedures to protect patient data).

as key components of safeguarding sensitive user information under the principle of privacy by design.⁵⁴

Given these considerations, regulators and courts may view ephemeral messaging as facilitating corporate compliance with data protection laws, including the GDPR. Satisfying these laws is an increasingly significant business imperative given the growing importance of privacy—both internationally and domestically—for organizations and individuals.

Regulators and courts may also consider the benefits surrounding ephemeral messaging in connection with four principal areas of information governance: recordkeeping, data preservation, regulatory scrutiny, and cross-border data transfers.⁵⁵ Concerns over the interplay of ephemeral messaging and these four areas can impact a party's legal interests as well as its reputation.⁵⁶ This is particularly the case where regulators and courts may be inclined to presume that ephemeral messaging is a means to conceal improper conduct. While ephemeral messaging—like phone calls, email, and

54. Cf. Federal Trade Commission Staff Report, *Internet of Things: Privacy & Security in a Connected World* (January 2015), at 33 *et seq.*; GDPR, *supra* note 1, art. 1(c).

55. See Section III.A.1, *supra*.

56. Robert Mueller observed in his report regarding interference into the 2016 U.S. presidential election that certain witnesses “deleted relevant communications or communicated during the relevant period using applications that feature encryption or that do not provide for long-term retention of data or communications records” and thereby prevented the corroboration of witness statements through contemporaneous communications or the use of such communications to “shed additional light on (or cast in a new light) the events described in the report.” U.S. Department of Justice, *Report On The Investigation Into Russian Interference In The 2016 Presidential Election*, Special Counsel Robert S. Mueller, III, at *10 (Mar. 2019), <https://www.justice.gov/storage/report.pdf>.

traditional messaging apps—has been used for improper purposes, such a perception may be tempered as regulators and courts consider the business purposes served by this technology.⁵⁷

Perhaps reflecting such understanding, the U.S. DOJ's 2019 FCPA Corporate Enforcement Policy recently abandoned its express prohibition against ephemeral messaging by organizations seeking cooperation credit. Instead, the U.S. DOJ ostensibly provides organizations more latitude to adopt ephemeral messaging and other technologies to further information retention policies and practices that satisfy business objectives. This change to the U.S. DOJ's FCPA Enforcement Policy may be viewed as recognizing the increasing importance of ephemeral messaging to advance those objectives.

Regulators and courts may consider evaluating the attendant circumstances surrounding an organization's use of ephemeral messaging. This includes the various technological

57. *See generally* *Arthur Andersen LLP v. United States*, 544 U.S. 696, 704 (2005) (“‘Document retention policies,’ which are created in part to keep certain information from getting into the hands of others, including the Government, are common in business.”); *Micron Tech., Inc. v. Rambus Inc.*, 645 F.3d 1311, 1322 (Fed. Cir. 2011) (“where a party has a long-standing policy of destruction of documents on a regular schedule, with that policy motivated by general business needs, which may include a general concern for the possibility of litigation, destruction that occurs in line with the policy is relatively unlikely to be seen as spoliation.”); *Phillip M. Adams & Assoc., L.L.C. v. Dell, Inc.*, 621 F. Supp. 2d 1173, 1193 (D. Utah 2009) (“A court—and more importantly, a litigant—is not required to simply accept whatever information management practices a party may have. A practice may be unreasonable, given responsibilities to third parties. While a party may design its information management practices to suit its business purposes, one of those business purposes must be accountability to third parties.”).

aspects of ephemeral messaging applications. The most widely used applications allow the user to set whether messages will be deleted after a certain time or after being read. Others, aimed at enterprise level applications, provide more centralized control by the organization. As detailed more fully in Guideline Three of this *Commentary*, the features that ephemeral messaging technologies offer are important considerations when organizations select their communication platforms and develop their approach to an ephemeral messaging policy. Regulators and courts need not defer to an organization's use of ephemeral messaging where the selected application cannot be configured to align and satisfy its obligations for information retention.

When organizations have implemented an ephemeral messaging program consistent with the recommendations memorialized in Guideline Two and Guideline Three of this *Commentary*, regulators and courts may consider that such a program (absent contrary circumstances) is both reasonable and executed in good faith.

B. Guideline Two: Organizations Should Take Affirmative Steps to Manage Ephemeral Messaging Risks

Organizations should be aware that communication channels leaving no evidence of wrongdoing may be favored by those engaging in secretive activity for an improper purpose.⁵⁸ Organizations should also understand that

58. In the U.S., an axiom among political insiders states that one should never send an email when a phone call suffices; never make a call when an in-person meeting is possible; and never say something when a nod can get the point across. Similarly, traders at a prominent financial services company once devised the abbreviation "LDL" (let's discuss live) as a way to take an email exchange into a phone conversation to avoid creating an incriminating trail. See Virginia Heffernan, *The Trouble With E-Mail*, THE

ephemeral messaging can provide an effective means for misconduct by enabling more communication than would be possible by phone or even in person and by allowing the sharing of documents or other data. Ephemeral messaging also facilitates the disappearance of a communication (including its metadata) after it is read by the recipient. This may not be possible with a telephone call or in-person meetings, particularly with the technology now available for tracking the use of mobile phones.

As a result, organizations should carefully select and evaluate their use of ephemeral messaging. As described in Guideline Three of this *Commentary*, an organization's data preservation policy and communications, including any information retention directive, should address the use of ephemeral messaging. This may include extending the duty to preserve (where applicable) to records generated by an ephemeral messaging application. It may also include addressing all forms of sanctioned and nonsanctioned use of such applications for both legal and improper purposes.

If an application does not have legal-hold capability that can retain communications in the event of a data retention directive, the organization should consider reasonable alternatives for addressing retention, including a possible prohibition on the use of ephemeral messaging. A defined policy and evidence of compliance should provide strong support if an organization is called upon by regulators or courts to demonstrate the reasonableness of its ephemeral messaging program. This policy should contemplate the opportunities for misconduct both within the selected enterprise ephemeral application as well as by consumer-

grade, nonsanctioned ephemeral messaging tools that employees may use.

Even with appropriate technology selections and policy implementation, there may be instances where the potential benefits of ephemeral messaging do not outweigh the risks. For example, organizations in highly regulated industries that use purely ephemeral messaging to communicate about regulated aspects of their business may not be able to satisfy recordkeeping requirements or regulatory audits or examinations. The risks associated with regulatory noncompliance or adverse findings from a regulatory examination may exceed the potential benefits gained through data minimization activities, enhanced security, or other organizational benefits.

C. Guideline Three: Organizations Should Make Informed Choices and Develop Comprehensive Use Policies for Ephemeral Messaging Applications

Organizations should consider evaluating which ephemeral messaging applications best address their regulatory, litigation, and business needs. Available technologies offer a range of applications depending on an organization's industry, size, global presence, litigation profile, and appetite for risk.

An organization contemplating the use of ephemeral messaging may consider engaging in a structured approach to selecting an ephemeral messaging technology. Such an approach could involve identifying stakeholders within the organization who can evaluate the appropriate features for such an application. After reaching a determination of those features, the stakeholders could then recommend applications or technologies that best meet the organization's needs. Depending on the size of the organization and the nature of

the process sought to be followed, the stakeholders might include representatives from legal, IT, information security, data privacy, document management, and appropriate business units.⁵⁹

An integral aspect of the stakeholders' work could include the preparation of a written policy addressing the use of ephemeral messaging within the organization. For many organizations, a comprehensive policy that identifies the benefits of ephemeral messaging, the corresponding risks, and actionable risk mitigation measures may be essential for demonstrating the business-use case of this technology to skeptical insiders and outsiders.⁶⁰ Depending on the nature of the organization and its industry profile, this may include company executives, shareholders, regulators, litigation adversaries, courts, and the public.

Depending on its needs and the type of application selected, the organization may decide that acceptable uses should be limited to logistical communications (scheduling calls or meetings) or a slightly broader category of nonsubstantive communications. Alternatively, acceptable uses may include specific types of business communications or other special circumstances. For example, in the incident response field, using out-of-band communications has long been an accepted and highly recommended practice.⁶¹

59. Having a member of the executive team on the committee will help ensure senior management support for this effort and can promote acceptance of ephemeral messaging application(s) and associated policies and practices.

60. See Favro, *supra* note 44, at *6.

61. Out-of-band communication should be reliable and secure in the event that a cyber intruder is monitoring email systems. Ephemeral messaging is ideal for this scenario as it allows for speed in response and security and enhances the openness of the team in communicating

Ephemeral messaging may also be advisable for certain internal investigations involving cross-border matters where counsel is seeking to protect information from third parties to better ensure that the matter is addressed with strict confidentiality. Finally, ephemeral messaging might be used for one-way communication from the organization to recipients where, at the same time, a backend system would store the substance and metadata of the communication.⁶²

In drafting the policy, organizations should understand that the most important information governance factors related to ephemeral messaging are legal-hold capabilities and the availability of customizable retention periods. Organizations that prefer to keep data for longer periods may value the security features of ephemeral messaging more than the opportunities for data minimization. Those organizations will therefore select an application with longer retention periods and the ability to effect legal-hold functionality when the need arises. Other organizations may prioritize minimizing the

information. *See, e.g.*, The Sedona Conference, *Incident Response Guide*, 21 SEDONA CONF. J. 125, 157–60 (2020) (“In the event of a significant cybersecurity incident or intrusion . . . it is essential to have reliable communication channels available to keep key players and essential stakeholders informed, and to lead and manage the incident response. In some cases, this may require alternative (and secure) communications channels. As with other incident response preparations, alternative communications channels should be planned and provisioned in advance to handle situations where corporate communications systems have been completely disrupted.”).

62. *C.f.* *Toftely v. Qwest Commc’ns Corp.*, No. C3-02-1474, 2003 WL 1908022, at *1 (Minn. App. Apr. 22, 2003) (denying plaintiff employment benefits because she was discharged for violating the company’s confidentiality policy by disclosing to a third party a confidential litigation hold instruction with an embedded “electronic tracer” that allowed defendant to monitor whether the message was forwarded outside the company).

volume of data retained and may instead choose a technology with shorter retention periods, disabling the application entirely once a legal hold is implemented. Organizations may alternatively select a middle ground, allowing employees to communicate with ephemeral messaging until a legal-hold obligation arises, at which time use of the application by key custodians of relevant information may be disabled or otherwise prohibited for any communications related to the subject matter of the hold.

An organization may also choose to adopt more than one ephemeral messaging application to maximize the possible number of acceptable uses. For example, one application could be permitted for all employees, but limited to logistical communications. Another application could be designated for specific departments relating to limited types of communications. Irrespective of the technology selected, organizations should consider the benefits of forbidding employees from using consumer applications for individual, unstructured, or one-off business purposes.

Once implemented, the ephemeral messaging policy should be followed by employee education and training, together with periodic auditing of use and rule observance to better ensure compliance.

In adopting an ephemeral messaging program, the organization should consider undertaking a thorough data mapping exercise to allow data managers to understand how the ephemeral messaging application interacts with other data systems.

With sufficient documentation of acceptable uses and data retention requirements, and selection of appropriate technologies tailored to their requirements, organizations can

better assess and manage risk in taking advantage of the benefits of ephemeral messaging.

D. Guideline Four: Regulators, Courts, and Organizations Should Consider Practical Approaches, Including Comity and Interest Balancing, to Resolve Cross-Jurisdictional Conflicts over Ephemeral Messaging

Conflicts with legal or regulatory requirements may arise where the use of ephemeral messaging fulfills applicable requirements in one jurisdiction while simultaneously conflicting with obligations in another jurisdiction. This is particularly the case with cross-border data transfers where the understanding and priority accorded data privacy and information retention differ between jurisdictions and where conflicts may arise between data retention and data minimization requirements.⁶³ To address these issues, regulators, courts, and organizations may find notions of comity, interest balancing, or other accommodations useful for resolving cross-jurisdictional conflicts over corporate uses of ephemeral messaging.

One accommodation that government regulators might consider is modeling enforcement policies after the U.S. DOJ's 2019 FCPA Corporate Enforcement Policy. Such an approach

63. Compare *Behrens v. Arconic, Inc.*, No. 19-2664, 2020 WL 1250956 (E.D. Pa. Mar. 13, 2020) (citing comity for the French Blocking Statute as a key basis for denying plaintiffs' motion to compel the production of documents pursuant to the Federal Rules of Civil Procedure from the French subsidiary of a defendant rather than resorting to Hague Convention procedures) with *In re Mercedes-Benz Emissions Litig.*, No. 16-cv-0881 (KM) (ESK), 2020 WL 487288 (D.N.J. Jan. 30, 2020) (reasoning that the GDPR and "considerations of international comity" did not relieve defendants from their duty to produce employee names, titles, dates of employment, organizational charts, and other relevant information).

would shift the focus from outright proscription to examining the basis for the organization's implementation of ephemeral messaging, along with related guidance and controls. This would allow the use of ephemeral messaging systems in appropriate cases while also addressing regulatory concerns about unlawful conduct facilitated by ephemeral messaging.

Courts and parties might also consider accommodations to address inconsistent obligations arising from the conflict of international data protection laws and preservation and production requirements in common law litigation over ephemeral messaging data.⁶⁴ If a conflict is found, the parties—and if needed, the court—could define the appropriate scope of preservation and production by balancing the competing needs of the litigation, the consequences of any potential violations of applicable data protection laws, the impact on affected data subjects, and other pertinent considerations.⁶⁵

64. See generally The Sedona Conference, *International Principles on Discovery, Disclosure & Data Protection in Civil Litigation* (Transitional Edition) (2017), https://thesedonaconference.org/publication/International_Litigation_Principles (describing tension between U.S. discovery and preservation obligations and non-U.S. data protection laws). See also Loi 80-538 du 16 juillet 1980 [French Penal Law No. 80-538 of July 16, 1980], JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE [OFFICIAL GAZETTE OF FRANCE] (blocking statute prohibiting the transfer of data for the purpose of discovery in foreign litigation); *In re Advocat "Christopher X,"* Cour de cassation [Cass.] Paris, crim., Dec. 12, 2007, No. 07-83228 (enforcing blocking statute by fining French lawyer €10,000 for obtaining evidence from a French insurer for use in civil litigation pending in the United States).

65. Compare *Salt River Project Agric. Improvement and Power Dist. v. Trench France SAS*, 17-cv-01468-DGC, 2018 WL 1382529 (D. Ariz. Mar. 19, 2018) (refusing to order defendant to immediately produce relevant documents stored in France outside the bounds of Hague Convention procedures on cross-border discovery) with *In re Mercedes-Benz Emissions Litig.*, No. 16-cv-881 (KM) (ESK), 2020 WL 487288 (D.N.J. Jan. 30, 2020)

Courts in common law jurisdictions may consider allaying preservation and production requirements for ephemeral messaging data where organizations use ephemeral messaging to comply with data minimization principles of cross-border data protection laws.⁶⁶ Those same considerations should also apply when conflicts arise relating to an organization's use of ephemeral messaging to meet domestic data privacy requirements or satisfy other business objectives.

Organizations may also need to consider how to implement approaches to address discovery and data minimization conflicts. One option could include building in accommodations for evaluating whether to deploy ephemeral messaging applications in limited geographic regions (localization) or for specific company divisions. The organization also could implement applications that have technological features allowing otherwise ephemeral messages to be kept in circumstances where a preservation duty is triggered.⁶⁷

E. Guideline Five: Reasonableness and Proportionality Should Govern Discovery Obligations Relating to Ephemeral Messaging Data in U.S. Litigation

Ephemeral messaging data that is stored temporarily is electronically stored information (ESI), even if it may not be reasonably accessible in certain circumstances.⁶⁸ ESI that does

(ordering the production of documents with employee names, titles, employment dates, organization charts, and other materials reflecting personal data and holding that a protective order would sufficiently safeguard such information for GDPR purposes).

66. See *Salt River*, 2018 WL 1382529, at *3-4.

67. See Section IV.C, *supra*.

68. See FED. R. CIV. P. 34(a)(1) advisory committee note to 2006 amendment ("Rule 34(a)(1) is expansive and includes any type of

not exist at the time a preservation duty triggers is not subject to a preservation obligation.⁶⁹

For prospective preservation obligations (i.e., where information is created after the duty to preserve attaches), preservation of relevant ephemeral messaging data may be required, though it will be limited by considerations of reasonableness. For example, it is generally recognized that the preservation obligation requires reasonable, good-faith efforts

information that is stored electronically . . . ‘in any medium,’ to encompass future developments in computer technology.”); *Columbia Pictures, Inc. v. Bunnell*, 245 F.R.D. 443, 446 (C.D. Cal. 2007) (holding that temporarily stored information is electronically stored information under Rule 34). If ephemeral messaging data is truly not “stored in any medium from which information can be obtained,” then such data would not qualify as “electronically stored information” for the purposes of discovery. FED. R. CIV. P. 34(a)(1) advisory committee note to 2006 amendment.

69. *See, e.g.*, FED. R. CIV. P. 37(e) advisory committee note to 2015 amendment (“court decisions hold that potential litigants have a duty to preserve relevant information when litigation is reasonably foreseeable”). For example, courts have not sanctioned parties that configured instant messaging systems to not retain messages. *See, e.g.*, *Williams v. UnitedHealth Group*, No. 2:18-cv-2096, 2020 WL 528604 (D. Kan. Feb. 3, 2020) (finding that defendant did not violate its preservation or production duties by configuring its Cisco Jabber instant messaging system to not retain instant messages); *King v. Catholic Health Initiatives*, No. 8:18-cv-326, 2019 WL 6699705 (D. Neb. Dec. 9, 2019) (holding that defendant did not have a preservation or production obligation relating to instant messages generated by its Microsoft Lync instant messaging system where it designed that system to not retain instant messages). *But see* *Franklin v. Howard Brown Health Ctr.*, No. 1:17 C 8376, 2018 WL 4784668 (N.D. Ill. Oct. 4, 2018); *report and recommendation adopted*, 2018 WL 5831995 (N.D. Ill. Nov. 7, 2018) (imposing sanctions on defendant for failing to preserve relevant messages from its instant messaging system where defendant configured the system to keep messages for up to two years).

as opposed to perfection.⁷⁰ The determination of this issue could largely depend on the preservation capabilities of the particular application used.⁷¹

Spoliation may occur when a party fails to take reasonable steps to preserve data that is lost and cannot be restored or replaced through additional discovery.⁷² Nevertheless, courts in U.S. litigation should be aware that organizations—particularly with cross-border operations—may use ephemeral messaging to comply with international and domestic privacy norms, along with other corporate objectives.⁷³ As a result, courts should not reflexively presume that ephemeral messaging has been implemented to avoid common law preservation obligations.⁷⁴

70. FED. R. CIV. P. 37(e) advisory committee note to 2015 amendment (“This rule recognizes that “reasonable steps” to preserve suffice; it does not call for perfection.”); *DR Distribs. v. 21 Century Smoking, Inc.*, --- F. Supp. 3d ---, 2021 WL 185082, at *54 (N.D. Ill. Jan. 19, 2021) (“Though a party need not preserve all documents in its possession—again, perfection is not the standard—it must preserve what it knows and reasonably ought to know is relevant to possible litigation and is in its possession, custody, or control.”); The Sedona Conference, *The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 SEDONA CONF. J. 1, 108, 111 (2018) (providing that “the obligation to preserve normally requires reasonable and good faith efforts” and that a “party’s preservation obligation does not require ‘freezing’ of all ESI”).

71. See Section IV.C, *supra*.

72. FED. R. CIV. P. 37(e).

73. Guideline Five focuses on U.S. litigation in federal courts. Nevertheless, the principles discussed in Guideline Five would be applicable to U.S. state courts or investigatory litigation as well.

74. *Contra WeRide Corp. v. Kun Huang*, No. 5:18-cv-07233, 2020 WL 1967209 (N.D. Cal. Apr. 24, 2020) (imposing terminating sanctions against defendants for, among other things, deploying an enterprise grade ephemeral messaging application (DingTalk) ostensibly to circumvent a

Instead, courts should examine the nature and use of ephemeral messaging against the recommendations memorialized in Guideline Two and Guideline Three of this *Commentary*. In the absence of contrary circumstances, courts may consider a litigant's use of ephemeral messaging that accords with Guideline Two and Guideline Three as being reasonable and executed in good faith. In contrast, it may be appropriate for courts to infer culpable intent with respect to prospective preservation obligations if a litigant's key custodians of relevant information begin to use or continue using ephemeral messaging *after* a duty to preserve has triggered.⁷⁵

As with all preservation obligations, the parties and the court must also consider proportionality factors.⁷⁶ Factors particularly applicable to the preservation of relevant ephemeral messaging data include the accessibility of the information, the relative burdens and costs of the preservation effort, and the probative value of the information.⁷⁷ Privacy considerations, along with the other proportionality

preservation order and to prevent the discovery of relevant communications).

75. See *id.*; *Herzig v. Arkansas Found. for Med. Care, Inc.*, No. 2:18-cv-02101, 2019 WL 2870106 (W.D. Ark. July 3, 2019).

76. See FED. R. CIV. P. 26(b)(1) and 37(e), including advisory committee's note to 2015 amendment: "[T]he routine, good-faith operation of an electronic information system would be a relevant factor for the court to consider in evaluating whether a party failed to take reasonable steps to preserve lost information." The Sedona Conference, *Commentary on Legal Holds, Second Edition: The Trigger & The Process*, 20 SEDONA CONF. J. 341, 367 (2019) (discussing at Guideline 6 that "[f]ulfilling the duty to preserve involves reasonable and good-faith efforts . . . applied proportionately.").

77. *Commentary on Legal Holds, Second Edition*, *supra* note 76, at 367 (Guideline 7). See also The Sedona Conference, *Commentary on Proportionality in Electronic Discovery*, 18 SEDONA CONF. J. 141, 150 (2017).

standards—the importance of the issues at stake in the action, the amount in controversy, and the parties’ respective available resources for discovery—are other factors that may merit consideration by the parties and court.⁷⁸

Even if ephemeral messaging technologies enable the preservation of relevant data, a blanket requirement to create records of ephemeral messaging content—thereby converting such content to non-ephemeral information—while litigation is pending could be too onerous.⁷⁹ This is particularly the case where organizations have implemented ephemeral messaging to advance business imperatives such as data minimization, security, and confidentiality.⁸⁰ Instead, it could be appropriate

78. Compare *Henson v. Turn, Inc.*, No. 15-cv-01497-JSW (LB), 2018 WL 5281629 (N.D. Cal. Oct. 22, 2018) (forbidding unfettered discovery of plaintiffs’ web browsing and related social media history given their privacy interests in certain information) with *In re Mercedes-Benz Emissions Litig.*, No. 16-cv-0881 (KM) (ESK), 2020 WL 487288 (D.N.J. Jan. 30, 2020) (finding that a protective order would adequately protect cross-border privacy interests during discovery in U.S. litigation). See also *Commentary on Proportionality in Electronic Discovery*, *supra* note 77, at 168–73 (explaining that privacy rights should be taken into account when determining the application of proportionality standards); Agnieszka A. McPeak, *Social Media, Smartphones, and Proportional Privacy in Civil Discovery*, 64 U. KAN. L. REV. 235 (2015) (arguing that privacy should be a factor in the proportionality analysis).

79. See *Commentary on Legal Holds, Second Edition*, *supra* note 76, at 395–96 (“Absent a showing of special need, *The Sedona Principles, Third Edition* states that a responding party should not be required to ‘preserve, review, or produce deleted, shadowed, fragmented, or residual [ESI].’”).

80. Notably, preservation of ephemeral messaging data may be unnecessary. Regulatory requirements may already mandate creation and retention of certain business records, and ephemeral communications are unlikely to be used for business records to which other retention requirements already apply. See, e.g., Home Mortgage Disclosure Act of 1975, 12 U.S.C. 2801 (1976) (requiring retention of certain information about

to treat an ephemeral message like a phone call rather than an email and refrain from imposing a duty to create and maintain records of *all* ephemeral messaging data.⁸¹ At the same time, organizations should be cognizant that the adoption and use of ephemeral messaging carries risks both in civil litigation and regulatory investigations.⁸²

If ephemeral messaging data satisfies notions of relevance and proportionality,⁸³ a court may then need to determine whether the data is reasonably accessible.⁸⁴ In connection with its analysis of this issue, a court may examine the nature of the ephemeral messaging applications at issue. For applications that do not have the technical functionality to preserve and in fact do not retain an active version of the data, a court may then consider whether such data is either not reasonably accessible because of undue burden or cost, or completely

mortgage applications for three years); Occupational Safety and Health Standards, 29 C.F.R. pt. 1910 (1993) (applying specific retention periods for payroll records, tax forms, human resource records, and other employee files); Federal Deposit Insurance Corporation Record Retention Requirements, 12 C.F.R. pt. 380.14 (2016) (mandating retention of internal company retention policies); Health Care Portability and Accountability Act, 45 C.F.R. pt. 160 (2007) (requiring maintenance of certain records under the “security rule”).

81. Although there is no duty to create a recording of a phone call, for example, a company that already records conversations for business purposes would have a duty to preserve those recordings. *See E*Trade Secs. LLC v. Deutsche Bank AG*, 230 F.R.D. 582, 590 (D. Minn. 2005).

82. *See* Section IV.B, *supra*.

83. *See* FED. R. CIV. P. 26(b)(1).

84. *See* FED. R. CIV. P. 26(b)(2)(B). The limits under Federal Rule of Civil Procedure 26(b)(2)(C) would also apply, including whether the discovery is unreasonably cumulative or duplicative, or can be obtained from some other source that is more convenient, less burdensome, or less expensive. FED. R. CIV. P. 26(b)(2)(C)(i)-(iii).

inaccessible. For purely ephemeral messaging applications with automated deletion and E2E encryption that eliminates encryption keys, any remnants of that content will likely be completely inaccessible and beyond recovery. In contrast, data from quasi-ephemeral messaging applications may be recoverable as not-reasonably-accessible data, depending on the nature of an application's storage and encryption features.⁸⁵

85. Although certain ephemeral messaging applications give users the ability to save some data, the mere existence of such settings should not convert ephemeral messages to "reasonably accessible" data given the burden that retention may impose in the face of data protection regulations and security considerations. *See* Section III.A.1, *supra*.

THE SEDONA CONFERENCE COMMENTARY ON
QUANTIFYING VIOLATIONS UNDER U.S. PRIVACY LAWS

*A Project of The Sedona Conference Working Group
on Data Security and Privacy Liability (WG11)*

Author:

The Sedona Conference

Editor-in-Chief:

James J. Pizzirusso

Contributing Editors:

Mark Bailey

Stephen Y. Chow

Ross M. Gotler

Amy E. Keller

Timothy R. Murphy

Kaleigh N. Powell

Jonathan M. Wilan

Steering Committee Liaison:

Al Saikali

Staff Editors:

David Lumia

Michael Pomarico

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 11. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

We thank all of our Working Group Series Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors, just click on the “Sponsors” navigation bar on the homepage of our website.

This publication may be cited as follows:

The Sedona Conference, *Commentary on Quantifying Violations under U.S. Privacy Laws*, 22 Sedona Conf. J. 489 (2021).

PREFACE

Welcome to the July 2021 final version of The Sedona Conference *Commentary on Quantifying Violations under U.S. Privacy Laws* (“*Commentary*”), a project of The Sedona Conference Working Group 11 on Data Security and Privacy Liability (WG11). This is one of a series of Working Group commentaries published by The Sedona Conference, a 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, intellectual property rights, and data security and privacy law. The mission of The Sedona Conference is to move the law forward in a reasoned and just way.

The mission of WG11 is to identify and comment on trends in data security and privacy law, in an effort to help organizations prepare for and respond to data breaches, and to assist attorneys and judicial officers in resolving questions of legal liability and damages.

The Sedona Conference acknowledges Editor-in-Chief James Pizzirusso for his leadership and commitment to the project. We also thank Contributing Editors Mark Bailey, Stephen Chow, Ross Gotler, Amy Keller, Tim Murphy, Kaleigh Powell, and Jonathan Wilan for their efforts, and Al Saikali for his contributions as Steering Committee liaison to the project. We thank Andrew Lucking for his contributions.

In addition to the drafters, this nonpartisan, consensus-based publication represents the collective effort of other members of WG11 who reviewed, commented on, and proposed edits to early drafts of the *Commentary* that were circulated for feedback from the Working Group membership. Other members provided feedback at WG11 annual and midyear meetings where drafts of the *Commentary* were the subject of the dialogue. The publication was also subject to a period of public comment.

On behalf of The Sedona Conference, I thank both the membership and the public for all of their contributions to the *Commentary*.

We encourage your active engagement in the dialogue. Membership in The Sedona Conference Working Group Series is open to all. The Series includes WG11 and several other Working Groups in the areas of electronic document management and discovery, cross-border discovery and data protection laws, international data transfers, patent litigation, patent remedies and damages, and trade secrets. The Sedona Conference hopes and anticipates that the output of its Working Groups will evolve into authoritative statements of law, both as it is and as it should be. Information on membership and a description of current Working Group activities is available at <https://thesedonaconference.org/wgs>.

Craig Weinlein
Executive Director
The Sedona Conference
July 2021

TABLE OF CONTENTS

I.	INTRODUCTION.....	495
II.	BACKGROUND ON DATA PRIVACY LAWS.....	497
	A. Ambiguity in Data Privacy Laws.....	497
	1. Classifying Violations.....	497
	2. Quantifying Violations.....	498
	B. Clear Context in Privacy Laws.....	502
	1. Telephone Consumer Protection Act of 1991	502
	2. State Data Security Breach Notification Laws.....	504
III.	POSSIBLE METHODOLOGIES FOR CALCULATING VIOLATIONS.....	506
	A. Option One: Calculation of Violations Based Singularly on Defendant’s Failure to Comply, Regardless of Number of Impacted Consumers or Parts of the Law Violated	506
	B. Option Two: Calculation of Violations Based on the Number of Parts of the Statute Violated	509
	C. Option Three: Calculation of Violations Based on the Number of Consumers Impacted	510
	D. Option Four: Calculation of Violations Based on the Number of Pieces of Personal Information Impacted By Failure to Comply	515
	E. Option Five: Calculation of Violations Based on the Number of Days Violation Occurred.....	516
	F. Due Process Concerns and Their Role in a “Per Violation” Analysis	517
IV.	ARTICULATING A STANDARD FOR THE MEANING OF PER VIOLATION	522

A. Scenario One: California Consumer Privacy Act.....	525
B. Scenario Two: Colorado Security Breach Notification Law.....	527
C. Scenario Three: Illinois Biometric Information Privacy Act.....	528
V. CONCLUSION.....	531

I. INTRODUCTION

Some privacy laws in the United States allow for enforcement authorities and plaintiffs in private actions to seek damages or statutory penalties based on certain violations. Many of these laws, however, do not clearly define how a “violation” should be calculated. This can lead to confusion at best—and due process concerns at worst—when authorities and courts seek to quantify damages or penalties. After an incident that leads to a violation of a U.S. data privacy law that may impact a significant number of victims, should calculations be assessed based on one violation of the law, or is there some other way to measure incidents or violations? For example, should the calculation be based on adding up the total number of consumers affected by the business’s conduct, the number of statutory sections the business violated, the number of days the violations in a particular incident occurred, or some combination thereof?

As data privacy receives more attention in the United States and elsewhere—and as new laws in the U.S. take shape and are enacted—The Sedona Conference Working Group 11 (WG11) recognizes that a consistent approach to quantifying violations under U.S. privacy laws could be helpful to impacted parties, courts, authorities, and practitioners, not to mention the general public. With the various jurisdictions and enforcement authorities involved in current and future enforcement of such data privacy laws, however, such consistency can be challenging to reach. WG11 hopes, however, that this *Commentary* will be of use to stakeholders in reaching a fair interpretation of the meaning of a “per violation” measure of damages.

The first section of this *Commentary* reviews at a high level the landscape of existing privacy laws in the United States, addresses certain ambiguities regarding the calculation of penalties and damages that may arise under such laws, and examines

the way in which other somewhat analogous statutes have been enforced across the country. The second section examines possible ways in which violations of privacy laws could be quantified given statutory construction and existing case law. Finally, the last section endeavors to provide a useful test courts can use to evaluate the meaning of a “per violation” measure of damages in the context of data privacy violations in a way that benefits consumers and provides deterrent value to regulators but is fair and provides due process to potential violators.

II. BACKGROUND ON DATA PRIVACY LAWS

The United States has no overarching and preemptive national “privacy law” or “data security law” in place. As a result, different states have passed different laws—some of which provide for significant statutory penalties or damages when the laws at issue are violated. Given this patchwork approach to privacy and security, there is no singular interpretation as to what constitutes a “violation” of any given law. Consumers and regulators often approach these issues on an ad hoc basis through lawsuits in the court system, leaving organizations with little guidance.

A. *Ambiguity in Data Privacy Laws*

Various U.S. privacy laws permit damages for each “violation” of the law. These statutes, and judicial interpretations thereof, present discrepancies and ambiguities in how to classify and quantify a “violation” upon a failure to comply with a statute, in whole or in part.

1. Classifying Violations

As explained in further detail below, some statutes are explicit in how they are “violated”—for example, by failing to comply with a particular provision of the act.¹ But where statutes are not explicit, how to classify a “violation” becomes a matter of statutory interpretation. Does “violation” mean failure to comply with the title itself, as opposed to some particular provision? Does it mean the number of consumers impacted, or the number of pieces of personal information that are

1. See, e.g., Migrant and Seasonal Agricultural Workers Protection Act, 29 U.S.C. §§ 1801–72 (1983); District of Columbia Consumer Protection Procedures Act, D.C. CODE §§ 28-3901–13 (1975).

implicated? Is there a “violation” for every day that a defendant fails to comply with the statute? The answers may have significant damages implications for potential plaintiffs and due process implications for potential defendants.

2. Quantifying Violations

Further complicating the analysis is how to quantify a violation even where it can be classified. For example, the California Consumer Protection Law (CCPA), which went into effect on January 1, 2020, provides that (1) “[a] person that violates this title shall be . . . liable for a civil penalty of not more than two thousand five hundred dollars (\$2,500) for *each violation* or seven thousand five hundred dollars (\$7,500) for *each intentional violation*” in a civil action brought by the California attorney general.² Under the private right of action provided by the Illinois Biometric Information Protection Act (BIPA), “[a] prevailing party may recover for *each violation* [among other remedies] (1) against a private entity that negligently violates a provision of this Act, liquidated damages of \$1,000 [or] (2) against a private entity that intentionally or recklessly violates a provision of this Act, liquidated damages of \$5,000”³

These statutes present ambiguities in how one measures a “violation,” especially where there are different types of “violations” covered by the prescribed (statutory) damages—in

2. CAL. CIVIL CODE § 1798.155(b) (West 2020) (emphasis added). A separate section of the CCPA provides for statutory damages “in an amount not less than one hundred dollars (\$100) and not greater than seven hundred and fifty (\$750) *per consumer per incident*” for certain security breaches (emphasis added). CAL. CIVIL CODE § 1798.150(1)(a) (West 2020). The CCPA provides more guidance for what constitutes a violation under the private right of action, which will be addressed in later parts of this paper.

3. 740 ILL. COMP. STAT. ANN. 14/20 (West 2020).

contrast to the more particularized “incident” of a breach of security for which the CCPA allows a limited private right of action.⁴ There is a question under the CCPA or the BIPA whether

4. CAL. CIVIL CODE § 1798.150(1)(a) (West 2020) (emphasis added). The next subsection calls these “statutory damages” subject to mandatory consideration of factors:

“In assessing the amount of statutory damages, *the court shall consider* any one or more of the relevant circumstances presented by any of the parties to the case, including, but not limited to, the nature and seriousness of the misconduct, *the number of violations*, the persistence of the misconduct, the length of time over which the misconduct occurred, the willfulness of the defendant’s misconduct, and the defendant’s assets, liabilities, and net worth.” *Id.* § 1798.150(1)(b) (emphasis added).

The reference to “number of violations” to be considered in the prescribed remedy for an “incident” suggests reference to prior “violations.” Compare the Uniform Law Commission’s provision in one privacy law directed to unauthorized internet distribution of “intimate images”:

“[S]tatutory damages not to exceed \$[10,000] against each defendant found liable under this [act] for all disclosures and threatened disclosures by the defendant of which the plaintiff knew or reasonably should have known when filing the action or which became known during the pendency of the action. In determining the amount of statutory damages under subsection (a)(1)(B), *consideration must be given to* the age of the parties at the time of the disclosure or threatened disclosure, *the number of disclosures or threatened disclosures made by the defendant*, the breadth of distribution of the image by the defendant, and other exacerbating or mitigating factors[.]” UNIF. CIVIL REMEDIES FOR UNAUTHORIZED DISCLOSURE OF INTIMATE IMAGES ACT § 6(a)(1)(B) (UNIF. LAW COMM’N 2018) (emphasis added) (subsection 6(a)(3) allows for punitive damages under other law of the state). The comments explain the structure:

“The statutory damages provision is unusual in that it

the measuring “violation” is an aggregate or general violation of a particular statutory provision, or a particular action (unauthorized collection of information, single or series of failure to comply with consumer requests, etc.).

Adding to the ambiguity is the blurring in the statutes of the traditional distinction between statutory and liquidated damages as compensatory versus punitive or exemplary damages.⁵

suggests a range of damages rather than a fixed amount, and is limited to *one statutory recovery for all disclosures by the defendant occurring within a certain time period*. This is due to the unique nature of the problem addressed by this act. *Technology makes it possible for the number of unauthorized disclosures of intimate images to range in the thousands, even millions*. This potential for vast proliferation makes it advisable to establish upper and lower boundaries. . . .” *Id.* § 6 cmt. (emphasis added).

In another privacy act, the Commission did not provide for statutory damages in private suits but allowed (optionally, according to the legislature) the attorney general to seek “a civil penalty of up to \$[1000] for *each violation, but not exceeding* \$[100,000] for all violations caused by the same event.” UNIF. EMPLOYEE & STUDENT ONLINE PRIVACY PROTECTION ACT § 5(a)(2) (UNIF. LAW COMM’N 2016) (emphasis added).

5. In international recognition of foreign money judgments, recognition of noncompensatory awards may be limited to the availability of such awards in the State of enforcement. *See, e.g.*, INTELLECTUAL PROPERTY: PRINCIPLES GOVERNING JURISDICTION, CHOICE OF LAW, AND JUDGMENTS IN TRANSNATIONAL DISPUTES § 411 (AM. LAW INST. 2007). “In the United States . . . statutory damages are awarded in lieu of actual damages and profits in copyright cases [and] the enforcement court should enforce the full amount of the damages.” *Id.* cmt. b. Relative to “liquidated damage,” traditionally contracted, “unless the rendering court specifically characterizes all or part of the liquidated damages as exceeding the amount necessary to compensate, these awards should be regarded as compensatory and fully enforceable.” *Id.* cmt. d. It is possible that certain violations of privacy rights in personal information may be compensated under a theory restitution for “use value” as recognized in “reasonable royalties” as statutory damages in

Thus the factors to be considered for some statutory damages awards (for example, as explained in note 4, *supra*) include consideration of defendant conduct relative to third parties, rather than strictly damage to the plaintiff, including unjust enrichment.

Issues also arise in aggregate (class action) litigation:

Statutes sometimes entitle persons to sue for liquidated or minimum damages—also known as statutory damages—for technical violations of law that result in either no actual loss or an actual loss too small to warrant conventional litigation. . . .

[B]ecause conduct regulated by statutes with minimum-damages provisions often affects large populations, technical violations can foster lawsuits with enormous potential damage awards if aggregation is permitted. . . .

Difficulties arise when statutes providing for minimum damages make no reference to aggregate procedures. In cases brought under such silent statutes, judges have tried to mediate between the risk of under-deterrence, which a denial of aggregation might cause, and the risk of over-compensation and over-deterrence, which a decision allowing aggregation would encourage. . . .⁶

patent infringement. See RESTATEMENT (THIRD) OF RESTITUTION AND UNJUST ENRICHMENT § 42 (AM. LAW INST. 2011). Reasonable royalties awards are also statutorily available for trade secrets misappropriation. UNIF. TRADE SECRETS ACT § 3(a) (UNIF. LAW COMM'N 1985); 18 U.S.C. § 1836(b)(3)(B)(ii). Trade secrets misappropriation may be characterized as invasion of commercial information privacy.

6. PRINCIPLES OF THE LAW OF AGGREGATE LITIGATION § 1.03 cmt. e (AM.

The ambiguity on the measurement of “violation” affects appropriate aggregation.

B. Clear Context in Privacy Laws

Although rare, there are U.S. and state laws concerning privacy-type issues that provide clear guidance in quantifying the defendants’ exposure following a violation of the law. Nevertheless, in some cases, courts have reduced the statutorily mandated “per violation” damages on other grounds such as due process. The following are examples of statutes that explicitly define how “each violation” is calculated or totaled.

1. Telephone Consumer Protection Act of 1991

Enacted in 1991, the Telephone Consumer Protection Act of 1991 (TCPA) was a response by Congress to the reactions of American consumers over intrusive and unwanted phone calls to their homes.⁷ The TCPA contains a number of restrictions on the use of automated telephone equipment, including prohibiting the “initiat[ion of] any telephone call to any residential telephone line using an artificial or prerecorded voice to deliver a message without the prior express consent of the called party.”⁸ This subsection of the TCPA includes an express private right of action and statutory damages, permitting “an action to recover for actual monetary loss from such a violation, or to receive \$500 in damages for each such violation, whichever is greater.”⁹ If the court finds that the defendant willfully or

LAW INST. 2010).

7. 47 U.S.C. § 227 (1991).

8. *Id.* § 227(b)(1)(B).

9. *Id.* § 227(b)(3)(B). *See also* Section 227(c)(5) of the TCPA, which provides a private right of action on behalf of “[a] person who has received more than one telephone call within any 12-month period by or on behalf of the

knowingly violated, the court may, in its discretion, increase the amount of the award to \$1,500 per violation.¹⁰

The TCPA's plain text makes the defendant strictly liable for any violative calls and can lead to windfall verdicts in a class action. In *Wakefield v. ViSalus Inc.*,¹¹ for example, an Oregon federal jury returned a verdict finding ViSalus violated Section 227 of the TCPA by placing 1,850,440 calls using an artificial or pre-recorded voice without prior express consent of the class members.¹² Since the TCPA provides for statutory damages of \$500 per call, the verdict resulted in a total monetary award of more than \$925 million.¹³

Though the statutory damages per violation under Section 227 of the TCPA are clear, district courts have reduced the statutory mandated award on other grounds.¹⁴ One of those

same entity in violation of the regulations prescribed under this subsection . . . an action to recover for actual monetary loss from such a violation, or to receive *up to \$500* in damages for each such violation . . ." *Id.* § 227(c)(5)(B) (emphasis added).

10. *Id.* § 227(b)(3)(C).

11. No. 15-cv-1857, 2019 WL 2578082, at *1 (D. Or. June 24, 2019) (denying plaintiffs' claim for additional trebled damages).

12. *Id.*

13. *Id.*

14. See *Texas v. American Blastfax, Inc.*, 164 F. Supp. 2d 892, 900–01 (W.D. Tex. 2001) (finding it would be inequitable and unreasonable to award \$500 for each violation); *Maryland v. Universal Elections, Inc.*, 862 F. Supp. 2d 457, 465 (D. Md. 2012) (holding the penalty was disproportionate to the size of the company and the defendants' presumptive ability to pay); *United States v. Dish Network LLC*, 256 F. Supp. 3d 810, 906 (C.D. Ill. 2017) (awarding civil penalties and statutory damages of \$280,000,000—approximately 20 percent of the defendant's after-tax profits for 2016—finding this amount was "appropriate and constitutionally proportionate, reasonable, and consistent with due process").

grounds—due process—is discussed in further detail below.

2. State Data Security Breach Notification Laws

As of March 28, 2018, all 50 states had enacted breach notification laws requiring notification to individuals where there is an unauthorized access or acquisition of the individual's personally identifiable information.¹⁵ While most breach notification statutes do not make clear what "per violation" means, some articulate the overall liability in the enforcement section of the notification statute.

Unlike the CCPA, for example, Florida's breach notification statute crystalizes that civil penalties apply per breach and not per individual affected by the breach.¹⁶ Specifically, an entity that violates the provisions regarding notification of affected individuals or notification to the Florida Department of Legal Affairs is liable for a civil penalty of \$1,000 per day up to 30 days following any violation and \$50,000 per 30-day period thereafter, up to a maximum total of \$500,000.¹⁷ Virginia's Personal Information Breach Notification Statute also caps the civil penalty that the Virginia Attorney General can recover at \$150,000 per breach of the security of the system or a series of breaches of a similar nature that are discovered in a single investigation.¹⁸

Other statutes make apparent that the civil penalty is calculated on a per-resident basis. The District of Columbia's notification statute allows for the Attorney General to recover a modest civil penalty not to exceed \$100 for each resident who was

15. See Daniel J. Marcus, *The Data Breach Dilemma: Proactive Solutions for Protecting Consumers' Personal Information*, 68 DUKE L.J. 555, 577 (2018).

16. FLA. STAT. ANN. § 501.171(9)(b)(2) (West 2020).

17. FLA. STAT. ANN. § 501.171(9)(b)(1)-(2) (West 2020).

18. VA. CODE ANN. § 18.2-186.6(I) (West 2020).

not provided notice.¹⁹ In South Carolina, a person who is found to have knowingly and willfully violated the state's notification statute is subject to an administrative fine of \$1,000 per South Carolina resident affected by a breach.²⁰

19. D.C. CODE § 28-3853(b) (West 2020).

20. S.C. CODE ANN. § 39-1-90(H) (2013).

III. POSSIBLE METHODOLOGIES FOR CALCULATING VIOLATIONS

There are several ways to calculate “violations” of the law—below are some of the most common methodologies for calculating violations.

A. Option One: Calculation of Violations Based Singularly on Defendant’s Failure to Comply, Regardless of Number of Impacted Consumers or Parts of the Law Violated

Under this approach—which has an attractive simplicity—“violation” requires only one finding: that the defendant failed to comply with the title, regardless of the number of impacted consumers, the length of the breach, the amount of data exposed, or the number of failures.

But this approach almost certainly undermines the purpose of the inclusion of a statutory damages provision at all. Most courts recognize that statutory damages can serve “both a compensatory and punitive purpose,” depending on the statutory structure.²¹ They can also incentivize private suits to vindicate

21. See *Los Angeles News Serv. v. Reuters Television Int’l, Ltd.*, 149 F.3d 987, 996 (9th Cir. 1998); see also *Bateman v. Am. Multi-Cinema, Inc.*, 623 F.3d 708, 718 (9th Cir. 2010) (“We further note that Congress provided for punitive damages in addition to any actual or statutory damages, . . . which further suggests that the statutory damages provision has a compensatory, not punitive, purpose.”); *Schnall v. Amboy Nat’l Bank*, 279 F.3d 205, 216 (3d Cir. 2002) (“But the structure of § 4310, which permitted a plaintiff to recover both actual damages and statutory damages, suggests that this provision served the dual purpose of both compensating plaintiffs who have been misled and deterring banks [committing allegedly harmful conduct].”); *Dryden v. Lou Budke’s Arrow Fin. Co.*, 661 F.2d 1186, 1191 (8th Cir. 1981) (“[A]lthough we may disagree with Congress’s wisdom in providing for statutory damages in an instance such as this, we are bound to recognize the remedial purpose of the act.”); *Williams v. Pub. Fin. Corp.*, 598 F.2d 349, 356

the public interest.²²

If “violation” means only the failure to comply with the title, statutory damages are exceedingly (and likely inappropriately) limited. Whether statutory damages provisions are designed to deter or to compensate victims (or both),²³ such a limited interpretation undermines the statute’s likely purpose: to force a defendant to pay an amount that would deter wrongful conduct in the future or to compensate victims who might otherwise have trouble quantifying their damages.²⁴

The statutory language across provisions, moreover, may suggest that “violation” means something other than violation of the title only. The portion of the CCPA authorizing private causes of action, for example, contemplates plural “*violations of*

(5th Cir. 1979) (“The remedial scheme in the [Truth in Lending Act] is designed to deter generally illegalities which are only rarely uncovered and punished, and not just to compensate borrowers for their actual injuries in any particular case.”).

22. See *Perrone v. Gen. Motors Acceptance Corp.*, 232 F.3d 433, 436 (5th Cir. 2000) (“The caselaw confirms that statutory damages may be imposed as a means to encourage private attorneys general to police disclosure compliance even where no actual damages exist.”); *Schnall*, 279 F.3d at 217 (“... Congress may have deemed it more cost-effective to cede [Truth in Savings Act] enforcement to individuals in the private sector who stand to profit from efficiently detecting and prosecuting [Truth in Savings Act] violations.”).

23. For the California Consumer Privacy Act, CAL. CIV. CODE § 1798.155(b) (West 2020), the description of the statutory damages provision as a “penalty” in this context on the one hand suggests that the aim is deterrence. The collected penalty, however, is “deposited into the Consumer Privacy Fund . . . with the intent to fully offset any costs incurred by the state courts and the Attorney General in connection with this title,” § 1798.155(c), which suggests a remedial aim at least for the public at large.

24. See *Capital Records, Inc. v. Thomas-Rasset*, 692 F.3d 899, 908 (8th Cir. 2012) (“[S]tatutory damages are designed precisely for instances where actual harm is difficult or impossible to calculate.”).

this title” by a singular defendant.²⁵ Indeed, in “assessing the amount of statutory damages,” the court is to consider, among other factors, “the number of violations” a defendant made.²⁶ If “violation” means only a failure to comply with the title as a whole, it is difficult to see how a defendant could have engendered multiple violations.

Ironically, such a limited interpretation may also work against defendants trying to invoke federal court jurisdiction.²⁷ As discussed further below, courts—and many defendants—often assume that “violation” implies a per-person basis even in the data breach context. In *Attias v. CareFirst, Inc.*,²⁸ for example, when evaluating whether the plaintiffs had met the amount-in-controversy requirement under the Class Action Fairness Act in a data breach case, the court explained:

[P]laintiffs have brought claims under the District of Columbia Consumer Protection Procedures Act, D.C. Code Ann. § 28-3901 *et seq.*, which provides statutory damages of \$1,500 *per violation*, and the Virginia Consumer Protection Act (“VCPA”), Va. Code Ann. § 59.1-196 *et seq.*, which entitles successful plaintiffs up to \$ 500 to \$ 1,000 *per violation*. . . . Although plaintiffs do not provide a breakdown of the numbers in each

25. See CAL. CIV. CODE § 1798.150(b) (West 2020).

26. *Id.* § 1798.150(a)(2).

27. See, e.g., *Peatry v. Bimbo Bakeries USA, Inc.*, 393 F. Supp. 3d 766, 769 (N.D. Ill. 2019) (explaining in class action under the Illinois Biometric Information Privacy Act, “[f]or jurisdictional purposes, the parties’ positions are reversed, with [plaintiff] seeking to limit the potential damages and [defendant] arguing that the complaint provides the possibility of almost unlimited damages against it.”).

28. 365 F. Supp. 3d 1, 8 (D.D.C. Jan. 30, 2019) (appeal pending).

subclass, it's hard to imagine a distribution that would not satisfy the amount-in-controversy requirement based solely on these statutory claims.²⁹

B. Option Two: Calculation of Violations Based on the Number of Parts of the Statute Violated

“Violation” could also mean each failure to comply with a particular provision within the statute. Arguably, though, if that was the intent of the legislature in enacting a particular statute, it could have said so. Multiple statutes contemplating statutory damages provisions on a “per violation” basis describe when “per violation” means “per provision violated.” As noted above, under the Migrant and Seasonal Agricultural Workers Protection Act,³⁰ for example, statutory damages are available on a per plaintiff “per violation” basis, and the statute expressly contemplates that violations of the same provision constitute one violation—thereby implying that violating different provisions of the same title amounts to different violations.³¹

Likewise, the District of Columbia Consumer Protection Procedures Act provides that in an action by the Department of

29. See also *Edoff v. T-Mobile Ne. LLC*, 2019 WL 1459046, No. ELH-18-3777, at *2 (D. Md. Apr. 2, 2019) (explaining in a data breach case that in case involving “approximately 15,280 Maryland residents,” amount in controversy requirement met for Class Action Fairness Act where plaintiffs sought “statutory damages of \$1,000 ‘per first-time violation.’”).

30. U.S.C. § 1854(c)(1) (1983).

31. See *id.* (“[M]ultiple infractions of a single provision of this chapter or of regulations under this chapter shall constitute only one violation for purposes of determining the amount of statutory damages due a plaintiff.”); *Elizondo v. Podgorniak*, 100 F. Supp. 2d 459, 462 (E.D. Mich. 2000) (“It is also clear that violations of separate provisions of [the Migrant and Seasonal Agricultural Workers Protection Act] are evaluated separately.”).

Consumer and Regulatory Affairs, “[a]ny person found to have executed a trade practice in violation of a law of the District within the jurisdiction of the Department may be liable for a civil penalty not exceeding \$1,000 for *each failure to adhere to a provision* of an order described in subsection (f), (g), or (j) of this section, or a consent decree described in subsection (h) of this section.”³² Thus if “per violation” meant “per provision violated,” the legislature could have said so.

But some statutes may imply, by their language, that “violation” is in fact based on a “per provision” understanding. As noted above, the CCPA, for example, uses the term “violation” not just in the provision authorizing attorney general action, but also in the section authorizing a private right of action.³³ The statute’s use of “violation” in Section 1798.150 refers specifically to violations of provisions: before bringing an action for statutory damages, a consumer must provide 30 days’ written notice “identifying the specific provisions of this title the consumer alleges have been or are being violated.” Since courts are supposed to give words used across a statute the same meaning,³⁴ one could argue that “violation” for the purpose of the civil penalty provision similarly means each provision violated.

C. Option Three: Calculation of Violations Based on the Number of Consumers Impacted

Courts may also look to the number of consumers impacted by a defendant’s failure to comply with the statute as a separate “violation.” This approach is consistent with the provision for

32. D.C. CODE § 28-3905(i)(3)(A) (1975) (emphasis added).

33. See CAL. CIV. CODE § 1798.150 (West 2020).

34. See *Miranda v. Nat’l Emergency Servs., Inc.*, 35 Cal. App. 4th 894, 905 (Cal. Ct. App. 1995).

damages in some other consumer protection statutes, including the TCPA.³⁵ Even the CCPA's private right of action provision provides for a fine of not less than \$100 "per consumer per incident."³⁶ It may also provide some certainty when assessing damages for settlement purposes.³⁷

A recent Pennsylvania case is instructive for proponents of this approach: *Taha v. Bucks County Pennsylvania*.³⁸ In 2011, Bucks County launched an internet-accessible database of individuals who had been incarcerated in the county from 1938 onward—for a total of 66,799 people.³⁹ The plaintiff had been arrested and processed by the county but had been released the following day and had his arrest record expunged. He alleged that the database was a violation of Pennsylvania's Criminal History Record Information Act (PCHRIA),⁴⁰ which authorizes plaintiffs to bring suit for its violation.⁴¹

Notably, the PCHRIA's private right of action section

35. 47 U.S.C. § 227(b)(3)(B) (1991) (providing for the greater of actual damages or "\$500 in damages for each such violation").

36. CAL. CIV. CODE § 1798.150(a)(1) (West 2020).

37. Cf. Marcello Antonucci et al., *Post-Spokeo, Data Breach Defendants Can't Get Spooked*, FIRST QUARTER 2017 PLUS JOURNAL, available at https://www.wiley.law/media/publication/271_Post-Spokeo-Data-Breach-Defendants-Cant-Get-Spooked-They-Should-Stand-Up-to-the-Class-Action-Plaintiff-Boogeyman.pdf). Notably, some states specifically preclude this type of calculation. See, e.g., N.H. REV. STAT. § 358-A:4(III)(b) (1997) (providing for "civil penalties up to \$10,000 for each violation of this chapter" but providing that "the court shall determine the number of unlawful acts or practices which have occurred without regard to the number of persons affected thereby").

38. 367 F. Supp. 3d 320 (E.D. Pa. 2019).

39. *Id.*

40. 18 PA. CON. STATS. ANN. § 9101 *et seq.* (West 1979).

41. *Taha*, 367 F. Supp. 3d at 323.

provides that:

A person found by the court to have been aggrieved by a violation of this chapter or the rules or regulations promulgated under this chapter, shall be entitled to actual and real damages of not less than \$100 for each violation and to reasonable costs of litigation and attorney's fees. *Exemplary and punitive damages of not less than \$1,000 nor more than \$10,000 shall be imposed for any violation of this chapter, or the rules or regulations adopted under this chapter, found to be willful.*⁴²

After the district court certified a class action of individuals whose information was released in the database, the county appealed to the Third Circuit.⁴³ The county argued, in part, that the district court had improperly certified a punitive damages class under the statute because the named plaintiff had no actual damages.⁴⁴ The Third Circuit disagreed and—in reaching its conclusion—noted specifically that “the District Court has not made any decision regarding what conduct constitutes a violation or violations” for the purposes of the PCHRIA’s statutory damages provision.⁴⁵

The district court made its ruling on that score on remand.⁴⁶ Unsurprisingly, the county argued that “violation” under the statute meant only the dissemination of the database itself—“and therefore punitive damages must be capped at \$10,000.”⁴⁷

42. 18 PA. CON. STATS. ANN. § 9183(b)(2) (emphasis added) (West 1979).

43. *See Taha v. Cty. of Bucks*, 862 F.3d 292 (3d Cir. 2017).

44. *Id.* at 303.

45. *Id.* at 305.

46. *See Taha*, 367 F. Supp. 3d at 333–34.

47. *Id.* at 333.

The plaintiff, on the other hand, argued that “each release of criminal history record information—that is, the releases as to each of the 66,799 class members—constituted a violation of the statute.”⁴⁸

The district court agreed with the plaintiff. According to the district court, the defendants’ argument was based on a flawed assumption: “that a ‘violation’ is synonymous with an ‘act.’”⁴⁹ But violation, according to the court, is more appropriately considered with reference to the number of people whose rights have been violated. It provided this example: “If a tortfeasor breaks into a single computer, obtains private information relating to five different people, and publishes that information, the tortfeasor has violated five different peoples’ rights and could give rise to five different causes of action, despite only engaging in one act.”⁵⁰

It also distinguished between its case and *Tomasello v. Rubin*,⁵¹ which addressed a violation of the Privacy Act of 1974.⁵² In *Tomasello*, the defendant faxed one letter to 4,500 people.⁵³ The *Tomasello* plaintiff argued he was entitled to statutory damages for each letter.⁵⁴ The D.C. Circuit disagreed, finding that—consistent with the concept that waivers of sovereign immunity, as the Privacy Act in this case was, should be narrowly construed—the sending of the letter was the failure for which the

48. *Id.*

49. *Id.*

50. *Id.*

51. 167 F.3d 612 (D.C. Cir. 1999).

52. 5 U.S.C. § 552a (1974).

53. *Tomasello*, 167 F.3d at 616.

54. *Id.* at 617.

defendant was liable.⁵⁵ According to the *Taha* court, however, the appropriate analysis under the PCHRIA was the inverse: “Plaintiff’s claim does not turn on the number of people to whom private information was impermissibly sent, but rather on the number of class members whose information was published.”⁵⁶ Ultimately, the jury awarded—and the district court upheld—a statutory damages award of \$1,000 per class member, totaling over \$60 million in damages.⁵⁷

This approach may raise due process concerns, however. In *Taha*, the district court ruled that due process did not apply because the defendant was a governmental entity.⁵⁸ But for private defendants, the calculus is likely different.⁵⁹ Indeed, the potential for large damages awards may also make courts reluctant to certify classes.⁶⁰ The role due process plays in selecting among

55. *Id.* at 617–18.

56. *Taha v. Bucks Cty, Pa.*, 367 F. Supp. 3d 320, 334.

57. *Taha v. Bucks Cty. Pa.*, 408 F. Supp. 3d 628, 646–47.

58. *Id.* at 648–49 (“[T]he Due Process Clause protects persons, not governmental entities such as Bucks County.”).

59. *See, e.g., Golan v. FreeEats.com, Inc.*, 930 F.3d 950 (8th Cir 2019) (holding that \$1.6 billion in statutory damages for an “innocent” violation violated the Due Process Clause); J. Gregory Sidak, *Does the Telephone Consumer Protection Act Violate Due Process as Applied?*, 68 FLA. L. REV. 1403 (2016) (calculating that a TCPA violation causes only approximately \$.70 of harm per violation); *Maryland v. Universal Elections, Inc.*, 862 F. Supp. 2d 457 (D. Md. 2011); *see also Larson v. Harman-Mgmt. Corp.*, No. 1:16-cv-00219-DAD-SKO, 2019 U.S. Dist. LEXIS 219294 (E.D. Cal. Dec. 18 2019) (approving settlement agreement of TCPA, in part, because “likelihood that an award of damages in the billions would be deemed unconstitutional”).

60. *See, e.g., Parker v. Time Warner Entm’t Co., L.P.*, 331 F.3d 13 (2d Cir. 2002) (acknowledging due process concerns for large damages awards in class cases and noting that “[i]t may be that the aggregation in a class action of large numbers of statutory damages distorts the purpose of both statutory damages and class actions.”).

the possible methodologies for quantifying violations is set forth below.

D. Option Four: Calculation of Violations Based on the Number of Pieces of Personal Information Impacted By Failure to Comply

A fourth approach would be to treat each piece of personal information, for each consumer, affected by a violation of a statute as a “violation” under the civil penalties cap. In a way, a version of this approach has been adopted in state unfair competition laws insofar as those laws sometimes focus on the pieces of information disseminated for false advertising purposes.⁶¹

This approach has intuitive appeal. If each piece of personal information is treated as a discrete “thing,” and conduct that results in a violation as to a single piece of personal information is a violation, then it makes sense that conduct that results in violations as to ten pieces of personal information would be treated as ten violations. But defining a “piece” of personal information

61. See, e.g., *Commonwealth v. Fall River Motor Sales, Inc.*, 565 N.E.2d 1205, 1213 (Mass. 1991) (holding that each advertisement disseminated constituted violation of a consent judgment despite the fact that the advertisements were identical and paid for in a single transaction); *In re Miss. Medicaid Pharm. Average Wholesale Price Litig.*, 190 So. 3d 829, 847 (Miss. 2015) (upholding statutory damages based on the number of falsely reported average wholesale prices of medications); *State ex rel. Wilson v. Ortho-McNeil-Janssen Pharm., Inc.*, 777 S.E.2d 176, 204 (S.C. 2015) (reducing per violation damages but upholding application of uniform civil penalty to each “sample box” defendant distributed in violation of state unfair trade practices act); *State v. Ralph Williams’ N. W. Chrysler Plymouth, Inc.*, 553 P.2d 436 & n.12 (1976) (upholding a per-misrepresentation civil penalty and noting that “[a] single advertisement may include a number of misrepresentations . . . [e]ach of these acts is a separate violation”); *State v. Going Places Travel Corp.*, 864 N.W.2d 885, 898 (2015) (violations calculated by multiplying the number of misrepresentations by the number of consumers).

may be difficult. For example, would a record of a visit to web page that is tied to an IP address be a “piece”? Would the IP address and the address of the web page be “pieces”?

And as with the consumer/statutory section approach, this approach may tend to result in damages calculations that may violate the Due Process Clause. Depending on the definition of “piece” of personal information, amounts calculated under this method may easily be ten, one hundred, or one thousand times amounts calculated under the consumer/statutory section approach, which creates enormous theoretical exposures that may result in overdeterrence or an inefficient overspend on compliance.

E. Option Five: Calculation of Violations Based on the Number of Days Violation Occurred

Finally, “violation” might mean that each day a statutory violation continues after a demand to cease is treated as a separate “violation” for civil penalties purposes. At least one consumer protection statute explicitly provides for this sort of calculation—though with a limit. The Cable Privacy Act⁶² permits the court to award actual damages, though those damages cannot be less than the statutory damages of \$100 for each day of violation or \$1,000, whichever is greater.⁶³

62. 47 U.S.C. § 551(f)(2)(A) (1984).

63. See also 18 U.S.C. § 2520(c) (2018) (providing for “statutory damages of whichever is the greater of \$100 a day for each day of violation or \$10,000” under the federal Wiretap Act); 33 U.S.C. § 1319(d) (2019) (providing for civil penalty “not to exceed \$25,000 per day for each violation” of the Clean Water Act); WASH. REV. CODE ANN. § 42.56.550(5) (West 2017) (permitting trial court to award “an amount not to exceed one hundred dollars for each day that he or she was denied the right to inspect or copy” public records under the public records act).

In theory, unlike the approaches outlined above, this approach creates incentive for violators to cure quickly, because exposure increases linearly with time—which, in turn, brings exposure in line with the public interest in cessation of violations.

But adding a violation for each day the defendant fails to cure would likely further increase civil damages, especially if it is coupled with other “high exposure” methods like per-consumer or per-piece of information. That would likely further deepen due process concerns and a hesitancy to certify class actions.

F. Due Process Concerns and Their Role in a “Per Violation” Analysis

Due process concerns are present in any evaluation of the methodologies set forth above. In many ways statutory damages seem comparable to punitive damages—which are often challenged on due process grounds—especially insofar as both may be disconnected from compensatory damages. As to punitive damages, in the seminal cases of *State Farm Mutual Auto Insurance Co. v. Campbell* and *BMW of North America, Inc. v. Gore*, the Supreme Court instructed courts to consider various factors in determining whether an award of punitive damages comports with due process: “(1) the degree of reprehensibility of the defendant’s misconduct; (2) the disparity between the actual or potential harm suffered by the plaintiff and the punitive damages award; and (3) the difference between the punitive damages awarded by the jury and the civil penalties authorized or imposed in comparable cases.”⁶⁴

64. *State Farm Mut. Auto Ins. Co. v. Campbell*, 538 U.S. 408, 418 (2003); see also *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559 (1996).

But courts have also held that the “guideposts” the Supreme Court imposed on punitive damages in the *Campbell* and *Gore* do not apply to statutory damage awards.⁶⁵ According to these courts, due process prohibits excessive punitive damages awards because the defendant lacks fair notice of the severity of the penalty it may face for its conduct.⁶⁶

These courts instead follow the Supreme Court’s ruling in *St. Louis, I.M. & Southern Railway Co. v. Williams*:⁶⁷ A statutory damages award violates due process only when the award is “so severe and oppressive as to be wholly disproportioned to the offense and obviously unreasonable.”⁶⁸ The standard is “extraordinarily deferential—even more so than in cases applying

65. See, e.g., *Capitol Records, Inc. v. Thomas-Rasset*, 692 F.3d 899, 907–08 (8th Cir. 2012); see also *Sony BMG Music Entm’t v. Tenenbaum*, 719 F.3d 67, 70–71 (1st Cir. 2013) (“[W]e conclude, as have other courts, that the standard articulated in *Williams* governs the review of an award of statutory damages under the Copyright Act.”); *Zomba Enters., Inc. v. Panorama Records, Inc.*, 491 F.3d 574, 587 (6th Cir. 2007) (“We know of no case invalidating [an award of statutory damages] under *Gore* or *Campbell*, although we note that some courts have suggested in dicta that these precedents may apply to statutory-damage awards.”).

66. See *Capitol Records*, 692 F.3d at 907; *Sony BMG*, 719 F.3d at 70 (“The concerns regarding fair notice to the parties of the range of possible punitive damage awards, which underpin *Gore*, are simply not present in a statutory damages case where the statute itself provides notice of the scope of the potential award.”).

67. 251 U.S. 63 (1919).

68. *Zomba*, 491 F.3d at 587 (quoting *Williams*, 251 U.S. at 66–67); see also *Golan v. FreeEats.com, Inc.*, 930 F.3d 950, 962 (8th Cir. 2019) (applying *Williams* standard to TCPA claim and upholding finding that \$1.6 billion statutory damages award violated due process); *Perez-Farias v. Global Horizons, Inc.*, 499 F. App’x 735, 737 (9th Cir. 2012) (applying *Williams* standard to claim under the Washington Farm Labor Contractors Act and finding statutory damages did not violate due process).

abuse of discretion review.”⁶⁹ Thus, statutory damages may be even more disconnected from compensatory damages than punitive damages.⁷⁰ And when deciding whether the statutory damages award fails to comport with due process, some courts look to the award as a whole—not the awards for individual “violations.” That is, “[t]he absolute amount of the award, not just the amount per violation, is relevant to whether the award” violates due process under the reasoning in *Williams*.⁷¹

One district court, for example, slashed a TCPA-mandated statutory damages award of \$1.6 billion to \$32 million.⁷² In a post-trial motion for reduction of excessive damages, the defendant argued that the statutory damages of \$500 per call for 3,242,493 calls—totaling \$1,621,246,500—was so excessive it violated the Due Process Clause of the Fifth Amendment.⁷³ The district court agreed, calling the required damage award “obviously unreasonable and wholly disproportionate to the offense” and awarded the plaintiffs the amount of \$10 per call.⁷⁴

On appeal, the class members argued that the statutory damages of \$500 per call do not violate the Due Process Clause and should not have been reduced.⁷⁵ Although the circuit court agreed with the class members that nothing in the relevant provision of the TCPA allows for the reduction of statutory

69. *Zomba*, 491 F.3d at 587.

70. *See, e.g., id.* at 588 (upholding a 44:1 ratio of statutory to compensatory damages); *Williams*, 251 U.S. at 67 (upholding what amounted to a 113:1 ratio of statutory to compensatory damages).

71. *See Capitol Records*, 692 F.3d at 910.

72. *Golan v. Veritas Entm’t, LLC*, No. 4:14CV00069 ERW, 2017 WL 3923162, at *1 (E.D. Mo. Sept. 7, 2017).

73. *Id.*

74. *Id.* at *4.

75. *Golan v. FreeEats.com, Inc.*, 930 F.3d 950, 962 n.11 (8th Cir. 2019).

damages, it held that the district court did not err in concluding the statutory damages of \$1.6 billion violated the Due Process Clause.⁷⁶ It concluded: “[u]nder [the] facts [of this case], \$1.6 billion is ‘so severe and oppressive as to be wholly disproportioned to the offense and obviously unreasonable.’”⁷⁷

Not all courts follow this reasoning, however. The district court in *Wakefield v. ViSalus, Inc.*,⁷⁸ for example, explicitly rejected this line of thinking. In *Wakefield*, as explained above, a jury found that the defendant ViSalus had violated the TCPA 1,850,440 times, for a total damages award of \$925,220,000. The defendant challenged the award as excessive and thus unconstitutional under the standard in *Williams*.⁷⁹ The district court—while noting that the Ninth Circuit had not decided the issue—concluded that due process does not require reducing aggregate statutory damages.⁸⁰ Because *Williams* analyzed only “the penalty for a *single* statutory violation,” according to the district court, it implies that “the Supreme Court construed ‘penalty’ to mean the fine for a single statutory violation, not for the aggregate amount of damages.”⁸¹ And because the TCPA’s \$500 per violation statutory damages was not so unreasonable or oppressive as to violate due process, it was constitutional—all that was left was the “arithmetic” of multiplying the number of violations by the minimum statutory penalty for each violation.⁸²

76. *Id.* at 963.

77. *Id.* (quoting *St. Louis, I.M. & S. Ry. Co. v. Williams*, 251 U.S. 63, 67 (1919)).

78. No. 3:15-cv-1857-SI, 2020 WL 4728878 (D. Or. Aug. 14, 2020).

79. *Id.* at *2.

80. *Id.* at *3.

81. *Id.* (emphasis original).

82. *Id.* at *4.

Anything else, according to the district court, would be at odds with *Williams* “and would effectively immunize illegal conduct if a defendant’s bad acts crossed a certain threshold.”⁸³ Quoting the Seventh Circuit, the *Wakefield* court concluded: “Someone whose maximum penalty reaches the mesosphere only because the number of violations reaches the stratosphere can’t complain about the consequences of its own extensive misconduct.”⁸⁴

83. *Id.*

84. *Id.* (quoting *United States v. Dish Network, LLC*, 954 F.3d 970, 979–80 (7th Cir. 2020)).

IV. ARTICULATING A STANDARD FOR THE MEANING OF PER VIOLATION

The statutory landscape and applicable case law suggest that there is no one-size-fits-all answer to how violations should be quantified. Rather, the calculation of violations depends on the language and purpose of the statute and the nature of the conduct. As a result, it is quite possible that the same exact language could be subject to different interpretations as used in different laws, jurisdictions, or fact patterns.

As an initial matter, courts faced with a statutory damages or penalties provision will apply familiar principles of statutory interpretation, which are not addressed extensively in this paper. These will generally include looking initially at the plain meaning of the statute, and if that does not provide the answer, applying additional tools such as legislative history, a comparison to other language in the statute, and legislative intent.⁸⁵ Notably, in the context of damages provisions, which, depending on the mathematical calculation, could quickly lead to results in the billions of dollars, courts will seek to avoid interpreting the statute in a way that leads to absurd results or in a way that is inconsistent with due process.⁸⁶ Courts will also look to determine the legislative intent, which in the case of privacy damages and penalties provisions may include both deterrence and compensation elements.⁸⁷

85. See *United States v. LKAV*, 712 F.3d 436 (9th Cir. 2013).

86. See *Sloan v. Soul Circus, Inc.*, No.: 15-01389 (RC), 2015 WL 9272838, at *8 n.8 (D.D.C. 2015) (noting in the context of a remand petition that “[i]n statutory interpretation it is a given that statutes must be construed reasonably so as to avoid absurdities’ The Court cannot adopt the Circus’s damages theory when such absurd consequences might follow.” (quoting *In re Nofziger*, 925 F.2d 428, 434 (D.C. Cir. 1991) (*per curiam*))).

87. See, e.g., *Cabell v. Markham*, 148 F.2d 737, 739 (2d Cir. 1945), *aff’d*, 326

As noted above, some statutes provide courts with more specific direction on how to assess the number of violations and calculate a penalty or civil damage award.⁸⁸ However, for those statutes that simply authorize a penalty or damages award “per violation,” the case law, taking California as an example, suggests that the determination of the number of violations may depend on the circumstances of the case.⁸⁹ In *People v. Witzerman*, for example, the court upheld the trial court’s decision to assess penalties for false advertising “roughly on a per victim rather than per culpable statement made basis.”⁹⁰ The court held that “[w]hat constitutes a single violation . . . depends on the type of violation involved, the number of victims and the repetition of the conduct constituting the violation—in brief, the circumstances of the case.”⁹¹

In subsequent cases, the California Court of Appeals has continued in this vein, deferring to the trial court’s application of the facts in determining the number of violations.⁹²

While varying circumstances will lead to different results,

U.S. 404 (1945) (“[R]emember that statutes always have some purpose or object to accomplish, whose sympathetic and imaginative discovery is the surest guide to their meaning.”).

88. See Section II.B, *supra*.

89. See *People v. Witzerman*, 29 Cal. App. 3d 169, 181 and n.8 (Cal. Ct. App. 1972).

90. *Id.* at 180.

91. *Id.* at 171.

92. See *People v. Overstock.com, Inc.*, 219 Cal.Rptr.3d 65, 85, (Cal. Ct. App. 2017), as modified (June 23, 2017) (noting that the trial court considered determining the number of violations “by the number of Californians who saw the offending advertisements, by the number of sales made through the offending pages, and by the number of days Overstock violated the statutes,” and affirming the trial court’s decision to calculate penalties on a per-day basis).

the number of violations should be calculated with reference to the specific facts the plaintiff proves in connection with the alleged statutory violations. This rule is illustrated in *State v. Ralph Williams' North West Chrysler Plymouth, Inc.*,⁹³ in which the Washington Attorney General alleged that a car dealership violated the Washington Consumer Protection Act by making ten different categories of misrepresentations to prospective car buyers.⁹⁴ The court concluded that each misrepresentation could constitute a separate violation, so long as “[e]ach cause of action required [respondent] to prove divergent facts to establish a violation.”⁹⁵

Thus, in evaluating the meaning of “per violation” measure of damages where the statute provides no further guidance, the following test can be articulated:

In the absence of clear statutory language or legislative history to the contrary, each violation is considered a separate and distinct violation when divergent facts are required to establish such a violation.

This analysis will be backstopped by the due process limitations discussed in detail in the previous section. In particular, in the first instance, courts will look to avoid interpretations of the statute that will lead to significant constitutional concerns “where the text fairly admits of a less problematic construction.”⁹⁶ In determining an appropriate level of damages, courts may also look to common law principles that have evolved in particular in the area of consumer protection laws to provide

93. 553 P.2d 423 (1976).

94. *Id.* at 430–31, 436.

95. *Id.* at 436.

96. *Pub. Citizen v. U.S. Dep’t of Justice*, 491 U.S. 440, 455 (1989).

additional factors that they (or jurors) may apply. These factors include the good or bad faith of the defendant, the injury to the public, the defendant's ability to pay, and the desire to eliminate the benefits derived from the legal violations.⁹⁷

Below are three scenarios that illustrate how the number of violations can be determined by looking to the specific, divergent facts the plaintiff has proved.

A. Scenario One: California Consumer Privacy Act

The CCPA grants consumers the right to direct organizations not to sell their personal information.⁹⁸ "A business that has received direction from a consumer not to sell the consumer's personal information . . . shall be prohibited, pursuant to paragraph (4) of subdivision (a) of Section 1798.135, from selling the consumer's personal information after its receipt of the consumer's direction."⁹⁹ As stated before, under California Civil Code Section 1798.155(b), "[a]ny business, service provider, or other person that violates this title shall be . . . liable for a civil penalty of not more than two thousand five hundred dollars (\$2,500) for each violation or seven thousand five hundred dollars (\$7,500) for each intentional violation."

Assume a company called The Data Guys collects and sells personal information of California consumers. The California Attorney General brings an action and proves the following facts: 500 consumers sent an opt-out notice to The Data Guys. After receiving these notices, The Data Guys sold personal information of 250 of the consumers to Company A. The Data

97. See *State ex rel. Woodard v. May Dep't Stores Co.*, 849 P.2d 802, 810 (Colo. App. 1992).

98. See CAL. CIV. CODE § 1798.120(a) (West 2020).

99. *Id.* § 1798.120(d).

Guys then sold the personal information of all 500 consumers to Company B. Later, the Data Guys sold the personal information of 150 of the consumers to Company C.

The Data Guys may argue that there can only be three violations, one for each of its sales to Companies A, B, or C. Or The Data Guys could argue that there were only 500 violations—one per each consumer. However, the California Attorney General has arguably proved 900 violations (one violation per customer per illegal sale—250 plus 500 plus 150). This approach appears sensible. There is no constitutional concern with multiple punishments for the same conduct. And the number of violations is tied to specific acts that must be proved with individualized evidence, each of which causes a distinct harm to the privacy interest of the affected consumers.

What if The Data Guys sold multiple pieces of personal information relating to each consumer? The CCPA's definition of "personal information" is quite broad and includes, for example, "[i]nternet or other electronic network activity information, including, but not limited to, browsing history, search history, and information regarding a consumer's interaction with an internet website, application, or advertisement."¹⁰⁰ With a little creativity, the Attorney General might be able to identify and prove hundreds or even thousands of "divergent facts"—i.e., distinct pieces of personal information sold by The Data Guys—potentially adding an exponential multiplier to the number of violations. In this instance, the trial court would retain discretion to determine the number of violations in a manner that is reasonable given the circumstances.¹⁰¹

100. *Id.* § 1798.140(o).

101. *See* *People v. Overstock.com, Inc.*, 219 Cal.Rptr.3d 65, 85-86 (Cal. Ct. App. 2017) (affirming the trial court's use of per-day methodology for determining the number of violations where other approaches "would result in

B. Scenario Two: Colorado Security Breach Notification Law

Colorado law requires organizations that maintain, own, or license personal information about Colorado residents to provide notice to the affected residents when a security breach results or could result in the misuse of their personal information.¹⁰² Penalties may be applied for each violation of Colorado Revised Statutes Section 6-1-716.¹⁰³

Assume a company called The Open Network is hacked by cybercriminals. After gaining access to The Open Network's computer system, the attackers obtain the credentials of an employee and begin emailing The Open Network's unencrypted files to the attacker's account. When all is said and done, the hackers have stolen the names and social security numbers of 10,000 Colorado residents. The Open Network does not provide notice until a whistleblower threatens to inform the Colorado Attorney General. At this point, The Open Network provides notice, but 180 days have passed since the time that The Open Network should have provided notice under C.R.S. § 6-1-716. The Attorney General subsequently brings suit for failure to provide timely notice.

The Attorney General proves the following facts: Ten thousand Colorado residents had their information stolen, and The Open Network didn't provide notice to any of them. The Open Network's failure to provide timely notice lasted 180 days.

excessive penalties of at least hundreds of millions of dollars"); *People v. Witzerman*, 29 Cal. App. 3d 169, 180 (Cal. Ct. App. 1972) (finding no fault in the trial court's failure to "exhibit[] mathematical exactitude" and affirming the court's decision to apply a penalty for only a subset of the violations the court found).

102. COLO. REV. STAT. ANN. § 6-1-716 (West 2018).

103. See COLO. REV. STAT. ANN. § 6-1-112 (West 2019).

The Open Network might argue that the security breach was an isolated incident and that its failure to provide timely notice was therefore just a single violation. However, because the statute requires notice to all affected consumers, C.R.S. § 6-1-716(2), and the Attorney General proved a failure to provide notice to each one of them of them, the court could find 10,000 violations—one for each Colorado resident who did not receive the required notice.

The Attorney General might argue that there was a violation for each day that each consumer did not receive the required notice. According to the Attorney General, the number of violations would be 1,800,000 (10,000 times 180). However, the facts the Attorney General has proved are that notice was given 180 days late to 10,000 residents of State X. There are no “divergent facts” that establish a separate violation for each of the 10,000 residents for each day. Moreover, courts may be reluctant to read a “per day” component into the provision when it is entirely absent from the language of the statute, whereas other statutes explicitly incorporate a “per day” element.¹⁰⁴ The better result is a finding of 10,000 violations, one for each consumer the Attorney General proved was entitled to notice and did not receive it.

C. *Scenario Three: Illinois Biometric Information Privacy Act*

The Illinois Biometric Information Privacy Act (BIPA) imposes requirements on businesses that collect or possess biometric information (for example, retina or iris scans, fingerprints, or scans of hand or face geometry).¹⁰⁵ One requirement is

104. See, e.g., FLA. STAT. ANN. § 501.171(9) (West 2019) (authorizing civil penalties of \$1,000 per day for the first 30 days and \$50,000 per day for days 31 to 180).

105. 740 ILL. COMP. STAT. ANN. § 14/1, *et seq.* (West 2008).

that a covered business may not “collect, capture, . . . or otherwise obtain a person’s or a customer’s biometric identifier or biometric information” without first obtaining informed consent from the person.¹⁰⁶ As noted above, among other remedies, BIPA provides for liquidated damages of \$1,000 “for each violation” or \$5,000 for “intentional violations.”¹⁰⁷

Assume that a hotel chain decides to replace keys with iris scans for controlling entry to its hotel rooms. At check-in, guests are asked to provide an iris scan after showing their ID. Under this new system, the guest’s room door opens automatically when the guest approaches the door. Each time a guest enters her room, the hotel collects and retains the scan to improve its scanning technology. While guests are informed of the new procedure at check-in, the hotel fails to obtain the guests’ informed, written consent as required by BIPA. Applying the statutory language of BIPA to this conduct, the court would be justified in finding a violation for each time that a guest had his or her iris scanned.

In the end, external consensus around how to calculate “per violation” damages is challenging, as the answer can lead to outsized results one way or the other. Ideally, legislatures would do a better job of answering this question explicitly in the first instance. If the intent is to punish conduct on a “per incident” basis, on a daily basis, or on a per consumer basis, this would be easy enough to incorporate into the language of the statute itself, and there are multiple examples of where legislatures have done the hard work to incorporate more concrete and explicit language along these lines in any number of different contexts. In the absence of such concrete language, courts will

106. *Id.* § 14/15(b).

107. *Id.* § 14/20.

be left to interpret the language that is there, and as discussed above, will do so using the tools that they always use. The law suggests that courts have a certain degree of flexibility in undertaking this analysis, and rigid calculations, especially those that lead to absurd or even unconstitutional results, will not prevail. Rather, courts will likely consider the legislative and remedial intent and look to avoid extreme outcomes on either side of the range of potential answers to the question.

V. CONCLUSION

Although the country's statutory framework for privacy litigation provides some uncertainty concerning how violations of certain statutes should be quantified, existing case law provides guidance when the statutes are ambiguous. Although greater certainty in the construction of privacy statutes can better alleviate uncertainty, application of the above flexible analysis can provide clarity for violators, certainty for regulators, and protection for consumers.

FROM THE KKK TO GEORGE FLOYD: THREE JUDGES EXPLORE QUALIFIED IMMUNITY

*Hon. Cathy Bissoon, Hon. Benita Y. Pearson & Hon. David A. Sanders**

INTRODUCTION

This article's purpose is to place the often-used defense of qualified immunity in a historical context and highlight the primary opposition to its continued utilization, focusing on its role in litigation of excessive force claims. This article declines itself to render any specific judgment on the qualified immunity defense, as the authors here, all members of federal district courts, do not view that as the proper role of the judiciary, particularly the lower courts. The authors endeavor to explain the history, practical implications, and judicial and scholarly criticism of the doctrine in an accessible manner.

* Judge Cathy Bissoon is a federal district court judge for the United States District Court for the Western District of Pennsylvania. Judge Benita Y. Pearson is a federal district court judge for the United States District Court for the Northern District of Ohio. Judge David A. Sanders is a federal magistrate judge for the United States District Court for the Northern District of Mississippi. The authors acknowledge and appreciate the assistance of Catherine Dowie, particularly, in the finalization of this article.

I. HISTORY AND BACKGROUND

Following the Civil War and with Reconstruction under way, it became clear to Congress that the Southern states could not be trusted with caring for their citizens in a fair and equal manner.¹ In early 1866, Senator Lyman Trumbull of Illinois introduced the bill that would become the first Civil Rights Act.² This original bill was drafted essentially to make clear that “all persons born in the United States . . . are hereby declared citizens of the United States.”³ The act seemed to follow logically from the conclusion of the Civil War and the enactment of the Thirteenth Amendment the year before.⁴ Nevertheless, President Andrew Johnson vetoed the bill twice.⁵ Johnson’s reasoning was both racist and political; he believed the act favored Blacks over Whites, and that the act would set off a move toward centralization of the federal government.⁶ Nevertheless, on April 5, 1866, the Senate overrode Johnson’s veto and on April 9, the House of Representatives did the same.⁷

During this period, Congress was focused on how best to handle the recently defeated Southern states. While President Abraham Lincoln’s 10 percent plan was seen as a moderate one, upsetting many radical Republicans in Congress at the time,

1. *Jones v. Alfred H. Mayer Co.*, 392 U.S. 409, 426–27 (1968) (collecting authorities).

2. *Id.* at 431.

3. *Id.* at 422.

4. *Id.* at 437–39.

5. *Id.* at 435.

6. *Mahone v. Waddle*, 564 F.2d 1018, 1028–29 (3d Cir. 1977); Rogers M. Smith, *Legitimizing Reconstruction: The Limits of Legalism*, 108 YALE L.J. 2039, 2071 (1999); see also Randall Kennedy, *Persuasion and Distrust: A Comment on the Affirmative Action Debate*, 99 HARV. L. REV. 1327, 1342 n.51 (1986).

7. *Jones*, 392 U.S. at 435.

Johnson's plan became far more accommodating to the South.⁸ Indeed, Johnson was "willing to accept the South back into the Union so long as the Southern states recognized merely that the formal institution of slavery was a thing of the past. He was willing to leave the treatment of the freedmen in the hands of the southern whites."⁹ Doing so, however, led several states, including South Carolina and Mississippi, to draft "Black Codes."¹⁰ These state statutes were drafted to make certain freedmen did not enjoy the same rights and privileges held by Whites.¹¹ Illustrative examples of sections provided in Black Codes included

8. Keith E. Whittington, *Bill Clinton Was No Andrew Johnson: Comparing Two Impeachments*, 2 U. PA. J. CONST. L. 422, 427–28 (2000). Specifically, "[o]n the question of readmitting the seceded states to the Union, Lincoln clashed with Congress over his 'Ten Percent Plan' and the Radical Republicans' Wade-Davis bill." Hon. Henry S. Cohn, *Book Review*, THE FEDERAL LAWYER, Sept. 2014, at 87 (reviewing JOHN C. RODRIGUE, LINCOLN AND RECONSTRUCTION (2013)). "The dispute centered on the percentage of male citizens in a state that had seceded that would have to sign a loyalty oath before the state could rejoin the Union." *Id.* "Lincoln's plan required only 10 percent, whereas the Wade-Davis bill would have required a majority of voters to take an oath and included other requirements that no Confederate state (except perhaps Tennessee) could have met." *Id.* "Lincoln pocket-vetoed the Wade-Davis bill." *Id.*

9. Sanford V. Levinson, *New Perspectives on the Reconstruction Court*, 26 Stan. L. Rev. 461, 467 (January 1974) (reviewing Charles Fairman, *The Oliver Wendell Holmes Devise History of the Supreme Court of the United States Volume VI: Reconstruction and Reunion: 1864–1888, Part I* (1971) and citing L. Cox & J. Cox, *Politics, Principle, and Prejudice, 1865–66: Dilemma of Reconstruction America* 151–71 (1963)). Professor Levinson also points out that the southern states had almost two years to show they would treat the freedmen in good faith and "failed overwhelmingly to do so." *Id.*

10. Katesha Long, *Debunking the Broken Windows Theory in Policing: An Incident and Badge of Slavery*, 4 HOW. HUM. & C.R. L. REV. 77, 89 (2020).

11. Hon. Bernice B. Donald and Pablo J. Davis, "To This Tribunal the Freedman has Turned": The Freedmen's Bureau's Judicial Powers and the Origins of the Fourteenth Amendment, 79 LA. L. REV. 1, 21 (Fall 2018).

North Carolina's provision requiring Blacks to have a White person as a witness when they contracted, or Mississippi's apprenticeship provision allowing "former owners" to have young Blacks as apprentices, and if the apprentices should "escape," the "former owners" were allowed to recapture them and bring them before a justice of the peace.¹²

With passage of the Civil Rights Act of 1866, Congress took active steps toward eliminating Black Codes and began the long road to protecting the civil rights of all Americans. The Act extended a federal guarantee of the basic rights to own and convey property and to use the civil courts to vindicate property rights.¹³ To be sure, acceptance came slowly. During this time, virtually all the states bristled at overarching federal oversight. Even many Republicans, while accepting "the enhancement of national power resulting from the Civil War . . . did not believe the legitimate rights of the states had been destroyed, or the traditional principles of federalism eradicated."¹⁴ Not surprisingly, however, the Southern states put up the strongest resistance, and at times, that resistance was violent. Not long after the surrender of the Confederate Army at the Appomattox Courthouse and in reaction to Reconstruction plans being put into place by Congress, Southerners founded the Ku Klux Klan in Pulaski,

12. Aremona G. Bennett, *Phantom Freedom: Official Acceptance of Violence to Personal Security and Subversion of Proprietary Rights and Ambitions Following Emancipation, 1865–1910*, 70 *CHI.-KENT L. REV.* 439, 455 (1994).

13. James R. Stoner, Jr., *From Magna Carta to the Montgomery March: Common Law and Civil Rights*, 6 *FAULKNER L. REV.* 49, 54–55 (Fall 2014).

14. ERIC FONER, *RECONSTRUCTION: AMERICA'S UNFINISHED REVOLUTION, 1863–1877*, 242 (2014). Indeed, as Professor Foner notes, "[i]nstead of envisioning continuous federal intervention in local affairs, [the Civil Rights Bill] honored the traditional presumption that the primary responsibility for law enforcement lay with the states, while creating a latent federal presence, to be triggered by discriminatory state laws." *Id.* at 245.

Tennessee, in 1865. The Klan was just one example of Southern Whites pushing back against many aspects of Reconstruction, chief among them rights being given to freedman living throughout the South. Shortly after the Civil War, Congress began receiving reports of widespread violence against freed slaves, and these attacks continued despite passage of the Civil Rights Act.¹⁵ As time passed, it became evident that Congress needed something with teeth to enforce the provisions of the Civil Rights Act and the newly enacted Fourteenth Amendment. Members of the Klan and others were making it extremely difficult for freedmen to vote or afraid to even attempt it. As a result, in 1870 and 1871, Congress passed what came to be known as the "Enforcement Acts." In all, there were three Enforcement Acts, but the third Act provided what would later become Section 1983 of the Civil Rights Act. Specifically, that Act provided:

That any person who, under color of any law, statute, ordinance, regulation, custom, or usage of any State, shall subject, or cause to be subjected any person within the jurisdiction of the United States to the deprivation of any rights, privileges, or immunities secured by the Constitution of the United States, shall, any such law, statute, ordinance, regulation, custom, or usage of the State to the contrary notwithstanding, be liable to the party injured in any action at law, suit in equity, or other proper proceeding for redress; such proceeding to be prosecuted in the several district of circuit courts of the United States with and subject to the same rights of appeal, review upon error,

15. John Montoya, *Defying Congressional Intent: Justices Miller and Bradley Alter the Course of Reconstruction*, 10 *COLUM. J. RACE AND L.* 82, 83 (2020).

and other remedies provided in like cases in such courts, under the provisions of the act of the ninth of April, eighteen hundred and sixty-six, entitled "An act to protect all persons in the United States in their civil rights, and to furnish the means of their vindication"; and other remedial laws of the United States which are in their nature applicable in such cases.

This third Act, known as the Ku Klux Klan or KKK Act, succeeded to an extent, undermining the organized violence of the Klan. However, the Supreme Court in *United States v. Reese*¹⁶ and *United States v. Cruikshank*¹⁷ greatly weakened the Act, holding that voting rights were better handled by the states without federal intervention. Following those decisions, the Civil Rights Act, and more specifically Section 1983, was practically ignored. It was not until almost a century later in *Monroe v. Pape*¹⁸ that litigation against government officials and agencies began to increase.¹⁹

In *Monroe v. Pape*, thirteen Chicago police officers broke into Pape's home in the early morning without a warrant. The officers got him out of bed and made him stand naked in his living room while they searched every room, emptying drawers and ripping mattress covers. They then took Pape to the station and held him for ten hours without letting him contact anyone while they interrogated him about a murder. Pape was finally released with no criminal charges filed, and he pursued an action under Section 1983, suing the officers and the city for their

16. 92 U.S. 214 (1875).

17. 92 U.S. 542 (1875).

18. 365 U.S. 167 (1978).

19. Michael K. Cantwell, *Constitutional Torts and the Due Process Clause*, 4 TEMP. POL. & CIV. RTS. L. REV. 317, 317-18 (Spring 1995).

actions. After examining the history surrounding the KKK Act, Justice William O. Douglas wrote for the Supreme Court that the “under color of law” language in the statute was intended to allow civil rights suits in cases where officials acted in a manner unauthorized by state law. This familiar holding has been seen as the case that “revitalize[ed] the Civil Rights Act of 1871.”²⁰ Prior to *Monroe*, there had been very few cases filed under Section One of the Civil Rights Act—the precursor to Section 1983. The United States Code Annotated notes only nineteen decisions under the Section in its first sixty-five years.²¹ As of 2011, the courts saw an average of 40,000 to 50,000 per year.²² With that growth in claims filed came, of course, defenses to those claims. One of the first defenses to arise was that of qualified immunity, which first appeared before the Supreme Court in 1967.

In *Pierson v. Ray*, a group of fifteen Black and White clergymen attempted to use facilities in a Jackson, Mississippi bus terminal marked “White Waiting Room Only.”²³ Jackson police arrested the clergymen and charged them with violating a state statute, which made it unlawful for anyone to congregate “with others in a public place under circumstances such that a breach of the peace may be occasioned thereby”²⁴ After being vindicated in the misdemeanor proceedings, the clergymen

20. James E. Robertson, *Fatal Custody: A Reassessment of Section 1983 Liability for Custodial Suicide*, 24 U. TOL. L. REV. 807, 810 (Summer 1993).

21. *Limiting the Section 1983 Action in the Wake of Monroe v. Pape*, 82 HARV. L. REV. 1486, 1486 n.4 (1969).

22. MARTIN A. SCHWARTZ, SECTION 1983 LITIGATION: CLAIMS AND DEFENSES, S1.01[B], at 1–5 (4th ed. supp. 2011-1). The 2020 supplement to Schwartz’s book, *id.* (4th ed. supp. 2020-2), asserts that the same range persists, citing data compiled in 2014.

23. 386 U.S. 547, 552 (1967).

24. *Id.* at 549.

brought a civil rights action against the police officers under Section 1983 and under common law that the officers were liable for false arrest and imprisonment. Following trial in the Southern District of Mississippi, a jury found for the plaintiffs, and the Court of Appeals for the Fifth Circuit affirmed as to the Section 1983 action, holding the Mississippi statute had been held unconstitutional in *Thomas v. Mississippi*.²⁵ While *Thomas* had been decided subsequent to the arrests at issue, the court felt compelled to affirm by the Supreme Court's ruling in *Monroe v. Pape*. As to the common law claims, however, the Fifth Circuit reversed, holding that Mississippi law did not require police officers to predict at their peril whether a Mississippi statute would subsequently be held unconstitutional.

On appeal, the Supreme Court addressed the Section 1983 claims and the Fifth Circuit's interpretation of *Monroe v. Pape* and explained that it in "no way intimated that the defense of good faith and probable cause was foreclosed by statute."²⁶ The Court went on to hold "that the defense of good faith and probable cause, which the court of appeals found available to the officers in the common law action for false arrest and imprisonment, is also available to them in the action under section 1983."²⁷ The Court continued, "that a police officer is not charged with predicting the future course of constitutional law."²⁸

Following *Pierson*, the Supreme Court set out to provide a clear, workable explanation of this qualified immunity it had

25. 380 U.S. 524 (1965).

26. *Pierson*, 386 U.S. at 556.

27. *Id.* at 557.

28. *Id.*

created. While the first attempts proved largely unhelpful,²⁹ the Court in *Wood v. Strickland* laid out a relatively clear explanation that included both objective and subjective factors.³⁰ Specifically, the Court held that qualified immunity would not be available to a party who knew or reasonably should have known that the action he took would violate someone's constitutional rights, or if he took action with the malicious intention to cause a deprivation of constitutional rights or other injury. It soon became apparent, however, that the test prescribed in *Wood* was incapable of addressing the concerns inherent in the new doctrine, namely to avoid "insubstantial lawsuits."³¹ Indeed, dismissal of "insubstantial lawsuits was at the heart of the Court's next decision affecting qualified immunity.³²

In *Harlow v. Fitzgerald*, the plaintiff argued that White House aides to former President Richard M. Nixon participated in a conspiracy to violate his constitutional and statutory rights.³³ The issue before the Court was the scope of immunity afforded to senior aides and advisors to the President of the United States. After a lengthy explanation as to why absolute immunity would not apply, the Court found that qualified immunity was

29. See, e.g., *Scheuer v. Rhodes*, 416 U.S. 232 (1974); see also, Alan K. Chen, *The Ultimate Standard: Qualified Immunity in the Age of Constitutional Balancing Tests*, 81 IOWA L. REV. 261, 288 n.160 (1995) ("The Court set forth vague parameters without explaining how courts should apply them.").

30. See *Wood v. Strickland*, 420 U.S. 308, 322 (1975).

31. *Butz v. Economou*, 438 U.S. 478, 507–08 (1978) (relying on an assumption that the *Wood* standard would permit insubstantial lawsuits to be quickly terminated).

32. *Harlow v. Fitzgerald*, 457 U.S. 800 (1982).

33. *Id.* at 802. It should be noted that *Harlow* was an implied constitutional cause of action—not a Section 1983 action; however, the Court extended its holding to 1983 actions because "it would be untenable to draw a distinction for purposes of immunity law." See *Baxter v. Bracey*, 140 S.Ct. 1862, 1863 (2020) (Thomas, J., dissenting) (quoting *Harlow*, 457 U.S. at 818 n.30).

the “best attainable accommodation”³⁴ The petitioners argued that should absolute immunity not be available, then a change needed to be made in the standard being applied for qualified immunity at the time. The Court described their argument as “persuasive” and explained that “dismissal of insubstantial lawsuits without trial—a factor presupposed in the balance of competing interests struck by our prior cases—requires an adjustment of the “good faith” standard established by our decisions.”³⁵ The Court then looked closely at the test articulated in *Wood* and found it was the subjective component applied that was causing the problem—that is, allowing insubstantial claims to proceed to trial. Specifically, following *Wood*, it became apparent that lower courts were finding an official’s subjective good faith to be a question of fact, thus defeating dispositive motions.³⁶ Consequently, the Court did away with the subjective component of the analysis and held that qualified immunity would be available to officials performing discretionary functions when their “conduct does not violate clearly established statutory or constitutional rights of which a reasonable person would have known.”³⁷

With that background, this article will look more closely at the law developed following *Harlow*, specifically with respect to cases brought alleging excessive force by police officers. While the cases continue to look to *Harlow* and its “clearly established rights” framework, the Supreme Court has addressed these cases, adding a bit more nuance and at times what appears to be a more demanding standard. At first blush, it appears to be a fairly straightforward exercise. A plaintiff filing a lawsuit under

34. *Id.* at 814.

35. *Id.* at 814–15.

36. *Id.* at 816.

37. *Id.* at 818.

Section 1983 for excessive force must show that the officer (1) violated his Fourth Amendment rights and (2) that the right was “clearly established.”

II. THE LAW OF SECTION 1983

To demonstrate that a Fourth Amendment violation has occurred, courts balance “the nature and quality of the intrusion on the individual’s Fourth Amendment interests’ against the countervailing governmental interests at stake.”³⁸ The Court continued that while a test for reasonableness under the Fourth Amendment was not capable of precise definition, “its proper application requires careful attention to the facts and circumstances of each particular case”³⁹

To determine whether that right is clearly established, it must be such that a reasonable official would understand that what he is doing violates that right[.]”⁴⁰ Furthermore, “existing precedent must have placed the statutory or constitutional question” confronted by the official “beyond debate.”⁴¹ Exactly what is meant by “beyond debate,” however, is less than clear.⁴²

38. *Graham v. Connor*, 490 U.S. 386, 396 (1989).

39. *Id.* The Court added factors that could be relevant to consider, including “the severity of the crime at issue, whether the suspect poses an immediate threat to the safety of the officers or others, and whether he is actively resisting arrest or attempting to evade arrest by flight.” *See also Tennessee v. Garner*, 471 U.S. 1, 8–9 (1985) (explaining courts consider totality of circumstances when deciding whether intrusion was reasonable).

40. *Anderson v. Creighton*, 483 U.S. 635, 640 (1987).

41. *See Plumhoff v. Rickard*, 572 U.S. 765, 779 (2014) (quoting *Ashcroft v. al-Kidd*, 563 U.S. 731, 741 (2011)).

42. Joanna C. Schwartz, *Qualified Immunity’s Boldest Lie*, 88 U. CHI. L. REV. 605, 613–14 (2021). Professor Schwartz points out that while the Supreme Court has held twice that a prior court opinion with similar facts is unnecessary to establish excessive force, all its other decisions have repeatedly required that plaintiffs identify court decisions to overcome a qualified immunity motion.

A. *Fundamentals of Excessive Force Claims Under 42 U.S.C. § 1983*

The Fourth Amendment⁴³ protects individuals “against unreasonable searches and seizures[.]”⁴⁴ The hopefully-now-familiar text of Section 1983 provides that:

Every person who, under color of any statute, ordinance, regulation, custom, or usage, of any State or Territory or the District of Columbia, subjects, or causes to be subjected, any citizen of the United States or other person within the jurisdiction thereof to the deprivation of any rights, privileges,

43. Excessive force typically arises in Fourth, Eighth, and Fourteenth Amendment claims. This article focuses on the Fourth Amendment standards but uses examples of Fourteenth and Eighth Amendment claims when discussing Qualified Immunity. Such claims are subject to distinct substantive standards:

We reject this notion that all excessive force claims brought under § 1983 are governed by a single generic standard. As we have said many times, § 1983 “is not itself a source of substantive rights,” but merely provides “a method for vindicating federal rights elsewhere conferred.” *Baker v. McCollan*, 443 U.S. 137, 144, n.3 (1979). In addressing an excessive force claim brought under § 1983, analysis begins by identifying the specific constitutional right allegedly infringed by the challenged application of force. See *id.*, at 140 (“The first inquiry in any § 1983 suit” is “to isolate the precise constitutional violation with which [the defendant] is charged”). In most instances, that will be either the Fourth Amendment’s prohibition against unreasonable seizures of the person, or the Eighth Amendment’s ban on cruel and unusual punishments, which are the two primary sources of constitutional protection against physically abusive governmental conduct.

Graham, 490 U.S. at 393–94 (footnote omitted).

44. U.S. CONST. amend. IV.

or immunities secured by the Constitution and laws, shall be liable to the party injured in an action at law, suit in equity, or other proper proceeding for redress[.]

As relevant to this article, *Bivens v. Six Unknown Named Agents of Federal Bureau of Narcotics*⁴⁵ provides a parallel remedy against federal officers for violations of the federal Constitution.⁴⁶

“[T]he right to make an arrest or investigatory stop necessarily carries with it the right to use some degree of physical coercion or threat thereof to effect it.”⁴⁷ The degree of physical coercion that law enforcement officers may use is not unlimited, however, and “*all* claims that law enforcement officers have used excessive force . . . in the course of an arrest, investigatory stop, or other ‘seizure’ of a free citizen should be analyzed under the Fourth Amendment and its ‘reasonableness’ standard”⁴⁸

The Supreme Court recently has reinforced that for a constitutional “seizure” to occur, an application of force must be effective—if a suspect evades the officer’s application of force in its entirety, the encounter is more properly classified as an attempted seizure, not necessarily subject to a Fourth

45. 403 U.S. 388 (1971).

46. The Eighth and Fourteenth Amendment rights of federal prisoners and pretrial detainees, respectfully, are complicated by the fact that *Bivens* remedies (specifically money damages) do not extend to suits against private prisons. *Corr. Servs. Corp. v. Malesko*, 534 U.S. 61 (2001). The nuances of that topic are beyond the scope of this article.

47. *Graham*, 490 U.S. at 396.

48. *Id.* at 395.

Amendment analysis. However, “brief seizures are seizures all the same[.]” and an individual may have a Fourth Amendment claim against officers even if that individual ultimately overcame the officer’s application of force and was not arrested during the initial encounter. Specifically, “[i]n addition to the requirement of intent to restrain, a seizure by force—absent submission—lasts only as long as the application of force.”⁴⁹

In defining the parameters of reasonableness, the Supreme Court has explained:

Determining whether the force used to effect a particular seizure is reasonable under the Fourth Amendment requires a careful balancing of the nature and quality of the intrusion on the individual’s Fourth Amendment interests against the countervailing governmental interests at stake Because the test of reasonableness under the Fourth Amendment is not capable of precise definition or mechanical application . . . its proper application requires careful attention to the facts and circumstances of each particular case, including the severity of the crime at issue, whether the suspect poses an immediate threat to the safety of the officers or others, and whether he is actively resisting arrest or attempting to evade arrest by flight.⁵⁰

The *Graham* Court continued:

The reasonableness of a particular use of force must be judged from the perspective of a

49. *Torres v. Madrid*, 141 S. Ct. 989, 999 (2021).

50. *Graham*, 490 U.S. at 396 (cleaned up); *see also* *Tennessee v. Garner*, 471 U.S. 1, 8–9 (1985) (“[T]he question [is] whether the totality of the circumstances justify[s] a particular sort of . . . seizure.”).

reasonable officer on the scene, rather than with the 20/20 vision of hindsight With respect to a claim of excessive force, the same standard of reasonableness at the moment applies: Not every push or shove, even if it may later seem unnecessary in the peace of a judge's chambers, violates the Fourth Amendment. The calculus of reasonableness must embody allowance for the fact that police officers are often forced to make split-second judgments—in circumstances that are tense, uncertain, and rapidly evolving—about the amount of force that is necessary in a particular situation.

As in other Fourth Amendment contexts, however, the reasonableness inquiry in an excessive force case is an objective one: the question is whether the officers' actions are objectively reasonable in light of the facts and circumstances confronting them, without regard to their underlying intent or motivation. An officer's evil intentions will not make a Fourth Amendment violation out of an objectively reasonable use of force; nor will an officer's good intentions make an objectively unreasonable use of force constitutional.⁵¹

B. Pre-Force Conduct by Law Enforcement

Broadly speaking, while "[t]he reasonableness of an officer's use of force must be judged by considering 'the totality of the circumstances,'"⁵² "several circuits have held that '[w]here a

51. *Graham*, 490 U.S. at 396–97 (citations and quotation marks omitted).

52. *Garner*, 471 U.S. at 8–9.

police officer unreasonably places himself in harm's way, his use of deadly force may be deemed excessive."⁵³

Excessive force claims can be complicated when a case involves concerning or even provocative and unconstitutional behavior by officers before the alleged excessive force at issue was applied. Courts have struggled and are divided on how to incorporate such pre-force behavior into their Fourth Amendment analysis.

While the Supreme Court in 2017 rejected a framework previously applied by the Ninth Circuit, it has not resolved the question of which alternate competing framework should be applied.⁵⁴ Three approaches are to evaluate the force (1) at the split second it was applied; (2) during a discrete period or "segment" of the encounter which may be longer than a split second, but less than the entire interaction and buildup thereto; and (3) under a totality of the circumstances analysis.

The Fourth, Fifth, and Eleventh Circuits evaluate an officer's use of force only at the instant it was applied, regardless of the preceding circumstances.⁵⁵ Some circuits, specifically the Sixth and Seventh, apply a segmenting approach, in which the reasonability of the officer's conduct is assessed "at each stage" or segment.⁵⁶ When applying this approach, the Sixth Circuit consider events in "close temporal proximity" and related to the identified violation. "[T]he court should first identify the 'seizure' at issue here and then examine 'whether the force used to

53. *Orn v. City of Tacoma*, 949 F.3d 1167, 1176 n.1 (9th Cir. 2020) (quoting *Kirby v. Duva*, 530 F.3d 475, 482 (6th Cir. 2008) and citing *Thomas v. Durasanti*, 607 F.3d 655, 667 (10th Cir. 2010); *Lytle v. Bexar County*, 560 F.3d 404, 413 (5th Cir. 2009); *Estate of Starks v. Enyart*, 5 F.3d 230, 234 (7th Cir. 1993)).

54. *Cty. of Los Angeles, Calif. v. Mendez*, 137 S. Ct. 1539, 1548 (2017).

55. *Anderson v. Russell*, 247 F.3d 125 (4th Cir. 2001).

56. *Dickerson v. McClellan*, 101 F.3d 1151, 1161 (6th Cir. 1996).

effect that seizure was reasonable in the totality of the circumstances, not whether it was reasonable for the police to create the circumstances.”⁵⁷ ‘Segmenting does not mean breaking down an incident into the smallest portion available,⁵⁸ but recognizing that there may be natural braking points between multiple actions.

The Seventh Circuit applies a similar approach in some cases, when such a division of the total interaction is reasonably justified by the circumstances. “[W]e carve up the incident into segments and judge each on its own terms to see if the officer was reasonable at each stage.”⁵⁹ “In some cases each discrete use of force must be separately justified. We think a sequential analysis is appropriate here[.]”⁶⁰

The First and Ninth Circuits look to the totality of the circumstances surrounding each claim before them.⁶¹ That is not to say that these circuits evaluate the entire interaction as a single

57. *Scozzari v. City of Clare*, 653 F. App’x 412, 419 (6th Cir. 2016) (quoting *Livermore ex rel. Rohm v. Lubelan*, 476 F.3d 397, 406 and quoting *Dickerson*, 101 F.3d at 1161).

58. For example, the Sixth Circuit has recently cautioned against using individual frames from footage of a rapidly evolving incident: The officer’s “perspective did not include leisurely stop-action viewing of the real-time situation that they encountered.” *Cunningham v. Shelby Cty., Tennessee*, No. 20-5375, 2021 WL 1526512, at *4 (6th Cir. Apr. 19, 2021). The value of footage is fact-specific. The Sixth Circuit has rejected factual findings made by a District Court upon reviewing video footage and concluding that an officer twice pepper-sprayed a prisoner who, according to the Sixth Circuit, “was not a threat.” *Anderson v. Sutton*, 717 F. App’x 548, 552 (6th Cir. 2017).

59. *Deering v. Reich*, 183 F.3d 645, 652 (7th Cir. 1999).

60. *Dockery v. Blackburn*, 911 F.3d 458, 467 (7th Cir. 2018) (citation omitted).

61. *Stamps v. Town of Framingham* (1st Cir. 2016); *S.R. Nehad v. Browder*, 929 F.3d 1125, 1132 (9th Cir. 2019); *Bryan v. MacPherson*, 630 F.3d 805, 823 (9th Cir. 2010).

claim. Each claim is analyzed individually, but these circuits do not limit or cabin consideration of pre-force conduct.

The Tenth Circuit also applies a totality of the circumstances approach. Its test has long expressly considered reckless or deliberate provocation by officers as a part of the totality of the circumstances to be analyzed: “The reasonableness of [officers’] actions depends both on whether the officers were in danger at the precise moment that they used force and on whether [their] own reckless or deliberate conduct during the seizure unreasonably created the need to use such force.”⁶²

C. Supreme Court Guidance on Pre-Force Conduct

The Ninth Circuit formerly utilized a “provocation rule,” which held that “an officer’s otherwise reasonable (and lawful) defensive use of force is unreasonable as a matter of law, if (1) the officer intentionally or recklessly provoked a violent response, and (2) that provocation is an independent constitutional violation.”⁶³ The Supreme Court rejected this framework: “We hold that the Fourth Amendment provides no basis for such a rule. A different Fourth Amendment violation cannot transform a later, reasonable use of force into an unreasonable seizure.”⁶⁴

The earlier Fourth Amendment violation in *Mendez* was a warrantless entry. Once they reached the Supreme Court, the *Mendez* plaintiffs did not attempt to defend the Ninth Circuit’s

62. *Bond v. City of Tahlequah, Oklahoma*, 981 F.3d 808, 816 (10th Cir. 2020) (quoting *Sevier v. City of Lawrence*, 60 F.3d 695, 699 (10th Cir. 1995) (footnote omitted)) (alterations by Bond Court).

63. This test is arguably distinct from that applied within the Tenth Circuit, which it has expressly maintained post-*Mendez*. *Pauly v. White*, 874 F.3d 1197, 1219 n.7 (10th Cir. 2017); *Cox v. Wilson*, 971 F.3d 1159, 1171 (10th Cir. 2020).

64. *Cty. of Los Angeles, Calif. v. Mendez*, 137 S. Ct. 1539, 1544 (2017).

provocation rule but attempted to defend the judgment below on a totality of the circumstances theory.⁶⁵ The Supreme Court declined to engage in such an analysis in the first instance.⁶⁶ The Court's ruling was thus a narrow one: "All we hold today is that *once* a use of force is deemed reasonable under *Graham*, it may not be found unreasonable by reference to some separate constitutional violation."⁶⁷

The Court expressly left open the argument that if an earlier constitutional violation *proximately caused* a plaintiff's damages, even if the application of force was reasonable under *Graham*, that a plaintiff may still recover for those damages in his or her claim for the initial violation, subject to standard defenses, including qualified immunity.⁶⁸ "[I]f the plaintiffs in this case cannot recover on their excessive force claim, that will not foreclose recovery for injuries proximately caused by *the warrantless entry*. The harm proximately caused by these two torts may overlap, but the two claims should not be confused."⁶⁹ How Courts handle such a proximate causation analysis would have significant impacts on a plaintiff's ability to recover for personal injuries but might create challenges under the frameworks used by some circuits. For example, the recognition of overlapping damages may not be entirely consistent with an approach predicated on segmenting constitutional claims.

D. Qualified Immunity

Not every Section 1983 case is decided on the merits. In addition to typical procedural safeguards, qualified immunity

65. *Mendez*, 137 S. Ct. at 1547 n.*.

66. *Id.*

67. *Id.*

68. *Id.* at 1548.

69. *Id.*

protects law enforcement from liability, and from litigation itself.⁷⁰ It is a strong protection, and officers are entitled to interlocutory appellate review if they are denied qualified immunity.⁷¹ As described above, the doctrine is judicially created, and based, in large part, on purely practical concerns and competing policy considerations. Indeed, the Supreme Court describes it as “as the best attainable accommodation of competing values[.]”⁷² “Qualified immunity shields an officer from suit when she makes a decision that, even if constitutionally deficient, reasonably misapprehends the law governing the circumstances she confronted.”⁷³

In the Fourth Amendment context, an officer is entitled to qualified immunity when “clearly established” precedent does not show that the search, seizure, or use of force violated the

70. *Saucier v. Katz*, 533 U.S. 194, 200–01 (2001).

In a suit against an officer for an alleged violation of a constitutional right, the requisites of a qualified immunity defense must be considered in proper sequence. Where the defendant seeks qualified immunity, a ruling on that issue should be made early in the proceedings so that the costs and expenses of trial are avoided where the defense is dispositive. Qualified immunity is “an entitlement not to stand trial or face the other burdens of litigation.” *Mitchell v. Forsyth*, 472 U.S. 511, 526 (1985). The privilege is “an immunity from suit rather than a mere defense to liability; and like an absolute immunity, it is effectively lost if a case is erroneously permitted to go to trial.” *Ibid.* As a result, “we repeatedly have stressed the importance of resolving immunity questions at the earliest possible stage in litigation.” *Hunter v. Bryant*, 502 U.S. 224, 227 (1991) (per curiam).

71. *Mitchell v. Forsyth*, 472 U.S. 511 (1985).

72. *Harlow v. Fitzgerald*, 457 U.S. 800, 814 (1982).

73. *Brosseau v. Haugen*, 543 U.S. 194, 198 (2004) (per curiam).

Fourth Amendment.⁷⁴ A court objectively evaluates the “reasonableness of the action, assessed in light of the legal rules that were clearly established at the time it was taken.”⁷⁵

“Clearly established” means that, at the time of the officer’s conduct, the law was “‘sufficiently clear’ that every ‘reasonable official would understand that what he is doing’” is unlawful. [*al-Kidd*, 563 U.S. at 741] (quoting *Anderson v. Creighton*, 483 U.S. 635, 640 (1987)). In other words, existing law must have placed the constitutionality of the officer’s conduct “beyond debate.” [*al-Kidd*, 563 U.S. at 741]. This demanding standard protects “all but the plainly incompetent or those who knowingly violate the law.” *Malley v. Briggs*, 475 U.S. 335, 341 (1986).

To be clearly established, a legal principle must have a sufficiently clear foundation in then-existing precedent. The rule must be “settled law,” *Hunter v. Bryant*, 502 U.S. 224, 228 (1991) (per curiam), which means it is dictated by “controlling authority” or “a robust ‘consensus of cases of persuasive authority,’” [*al-Kidd*, 563 U.S. at 741–742] (quoting *Wilson v. Layne*, 526 U.S. 603, 617 (1999)). It is not enough that the rule is suggested by then-existing precedent. The precedent must be clear enough that every reasonable official would interpret it to establish the particular rule the plaintiff seeks to apply. See [*Reichle v. Howards*, 566 U.S.

74. See *Pearson v. Callahan*, 555 U.S. 223, 243–44 (2009); *Anderson v. Creighton*, 483 U.S. 635, 641 (1987).

75. *Wilson v. Layne*, 526 U.S. 603, 614 (1999) (internal quotation marks omitted).

658, 666 (2012)]. Otherwise, the rule is not one that “every reasonable official” would know. *Id.*, at 664 (internal quotation marks omitted).⁷⁶

An aim of the doctrine is to “ensure that before they are subjected to suit, officers are on notice their conduct is unlawful.”⁷⁷ In conducting a qualified immunity analysis, courts operate under the assumption that officers are informed of relevant controlling precedent, as defined by the circuit in which they operate.⁷⁸ The inquiry into whether a right was clearly established is an objective one—an officer’s actual ignorance of controlling precedent is irrelevant to the Court’s analysis.⁷⁹

Between 2001 and 2009, federal courts were required to determine whether a constitutional violation had occurred, regardless of whether qualified immunity was granted in an

76. *District of Columbia v. Wesby*, 138 S. Ct. 577, 589–90 (2018).

77. *Hope v. Pelzer*, 536 U.S. 730, 739 (2002) (cleaned up).

78. Scholars have questioned whether this assumption was reasonable at its inception, or, more importantly, whether the Court’s enshrinement of this assumption has resulted in officers receiving training on relevant circuit precedent. Schwartz, *supra* note 42, at 629–30. A recent study, surveying the policies and practices of police departments throughout California, concluded that officers receive little, if any, training related to specific case law other than the general contours of *Graham* and *Garner*. *Id.*

79. Even under this objective analysis, at least one court of appeal granted qualified immunity where a right became clearly established two days before a subsequent constitutional violation occurred: “[I]t is beyond belief that within two days the government could determine . . . what new policy was required to conform to the ruling, much less communicate that new policy to the [relevant] officers.” *Bryan v. United States*, 913 F.3d 356, 363 (3d Cir. 2019) (“Within one or two days, neither [officer] could reasonably be expected to have learned of this development in our Fourth Amendment jurisprudence.”).

individual case.⁸⁰ Specifically, courts were directed to follow a two-step inquiry in a specific order:

First, a court must decide whether the facts that a plaintiff has alleged (*see* Fed. Rules Civ. Proc. 12(b)(6), (c)) or shown (*see* Rules 50, 56) make out a violation of a constitutional right. [*Saucier*, 533 U.S. at 201]. Second, if the plaintiff has satisfied this first step, the court must decide whether the right at issue was “clearly established” at the time of defendant’s alleged misconduct. *Ibid.*⁸¹

While this encouraged the clear establishment of substantive law, the requirement that courts always address whether or not a constitutional violation occurred was not without drawbacks. As the Supreme Court explained when it overturned *Saucier*, in *Pearson*, *Saucier*’s “rigid order of battle” compelled courts to devote substantial resources to “difficult questions that have no effect on the outcome of the case.”⁸² Alternately, when the merits question had little to nothing to do with the outcome of a case, the parties, or courts, could be inclined to address the issue in a cursory manner, meaning that judges had scant argument before them, or that future jurists were reviewing opinions with scant analysis in determining whether a principle had been clearly established.

Furthermore, “[r]igid adherence to the *Saucier* rule may make it hard for affected parties to obtain appellate review of constitutional decisions that may have a serious prospective effect on their operations.”⁸³ Where law enforcement is granted qualified immunity and is thus the prevailing party in a lower

80. *Saucier v. Katz*, 533 U.S. 194, 201 (2001).

81. *Pearson v. Callahan*, 555 U.S. 223, 232 (2009).

82. *Id.* at 236–37.

83. *Id.* at 240.

court, parties may lack an ability to appeal an adverse decision on the merits of the constitutional claim, further undermining the value of such decisions in the development of the law more broadly.⁸⁴

The Supreme Court recognized that full adherence to the *Saucier* two-step approach is “often, but not always, advantageous, [and] the judges of the district courts and the courts of appeals are in the best position to determine the order of decisionmaking that will best facilitate the fair and efficient disposition of each case.”⁸⁵ Since *Pearson*, lower courts have retained the discretion to not answer the merits of whether an act violates the constitution where it is granting qualified immunity. Courts have recognized that the *Pearson* approach presents its own challenges, particularly that it leaves important, and properly presented, aspects of constitutional law undeveloped, which has a dispositive impact on future cases. As one Judge has described this change: “No precedent = no clearly established law = no liability. An Escherian Stairwell. Heads government wins, tails plaintiff loses.”⁸⁶

Furthermore, “[o]n occasion, [some Courts of Appeal have] add[ed] a third prong to the *Saucier* test, examining ‘whether the plaintiff offered sufficient evidence to indicate that what the official allegedly did was objectively unreasonable in light of the clearly established constitutional rights.’”⁸⁷

84. *Id.*

85. *Id.* at 242.

86. *Zadeh v. Robinson*, 928 F.3d 457, 479–80 (5th Cir. 2019) (Willett, J, concurring in part, dissenting in part).

87. *Srisavath v. City of Brentwood*, 243 F. App’x 909, 912 (6th Cir. 2007) (quoting *Estate of Carter v. City of Detroit*, 408 F.3d 305, 311 n.2 (6th Cir.2005)); see also generally E. Lee Whitwell, *How Qualified Is Qualified Immunity: Adding A Third Prong to the Qualified Immunity Analysis*, 43 CAMPBELL L. REV. 403 (2021).

The Fourth Circuit has expressly adopted use of a third prong, and the First and Fifth Circuits engage in the same analysis, but on occasion treat the second and third prongs as independent subsets of a two-pronged analysis.⁸⁸

The Seventh Circuit has squarely rejected the use of a third prong.⁸⁹ The Second and Eighth Circuits, which employed a third prong for a time, have stepped back to two inquires in more recent cases.⁹⁰

In criticizing the three-pronged approach, then-Judge Sotomayor explained:

Our approach does not simply divide into two steps what the Supreme Court treats singly, asking first, whether the right is clearly established *as a general proposition*, and second, whether the application of the general right *to the facts of this case* is something a reasonable officer could be

88. *Gould v. Davis*, 165 F.3d 265, 273 (4th Cir. 1998); *Whalen v. Massachusetts Trial Court*, 397 F.3d 19, 27 n.9 (1st Cir. 2005) (“We note that, on occasion, we have combined the second and third prongs of the qualified immunity analysis into a single step.” (citations omitted)); *Hare v. City of Corinth, Miss.*, 135 F.3d 320, 326 (5th Cir. 1998) (“The second prong of the qualified immunity test is better understood as two separate inquiries[.]”).

89. *Estate of Escobedo v. Bender*, 600 F.3d 770, 779 n. 3 (7th Cir. 2010) (explaining why objective reasonableness of officers’ tactics in using force relates to first prong of qualified immunity analysis, not second).

90. *Bailey v. Pataki*, 708 F.3d 391, 404 n. 8 (2d Cir. 2013) (“There is some tension in our Circuit’s cases as to whether the qualified immunity standard is of two or three parts, and whether the “reasonable officer” inquiry is part of step two—the “clearly established” prong—or whether it is a separate, third step in the analysis”); *Feist v. Simonson*, 222 F.3d 455, 464 (8th Cir. 2000), overruled on other grounds by *Helseth v. Burch*, 258 F.3d 867 (8th Cir. 2001) (analyzing qualified immunity using three prongs); *Henderson v. Munn*, 439 F.3d 497, 501 (8th Cir. 2006) (“To determine whether an official is entitled to qualified immunity, we ask two questions[.]”).

expected to anticipate. Instead, we permit courts to decide that official conduct was “reasonable” even after finding that it violated clearly established law in the particularized sense. By introducing reasonableness as a separate step, we give defendants a second bite at the immunity apple, thereby thwarting a careful balance that the Supreme Court has struck “between the interests in vindication of citizens’ constitutional rights and in public officials’ effective performance of their duties.”⁹¹

Certain panels of the Sixth Circuit continue to employ a third prong, although other panels have criticized the practice.⁹² The Seventh Circuit has expressly rejected a three-pronged analysis.⁹³ The Ninth Circuit has created similar ambiguity.⁹⁴

The Supreme Court has not expressly addressed the use of a three-prong framework, although the potential circuit split has been raised by parties before it. Indeed, the issue was robustly briefed⁹⁵ by the parties in *Tolan v. Cotton*,⁹⁶ in which the Court

91. *Walczyk v. Rio*, 496 F.3d 139, 168–69 (2d Cir. 2007) (quoting *Anderson v. Creighton*, 483 U.S. 635, 639 (1987), in turn quoting *Davis v. Scherer*, 468 U.S. 183, 195 (1984)).

92. *Dunigan v. Noble*, 390 F.3d 486, 491 n.6 (6th Cir. 2004) (collecting disagreeing panels within the Sixth Circuit); *Robertson v. Lucas*, 753 F.3d 606, 615 (6th Cir. 2014) (acknowledging an ongoing in-circuit dispute over the precise contours of the analysis).

93. *Jones v. Wilhelm*, 425 F.3d 455, 460 (7th Cir. 2005).

94. *CarePartners, LLC v. Lashway*, 545 F.3d 867, 876 n. 6 (9th Cir. 2008) (“We have previously expressed the qualified immunity test as both a two-step test and a three-step test”).

95. The parties’ briefing is freely accessible at <https://www.scotusblog.com/case-files/cases/tolan-v-cotton/>.

96. 572 U.S. 650 (2014).

ultimately issued a per curiam opinion. While the opinion repeated that “[i]n resolving questions of qualified immunity at summary judgment, courts engage in a two-pronged inquiry[,]” the Court did not acknowledge lower-court disagreement regarding the use of a third prong, or splitting the second inquiry into two sub-inquiries.⁹⁷

E. But Which Courts Can Clearly Establish a Right?

“The Supreme Court has not expressly resolved the question of what authorities ‘count’ and how conflicting authorities should be evaluated when there is no binding Supreme Court precedent to ‘clearly establish’ the law.”⁹⁸ In *Elder v. Hollowal*, the Court reflected a permissive view, and instructed that a court should use its “full knowledge of its own [and other relevant] precedents.”⁹⁹

“[D]istrict court decisions—unlike those from the courts of appeals—do not *necessarily* settle constitutional standards or prevent repeated claims of qualified immunity[,]” and therefore “[m]any Courts of Appeals [] decline to consider district court precedent when determining if constitutional rights are clearly established for purposes of qualified immunity.”¹⁰⁰ Circuits take different approaches on whether out-of-circuit, unpublished or district court decisions can establish a right with sufficient clarity.¹⁰¹

97. *Id.* at 655.

98. Michael Avery, David Rudovsky, Karen M. Blum, & Jennifer Laurin, *Police Misconduct: Law and Litigation* § 3:9 (Thompson Reuters, 3rd ed. 2020) (footnotes omitted).

99. 510 U.S. 501, 516 (1994).

100. *Camreta v. Greene*, 563 U.S. 692, 709 n.7 (2011) (emphasis added).

101. David R. Cleveland, *Clear As Mud: How the Uncertain Precedential Status of Unpublished Opinions Muddles Qualified Immunity*

Concerns are compounded when a court decides a Section 1983 case on qualified immunity grounds, rather than merits, in circuits that take a narrow view of what prior opinions can clearly establish a right. A recent example brings those concerns, about the law remaining underdeveloped and not providing relief to potentially deserving plaintiffs, into sharp relief. In *Norris v. Hicks*,¹⁰² the Eleventh Circuit considered the execution of a search warrant on the wrong home. The officers arrived at the specific, and only, address they had a warrant to search, but, after throwing flash grenades into the building, determined that it was uninhabitable and abandoned. Because they understood the object of their search to be an inhabited residence, they assumed that the building they were at was not the location to be searched. Without engaging in any discussion of the issues, or seeking a warrant for an alternate location, they moved through the backyard of the first building, “and then forcibly entered a nearby yellow house whose lights were on[.]” A resident was apprehended in the house, and he brought suit against the officers.

The Court of Appeals for the Eleventh Circuit had previously affirmed a denial of qualified immunity where officers had a warrant to search a specific address, but entered another residence on the same block (173 Powerline Drive, rather than 133), despite the units being properly labeled with their respective numbers.¹⁰³ The Eleventh Circuit in *Norris* refused to

Determinations, 65 U. MIAMI L. REV. 45 (2010) (collecting and describing authorities from each circuit).

102. *Norris v. Hicks*, No. 20-11460, 2021 WL 1783114, at *1 (11th Cir. May 5, 2021).

103. *Treat v. Lowe*, 668 F. App'x 870 (11th Cir. 2016) (per curiam). The factual underpinnings of *Treat* are more robustly discussed in the District Court's decision: *Treat v. Lowe*, No. 1:14CV174, 2016 WL 1246406 (S.D. Ga. Mar. 24, 2016).

consider the impact of *Treat*, because the decision was unpublished, and unpublished decisions in the Eleventh Circuit cannot clearly establish constitutional law.¹⁰⁴ Because Norris is also unpublished, neither *Norris* nor *Treat* provide any meaningful benefit to future litigants or courts within the Eleventh Circuit. This result is somewhat ironic, given the Eleventh Circuit's policy on the publication of opinions is designed to *minimize* impairments to "the development of the cohesive body of law."¹⁰⁵

F. *Contours of a Clearly Established Right*

Most courts continue to analyze qualified immunity using the two steps expressly discussed in *Saucier*: (1) were a plaintiff's rights violated, and (2) was the violated right clearly established at the time of the violation. If a court finds for the officer on the first, question, the officer is entitled to judgment on the merits. If the officer prevails on the second question, he is entitled to judgment on the basis of qualified immunity.

A common criticism of the second prong relates to the level of granularity courts use to define whether a right was clearly established at the relevant time. The challenge facing courts is hardly surprising. Broadly speaking, a denial of qualified immunity asks individuals at the top levels of a profession premised on the ability to craft and advance arguments to concede that an issue—one argued by thoughtful counsel who might have encouraged their clients to settle if that client was without valid argument—is beyond debate. That can be a tall order.

104. *Norris*, 2021 WL 1783114, at *6 n.9.

105. FED. R. APP. P. 36, Eleventh Circuit Internal Operating Procedure 5, available at https://www.ca11.uscourts.gov/sites/default/files/courtdocs/clk/Rules_Bookmark_DEC20.pdf.

Some courts have responded to the clear-establishment principle through the use of what at least one scholar has called “Ultra-Particularity.” “‘Ultra-particularity’ is a clever tool used to invoke qualified immunity and shield officers and jailers from liability. The term is a combination of the words *ultra*, which is defined as ‘beyond what is ordinary, proper, or moderate; excessively; extremely,’ and particularity, which means ‘the quality or state of being particular as distinguished from the universal.’”¹⁰⁶ Some judges apply what Wallach would call ultra-particularity to only deny qualified immunity where nearly identical conduct was already found, in a binding manner, to violate the constitution.

While the Supreme Court has not granted certiorari and accepted full briefing and argument in many cases concerning the contours of “clear” establishment, a significant number of per curiam decisions in unargued cases have been devoted to the issue.¹⁰⁷ In each of these cases, the Supreme Court reversed a lower court which had denied an officer qualified immunity.

These per curiam qualified immunity opinions are somewhat atypical of cases decided by the Supreme Court. The Supreme Court’s rules state: “A petition for a writ of certiorari is rarely granted when the asserted error consists of erroneous factual findings or the misapplication of a properly stated rule of

106. Ian Wallach, The Use of ‘Ultra-particularity’ to Invoke Qualified Immunity: A Substantial Threat to Civil Rights Claims and a Potentially Fatal Blow to Eighth Amendment Claims, *THE CHAMPION*, Mar. 2021, at 42–47.

107. See, e.g., *City of Escondido, Cal. v. Emmons*, 139 S. Ct. 500 (2019) (per curiam); *Kisela v. Hughes*, 138 S.Ct. 1148 (2018) (per curiam); *White v. Pauly*, 137 S. Ct. 548 (2017) (per curiam); *Mullenix v. Luna*, 577 U.S. 7 (2015) (per curiam); *Taylor v. Barkes*, 575 U.S. 822 (2015) (per curiam); *Tolan v. Cotton*, 572 U.S. 650 (2014) (per curiam); *Stanton v. Sims*, 571 U.S. 3 (2013) (per curiam).

law.”¹⁰⁸ Cases involving questions of qualified immunity make up a fair number of these rare grants—about one a term for nearly a decade. Reaction, or what some scholars and practitioners might term overreaction, to cases like these may be a driving factor in the development of cases utilizing ultra-particularity. In these reversals, the Supreme Court has emphasized that it “has repeatedly told courts not to define clearly established law at a *high* level of generality.”¹⁰⁹

Before November 2020, each of these unargued per curiam decisions reversed a decision denying qualified immunity. Before November 2020, the Supreme Court had only reached a Plaintiff’s result *twice* where a case raised qualified immunity. This term, the Court has more than doubled that number.

One pre-November 2020 case centered on a claim of excessive force.¹¹⁰ In *Hope v. Pelzer*,¹¹¹ the Court was faced with a gruesome Eighth Amendment claim, in which a prisoner had been chained to a hitching post, shirtless, for seven hours in the hot Alabama sun while guards taunted him about his thirst and obvious agony.¹¹² The Court described this as an “obvious” violation of the Eighth Amendment.¹¹³ It explained that, even if he had at some point been disruptive or posed a safety concern—which was not at all clear—by the time Hope was restrained, any such fear had abated, making the act “punitive treatment amount[ing] to gratuitous infliction of “wanton and

108. SCOTUS Rule 10. The Supreme Court’s Rules are available at <https://www.supremecourt.gov/ctrules/2019RulesoftheCourt.pdf>.

109. *Emmons*, 139 S. Ct. at 503 (citation and quotation marks omitted, emphasis added).

110. The second, *Groh v. Ramirez*, 540 U.S. 551 (2004), involved the nonviolent execution of a facially deficient warrant.

111. 536 U.S. 730 (2002).

112. *Id.* at 734–35.

113. *Id.* at 738.

unnecessary” pain that our precedent clearly prohibits.”¹¹⁴ The Court denied the guards qualified immunity, even though the facts presented were “novel.”¹¹⁵

In November 2020, the Court deviated from its per curiam pattern—it reversed a *grant* of qualified immunity and remanded the unargued case for further proceedings.¹¹⁶ *Taylor* was a conditions-of-confinement case with “particularly egregious facts[, which] any reasonable officer should have realized . . . offended the Constitution.”¹¹⁷ The plaintiff Taylor was housed in a “pair of shockingly unsanitary cells” for six full days and was unable to eat or drink for four days due to contamination concerns.¹¹⁸

Then, less than six months later in *McCoy v. Alamu*, the Court reversed another grant of qualified immunity, without opinion, in light of *Taylor v. Riojas*.¹¹⁹ *McCoy* involved a guard using pepper spray on an incarcerated plaintiff, without cause or provocation. The Fifth Circuit concluded that it was not clearly

114. *Id.*

115. *Id.* at 741. In urging reform, an Assistant City Attorney in Mesa, Arizona, has argued that lower courts should more regularly follow the lead of *Hope*, and deny qualified immunity more regularly when confronted with obvious, if novel, constitutional violations. Alexander J. Lindvall, *Qualified Immunity and Obvious Constitutional Violations*, 28 GEO. MASON L. REV. 1047 (2021).

116. *Taylor v. Riojas*, 141 S. Ct. 52 (2020).

117. *Id.* at 54.

118. *Id.*

119. *McCoy v. Alamu*, No. 18-40856, 2021 WL 1279403, at *1 (5th Cir. Apr. 6, 2021). Filings on the Supreme Court’s docket in *McCoy* indicate that this decision was made *sua sponte*. The Supreme Court’s opinion in *Taylor v. Riojas* had not been cited by the parties following its issuance, although Plaintiff cited to an earlier, unrelated *Taylor* order from the Fifth Circuit. The parties’ briefing is freely accessible at <https://www.supremecourt.gov/search.aspx?filename=/docket/docketfiles/html/public/20-31.html>.

established that a *single* assault with pepper spray, even if it had no legitimate purpose and was entirely unprovoked, violated McCoy's rights, and thus that the assailant was entitled to qualified immunity, because it was not "beyond debate" that a single application of pepper spray was not a "*de minimus* use of physical force[.]"¹²⁰

Finally, on June 28, 2021, the Court summarily reversed a grant of qualified immunity in *Lombardo v. City of St. Louis, Missouri*.¹²¹ The plaintiff had been picked up by officers earlier that day for trespassing and failing to appear in court for a traffic ticket and was placed in a holding cell.¹²² He apparently made some effort to hang himself, and at least six officers responded to his cell and restrained him following some physical struggle.¹²³ After the plaintiff was prone, handcuffed, and shackled with leg irons, "officers held [plaintiff]'s limbs down at the shoulders, biceps, and legs. At least one other placed pressure on [plaintiff]'s back and torso. [Plaintiff] tried to raise his chest, saying, 'It hurts. Stop.'"¹²⁴ "*After 15 minutes of struggling in this position, [plaintiff]'s breathing became abnormal and he stopped moving.*"¹²⁵ Officers and medical personnel were unable to resuscitate him.¹²⁶

Only time will tell whether these cases represent a shift in the qualified immunity jurisprudence. They may be seen as a message to lower courts to shift away from ultra-particularity,

120. McCoy v. Alamu, 950 F.3d 226, 233 (5th Cir. 2020), *vacated*, 141 S. Ct. 1364 (2021).

121. No. 20-391, 2021 WL 2637856, at *1 (U.S. June 28, 2021).

122. *Id.*

123. *Id.*

124. *Id.*

125. *Id.* (emphasis added).

126. *Id.*

or perhaps they reflect some dichotomy in which the Court is more willing to reject qualified immunity defenses in the context of alleged misconduct behind prison walls. Many of the Supreme Court's grants of qualified immunity—all but one of the per curiam decisions cited in Footnote 107, *supra*—were Fourth Amendment claims involving interactions between officers and individuals who came into contact in homes or on the street. *Hope*, *Taylor*, and *McCoy* all involved Eighth Amendment claims by convicted and incarcerated plaintiffs, and *Lombardo* involved a pretrial detainee who was already inside of a holding cell at the time the force was applied.¹²⁷ If, and how, these recent cases will shift qualified immunity jurisprudence going forward is beyond the scope of this article.

127. The Court declined to specify whether his claim was properly viewed under the Fourth or Fourteenth Amendments. *Lombardo*, 2021 WL 2637856, at *1, n.2.

III. THE CRITICISM

Given this background, both scholars and jurists alike have challenged the continued efficacy of the jurisprudence in our current societal context. Those opponents argue that the qualified immunity defense undermines government accountability by shielding government officials from liability even in situations that appear to be, on their face and to the public, egregious examples of government overreach and abuse, particularly in the area of alleged police misconduct. “Commentators have argued that the Court’s decisions have provided unclear and shifting guidance about how factually similar a case must be to clearly establish the law and which courts’ decisions can clearly establish the law.¹²⁸ Commentators have also argued that the ‘clearly established’ standard protects officers who have outrageously abused their power simply because no prior decision has declared that conduct unlawful.”¹²⁹

The late Justice Ruth Bader Ginsburg acknowledged the potential for abuse created by the Court’s current jurisprudence on qualified immunity in the context of a false arrest case arising

128. See, e.g., Karen Blum, Erwin Chemerinsky & Martin A. Schwartz, *Qualified Immunity Developments: Not Much Hope Left for Plaintiffs*, 29 *TOURO L. REV.* 633, 653–56 (2013) (describing shifting standards for clearly established law); Alan K. Chen, *The Intractability of Qualified Immunity*, 93 *NOTRE DAME L. REV.* 1937, 1948–51 (2018) (describing confusion about how factually analogous prior court decisions must be to clearly establish the law); John C. Jeffries, Jr., *What’s Wrong with Qualified Immunity?*, 62 *FLA. L. REV.* 851, 854–59 (2010) (describing confusion about which sources can clearly establish the law and how factually analogous prior cases must be to clearly establish the law).

129. Schwartz, *supra* note 42 at 608 (footnotes in original). See Jeffries, *supra* note 128, at 854–58, 863–66; John C. Jeffries, Jr., *The Liability Rule for Constitutional Torts*, 99 *VA. L. REV.* 207, 256–58 (2013); see also Michael L. Wells, *Qualified Immunity After Ziglar v. Abbasi: The Case for a Categorical Approach*, 68 *AM. U. L. REV.* 379, 436–38 (2019).

under the Fourth Amendment. In *District of Columbia v. Wesby*, Justice Ginsburg, while concurring in the judgment based on precedent, observed:

The Court's jurisprudence, I am concerned, sets the balance too heavily in favor of police unaccountability to the detriment of Fourth Amendment protection. A number of commentators have criticized the path we charted in *Whren v. United States*, 517 U. S. 806 (1996), and follow-on opinions, holding that "an arresting officer's state of mind . . . is irrelevant to the existence of probable cause," *Devenpeck v. Alford*, 543 U. S. 146, 153 (2004). See, e.g., 1 W. LaFare, *Search and Seizure* §1.4(f), p. 186 (5th ed. 2012) ("The apparent assumption of the Court in *Whren*, that no significant problem of police arbitrariness can exist as to actions taken with probable cause, blinks at reality."). I would leave open, for reexamination in a future case, whether a police officer's reason for acting, in at least some circumstances, should factor into the Fourth Amendment inquiry.¹³⁰

That same year in *Kisela v. Hughes*,¹³¹ one of the per curiam decisions referenced above, Justice Sonia Sotomayor, in an impassioned dissent, challenged the majority's view that a police officer was entitled to qualified immunity for shooting a woman who was alleged to have been engaging in "erratic behavior" with a knife. Justice Sotomayor observed that the majority opinion:

130. 138 S. Ct. 577, 594 (2018) (Ginsburg, J, concurring).

131. 138 S. Ct. 1148 (2018).

is symptomatic of “a disturbing trend regarding the use of this Court’s resources” in qualified-immunity cases. [*Salazar-Limon v. Houston*, 137 S.Ct. 1277, 1282 (2017) (Sotomayor, J., dissenting from denial of certiorari)]. As I have previously noted, this Court routinely displays an unflinching willingness “to summarily reverse courts for wrongly denying officers the protection of qualified immunity” but “rarely intervene[s] where courts wrongly afford officers the benefit of qualified immunity in these same cases.” *Id.*, at [1282–83]; see also Baude, *Is Qualified Immunity Unlawful?*, 106 CAL. L. REV. 45, 82 (2018) (“[N]early all of the Supreme Court’s qualified immunity cases come out the same way—by finding immunity for the officials”); Reinhardt, *The Demise of Habeas Corpus and the Rise of Qualified Immunity: The Court’s Ever Increasing Limitations on the Development and Enforcement of Constitutional Rights and Some Particularly Unfortunate Consequences*, 113 MICH. L. REV. 1219, 1244–1250 (2015). Such a one-sided approach to qualified immunity transforms the doctrine into an absolute shield for law enforcement officers, gutting the deterrent effect of the Fourth Amendment.

The majority today exacerbates that troubling asymmetry. Its decision is not just wrong on the law; it also sends an alarming signal to law enforcement officers and the public. It tells officers that they can shoot first and think later, and it tells the public that palpably unreasonable conduct will go unpunished. Because there is nothing right

or just under the law about this, I respectfully dissent.¹³²

Consistent with Justice Sotomayor's observations, Professor Erwin Chemerinksi, Dean of the University of California Berkeley School of Law, recently explained that "[i]n case after case, the Supreme Court found officers were protected by qualified immunity under [the Court's Section 1983 standing jurisprudence]. From 1982 to 2020, the court dealt with qualified immunity in 30 cases. The plaintiffs prevailed in only two: [*Hope v. Pelzer*, 536 U.S. 730 (2002) and *Groh v. Ramirez*, 540 U.S. 551 (2004)]."¹³³

Professor Kit Kinports, in her article *The Supreme Court's Quiet Expansion of Qualified Immunity*, likewise noted that, as of the time of her 2016 article, of the eighteen Section 1983 cases before the Supreme Court in the fifteen preceding years, the Court found that qualified immunity applied in sixteen of them.¹³⁴ She blames this result, in part, on the Court engaging "in a pattern of covertly broadening the defense, describing it in increasingly generous terms and inexplicably adding qualifiers to precedent that then take on a life of their own."¹³⁵ Indeed, Kinports suggests that the Court has all but lost sight of one of the of the countervailing interests first acknowledged by it in *Harlow*: "vindicating constitutional rights and compensating victims of constitutional injury."¹³⁶

132. *Id.* at 1162 (Sotomayor, J., dissenting).

133. Erwin Chemerinksi, *SCOTUS hands down a rare civil rights victory on qualified immunity*, ABA JOURNAL (February 1, 2021), available at <https://www.abajournal.com/columns/article/chemerinsky-scotus-hands-down-a-rare-civil-rights-victory-on-qualified-immunity>.

134. Kit Kinports, *The Supreme Court's Quiet Expansion of Qualified Immunity*, 100 MINN. L. REV. 62, 63 (2016).

135. *Id.* at 64.

136. *Id.* at 68.

In yet another of the Court's per curiam decisions, *Mullenix v. Luna*,¹³⁷ for example, a case involving alleged excessive force by a state trooper who shot and killed a motorist who was allegedly fleeing from arrest, Kinports observes that the Court's recitation of the governing qualified immunity standard is as follows:

The doctrine of qualified immunity shields officials from civil liability so long as their conduct "does not violate clearly established statutory or constitutional rights of which a reasonable person would have known." A clearly established right is one that is "sufficiently clear that every reasonable official would have understood that what he is doing violates that right." "We do not require a case directly on point, but existing precedent must have placed the statutory or constitutional question beyond debate." "Put simply, qualified immunity protects all but the plainly incompetent or those who knowingly violate the law."¹³⁸

Kinports notes the absence in this updated standard of any acknowledgment of the rights of the victim to redress for the alleged harms inflicted—a departure, she claims, from the jurisprudence in *Harlow* and its progeny.¹³⁹

In his article *Is Qualified Immunity Unlawful?*, Professor William Baude, an outspoken critic of the defense, argues that in addition to the broadening of the defense, the qualified immunity defense has been bolstered in a more fundamental way:

137. 577 U.S. 7 (2015).

138. Kinports, *supra* note 134, at 67–68.

139. *Id.* at 68.

Over the past several decades, the Court has been slowly changing the doctrinal formula for qualified immunity. Most recently, it has begun to strengthen qualified immunity's protection in another way: by giving qualified immunity cases pride of place on the Court's docket. It exercises jurisdiction in cases that would not otherwise satisfy the certiorari criteria and reaches out to summarily reverse lower courts at an unusual pace. Essentially, the Court's agenda is to especially ensure that lower courts do not improperly deny any immunity.¹⁴⁰

Scholars like Baude argue that the qualified immunity defense, which he claims was intended to serve three Court-proffered purposes—creating a good-faith exception to alleged constitutional wrongs suffered at the hand of the state; correcting for the erroneous overinclusion of actions recognized under Section 1983; and providing a warning against future violations of like kind—does not truly serve those purposes and, even if it did, there are better alternatives to the doctrine.¹⁴¹

Baude explains that the contemporary expansion of qualified immunity suggests that the statute itself—Section 1983—demands this “good faith” exception to the deprivation of rights by a state actor.¹⁴² Baude rejects this both facially, insofar as the statute itself does not provide for it, and historically, as having

140. William Baude, *Is Qualified Immunity Unlawful?*, 106 CAL. L. REV. 45, 48 (2018).

141. See, e.g., *id.*; Joanna C. Schwartz, *After Qualified Immunity*, 120 COLUM. L. REV. 309 (2020); Scott Michelman, *The Branch Best Qualified to Abolish Immunity*, 93 NOTRE DAME L. REV. 1999 (2018); Joanna C. Schwartz, *The Case Against Qualified Immunity*, 93 NOTRE DAME L. REV. 1797 (2018).

142. Baude, *supra* note 140, at 55–58.

been previously rejected in the Court's jurisprudence.¹⁴³ Moreover, Baude explains that this expansion is not lost on the High Court. He cites to both Justice Anthony Kennedy's and Justice Clarence Thomas's specific acknowledgements that the current jurisprudence of qualified immunity has strayed far from its historical roots and far from analogous common law immunities.¹⁴⁴

As the Court's qualified immunity expansion is grounded in neither the statutory framework of Section 1983 nor in its historical roots, Baude explores an alternative theory that the broadening of the qualified immunity defense was a course correction for the Court's expansive view of recovery and actionability under the statute. Baude cites to Justice Antonin Scalia's dissent in *Crawford-El v. Britton*,¹⁴⁵ in which Justice Scalia explains, in part:

Monroe [v. Pape] changed a statute that had generated only 21 cases in the first 50 years of its existence into one that pours into the federal courts tens of thousands of suits each year, and engages this Court in a losing struggle to prevent the Constitution from degenerating into a general tort law. (The present suit, involving the constitutional violation of misdirecting a package, is a good enough example.) Applying normal common-law rules to the statute that *Monroe* created would carry us further and further from what any sane Congress could have enacted.

We find ourselves engaged, therefore, in the essentially legislative activity of crafting a sensible scheme of qualified immunities for the statute we have invented—rather than applying the common

143. *Id.* at 51–60.

144. *Id.* at 61.

145. 523 U.S. 574 (1998).

law embodied in the statute that Congress wrote.¹⁴⁶

This explanation, however, is found inadequate by Professor Baude, as it is premised on the notion that *Monroe* was wrongly decided because it was based upon an erroneous interpretation of what it means to commit an act “under color of law.”¹⁴⁷ Baude goes on to explain that historically, the interpretation offered by the majority in *Monroe* was historically grounded and, therefore, accurate.¹⁴⁸

Additionally, Baude explained that if Justice Scalia’s compensation theory—in essence that two wrongs make a right, and *Monroe* was one of those wrongs—were correct, the resulting immunity doctrine should be the opposite of the immunity doctrine that currently exists.¹⁴⁹ Baude writes:

Section 1983 fills in a remedial gap: it provides a federal forum for conduct legalized or immunized by the state. Yet qualified immunity entirely ignores both state liability and state immunity. . . . That would mean denying immunity in cases where states grant it, while granting immunity only in cases where states deny it. Yet modern qualified immunity doctrine looks nothing like this.¹⁵⁰

Finally, Baude tested the theory that the purpose of the expansive qualified immunity doctrine is to give “fair warning” to

146. Baude, *supra* note 140 at 62–63, quoting *Britton*, 523 U.S. at 611–12 (Scalia, J., dissenting).

147. *Id.* at 63–64.

148. *Id.* at 64–65.

149. *Id.* at 66.

150. *Id.* at 68.

a potential wrongdoing state actors that a yet-to-be committed act is contrary to the Constitution—the concept familiar in the criminal law, known as “lenity.”¹⁵¹ Baude rejects this explanation, citing the differential treatment a criminal defendant receives in response to a lenity defense when compared to that raised by a state actor in response to a Section 1983 claim: “The Justices regularly empathize with officials subject to suit, asking if the official can really be expected to anticipate constitutional rulings that even federal appellate judges did not. But one rarely sees a similar empathy for regular criminal defendants, and indeed the Court’s decisions do not bear it out.”¹⁵² Baude posits that even if the concept of lenity were the driving force behind the Court’s qualified immunity defense jurisprudence, lenity “seems to justify a much more modest immunity doctrine than the one we have, one that at most, tracks the modest defenses available to real criminal defendants.”¹⁵³

In the end, Professor Baude answers his titular question in the affirmative, concluding that qualified immunity is unlawful.¹⁵⁴ For Baude, the doctrine is neither founded in the statutory language of Section 1983, nor authorized under any appropriate theory of statutory interpretation.

Echoes of Professor Baude’s conclusions are whispered in the halls of the Supreme Court. In addition to the above-referenced critiques from Justices Ginsburg and Sotomayor, Justice Thomas has offered one of the most recent embraces of Baude’s findings in a dissent from the denial of certiorari in *Baxter v. Bracey*:¹⁵⁵

151. *Id.* at 69.

152. *Id.* at 77.

153. *Id.*

154. *Id.* at 80.

155. 140 S. Ct. 1862, 1865 (2020).

There likely is no basis for the objective inquiry into clearly established law that our modern cases prescribe. Leading treatises from the second half of the 19th century and case law until the 1980s contain no support for this “clearly established law” test. Indeed, the Court adopted the test not because of “‘general principles of tort immunities and defenses,’” . . . but because of a “balancing of competing values” about litigation costs and efficiency. . . .

Regardless of what the outcome would be, we at least ought to return to the approach of asking whether immunity “was ‘historically accorded the relevant official’ in an analogous situation ‘at common law.’”¹⁵⁶

Still, others like Professors Hillel Levin and Michael Wells of the University of Georgia School of Law in their article, *Qualified Immunity and Statutory Interpretation: A Response to William Baude*, disagree with Baude’s ultimate answer and suggest that his statutory interpretation argument is flawed insofar as it relies on faulty methodology.¹⁵⁷ Nevertheless, even both Levin and Wells appear to agree that the defense has been subject to abuse and should be subject to “adjustment,”¹⁵⁸ writing “[t]here is much to criticize about the Court’s § 1983 jurisprudence, including the expansive qualified immunity doctrine it has

156. *Id.* at 1864 (citations omitted) (Thomas, J., dissenting).

157. Hillel Y. Levin & Michael L. Wells, *Qualified Immunity and Statutory Interpretation: A Response to William Baude*, 9 CAL. L. REV. ONLINE 40, 41–43 (2018).

158. *Id.* at 41.

developed. We share many of Professor Baude's apparent policy preferences, but we think his methodology is wrong."¹⁵⁹

Even in the face of this more academic disagreement concerning the origin of the qualified immunity defense and whether its development is grounded in sound statutory interpretation, the strains of the defense itself are best viewed through the lens of those on the front line of its application—the lower federal courts. While only few of these cases ever receive significant press coverage, and even fewer get to the Supreme Court, thousands of Section 1983 cases are filed each year in our nation's trial courts.¹⁶⁰ It is district and circuit court judges who must grapple first with what the law forbids and what will go uncompensated.

In recent years, frustration has crept into the jurisprudence of some lower federal court judges forced to apply the qualified immunity doctrine in the face of what might appear to be unjustified police conduct. In *Jamison v. McClendon*,¹⁶¹ District Judge Carlton Reeves of the Southern District of Mississippi was forced to grapple with qualified immunity in a case involving a Black welder in South Carolina who was stopped and searched for approximately two hours by police seemingly because he was driving a Mercedes.

Judge Reeves began his opinion with a gut-wrenching recitation of cases involving Black men and women who had been stopped, searched and, largely, killed, by police officers.¹⁶² Finding that despite what he believed to be outrageous and unjustified conduct by the police officer who stopped Mr. Jamison, the

159. *Id.* at 70.

160. *See* page 539, *supra*, (estimating the number of cases at between 40,000 and 50,000).

161. 476 F. Supp. 3d 386 (S.D. Miss. 2020).

162. *Id.* at 390–91.

officer was entitled to qualified immunity, Judge Reeves explained:

The Constitution says everyone is entitled to equal protection of the law—even at the hands of law enforcement. Over the decades, however, judges have invented a legal doctrine to protect law enforcement officers from having to face any consequences for wrongdoing. The doctrine is called “qualified immunity.” In real life it operates like absolute immunity.

Tragically, thousands have died at the hands of law enforcement over the years, and the death toll continues to rise. Countless more have suffered from other forms of abuse and misconduct by police. Qualified immunity has served as a shield for these officers, protecting them from accountability.

This Court is required to apply the law as stated by the Supreme Court. Under that law, the officer who transformed a short traffic stop into an almost two-hour, life-altering ordeal is entitled to qualified immunity. The officer’s motion seeking as much is therefore granted.

But let us not be fooled by legal jargon. Immunity is not exoneration. And the harm in this case to one man sheds light on the harm done to the nation by this manufactured doctrine.¹⁶³

163. *Id.* at 391–92.

Similarly, in *Ventura v. Rutledge*,¹⁶⁴ District Judge Dale Drozd, offered his own perspectives on qualified immunity in a case involving the fatal shooting of an individual by police, who was alleged not to have posed an immediate threat to himself or others:

In legal circles and beyond, one of the most debated civil rights litigation issues of our time is the appropriate scope and application of the qualified immunity doctrine, particularly in cases of deaths resulting from police shootings. . . . While there is so much more that could, and perhaps should, be said about the current state of this judicially created doctrine, the undersigned will stop here for today. In short, this judge joins with those who have endorsed a complete reexamination of the doctrine which, as it is currently applied, mandates illogical, unjust, and puzzling results in many cases. However, the Supreme Court's decision in *Kisela* is, of course, binding on this court.¹⁶⁵

At least one federal judge has done what would normally be unthinkable—sounded the alarm in the media. On the heels of a string of tragic killings by police officers in 2020, Circuit Judge James Wynn Jr. of the U.S. Court of Appeals for the 4th Circuit took to the Washington Post to air his grievances with qualified immunity, writing in an opinion piece:

The judge-made law of qualified immunity subverts the Civil Rights Act of 1871, which Congress intended to provide remedies for constitutional violations perpetrated by state officers.

164. 398 F. Supp. 3d 682 (E.D. Cal. 2019).

165. *Id.* at 687 n.6 (citations omitted).

Eliminating the defense of qualified immunity would improve our administration of justice and promote the public's confidence and trust in the integrity of the judicial system.¹⁶⁶

Notwithstanding these expressions of frustration—including frustrations expressed by the justices themselves—Supreme Court jurisprudence on qualified immunity for alleged police misconduct remains obdurate.

* * *

Against this backdrop, on May 25, 2020, an African American man entered a grocery store in Minneapolis, Minnesota. The man was alleged to have used a counterfeit \$20 bill to pay for a purchase. Police were called. The man was arrested for this crime, and during that arrest was murdered by a police officer who restrained the man by kneeling on the man's neck while the man pleaded for his life. What followed was an unprecedented level of protest activity focused on police brutality. The cries for justice for the man—George Floyd—and others who had died at the hands of police, could be heard throughout the country—in cities and towns, big and small.

Faced with the growing chorus of outrage concerning alleged police misconduct, attention soon turned to the Congress to act. On February 24, 2021, the George Floyd Justice in Policing Act, H.R. 1280, was introduced in the United States House of Representatives. The law's purpose is to address police misconduct, including excessive force, and racial bias in policing. Relevant here, the bill seeks to limit qualified immunity as a

166. Hon. James Wynn Jr., *Opinion: As a judge, I have to follow the Supreme Court. It should fix this mistake.*, WASHINGTON POST, June 12, 2020, available at <https://www.washingtonpost.com/opinions/2020/06/12/judge-i-have-fol-low-supreme-court-it-should-fix-this-mistake/>.

defense to liability in private civil actions against a law enforcement officer.

On March 3, 2021, H.R. 1280, passed the House by a narrow margin and was sent to the United States Senate. Indicative or reflective of the same jurisprudential paralysis detailed above—changes stymied by the acknowledgment of the realities and dangers of police work in the face of a stream of police killings of predominantly Black men—the primary points of contention surrounding the bill concern the availability of the qualified immunity defense. As of this writing, despite a good deal of negotiation, the Senate has yet to bring the George Floyd Justice in Policing Act to vote.

IMPLICIT BIAS: THE SCIENCE, INFLUENCE, AND IMPACT ON JUSTICE

*Hon. Bernice B. Donald**

Walking, answering the phone, drinking a hot beverage, driving a car, eating out—every day we do many of these things with little conscious effort. When we see steam coming from a hot beverage, it takes little time to process the information to determine its meaning; we know from past experiences that steam coming from a beverage means we should proceed with caution. Phones come in many forms, yet we know when the landline phone on our desk, or the almost-obsolete wall-mounted phone in our kitchen, or the computer-like, rectangular, handheld device rings, we should answer. When walking, we need not analyze each obstacle on the sidewalk to determine how to proceed. When driving, our brain has milliseconds to process information and tell our body to react to avoid collision. When we see something barreling toward us, we know instantly to avoid the object. There is no time to consciously think about what the object may be, how fast it is traveling, or where it came from; we act immediately. Whether at a fast-food restaurant or an elegant establishment with refined cuisine, we have a general idea how to act when we walk in. Does one go directly to the counter or wait to be seated? In each of these scenarios, we know how to respond each time—without thinking—based on our past experiences dating back as far as early childhood.

* Judge Bernice B. Donald is a United States appellate court judge for the U.S. Court of Appeals, Sixth Circuit. The author wishes to thank her externs Alexxas Johnson, Juedon Kebede, Alex McWhirter, Hailey Townsend; her law clerk Naira Umarov; and her judicial assistant Amy Dueñes for their contributions.

The human body sends 11 million bits of information per second to the brain for processing, yet the conscious mind can process a mere 50 of those bits in the same amount of time.¹ What happens to the 10,999,950 bits of information that our conscious mind does not process? Researchers conclude that the vast majority of processing is accomplished outside of the conscious mind and the body's direct conscious control.²

According to the American Academy of Family Physicians, automatic cognitive processes shape human behavior, beliefs, and attitudes from a very young age.³ As we grow, the processes transform according to personal life experiences, family upbringing, and information absorbed through media. These cognitive processes help determine how humans filter perceptions, decision-making, and systematic errors in judgment.⁴ The cognitive process also results in a preferential ranking and grouping of our peers and others in our community.

Attitudes or stereotypes that affect our understanding, actions, and decisions in an unconscious way are defined as implicit or unconscious bias.⁵ Mahzarin Banaji and Anthony Greenwald coined the term "implicit bias" in 1995. They argued that social behavior is largely influenced by these unconscious associations and judgments—those other 10,999,950 bits of

1. BRITANNICA.COM, <https://www.britannica.com/science/information-theory/Physiology> (last visited Apr. 29, 2021).

2. *Id.*

3. Jennifer Edgoose, Michelle Quiogue & Kartik Sidhar, *How to Identify, Understand, and Unlearn Implicit Bias in Patient Care*, FAM. PRAC. MGMT., (July/August 2019), <https://www.aafp.org/fpm/2019/0700/p29.html>.

4. *Id.*

5. Charlotte Ruhl, *Implicit or Unconscious Bias*, SIMPLY PSYCHOLOGY (July 1, 2020), <https://www.simplypsychology.org/implicit-bias.html>.

information (per *second*) that our conscious brain is not capable of processing.⁶

Cognitive science research reveals that our automatic nervous system triggers unconscious frameworks of thinking that, in turn, influence our otherwise neutral, logical, and reasoned judgments.⁷ The brain processes information via schemas, which are templates of knowledge that assist us with organizing data into broader categories.⁸ For example, “when we see a figure with four equal sides, we quickly recognize that figure to be a square without giving much thought.”⁹

These schemas are “important and helpful because they allow us to function without unnecessarily expending mental resources.”¹⁰ Schemas apply not only to objects, shapes, or behaviors, but also to human beings.¹¹ Our brains naturally assign people into various categories “divided by salient and readily accessible traits, such as age, gender, and race.”¹² Just as schemas help us walk and drive, our brains create schemas and implicit social cognition, which can guide our thinking and action.¹³ These schemas develop not at once and not from one source, but rather over time through culture, direct or indirect

6. *Id.*

7. Ronald Chen & Jon Hanson, *Categorically Biased: The Influence of knowledge Structures on Law and Legal Theory*, 77 S. CAL. L. REV. 1103, 1128 (2004).

8. See Alfred Ray English, *Understanding Implicit Bias*, 55 ARIZ. ATT’Y 10 (2019).

9. See *id.*

10. *Id.*

11. See MAHZARIN BANAJI & ANTHONY G. GREENWALD, *BLINDSPOT* (2013).

12. See Shawn C. Marsh & Diane C. Marsh, *Being Explicit about Implicit Bias Education for the Judiciary*, 56 CT. REV. 92 (2020).

13. *Id.*

messaging, and past experiences.¹⁴ The sources of these schemas can be our parents, family, friends, school, and media, among infinite other sources.¹⁵

Beyond relying on schemas for daily activities, research on implicit bias identifies several other conditions in which individuals are likely to rely on their unconscious behaviors. These include situations that involve ambiguous or incomplete information; the presence of time constraints; and circumstances in which our cognitive control may be compromised, such as when we are fatigued or have too many other things on our mind.¹⁶

We are continuously exposed to certain identity groups paired with specific characteristics, and we begin to automatically and unconsciously associate the identity with the characteristic, whether or not that association finds any basis in reality.¹⁷ Without schemas, we would not be able to process as efficiently or effectively the “vast amount of sensory data” we obtain on a daily basis.¹⁸ Reliance on schemas our brain has created from past experiences or other sources is a natural occurrence, though this reliance can (and often does) lead to inaccurate and biased judgments.¹⁹ We are taught to be aware of our surroundings when walking alone or at night, and we therefore might react with caution when we see someone approaching us, but do we act differently depending upon what *type* of person approaches us? White, black, male, female, tall, short, old, young, person with a disability—do we change our reaction

14. *Id.* at 92.

15. *Id.*

16. See Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 UCLA L. REV. 1124, 1129 (2012).

17. *Id.* at 1130.

18. English, *supra* note 8, at 12.

19. *Id.*

based on any of these characteristics? For many, the answer is yes. Though unfortunate, these differing reactions are entirely human. Implicit bias is a result of those learned schemas from our environment, society, media, and other sources. How would one describe a drug dealer from a movie? What type of person comes to mind in the first split second? What about a professional football player, astronaut, or doctor? Our past experiences continuously and unrelentingly shape our unconscious decisions.

A person's actions or comments based on implicit bias may be discriminatory but not necessarily intentional.²⁰ Explicit biases are attitudes and stereotypes that are consciously accessible through one's own conscious, while implicit biases are not consciously accessible and are experienced without awareness.²¹ Explicit bias can be somewhat easy to recognize because it is "deliberately generated and consciously experienced as one's own belief."²² Common examples of explicit biases can be overt acts of racism and racist comments.²³

Implicit bias, however, does not require animus but instead only familiarity with some stereotype.²⁴ Nevertheless, implicit bias can be just as problematic as explicit bias because both can cause prejudice against a marginalized community.²⁵ With

20. *Id.*

21. Kang, *supra* note 16, at 1132.

22. J. BERNICE B. DONALD & SARAH E. REDFIELD, *ENHANCING JUSTICE: REDUCING BIAS*, Ch. 2 *Framing the Discussion*, 5, 14, (Sarah E. Redfield ed., 2017).

23. See Michele Benedetto Neitz, *Pulling back the Curtain: Implicit Bias in the Law School Dean Search Process*, 49 SETON HALL L. REV. 629, 655 (2019).

24. See B. Keith Payne, Heidi A. Vuletich & Kristjen B. Lundberg, *The Bias of Crowds: How Implicit Bias Bridges Personal and Systemic Prejudice*, 28 PSYCHOL. INQUIRY 233, 238 (2017).

implicit biases, individuals may not be mindful that their biases—rather than the reality of a situation—influence their decision-making.²⁶ By way of common example, implicit bias might make police officers automatically suspicious of two young African American males driving in a neighborhood where few African Americans reside.²⁷ While much education on implicit bias has centered on race and ethnic backgrounds, it is important to note that there are many other implications of the unconscious judgment, such as gender, body type, and age.²⁸

The social science on implicit bias has grown tremendously, becoming a popular topic in judicial education.²⁹ In the judicial context, education regarding implicit bias is critical because evidence from fields such as cognitive psychology suggests “that people can and do make decisions about others via cognitive mechanisms operating outside of their awareness.”³⁰ Since a judge’s primary role is to make decisions impacting others while sustaining objectivity, it is essential that judges understand both the existence of implicit biases and ways to counteract them.³¹

25. See Neitz, *supra* note 23, at 656.

26. *Id.*

27. See Meera E. Deo, *Faculty Insights on Education Diversity*, 83 FORDHAM L. REV. 3115 (2015).

28. See Marsh & Marsh, *supra* note 12, at 92.

29. See Catie Wheatley, *Honesty is the Best Policy: Addressing Implicit Bias in the Judiciary*, 9 IND. J.L. & SOC. EQUAL. 94, 96 (2021).

30. See Marsh & Marsh, *supra* note 12, at 92.

31. Justin D. Levison, Mark W. Bennett & Koichi Hioki, *Judging Implicit Bias: A National Empirical Study of Judicial Stereotypes*, 69 FLA. L. REV. 63, 65 (2017).

IMPLICIT BIAS IN CIVIL LAW

Civil cases make up the majority of cases in courts.³² But despite the prevalence of civil cases, the criminal system influences laypersons' perceptions of the civil system.³³ On the civil side, potential plaintiffs might forgo a lawyer's assistance in their case or even forgo filing suit entirely. In some cases, as many as three-quarters of low-income individuals, mostly minorities, did not even seek an attorney's service for their legal issues.³⁴ Perceived mistreatment and bias in the *criminal* system leads to a strong sense of disenfranchisement among minority groups even in the *civil* system.

In a perfect world, all parties to these civil cases would make rational decisions free from any biases or undue influence. Judges in particular, with their experience and knowledge of the law, are expected to look beyond any biases and extraneous influences that might alter their decision-making.³⁵ However, even those with a mind trained towards equality can hold biases against others.³⁶ In other words, even individuals who are

32. See *Federal Judicial Caseload Statistics 2018*, U.S. COURTS, <https://www.uscourts.gov/statistics-reports/federal-judicial-caseload-statistics-2018>, (last visited Mar. 20, 2021) (displaying how there were 81,553 criminal cases and 277,010 civil cases in 2018).

33. *Report to the United Nations on Racial Disparities in the U.S. Criminal Justice System*, THE SENTENCING PROJECT (Apr. 19, 2018) <https://www.sentencingproject.org/publications/un-report-on-racial-disparities/> (noting the over-representation of persons of color in the U.S. prisons and the "prevalence of bias in the criminal justice system").

34. Sara Sternberg Greene, *Race, Class, and Access to Civil Justice*, 101 IOWA L. REV. 1263, 1265 (2016).

35. Melissa L. Breger, *Introducing the Construct of the Jury into Family Violence Proceedings and Family Court Jurisprudence*, 13 MICH. J. GENDER & L. 1, 25 (2006).

36. *U.S. Supreme Court Recognizes Role of Unconscious Bias in Disparate Treatment*, ASS'N FOR PSYCHOL. SCI. (July 1, 2015), <https://www.psychological>

trained to treat everyone equally can still attribute negative characteristics, such as “poverty, aggression, and even crime,” with certain demographics.³⁷

In 2013, the decision by the United States Supreme Court in *Shelby County v. Holder*³⁸ struck down Section 4(b) of the Voting Rights Act of 1965 containing a coverage formula that determined which state and local jurisdictions are subject to federal preclearance based on their histories of discrimination in voting. *Shelby County v. Holder* is an example of how even the pinnacle of the American judiciary is not exempt from these biases.³⁹ Writing for the majority, Chief Justice John Roberts noted the level of progress made since the enactment of the Voting Rights Act of 1965.⁴⁰ The law was “one of the most consequential, efficacious and amply justified exercises of the federal legislative power in our Nation’s history.”⁴¹ States could no longer pass blatantly discriminatory laws that made it difficult or almost impossible for minorities to vote.⁴² Minority populations had a voice in local and federal politics after years of being unduly silenced.

Many believe the *Shelby County* decision allowed voter suppression efforts in various states to occur—an issue we still face

science.org/news/releases/us-supreme-court-recognizes-role-of-unconscious-bias-in-disparate-treatment.html.

37. *Id.*

38. 570 U.S. 529 (2013).

39. See, e.g., Adam Bolotin, *Out of Touch: Shelby v. Holder and the Callous Effects of Chief Justice Roberts’s Equal State Sovereignty*, 49 J. MARSHALL L. REV. 751 (2016); Bridgette Baldwin, *Backsliding: The United States Supreme Court, Shelby County v. Holder and the Dismantling of the Voting Rights Act of 1965*, 17 BERKELEY J. AFR.-AM. L. & POL’Y 251 (2015).

40. *Shelby County*, 570 U.S. at 551.

41. *Id.* at 561 (Ginsburg, J., dissenting).

42. Baldwin, *supra* note 39, at 251.

today.⁴³ In 2017 alone, minority voters were more than four times more likely to experience discrimination or voter disenfranchisement measures than white voters.⁴⁴ Considering the disparity between experiences at polling places, it is no stretch to imagine a scenario in which a bench with more minority voices might have viewed the case differently.⁴⁵ Perhaps someone who has personally dealt with discrimination might not view the progress under the Voting Rights Act of 1965 as complete.⁴⁶ A diverse judiciary⁴⁷ that is aware of biases can better ensure all voices are heard fairly in American courts.⁴⁸

43. See, e.g., Vann R. Newkirk, II, *How Shelby County v. Holder Broke America*, THE ATLANTIC (July 10, 2018), <https://www.theatlantic.com/politics/archive/2018/07/how-shelby-county-broke-america/564707/>.

44. *Discrimination in America*, ROBERT WOOD JOHNSON FOUND., <https://www.rwjf.org/en/library/research/2017/10/discrimination-in-america-experiences-and-views.html> (last visited Mar. 19, 2021).

45. “The grand aim of the Act is to secure to all in our polity equal citizenship stature, a voice in democracy undiluted by race.” *Shelby County*, 570 U.S. at 592 (Ginsburg, J., dissenting). Justice Ginsburg further noted in her dissent that getting rid of the Voting Rights Act of 1965 because of the progress it has made is like “throwing away your umbrella in a rainstorm because you are not getting wet.” *Id.* at 590 (Ginsburg, J., dissenting).

46. See Pat K. Chew & Robert E. Kelley, *The Realism in Judicial Decision Making: An Empirical Analysis of Plaintiffs’ Race and Judge’s Race*, 28 HARV. J. RACIAL & ETHNIC JUST. 91, 105 (2012) (noting how judges of different races and backgrounds can perceive things differently and might not be sensitive to racial harassment if never having personally experienced it).

47. Since President Reagan, every president has increased the racial diversity of the federal judiciary from his party’s predecessor. President Trump is the only president in that time to break the trend. John Gramlich, *Trump has appointed a larger share of female judges than other GOP presidents, but lags Obama*, PEW RES. CTR. (Oct. 2, 2018), <https://www.pewresearch.org/fact-tank/2018/10/02/trump-has-appointed-a-larger-share-of-female-judges-than-other-gop-presidents-but-lags-obama/>.

“[E]xperience and common sense” can lead to disparate conclusions based on otherwise identical information.⁴⁹ Studies indicate that immigration and discrimination claims before Asian American judges have a higher success rate than those before white judges.⁵⁰ The shared, firsthand experiences of immigration and discrimination give the Asian American judge a more sympathetic interpretation.⁵¹ Likewise, African American judges tend to view Fourth Amendment cases, prohibiting unreasonable searches and seizures, more favorably than white judges.⁵² Moreover, workplace discrimination and harassment cases are more likely to succeed on their claims when before a judge of the same race as the plaintiff.⁵³

Religion and gender can affect the outcome of judgments as well. With regard to religion, studies indicate that Jewish judges have a tendency to side in favor of minority religions, likely due to their belonging to a religion that has suffered much persecution.⁵⁴ Alternatively, Catholic and Evangelical judges are more likely to disfavor LGBTQ plaintiffs in their cases, and also

48. See Dana Leigh Marks, *Who, Me? Am I Guilty of Implicit Bias?*, 54 JUDGES' J. 20 (2015) (detailing one judge's need to remind herself that one culture's view of something as simple as direct eye contact or storytelling may be different).

49. Elizabeth Thornburg, *(Un)Conscious Judging*, 76 WASH. & LEE L. REV. 1567, 1582 (2019).

50. Josh Hsu, *Asian American Judges: Identity, Their Narratives, & Diversity on the Bench*, 11 ASIAN PAC. AM. L.J. 92, 108 (2006).

51. *Id.*

52. Nancy Scherer, *Diversifying the Federal Bench: Is Universal Legitimacy for the U.S. Justice System Possible?*, 105 NW. L. REV. 587, 606 (2011).

53. Chew & Kelley, *supra* note 46, at 105.

54. Jeffrey J. Rachlinski & Andrew J. Wistrich, *Judging the Judiciary by the Numbers: Empirical Research on Judges*, 13 ANN. REV. L. & SOC. SCI. 203, 206 (2017).

disfavor defendants in obscenity cases.⁵⁵ Female judges are far more likely to side with plaintiffs in sexual harassment and employment discrimination cases.⁵⁶ Likewise, women are more likely to deem statutes as violative of equal protection or LGBTQ rights.⁵⁷ Even age can influence a judge's decision; older judges tend to side with plaintiffs in discrimination cases.⁵⁸ They are more than twice as likely to favor the plaintiffs than their younger counterparts.⁵⁹

Lastly, life experiences and political party affiliation play a role in implicit biases. One study showed that judges identifying as Republican or Democrat were, depending on their political affiliation, more or less likely to discharge an individual's debts in simulated bankruptcy adjudications.⁶⁰ Beyond identity, life experiences have an impact on bias and perception. Judges with at least one daughter are more likely to side with plaintiffs in gender bias cases.⁶¹ Everything from education level, quality of education, military experience, or previous employment can influence a judge's perception.⁶² Racial identity is more readily identifiable as a way to explain a judge's naivete or sensitivity

55. *Id.*

56. Jennifer L. Peresie, *Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts*, 114 *YALE L. J.* 1759, 1776 (2005).

57. Fred O. Smith Jr., *Gendered Justice: Do Male and Female Judges Rule Differently on Questions of Gay Rights?*, 57 *STAN. L. REV.* 2087, 2123 (2005).

58. Rachlinski & Wistrich, *supra* note 54, at 208.

59. *Id.*

60. Jeffrey J. Rachlinski et al., *Inside the Bankruptcy Judge's Mind*, 86 *B.U. L. Rev.* 1227, 1229–30 (2006).

61. Adam N. Glynn & Maya Sen, *Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Women's Issues?*, 59 *AM. J. POL. SCI.* 37, 38 (2015).

62. Chew & Kelley, *supra* note 46, at 105.

to societal discrimination, but all facets of a judge's identity must be considered when examining potential implicit biases.⁶³

Pretrial rulings—those rulings that determine whether a case will even proceed to a jury—are also critically important and just as susceptible to the influence of biases.⁶⁴ While *Shelby County* might illustrate an example of implicit bias and its effects at large, *Ashcroft v. Iqbal*⁶⁵ shows how a decision can open the door for implicit bias solely within the judicial system.⁶⁶ In general, dismissals went from 46 percent to 61 percent following *Iqbal*'s heightened pleading standard.⁶⁷ One author notes that judges' decisions are now overly determinative at the pretrial stage, which increases the impact of their biases.⁶⁸ Even if a case survives a Federal Rules of Civil Procedure 12(b)(6) challenge, it is sure to face a motion for summary judgment. Plaintiffs at least have the benefit of an expanded narrative through the discovery phase when faced with a motion for summary judgment, but judges still have discretion regarding whether there is any "genuine dispute of material fact"⁶⁹ that is still subject to some degree of personal interpretation.⁷⁰ Thus, between a motion to dismiss and a motion for summary judgment, plaintiffs face two important challenges that allow for judicial discretion and possible implicit bias before a case even reaches the jury.

63. *Id.*

64. Thornburg, *supra* note 49, at 157.

65. 556 U.S. 662 (2009).

66. Tasha Hill, *Inmates' Need for Federally Funded Lawyers: How the Prison Litigation Reform Act, Casey, and Iqbal Combine with Implicit Bias to Eviscerate Inmate Civil Rights*, 62 UCLA L. REV. 176, 213 (2015).

67. *Id.*

68. See Elizabeth Thornburg, *Law, Facts, and Power*, 114 PENN STATIM 1, 2 (2009), <http://pennstatelawreview.org/114/114%20Penn%20Statim%201.pdf>.

69. FED. R. CIV. P. 56(a).

70. Kang, *supra* note 16, at 1164.

Not only does this show the negative impact implicit bias can have on claims in general, but it highlights the problem that implicit bias can perpetuate throughout the system. As more minority plaintiffs choose to proceed *pro se* because of their perception of the legal system,⁷¹ they are even less likely to succeed on their claims.⁷² When examining the race of judges with *pro se* plaintiffs, it becomes clear that white and African American judges differ in their application of *Iqbal* to race discrimination claims.⁷³ White judges dismissed such claims almost twice as often as African American judges.⁷⁴

Implicit bias can also manifest in the voir dire process in civil matters. Not only may prospective jurors give answers more likely to please the judge, but a judge might also unduly weigh the opinions of certain attorneys in the selection process.⁷⁵ Excessive involvement from judges in the voir dire process can result in a jury that conforms with a judge's personal narrative.⁷⁶ Recent scholarship suggests judges should take a more passive approach to the jury selection process.⁷⁷ This approach allows trial lawyers who are more familiar with the case and their client's interest to question and select the jury rather than the judge.⁷⁸

71. See Hill, *supra* note 66, at 213.

72. *Id.*

73. *Id.*

74. *Id.* (noting that white judges dismiss 57.5 percent of race-discrimination claims, while African American judges dismiss just 33.3 percent).

75. J. Mark W. Bennett, *Unraveling the Gordian Knot of Implicit Bias in Jury Selection: The Problems of Judge-Dominated Voir Dire, the Failed Promise of Batson, and Proposed Solutions*, 4 HARV. L. & POL'Y REV. 149, 160 (2010).

76. *Id.* at 165.

77. *Id.*

78. *Id.*

IMPLICIT BIAS IN CRIMINAL LAW

Judges, and the criminal justice system as a whole, must provide a defendant with a fair trial.⁷⁹ Implicit bias complicates this task.⁸⁰

A. *Adjudications and Judicial Response to IAT testing*

Indictments and the counts contained within those indictments can reveal implicit bias.⁸¹ As a judge, I often tell of my own experience during my early days on the bench. A prosecutor brought an indictment for felony possession of a firearm with eight counts to reflect the eight weapons possessed by the African American defendant. That same day, another prosecutor brought forward an indictment against a White defendant with the same charge but only two counts of felony possession of a firearm. This gave me pause because the White defendant also possessed eight weapons but only received one count for the eight weapons. I asked the prosecutor why the African American defendant had been punished so harshly for the same crime as a White defendant. The response: the prosecutor did not realize the disparity.

79. See U.S. CONST. amend VI (“In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed, which district shall have been previously ascertained by law, and to be informed of the nature and cause of the accusation; to be confronted with the witnesses against him; to have compulsory process for obtaining witnesses in his favor, and to have the Assistance of Counsel for his defense.”).

80. See Kang, *supra* note 16, at 1126 (explaining the presence of implicit bias in the criminal justice system from start to finish).

81. See Jeffery J. Rachlinski, et. al., *Does Unconscious Racial Bias Affect Trial Judges?*, 786 CORNELL L. FACULTY PUB. 1195 (2009) (explaining that implicit bias can interfere with a defendant’s right to a fair trial).

Given the nature of these biases, prosecutors are not always aware of their implicit biases when they decide what charges to bring against a particular defendant.⁸² When prosecutors have not been through implicit-bias training or are unaware of their biases, then the judge proves integral in filtering out implicit biases.

Judges are just as susceptible to implicit bias as prosecutors or any other participant in the criminal justice system. Jeffery Rachlinski, and his colleagues conducted a study on several judges from different regions across the United States.⁸³ The judges remained anonymous but were asked to divulge their race, gender, age, and political affiliation.⁸⁴

Rachlinski utilized the Implicit Association Test (“IAT”).⁸⁵ The IAT is a computerized priming test that measures implicit associations using pictures and words.⁸⁶ The test asks participants to pair an object, such as a racial group or gender group, with an evaluative dimension, such as “good” or “bad.”⁸⁷ Participants of the test quickly press a response key without

82. See Joseph J. Avery & Joel Cooper, *Racial Bias in Post-Arrest and Pretrial Decision Making: The Problem and a Solution*, 29 CORNELL J.L. & PUB. POL’Y 257, 263–64 (2019) (citing *McCleskey v. Kemp*, 481 U.S. 279, 312 (1987) (internal citation omitted)) (“Similarly, Radelet and Pierce reviewed prosecutorial decision making in over 1,000 Florida homicide cases and found a combination of harsher treatment of black defendants and more lenient treatment of white defendants. As Justice Powell acknowledged in *McCleskey v. Kemp*, ‘The power to be lenient is [also] the power to discriminate.’”).

83. Rachlinski, *supra* note 81, at 1209.

84. *Id.*

85. Marsh & Marsh, *supra* note 12, at 93.

86. DONALD & REDFIELD, *supra* note 22, at 14.

87. JUSTIN D. LEVISON ET AL., *ENHANCING JUSTICE: REDUCING BIAS*, Ch. 3 *Implicit Bias: A Social Science Overview* 43, 51 (Sarah E. Redfield ed., 2017).

thinking through their reactions.⁸⁸ The speed of the participant's response "indicates implicit and automatic attitudes and stereotypes."⁸⁹ Implicit measures such as the IAT have greater validity in predicting "spontaneous behaviors" like eye contact, seating distance, and other actions that could indicate discomfort.⁹⁰

Rachlinski's IAT test affirmed that judges in the criminal justice system harbor implicit bias.⁹¹ The White judges who participated in the study reported a preference for White defendants over African American defendants.⁹² The African American judges did not exhibit a preference either way during the initial IAT assessment.⁹³

Table 2: Results of Race IAT by Race of Judge⁹⁴

Race of Judge (Sample Size)	Mean IAT Score in Milliseconds (and standard deviation)*		Percent of Judges with Lower Average of Latencies on the White/good versus black/bad round
	Judges	Internet Sample	
White (85)	216 (201)	158 (224)	87.1
Black (43)	26 (208)	39 (244)	44.2

88. *Id.*

89. *Id.*

90. Anthony G. Greenwald & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations*, 94 CAL. L. REV. 945, 954–55 (2006).

91. Rachlinski, *supra* note 81, at 1210.

92. By preference, the authors found that the judges were more likely to be lenient in sentencing or bail hearings toward White defendants and more harsh to African American defendants.

93. Rachlinski, *supra* note 81, at 1210.

94. *Id.*

*Note: Positive numbers indicate lower latencies on the white/good versus black/bad round

After the IAT test, Rachlinski gave the judges three hypothetical scenarios with different crimes and variations regarding the defendant's and victim's race.⁹⁵

The data [showed] that when the race of the defendant is explicitly identified to judges in the context of a psychology study (that is, the third vignette [an offense of Battery]), judges are strongly motivated to be fair, which prompts a different response from White judges (who may think to themselves "whatever else, make sure not to treat the Black defendants worse") than Black judges (who may think "give the benefit of the doubt to Black defendants"). However, when race is not explicitly identified but implicitly primed (vignettes one and two [an offense of Shoplifting]), perhaps the judges' motivation to be accurate and fair is not on full alert.⁹⁶

The study concluded with three findings: (1) judges, like others in the criminal justice system, have implicit bias, especially with regard to race; (2) implicit bias affects judges' judgments when they are unaware of the need to monitor their decisions for implicit racial bias; and (3) "when judges are aware of a need to monitor their own responses for the influence of implicit racial biases, and are motivated to suppress that bias, they appear able to do so."⁹⁷

95. *Id.* at 1217–20.

96. Kang, *supra* note 16 at 1148.

97. Rachlinski, *supra* note 81, at 1221.

B. Plea Bargaining: The Elusive Situation for Implicit Bias

Judges have little leeway when it comes to plea bargains.⁹⁸ However, implicit bias can still infect these determinations. When assessing the validity of a plea bargain, judges are only allowed to consider whether the acceptance was knowing, voluntary, and uncoerced.⁹⁹ Prosecutors presenting these plea bargains, on the other hand, are certainly susceptible to implicit bias.¹⁰⁰

Less research has been done on implicit bias during plea bargaining as compared to sentencing. This is concerning because most criminal trials are resolved through the plea-bargaining process.¹⁰¹ Prosecutorial decision-making may be subject to implicit bias, and this reflects in their decisions of which charges to file, to enhance or reduce, or to drop altogether.¹⁰²

There is an unfortunate lack of research regarding prosecutorial decision-making during plea bargains and implicit bias.¹⁰³ However, the amount of discretion a prosecutor has when

98. See Avery & Cooper, *supra* note 82, at 269.

99. *Fogus v. United States*, 34 F.2d 97, 98 (4th Cir. 1929).

100. Robert J. Smith & Justin D. Levinson, *The Impact of Implicit Racial Bias on the Exercise of Prosecutorial Discretion*, 35 SEATTLE U. L. REV. 795, 798 (2012).

101. Michael Nasser Petegorsky, *Plea Bargaining in the Dark: The Duty to Disclose Exculpatory Brady Evidence During Plea Bargaining*, 81 FORDHAM L. REV. 3599, 3601 (2013) (citing Stephanos Bibas, *Incompetent Plea Bargaining and Extrajudicial Reforms*, 126 HARV. L. REV. 150, 151 (2012)).

102. *Research Finds Evidence of Racial Bias in Plea Deals*, EQUAL JUSTICE INITIATIVE (Oct. 26, 2017), <https://eji.org/news/research-finds-racial-disparities-in-plea-deals/>.

103. See Kang, *supra* note 16, at 1141 (“Unfortunately, we have very little data on this front. Indeed, we have no studies, as of yet, that look at prosecutors’ and defense attorneys’ implicit biases and attempt to correlate them with those individuals’ charging practices or plea bargains. Nor do we know as much as we would like about their implicit biases more generally.”).

determining bail, charges, and offering a plea deal invites associations infected by implicit bias.¹⁰⁴ Without oversight, implicit bias could run rampant during plea negotiations. As noted, the judge has the responsibility to ensure that a plea is accepted free from coercion and with knowledge and voluntariness; this responsibility should also extend to ensuring that implicit bias and its automatic responses are not infiltrating a prosecutor's plea offer to a defendant.¹⁰⁵

C. *Sentencing: The All-Star Proceeding Highlighting Implicit Bias*

Judicial impartiality is of utmost importance during the sentencing proceedings. "Some findings show that trial court judges 'rely extensively on intuition, more than deliberative judging, in deciding matters before the bench.'"¹⁰⁶ Scholars have observed that the fast-paced nature of the criminal courtroom creates the perfect storm for the influence of implicit biases: time pressure and quick decision-making.¹⁰⁷ Judges across America

104. *See id.* ("[T]here is no reason to presume attorney exceptionalism in terms of implicit biases. And if defense attorneys, who might be expected to be less biased than the population, show typical amounts of implicit bias, it would seem odd to presume that prosecutors would somehow be immune. If this is right, there is plenty of reason to be concerned about how these biases might play out in practice. As we explain in greater detail below, the conditions under which implicit biases translate most readily into discriminatory behavior are when people have wide discretion in making quick decisions with little accountability. Prosecutors function in just such environments.").

105. *Fogus v. United States*, 34 F.2d 97, 98 (4th Cir. 1929).

106. *See* Shawn C. Marsh, *The Lens of Implicit Bias*, UCONN SCH. OF L., <https://libguides.law.uconn.edu/implicit/courts> (last updated Feb. 16, 2021, 11:48 AM) (quoting Laura Connelly, *Cross-Racial Identifications: Solutions to the "They All Look Alike" Effect*, 21 MICH. J. OF RACE & L. 125 (2015)).

107. *See* L. Song Richardson, *Systemic Triage: Implicit Racial Bias in the Criminal Courtroom*, 126 YALE L.J. 862 (2017) (reviewing NICOLE GONZALEZ VAN

deal with these same hectic situations and are susceptible to blind spots and implicit bias. Justice Anthony Kennedy noted that “bias is easy to attribute to others and difficult to discern in oneself.”¹⁰⁸

Two IAT studies given to trial judges in conjunction with judicial sentencing showed the same or greater implicit racial biases as with the public.¹⁰⁹ While the Rachlinski study focused on the bias against African Americans, the Levinson, Bennett, and Hioki study focused on the sentencing biases against Jewish people and Asian Americans.¹¹⁰ The Levinson study randomly selected magistrates, district court judges, and state court judges from eight states.¹¹¹ The researchers found that the federal and state judges displayed strong to moderate implicit bias against Asian Americans as compared to White people on the stereotype IAT.¹¹² The team also discovered that federal and state judges exhibited strong to moderate implicit bias against Jewish people as compared to Christians on the stereotype IAT.¹¹³ Asian Americans and Jewish people were associated with negative moral stereotypes (i.e., greed, dishonesty, and scheming) and White and Christian people were associated with positive moral stereotypes (i.e., trustworthiness, honesty, and generosity).¹¹⁴

CLEVE, CROOK COUNTY: RACISM AND INJUSTICE IN AMERICA’S LARGEST CRIMINAL COURT (2016)).

108. See *Williams v. Pennsylvania*, 136 S. Ct. 1899, 1905 (2016).

109. Rachlinski, *supra* note 81, at 1210–11; Justin D. Levinson, Mark W. Bennett & Koichi Hioki, *Judging Implicit Bias: A National Empirical Study of Judicial Stereotypes*, 69 FLA. L. REV. 63 (2017).

110. *Id.*

111. *Id.*

112. *Id.* at 65–68.

113. *Id.*

114. *Id.*

The federal district judges gave longer sentences to Jewish defendants as opposed to the Christian defendants.¹¹⁵ However, the magistrate judges' sentences did not vary based on the defendant's group, and state judges sentenced White defendants to longer sentences than Asian American defendants. The Rachlinski study concluded that implicit biases among judges can influence their judgments.¹¹⁶ However, when judges are aware of these potential biases, they have the skill to avoid these biases when assessing sentences.¹¹⁷ "Awareness of implicit bias" and "doubting one's objectivity" are beneficial interventions to stop the spread of bias in sentencing.¹¹⁸

Factors such as skin tone can trigger an implicit bias response when a judge is sentencing a defendant.¹¹⁹ Other empirical studies suggest that skin tone, Afrocentric facial features, and sex can also trigger implicit bias in judges that result in longer sentencing.¹²⁰

Thus, it is not race alone, but Afrocentric features like darker skin tone, wider noses, coarser hair, darker eyes, and fuller lips that influence the length of a criminal sentence, because defendants with these characteristics are perceived as more likely displaying a Black stereotype of

115. *Id.*

116. Rachlinski, *supra* note 81, at 1225.

117. *Id.*

118. Richardson, *supra* note 107, at 887.

119. *Id.*

120. Mark W. Bennett, *The Implicit Racial Bias in Sentencing: The Next Frontier*, 126 YALE L.J. F. 391, 402–03 (2017) (citing Irene V. Blair et al., *The Influence of Afrocentric Facial Features in Criminal Sentencing*, 15 PSYCHOL. SCI. 674 (2004); Jill Viglione et al., *The Impact of Light Skin on Prison Time for Black Female Offenders*, 48 SOC. SCI. J. 250 (2011)).

aggressiveness, criminality, dangerousness, and recidivist law-breaking.¹²¹

As one researcher suggests, something as simple as removing the defendant's photograph from the initial sentencing report can help.¹²² Without awareness that Afrocentric features might be triggering a bias response, judges cannot control or correct the potential bias.¹²³

D. *Putting it All Together*

Throughout these studies conducted by scholars, social psychologists, and even other judges, one thing is clear—implicit bias is real. During the entirety of the criminal justice process, judges make decisions, and those decisions are vulnerable to implicit bias. Training on implicit bias can only prove beneficial by bringing awareness to a potential flaw in a judge's thinking.

One recommendation is that judges be educated not only on the presence of implicit bias, but the science behind it as well.¹²⁴ If judges do not recognize and understand implicit biases, the effects could be dire, even for a single defendant. A single defendant must go through policing, charging, bail, plea bargaining, pretrial motions, evidentiary hearings, determinations of witness credibility, guilt determinations, sentencing proceedings, and appeals.¹²⁵ Throughout this process, there is typically a single judge making the decisions. If those decisions are

121. Bennett, *supra* note 120, at 403.

122. *Id.* ("One of my suggestions in my training is to eliminate the photograph of the offender on the front page of the pre-sentence report. The photograph is a classic psychological prime that can easily trigger implicit bias in the judges' evaluation of the rest of the pre-sentence report.").

123. Kang, *supra* note 16, at 1150.

124. *Id.* at 1175.

125. See *Criminal Justice Process*, U.S. DEP'T OF JUSTICE, <https://www.justice.gov/enrd/criminal-justice-process> (last visited Mar. 31, 2021).

tainted by implicit bias from the beginning, the defendant's fate is sealed before making it to trial.¹²⁶ Judges must be aware throughout a criminal proceeding of when their own implicit bias is affecting their decision-making, and they should also be aware of techniques to counteract and mitigate against the automatic tendency to label certain persons in certain ways based on those biases.

Implicit-bias education alone was never intended to eliminate bias; instead, it was initially viewed as adding to a greater discussion surrounding race in a justice context.¹²⁷ Proponents of implicit-bias education articulate that people need to be aware of their biases through instruments, such as the IAT, and build critical steps to change behavior.¹²⁸ Implicit-bias training and education can not only generate substantial awareness on the issue but can also inspire "serious individual and system reflection as to how experiences, environment, culture, and system design can lead to biased decision making."¹²⁹

126. Kang, *supra* note 16, at 1151.

127. Marsh & Marsh, *supra* note 12, at 93.

128. *Id.*

129. *Id.*

TECHNIQUES TO COMBAT IMPLICIT BIAS

The first and most important step is to become more knowledgeable about bias in general. The IAT offers a way for judges to see biases in action and how they can become interwoven into their thoughts and decision-making process. Judges should also work to build a more detailed and complete narrative to better understand the entire issue before making decisions.¹³⁰ Implicit bias works its way into cases when judges must make inferences that serve as gap-fillers in an incomplete narrative.¹³¹ These additional facts could be key information that frames a scenario with experiences and perspectives a judge does not personally know.¹³²

Increased diversity in the judiciary has offered “heightened awareness” of the adversity faced by certain disadvantaged populations.¹³³ The American Bar Association recognizes that everyone has biases in some way, and judges are not immune.¹³⁴ Judges, as neutral arbiters and gatekeepers, must strive to make decisions without any cognitive shortcuts.¹³⁵

Mentorship is one technique that can help combat the effect of implicit bias. However, research regarding biases shows that stereotypes and assumptions can be harmful in mentoring due

130. Thornburg, *supra* note 49, at 1659–64.

131. *Id.*

132. *Id.* at 1661.

133. Hsu, *supra* note 50, at 108.

134. Karen Steinhauser, *Everyone Is a Little Bit Biased*, AM. BAR ASS'N (Mar. 16, 2020), https://www.americanbar.org/groups/business_law/publications/blt/2020/04/everyone-is-biased/.

135. Thornburg, *supra* note 49, at 1665.

to a phenomenon called “stereotype threat.”¹³⁶ Stereotype threat can occur when a group of people who regularly have a stereotype attributed to them suffer in performance because they *feel* as though others are using the stereotype against them. The pressure of worrying about being stereotyped itself actually creates enough mental “baggage” and negative feelings that individuals cannot focus as much energy on the tasks they need to perform.¹³⁷ This phenomenon can be seen not only in courtrooms, but classrooms and organizations across the world.

Yale Law School’s Cultural Cognition Project sought to study the impact of different backgrounds on judicial fact-finding.¹³⁸ Two sets of people were shown a video of a driver evading a police car.¹³⁹ Each group reached a different conclusion regarding the danger and fault of the suspect in the video.¹⁴⁰ Clearly, it is almost impossible to “allow the [evidence] to speak for itself” when there are so many different voices that can be heard.¹⁴¹

136. Christy Pettit, *Unconscious Bias in the Workplace: Managing Differences Through Mentoring*, POLLINATE (June 12, 2020), <https://pollinate.net/unconscious-bias-in-the-workplace-managing-differences-through-mentoring>.

137. *Id.*

138. Thornburg, *supra* note 49, at 1632.

139. See Dan M. Kahan et al., *Whose Eyes Are You Going to Believe?: Scott v. Harris and the Perils of Cognitive Illiberalism*, 122 HARV. L. REV. 837, 903 (2009) (detailing how the video came from the Supreme Court case *Scott v. Harris*, 550 U.S. 372 (2007), in which Justice Scalia and the majority believed there was only a single interpretation of the video).

140. *Id.*

141. *Scott v. Harris*, 550 U.S. 372, 378 n.5 (2007). See also Kahan, *supra* note 139, at 903 (“Whites and African Americans, high-wage earners and low-wage earners, Northeasterners and Southerners and Westerners, liberals and conservatives, Republicans and Democrats—all varied significantly in their perceptions of the risk that Harris posed, of the risk the police created by

Just two years after the decision in *Shelby County*, the Supreme Court took an important step by recognizing not only the existence but also the importance of implicit bias in *Texas Department of Housing and Community Affairs v. Inclusive Communities Project, Inc.*¹⁴² There, Justice Kennedy noted how the Fair Housing Act allowed “plaintiffs to counteract unconscious prejudices and disguised animus that escape easy classification as disparate treatment.”¹⁴³ With this he acknowledged we can continue our “historic commitment to creating an integrated society.”¹⁴⁴ Such a recognition is a crucial move toward creating a judiciary that benefits all members of our society. Techniques such as the IAT can further elucidate the biases that all judges contain and propel this work even further.¹⁴⁵

In 2012, Yale’s Horsley Laboratory conducted a large study on strategies that can help courts address implicit bias by surveying judges and judicial educators. Yale’s Laboratory emphasized that strategies used to combat implicit biases need to be concrete and applicable to an individual’s work to be truly effective.¹⁴⁶ In applying this logic to the judicial system, it is important to understand that although a majority of people may want to be fair in their judgment of others, they may nonetheless

deciding to pursue him, and of the need to use deadly force against Harris in the interest of reducing public risk.”).

142. 576 U.S. 519 (2015).

143. *Id.* at 521.

144. *Id.* at 546.

145. *U.S. Supreme Court Recognizes Role of Unconscious Bias in Disparate Treatment*, ASS’N FOR PSYCHOL. SCI. (July 1, 2015), <https://www.psychologicalscience.org/news/releases/us-supreme-court-recognizes-role-of-unconscious-bias-in-disparate-treatment.html>.

146. *Helping Courts Address Implicit Bias: Strategies to Reduce the Influence of Implicit Bias*, NAT’L CTR. FOR ST. CTS., (2012), https://horsley.yale.edu/sites/default/files/files/IB_Strategies_033012.pdf.

lack concrete and applicable strategies to counteract ways in which they are *not* fair and impartial.

Yale cites various triggers that can cause judicial professionals to rely on implicit bias rather than making more consciously cognitive decisions. These triggers include ambiguity, salient social categories, and lack of feedback. There are still underdeveloped areas of the law that call for ambiguity in a judge's decision-making, and when there is vagueness, there is a potential for biased judgments. Without more explicit, concrete criteria for decision-making, individuals tend to disambiguate the situation using whatever information is most easily accessible, including stereotypes.¹⁴⁷ The social categories in which people are placed come from a variety of influences, such as television, literature, and news reports. Yale emphasizes that by requiring judges, jurors, and court staff to become aware of easily placed stereotypes, they can correct their thoughts before making a decision infected by bias.¹⁴⁸ Lastly, providing periodic feedback to decision makers increases accountability. When organizations fail to provide feedback that holds decision makers accountable for their judgments and actions, individuals are less likely to remain vigilant for possible bias in their own decision-making processes.¹⁴⁹ People struggle to hold themselves accountable or change their own behavior if they receive little to no feedback. The judiciary and other legal organizations can preemptively combat negative effects of implicit bias by instituting periodic feedback sessions for employees.

When studying a foreign language, it is often said that the best way to learn is to expose oneself to the culture to understand what makes it unique; the same logic applies when

147. *Id.*

148. *Id.*

149. *Id.*

implementing techniques to combat implicit biases. To expose oneself in the educational sense means to become immersed in the topic. When people expose themselves to the first-person perspective of others, it can create a true impact in their understanding and treatment of others. Walking a mile in another's shoes can be as easy as taking the initiative to learn another person's life perspective. Gaining this exposure can come from spending time with groups of people outside of our own in-groups or immersing oneself in media (from movies to documentaries to virtual conferences) that allows the viewer to understand a different culture or point of view. We will not solve the negative effects of implicit bias overnight; rather, it will take years of increasing awareness, providing training and education, and enacting piecemeal changes that each solve one piece of the implicit bias puzzle.

At the 2015 Annual Meeting of the National Association of Bar Executives, Sharon E. Jones of Jones Diversity remarked that "you can disrupt your automatic pilot — which can lead you to act on your biases even if you do not intend to[.]"¹⁵⁰ What remains for us to do is understand more specific ways that we can repel these biases. According to Jones, microaggressions can slip into language, images, and daily habits when we do not intend them to, but by implementing and encouraging implicit-bias training and awareness and its effect as a dialogue within the legal profession, the level of accountability and awareness will rise, and when accountability and awareness rise, the negative effects of implicit bias in our legal system will fall.¹⁵¹

150. Marilyn Cavicchia, *How to Fight Implicit Bias? With conscious thought, diversity expert tells NABE*, AM. BAR ASS'N, https://www.americanbar.org/groups/bar_services/publications/bar_leader/2015-16/september-october/how-fight-implicit-bias-conscious-thought-diversity-expert-tells-nabe/ (last visited Mar. 30, 2021).

151. *Id.*



**MOVING THE LAW FORWARD
IN A REASONED & JUST WAY**

Copyright 2021, The Sedona Conference
All Rights Reserved.
Visit www.thesedonaconference.org