

The Sedona Conference Journal

Volume 26, Number 2

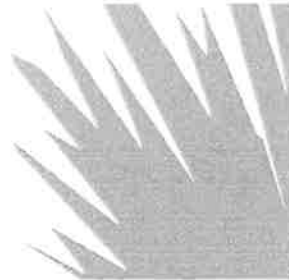
Forthcoming 2025

Commentary on Discovery of Collaboration Platforms Data

The Sedona Conference

October 2025

Final Post-Public-Comment Version



Recommended Citation:

The Sedona Conference, *Commentary on Discovery of Collaboration Platforms Data*,
26 SEDONA CONF. J. 627 (forthcoming 2025).

For this and additional publications see: <https://thesedonaconference.org/publications>.

THE SEDONA CONFERENCE COMMENTARY ON DISCOVERY OF COLLABORATION PLATFORMS DATA

*A Project of The Sedona Conference Working Group (WG1) on
eDiscovery*

Author

The Sedona Conference

Drafting Team Members

Stacey Blaustein	Michelle Briggs
Douglas Forrest	Adam Gajadharsingh
Hon. Jane Manning	Benson McGrath
Derek McNally	Jonathan Orent
Jonathan Swerdloff	Cristin Traylor

Drafting Team Leaders

Gareth T. Evans	Joseph P. Guglielmo
-----------------	---------------------

Steering Committee Liaisons

Tara Emory	Meghan Podolny
Maria Salacuse	

Staff Editor

Craig Morgan

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

Copyright 2025, The Sedona Conference.
All Rights Reserved.

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 Discovery of Collaboration Platforms Data*, 26 SEDONA CONF. J. 627 (forthcoming 2025).

PREFACE

Welcome to the October 2025 final version of The Sedona Conference's *Commentary on Discovery of Collaboration Platforms Data*, 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.

The mission of WG1, formed in 2003, is "to develop principles, guidance and best practice recommendations for information governance and electronic discovery in the context of litigation, dispute resolution, and investigations." The Working Group consists of members representing all stakeholders in eDiscovery and Information Governance.

In May 2022, the WG1 Steering Committee issued a "Call for Volunteers" for a Brainstorming Group on discovery of collaboration platforms. On July 19, 2022, the Brainstorming Group held their first meeting, and on October 27, 2022, the Brainstorming Group presented their outline for dialogue at the WG1 2022 Annual Meeting. Following the Annual Meeting, a drafting team was formed, and an initial draft was presented at the 2023 WG1 Annual Meeting. The draft *Commentary* was revised and expanded and submitted for dialogue at the WG1 2024 Midyear Meeting. A final draft was submitted to WG1 Steering Committee in September 2024. Additional member comments and edits were received and incorporated into a public comment draft published in April 2025. Those public

comments were subsequently reviewed and incorporated into this final version.

This *Commentary* represents the collective efforts of many individual contributors. On behalf of The Sedona Conference, I thank in particular the drafting team leaders Gareth T. Evans and Joseph P. Guglielmo. I also thank the drafting team members for their time and attention during this extensive drafting and editing process, including Stacey Blaustein, Michelle Briggs, Adam Gajadharsingh, Douglas Forrest, Hon. Jane Manning, Benson McGrath, Derek McNally, Jonathan Orent, Jonathan Swerdloff, and Cristin Traylor.

The drafting process for this *Commentary* has also been supported by the Working Group 1 Steering Committee, with Steering Committee members Tara Emory (now with WG13), Meghan Podolny, and Maria Salacuse acting as liaisons to the drafting team. I also thank the many contributors of public comments, which the drafting team found to be particularly helpful in improving this *Commentary*. 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.

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 artificial intelligence, 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.

2025] COMMENTARY ON COLLABORATION PLATFORMS DATA 631

Kenneth J. Withers
Executive Director
The Sedona Conference
October 2025

TABLE OF CONTENTS

I.	INTRODUCTION.....	633
II.	DISCUSSION	635
	A. Characteristics of Collaboration Platforms	635
	B. Common Electronic Discovery Issues	640
	1. Treatment as a Custodial or Noncustodial Data Source	640
	2. Preservation Challenges	643
	3. Family Relationships and Hyperlinks.....	650
	4. Collection Challenges	664
	5. Culling, Search, Review, and Production Challenges	667
	6. Evidentiary, Privilege, and Privacy Issues	676
	C. Information Governance Considerations	683
	D. GenAI Considerations	687
III.	CONCLUSION	688

I. INTRODUCTION

This *Commentary* is intended to provide organizations, lawyers, and the judiciary with foundational information regarding collaboration platforms, and to provide guidance on addressing legal issues and challenges that are likely to arise regarding the discovery of collaboration platform-related electronically stored information (ESI). The use of collaboration platforms has revolutionized the way organizations communicate and collaborate. They enable individuals and groups to work together, to share information, to communicate, and to coordinate tasks seamlessly. They also encompass a wide range of applications, including messaging, document sharing, and video conferencing. Examples at the time of publication of this *Commentary* include Slack, Microsoft Teams, and applications in Google Workspace.

Although communication and collaboration platforms had been around for years, and their use had steadily grown, the global COVID-19 pandemic of the early 2020s—which caused many workers to work from home—was accompanied by an explosion in their use. As collaboration platforms have become a primary means of communicating and working, they have increasingly become a source for discovery. As a source of relevant ESI, however, they can pose conceptual, technical, and practical challenges, some of which are highlighted below.

Similar to non-custodial sources such as databases, the traditional notion of a custodian—i.e., someone who has possession, custody or control over ESI—may be inapplicable to a group workspace on a collaboration platform, even though members of the group may be performing certain activities that can be associated with a specific custodian (such as messaging, as well as sharing and storing documents). Thus, information created, shared, or maintained on a collaboration

platform may not necessarily relate to or be identified with a traditional custodian, unlike the common conception of ESI collection, which revolved around the identification and collection of custodial information.

Identifying relevant shared workspaces on collaboration platforms for preservation can present challenges. Such shared workspaces can proliferate, and an organization may not possess a comprehensive list of such workspaces or their participants to readily identify or assess their relevance to litigation or an investigation. Additionally, private channels may exist within a shared workspace and, apart from administrators, only those who are members may have access to them or be aware of their existence. Furthermore, communications and data on collaboration platforms are often stored in various other applications outside of the platform. In Microsoft Teams, for example, documents and messages are often saved to SharePoint sites or to applications such as OneDrive. Thus, careful consideration should be given to identifying information for preservation.

Another feature of collaboration platforms is that documents may be shared by means of hyperlinks rather than as traditional attachments. Hyperlinked content may be located within the collaboration platform's environment or outside of the platform's environment, but within the organization's information systems. It may also be located outside of the organization's information system entirely. As the hyperlinked content is often stored outside of the collaboration platform, and may change over time, finding it and re-associating the specific version of the content to a contemporaneous communication may involve complications that are not present with traditional email and attachments.

Collaboration platforms also currently pose challenges for collection and processing in electronic discovery—e.g., locating relevant communications for collection, identifying hyperlinked documents, and associating the specific version of the hyperlinked documents (if versions even exist) with the contemporaneous message. Collaboration platforms may require different types of culling mechanisms, search strategies, review efforts, and production formats. For example, considerations may include whether to consolidate individual communications from a Slack channel into larger units for manageability purposes, such as a 24-hour period, or a particular number of chats.

Collaboration platforms may also present unique evidentiary issues. Because collaboration platforms may allow multiple users to view and edit a document simultaneously, authentication may be more complex. Additionally, information governance can play a significant role in facilitating organizations being able to fulfill their electronic discovery obligations in the event of litigation or a governmental investigation.

In the sections that follow, we discuss these and many other issues that arise in electronic discovery in connection with collaboration platforms, as well as explore potential solutions to these issues.

II. DISCUSSION

A. Characteristics of Collaboration Platforms

A collaboration platform is an application that provides multiple users with the ability to share information and to work collaboratively and often contemporaneously. Collaboration platforms may include the ability to draft, review, edit, and comment on work product simultaneously; to share information, documents, links, files, videos, and audio content; to

have live audio and video conferences; and to create and access recordings and transcripts of the same. The methods of communication in a collaboration platform's environment may include email, chat, video and audio conferencing, wikis, blogs, and social media, as well as comments on written work product. Examples of currently available collaboration applications include Microsoft Teams, Slack, Google Workspace (containing Google Drive, Google Chat, and Google Meet, all of which have collaboration features), Asana, and Trello.

Collaboration platforms may integrate email, text messaging, calendar, and document management tools and may be integrated with existing or legacy stand-alone software. Information created in or accessed through a collaboration platform is not necessarily stored in the platform itself. Rather, the collaboration platform may access and store the data through, or in, another application, so consideration should be given to where the data is stored and how it is accessed.

Discussions in this *Commentary* of features of specific collaboration platforms (e.g., Microsoft Teams, Slack, Google Workspace), as well as tools used to collect and to process their data, are based on the applications' characteristics at the time of publication, which may of course change over time. Similarly, the capabilities of particular platforms may depend on the type of license acquired. Such features, however, are discussed as examples illustrating more general points about the impacts that the settings, structure, and other characteristics of a collaboration platform may have on discovery from this type of data source.

Modern collaboration platforms generally have the following characteristics that can be used to characterize and understand them:

Not every document is a document: A feature of many collaboration platforms is that the content of a document can be stored in a different format or in multiple discrete parts than what is seen by a user. The content of what appears to the user to be a document may in fact be structured data.¹ Communications in a collaboration platform, such as chats, also may not fit neatly within traditional conceptions of a document, as different subjects may be discussed, dropped, and picked up again in a single thread, and individual messages within a thread may also be considered a document. In some instances, a document's history may be stored in time-stamped name/value pairs that are rendered viewable by a user in real time by the platform. Each collaboration platform has its own unique process for storing the data created in the platform.

The notion of a custodian may not be applicable: Electronic discovery processes often revolve around identifying custodians who are most likely to have information pertinent to a matter and then preserving and collecting relevant information in their possession, custody, or control. While custodians may participate in activities on collaboration platforms for activities that have commonly been associated with individual custodians, e.g., creating and editing documents, these platforms are—much like a database that a custodian may access—traditionally considered a noncustodial source of information. Thus, consideration should be given to how the ESI in a collaboration platform is accessed and used, and by whom, as the traditional concept of custodial-based ESI may not apply.

1. Structured data is “[d]ata stored in a structured format, such as databases or data sets according to specific form and content rules as defined by each field of the database.” See *The Sedona Conference Glossary: eDiscovery & Digital Information Management, Fifth Edition*, 21 SEDONA CONF. J. 263, 375 (2020).

Collaboration platforms typically provide shared workspaces and applications that teams or cohorts of users may utilize. Litigants may need to develop strategies for identifying, preserving, and collecting relevant data specific to each tool implemented in or accessed through the collaboration platforms.

Ability to communicate within documents: Collaboration platforms and the applications they incorporate often allow users to make notes and comments within documents, essentially having a conversation within them, much like word processing software allows for comments and responses to those comments. Litigants will need to consider and understand how these types of comments are stored and will be processed and reviewed. Some platforms may allow extraction and presentation of the data in separate fields, as extracted text, or to show the comments within the document itself. It can be important to understand where this data is stored, viewed, and produced, as that may impact issues such as privilege and confidentiality. If comments are present in the native, but not in the production images, or vice versa, that could cause issues for both the receiving and producing parties. If there is text that needs redaction, producing parties will need to know all the locations where that data resides so it can be redacted consistently. These considerations are unique to communications within documents on collaboration platforms.

Hyperlinked documents in lieu of attachments: Hyperlinked documents are not stored in the same way as traditional static attachments, and hyperlinks may be used instead of attachments in communications in collaboration applications. Traditional attachments and their parent emails are stored within a container such as a .pst file or an email server database that contains the parent emails with their attachments. A hyperlink sends a user to another location either within the current document or to another location on the network.

Documents referenced by hyperlinks generally are stored in a different location from the communications or documents in which the hyperlinks appear. The hyperlinked documents may reside in applications or sources outside of the collaboration application, but they are integrated into the collaboration platform (e.g., SharePoint Online or OneDrive). They may also reside entirely outside of an organization's information system. Because hyperlinked documents are shared and can be updated by one or more users, they may change over time as users revise them. Thus, it is important to understand how the hyperlinked document, and its versions (if any), were maintained, because the hyperlink may no longer point to the document as it existed at the time it was shared, or it may no longer be available.

Some systems may be configured to retain only the most current version of a document without any history of the prior iterations while other systems may maintain every version of a document to which a hyperlink points. Thus, it is important to understand the system, its settings, and its capabilities when discovery involves hyperlinks.

Portals to other applications: Collaboration platforms may appear to a user as a single, consolidated application providing various features, but they often operate as portals to one or more other applications. For example, users collaborating on drafting a document in a collaboration platform may be working on and storing the document in a separate application that the collaboration platform merely accesses without the users realizing that they are actually using the separate application, e.g., Dropbox or OneDrive. Similarly, instant messages, meeting notes, and calendar entries that appear to a user to be stored within a collaboration platform may actually be stored outside of the collaboration platform—for example in a cloud storage location or within a different application. Thus, it is

important to understand the various features of a collaboration platform and where ESI from those features may reside.

B. Common Electronic Discovery Issues

1. Treatment as a Custodial or Noncustodial Data Source

Collaboration platforms typically provide shared spaces and embedded applications or functionalities that teams or groups of users may utilize. While custodians may access and participate in shared spaces (e.g., teams or channels) on collaboration platforms, the platforms themselves are usually considered to be noncustodial sources of information. As stated in *The Sedona Principles*, “An organization’s networks or intranet may contain shared areas (such as public folders, discussion databases, and shared network folders) that are not regarded as belonging to any specific employee. Similarly, there may be no one ‘owner’ of the ESI for collaborative workspace areas within the organization.”²

Nevertheless, a custodian may individually use—i.e., not as part of a shared space—certain functionalities of a collaboration application. For example, it is common for users to utilize the chat function of some collaboration applications individually, much like an email account. In that circumstance, it may be appropriate to treat a custodian’s chats as a custodial source of information. Treating a collaboration platform as a custodial

2. The Sedona Conference, *The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 SEDONA CONF. J. 1, 116 (2018). See also *id.* at 103 (organizations “should assess the persons likely to have relevant information, and the sources of non-custodial relevant information . . . such as structured systems and databases, and other non-custodial sources such as collaboration tools [and] social media”).

or non-custodial source may impact the scope of preservation or collection of the platform's data.

Although a custodian's access to or use of a collaboration platform alone does not necessarily mean that all or a portion of the data in the platform must be preserved or searched, consideration should be given to the use of the platform and the information that may be collected and preserved. This will allow an organization to identify and understand what steps can be taken to preserve, search, and produce that information if it is later determined to be relevant in discovery. Custodians may be able to assist in identifying the existence of shared spaces in collaboration platforms that are relevant—i.e., those involving work product and communications relevant to the issues—because of their involvement in the matters underlying the litigation or investigation.

Understanding how a collaboration platform is being used may be helpful in identifying custodians and other relevant information. For example, most collaboration platforms allow users to collaborate on documents and to communicate in groups or one-on-one (e.g., Teams); some platforms are primarily communication tools (e.g., Discord); while still others are communication and file-sharing tools that integrate other cloud-based document management systems (e.g., Slack).³ Knowing at the outset the capabilities and configuration of a given platform will help to determine the scope of potentially relevant information.⁴

3. See, e.g., *Hayvin Gaming, LLC v. Workinman Interactive, LLC*, No. 23-CV-6172-FPG, 2023 WL 3748844, at *2 n.2 (W.D.N.Y. June 1, 2023) (describing Slack as “a collaboration tool that allows teams to message, share files, and archive information”).

4. An example of a specific functionality of a collaboration platform can be found in *In re Google Play Store Antitrust Litig.*, 664 F. Supp. 3d 981, 985

Once a relevant shared space within a collaboration platform is identified, it may be possible to identify other users (or collaborators) involved in the underlying matter who may be considered potential custodians or witnesses. Identifying those who actively participated in a relevant shared space can be particularly helpful in the early stages of litigation when the organization is investigating the issues to understand the facts and to determine relevant custodians and sources of relevant information within (or accessed through) the collaboration platform.

Additional potential custodians may be identified by reviewing who collaborated on a project or document, or who communicated in chats and channels, and the substance of those communications. Be aware that participants in chats may use pseudonyms (i.e., handles) instead of their true names, which may hide or obscure their identities.

In determining preservation and collection obligations, consideration should be given to who had access rights to a group or shared space on the collaboration platform, including whether a specific person was an active or passive participant. For example, an individual could have had the right to make edits to a document but have not made any. Likewise, an individual could have been a member of a group chat channel but not communicated on it. Moreover, it is possible an individual may have been given certain access rights without even knowing they were provided. By contrast, if an issue in the case relates to the user's state of mind or knowledge, for example, having access to certain documents or communications may be

(N.D. Cal. 2023) (Google Chat's default retention period for one-on-one chats with "history" turned off is 24 hours).

relevant.⁵ Collaboration platforms also may store documents in ways that have their own value as compilations, independent of whether the documents are duplicative of custodial productions.⁶

Finally, collaboration platforms often include applications or features that may be used outside of shared spaces. Chats, video calls, and audio calls may be made inside or outside of a shared space, and there may be ESI related to those calls (such as the date, time, topic, participants, a recording, and even an automatically generated transcript or summary of the call). Thus, consideration should be given to those other sources of information, how they are maintained, who maintains custody or control of the information, and whether they can be collected and produced.

2. Preservation Challenges

The principle that litigants have a duty to preserve discoverable information upon the commencement or reasonable anticipation of litigation or a governmental investigation applies to collaboration platforms, just as it applies to other sources of

5. See, e.g., *In re Uber Techs., Inc., Passenger Sexual Assault Litig.*, No. 23-md-03084-CRB (LJC), 2024 WL 1772832, at *2 (N.D. Cal. Apr. 23, 2024) (“contemporaneous versions of hyperlinked documents can support an inference regarding “who knew what, when.”).

6. See *In re Pfizer Inc. Sec. Litig.*, 288 F.R.D. 297, 317 (S.D.N.Y. 2013) (the defendant used a collaboration platform called “eRooms,” which allowed employees to share documents, exchange instant messages, and conduct informal polls. The court recognized that “[a]lthough the eRooms contain documents that may be largely duplicative of the custodial productions, they have a value in of themselves as compilations. The manner in which Pfizer and its employees internally organized documents is relevant because it allows Plaintiffs to draw connections and understand the narrative of events in a way not necessarily afforded by a custodial production.”).

ESI. As observed in the *Commentary on Legal Holds*, while the principle “is easy to state,” its “application in practice, however, often requires careful analysis and difficult decisions.”⁷ “Nonetheless, each day, organizations must apply the principle to real-world circumstances, first confronting the issue of whether an obligation is triggered, and then determining the scope of their obligation.”⁸

The trigger of the duty to preserve, of course, is no different with respect to collaboration platforms than as to other sources.⁹ Once a duty to preserve is triggered, litigants must take reasonable and proportional steps to preserve ESI that is subject to that duty.¹⁰

a. Determining the Scope of Preservation

The unique attributes of collaboration platforms should be considered when determining the scope of the duty to preserve, as these platforms may involve complexities associated with the multiple functionalities of these systems and the varying types of information that may be stored and collected. As an initial matter, in scoping a legal hold, one should determine

7. The Sedona Conference, *Commentary on Legal Holds, Second Edition: The Trigger & The Process*, 20 SEDONA CONF. J. 341, 351 (2019) (hereinafter *Commentary on Legal Holds*).

8. *Id.*

9. See *id.* at 354.

10. See FED. R. CIV. P. 37(e) (sanctions available only if relevant information was lost or destroyed because party failed to take reasonable steps to preserve it); see also The Sedona Conference, *Commentary on Legal Holds*, 20 SEDONA CONF. J. at 355 (“The proportionality principle applies to all efforts to plan and implement preservation, and in the assessment of those efforts.”), citing FED. R. CIV. P. 37(e) advisory committee’s note to 2015 amendment (a “factor in evaluating the reasonableness of preservation efforts is proportionality.”).

whether an organization uses one or more collaboration platforms and assess whether the information created in the platforms may be relevant to the matter. Then, one should determine whether the organization or the platform vendor maintains the data that is created by the users of the platform, as well as the forms in which and the locations where such information may be preserved.¹¹ Although not exhaustive, a good place to start to identify and determine what collaboration platforms are used by an organization and if they may possess relevant information is often with in-house counsel, IT personnel who have knowledge of an organization's information systems, and individuals who worked on or were associated with a specific project or program relevant to the litigation.

Additionally, investigating whether there may be relevant information on a collaboration platform may include questioning the use of the platforms in interviews of key employees or in questionnaires sent to potential custodians. These individuals may be asked to identify not only the collaboration platforms in use but also groups or shared spaces in the

11. See, e.g., *FTC v. Am. Future Sys., Inc.*, No. 2:20-CV-02266-JHS, 2023 WL 3559899, at *2–4, *7–8 (E.D. Pa. Mar. 28, 2023), adopted in part, modified in part by *FTC v. Am. Future Sys., Inc.*, No. 2:20-CV-02266-JHS, 2023 WL 355319 (E.D. Pa. May 17, 2023) (Defendant's employees used Slack for messaging and file sharing. Defendant, however, argued that it had no possession, custody, or control over its Slack data because the data was held by Slack, and that a court order to Slack would be required to export Slack data. Based on Slack's Customer Terms of Service agreement, which provided that the customer owns all its data held by Slack, the special master disagreed and ordered the production of Slack ESI. Although the special master rejected sanctions, he described defendant's failure to produce Slack data, failure to consider Slack as a potential source of relevant data, and failure to even reach out to Slack to determine what communications might exist as "troubling.").

collaboration platforms that may pertain to the events underlying the matter or that are otherwise relevant. Identifying relevant shared spaces or groups may lead to identifying other users with knowledge of relevant facts, who can then be asked if there are other shared spaces or groups that may be relevant to the matter.

Custodians may also confirm whether they may have used certain features of a collaboration platform either within or outside of a shared space. For example, the chat feature of a collaboration platform may be used in shared spaces, but it also may be used as custodians' primary chat application separate and apart from any group or channel. The chat function on the Microsoft Teams platform, for example, may serve as individuals' primary (or only) work chat application. Accordingly, consideration should be given to whether custodians' chats in a collaboration platform outside of shared spaces should be preserved, much like custodians' email accounts are often put on legal hold.

Custodian interviews and questionnaires and discussions with in-house counsel and IT personnel may not be sufficient alone to identify all relevant sources of information on a collaboration platform. For example, private channels may exist within a shared space that those who are not members of the private channel may not even be able to see. It may be helpful to employ other search methodologies, such as searching lists or indices of names of shared spaces (i.e., teams or channels) available to the IT department and sampling the content of spaces whose names appear to be potentially relevant to the issues in the litigation or investigation.

b. Determining the Nature and Sources of Relevant Collaboration Platform Data

Once a relevant shared space (e.g., a team or channel) is identified, it is important to determine how and in what form different types of communications and data are stored and can be preserved. It may be the case that the collaboration tool may limit the information that can be collected or preserved.¹²

Often, communications and other ESI are stored in various other applications outside of the platform. In Teams, for example, files uploaded into a channel may be stored in a folder in SharePoint. Files uploaded into a one-on-one or group chat may be stored in the Microsoft Teams Chat Files in a OneDrive for Business folder. Teams meeting recordings may be stored in either OneDrive or SharePoint. It will therefore be necessary to determine whether the files can be preserved in place in the locations where they are kept in the ordinary course of business, or whether they should be collected for preservation purposes. It is also important to recognize that where and how these platforms store data is constantly evolving and changing, so it is necessary to have people involved in the preservation process who are up to date on the technology.

Some collaboration platforms also contain dynamic documents, the access and control of which can vary by audience. A user may not see, know, or appreciate the full nature of the documents with which they are working. For example, certain users could see all information in a document, while others with less extensive permissions could see less. Special tools may be

12. See The Sedona Conference, *Commentary on ESI Evidence & Admissibility*, Second Edition, 22 SEDONA CONF. J. 91 at 139–140, nn.146–47 (2021) (discussing that a user’s ability to export certain information from Slack depends on the specific plan they purchase).

required to preserve the ESI and metadata required to appropriately generate these documents for use in the matter.

Additionally, in communications (such as chats) in collaboration applications, it is common for files to be referenced through hyperlinks to other sources, either within or outside of the organization's information system, rather than uploaded to the platform. It can be important, therefore, to consider the likely locations of hyperlinked documents, such as a document management system, and whether existing records retention in those locations will be sufficient for preservation purposes.

c. Retention Periods of Collaboration Platform Data
and Other Preservation Considerations

It is important to know the retention periods and location of collaboration platform data. That information can help guide how quickly targeted preservation measures must be implemented, whether relevant data can be preserved in place, or whether preservation must be effectuated through collection of the data. Thus, the retention settings for data on a collaboration platform may be relevant to the discovery being sought.¹³

13. See, e.g., *Drips Holdings, LLC v. Teledrip LLC*, No. 5:19-CV-2789, 2022 WL 4545233, at *3–5 (N.D. Ohio Sept. 29, 2022) (Defendant Teledrip, Inc. used Slack as a usual means of internal and external communication. Teledrip was found to have adjusted the retention settings of the company's Slack environment after learning of potential litigation. In particular, it changed the retention period from unlimited retention to only retaining data for seven days. It also deleted a prior export of Slack data. Teledrip argued that the change, despite the timing, was related to their understanding of the requirements of the California Consumer Privacy Act ("CCPA") and the associated International Standards Organization ("ISO") implementation. The court was not persuaded. Not only had Teledrip altered the retention period and destroyed a prior export, it also failed to update its retention settings to accommodate the legal hold, citing to the CCPA issue.

It is possible that some documents that relate to a collaboration platform may be downloaded to a user's account or personal computer—for example, where a user works on a local copy of a file. Accordingly, a custodian's local hard drive may contain unique copies of relevant information from a collaboration platform.

Further, in seeking to identify relevant ESI within a collaboration platform so that it may be considered for preservation, it is important to understand the collaboration platform's functionalities and the information that it or users may generate.¹⁴ For example, organizations are increasingly using the video and audio meeting features of collaboration platforms for online meetings, and the capabilities of those features are rapidly developing. Depending on the platform and the license level an organization purchases, it is now possible to obtain a great deal of information about a meeting within a collaboration platform. That information may include a complete recording of the meeting; a list of participants; a record of when participants entered and left the meeting; an AI-generated list of action items; an AI-generated transcript or summary of what

After rejecting the CCPA argument and characterizing certain assertions by the defendants as "blatantly false," the court ordered a mandatory adverse inference jury instruction as a sanction.). *See also* Adler v. McNeil Consultants, LLC, No. 3:19-CV-2025-K-BN, 2023 WL 2699511 (N.D. Tex. Feb. 15, 2023) (permissive adverse inference sanction imposed where defendants failed to suspend nine-day retention setting on Slack).

14. *See, e.g.,* Cisco Sys. v Chung, No. 19-CV-07562-PJH, 2023 WL 2622155, at *4–5 (N.D. Cal. Mar. 22, 2023) (discussing the defendant's failure to preserve SharePoint audit logs that would have demonstrated whether employees viewed alleged trade secrets, but declining to impose sanctions because evidence of whether the trade secrets were used—rather than just viewed—was relevant to the trade secrets claim and there was therefore no showing of relevance or prejudice).

was said and who said it; identification of documents or hyperlinks shared during the meeting; and a record of all the chat communications connected to the meeting.

Those capabilities are available at the time of writing of this publication and are likely to develop further over time. In the future, there are also likely to be other powerful capabilities that may generate relevant information potentially subject to a duty to preserve. It is important to be aware of the features of each collaboration platform to ensure that the organization has taken and will take reasonable steps to preserve relevant information related to the platform and its use.

3. Family Relationships and Hyperlinks

a. Location of Hyperlinked Documents

One of the challenges of dealing with collaboration platforms in electronic discovery is that documents and information are often shared by hyperlinks instead of as traditional attachments.¹⁵ With email, traditional attachments may be part of a .msg container file along with the parent email. The attachment is a static document from a specific point in time. Each

15. Various terms have been proposed for hyperlinks and the content to which hyperlinks point. "Pointers" and "modern attachments" are two examples, respectively, and they have engendered a fair amount of controversy. This *Commentary* uses the term "hyperlinks," as that term is defined in *The Sedona Conference Glossary* ("Hyperlink: A pointer in a hypertext document—usually appearing as an underlined or highlighted word or picture—that, upon selection, sends a user to another location either within the current document or to another location accessible on the network or internet."). See The Sedona Conference, *The Sedona Conference Glossary: Electronic discovery & Digital Information Management*, Fifth Edition, 21 SEDONA CONF. J. 263, 319 (2020). This *Commentary* also uses the terms "hyperlinked documents" (or files) and "referenced documents" (or files) for the documents or files to which the hyperlinks point.

email and its attachments are stored together and can be readily associated in document collection, review, and production.

Collaboration platforms and email applications using hyperlinks, by contrast, often reference content stored elsewhere, and the referenced content is generally not transferred with or stored with the message.¹⁶ At the time of publication of this *Commentary*, even traditional email applications are moving away from the paradigm of traditional attachments. Cloud-based email, such as email in Microsoft 365 and in Google Workspace, may utilize hyperlinks to content rather than attachments stored with the email message in a .msg file or another container file such as a .pst or .edb file.¹⁷

Discovery of hyperlinked documents can involve unique challenges and burdens. Hyperlinked content may be located within the collaboration platform's environment, such as in SharePoint or OneDrive in the Microsoft Teams environment. It may also be located outside the collaboration platform's environment, but within the organization's information system, such as a document management system (for example,

16. See, e.g., *Nichols v. Noom Inc.*, No. 20-CV-3677 (LGS) (KHP), 2021 WL 948646, *4 (S.D.N.Y. Mar. 11, 2021) (stating that the court did "not agree that a hyperlinked document is an attachment[.]"); *In re Insulin Pricing Litig.*, No. 23-md-3080, 2024 WL 2808083, at *7 (D.N.J. May 28, 2024) ("hyperlinks are not the same as traditional attachments"). But see *In re Uber Techs., Inc. Passenger Sexual Assault Litig.*, 2024 WL 1772832 (N.D. Cal. Apr. 23, 2024) (ordering the definition of attachments in the ESI protocol to include hyperlinked documents).

17. See, e.g., *Shenwick v. Twitter*, No. 16-CV-05314 (JST)(SK), 2018 WL 5735176, at *1 (N.D. Cal. Sept. 17, 2018) (observing that, rather than traditional attachments, emails in GSuite use hyperlinks to reference documents stored in the GSuite environment, "and sender and recipient can then modify the referenced document, which is stored centrally so that more than one person can access it.").

iManage). Hyperlinked content may also be located outside of an organization's information system entirely, such as hyperlinked content on a website. As the hyperlinked files are stored in separate locations and may change over time (including no longer being present in the target location, i.e., a broken link), following a hyperlink to the hyperlinked document and reassociating the contemporaneous version (if versions are available) may involve considerations and complications that are not present with traditional email and attachments.¹⁸ Additionally, hyperlinks can point to specific documents, but they can also point to an entire folder of documents, only some of which may be relevant.

Accordingly, traditional notions of document families may not necessarily apply to communications containing hyperlinks. Parties should take care not to reflexively treat hyperlinked documents in ESI protocols as the same as traditional attachments. Parties should investigate the specific facts associated with the collaboration platform's technology and any hyperlinked documents. Parties should consider the differences and similarities between hyperlinks and traditional attachments and should understand how such information was communicated, stored, and maintained and what limitations may exist to producing hyperlinked information in the particular IT environment and circumstances of their case. Doing so will

18. See, e.g., *Porter v. Equinox Holdings, Inc.*, No. RG19009052, 2022 WL 887242, at *2 (Cal. Super. Mar. 17, 2022) (“[L]inked documents can present unique challenges that make them different from email attachments.”); *In re Uber Techs., Inc. Passenger Sexual Assault Litig.*, 2024 WL 1772832, at *2 (“In sum, the briefing and evidence, as well as related case law, have made clear that cloud computing and document retention through Google Drive and Google Vault introduce a host of challenges to producing hyperlinked documents from Google Drive and other sources”) (citing *Nichols v. Noom Inc.*, 2021 WL 948646, at *2).

allow the parties to have a solid understanding and realistic expectations in attempting to reach informed agreements regarding the discoverability and production of hyperlinked information.¹⁹

b. Relevance and Proportionality of Hyperlinked Documents

Hyperlinked content may not always be independently relevant to the issues. Thus, requests for “all” hyperlinked content should be avoided and should be tailored to obtaining information relevant and proportional to the matter. An organization cannot, however, avoid production of any relevant information whatsoever in hyperlinked documents on the sole basis that the information appears in hyperlinked content.²⁰ For

19. See, e.g., *In re StubHub Refund Litig.*, No. 20-md-02951-HSG (TSH), 2024 WL 2305604, at *1–5 (N.D. Cal. May 20, 2024) (after having previously ordered the defendant to produce hyperlinked documents because it had agreed in the ESI protocol to treat hyperlinked documents the same as other document “family members” (see *In re StubHub Refund Litig.* 2023 WL 3092972, at *1–2 (N.D. Cal. Apr. 24, 2023)), the court subsequently granted motion to modify the ESI protocol to remove references to hyperlinks as family members after the defendant made an evidentiary showing that “the hyperlink requirement is technologically impossible to fulfill most of the time”) (also commenting that the defendant “carelessly stipulated to an ESI Order that [the defendant] later realized required it to do something that is usually impossible: produce the hyperlinked documents with the parent-child relationship with the original emails intact”); *In re Insulin Pricing Litig.*, 2024 WL 2808083, at *7–8 (D.N.J. May 28, 2024) (rejecting plaintiffs’ proposed language in ESI protocol that would have imposed the same production obligations regarding hyperlinked documents as those applicable to attachments and family members where defendants submitted declarations sufficiently proffering the technical infeasibility or undue burden of doing so).

20. See, e.g., *Porter v. Equinox Holdings, Inc.*, No. RG19009052, 2022 WL 887242, at *1 (Cal. Super. Mar. 17, 2022) (“parties generally have an

example, an organization must show that producing the relevant hyperlinked documents would be disproportionate or unduly burdensome to avoid production.

Hyperlinked documents may not appear relevant independently, but they may be relevant in the context of the communication containing the hyperlink.²¹ Similarly, a communication containing a hyperlink may not be independently relevant on its face, but may provide context to a relevant hyperlinked document, such as who was sent a document and when. As with discovery of other types of ESI, relevance, burden, and proportionality factors must be considered.²²

obligation to produce documents within the scope of discovery and that have been requested in discovery . . . litigants should generally produce relevant, responsive documents within their possession, custody, or control irrespective of whether the documents are stored in a cloud site, a company server, a laptop computer, a smartphone, or other location.”); *cf. id.* at *2 (“At first glance, this process [collecting and producing hyperlinked documents] may seem fairly straightforward and analogous to producing email attachments in family relationships. Nevertheless, linked documents can present unique challenges that make them different from email attachments.”) (listing challenges and also finding that, as a result, *IQVIA, Inc. v. Veeva Systems*, No. 2:17-cv-00177, 2019 WL 3069203 (D.N.J. Jul. 11, 2019), “which essentially held that linked documents were comparable to attachments[.] . . . is inapposite”).

21. *See, e.g.,* FED. R. CIV. P. 26(b)(1).

22. *See, e.g.,* *Shenwick v. Twitter*, 2018 WL 5735176, at *1, in which defendants presented evidence that producing a document referenced in a hyperlink in an email in the Google GSuite environment requires a multi-step, manual process. “The steps required are locating the document containing the link, clicking through the link to the source file, determining the file owner’s identity if a passcode is required and obtaining that passcode, manually identifying the date-stamped version of a linked GSuite document that corresponds to the referring electronic mail message, capturing the data in a manner that minimally affects the metadata, exporting the data to the vendor for processing, and producing the data in a manner that

Certain hyperlinked content also may not even have been intended to be a necessary part of the communication containing the hyperlink. The court in *Nichols v. Noom*, for example, stated that “[a] document might have a hyperlink shortcut to a SharePoint folder. The whole folder would not be an attachment.”²³

c. Versions of Hyperlinked Documents

A particularly problematic issue related to the identification and preservation of hyperlinked documents is “versioning.” Versioning is an automated feature that creates a new “version” of an electronic record when it changes in some way.²⁴ These changes can include modifications to file format, metadata, or content. On the one hand, collaboration platforms

matches it to the referring electronic mail message.” The court further stated that “Defendants claim that searching for and producing documents in the 725 hyperlinks identified by Plaintiffs will take six weeks and cost \$100,000.” The court, mindful of the burdens to defendants but also mindful of the plaintiffs’ need for the hyperlinked documents, ordered that plaintiffs could identify up to 200 hyperlinks for which they sought the referenced documents and that defendants must produce them. *Id.* See also UAB “Planner5D” v. Meta Platforms, Inc., No. 19-CV-03132-WHO (SK), 2024 WL 4190879, at *1 (N.D. Cal. Aug. 26, 2024) (“the Court agrees that hyperlinked documents are not the same as attachments to emails. Rather, the effort required to search for them is substantially greater than the effort required to produce attachments to emails.”).

23. See *Nichols v. Noom Inc.*, No. 20-CV-3677 (LGS) (KHP), 2021 WL 948646, at *4 (S.D.N.Y. Mar. 11, 2021).

24. See, e.g., *Shenwick v. Twitter*, 2018 WL 5735176, at *1 (defendants introduced evidence that hyperlinked documents within GSuite are not static but are instead “evolving,” and that producing a hyperlinked document within GSuite involves manually identifying the date-stamped version of the hyperlinked document that corresponds to the time of the email containing the hyperlink, which can involve a number of different steps).

and the use of hyperlinks can result in a linked document that has undergone numerous changes or edits from the time the hyperlink was originally used in an email or other communication to when ESI is being preserved or collected for a case. On the other hand, some linked documents—e.g., .pdf files or images—may either be effectively immutable or unlikely to be revised.

The version of a hyperlinked document that was produced may not be the same as the version circulated with the hyperlink. Using a hyperlinked document that is a different version than that which was contemporaneous with the message containing the hyperlink as if it were the same version of the document as sent in a message, risks creating confusion regarding the content of the hyperlinked document at the time of the message. If used in a deposition or presented as evidence, the parties should carefully consider whether and, if so, how to identify and treat a version of a hyperlinked document that is not contemporaneous with the communication in which the hyperlink appears.

Certain collaboration platforms may automatically save distinct versions of a document, providing snapshots at certain times, while others may not. For document management applications such as iManage, retaining versions of collaboration documents is essential to track stakeholder participation in the editing of the document. It may not be of equal significance for other types of applications. Indeed, some applications do not create distinct versions at all. Whether distinct versions are kept and are accessible may also depend on the licensing level of an application that an organization purchases. In litigation, the ability to capture all versions or a specific version of a document could be important to specific circumstances, such as where the document's evolution is itself relevant to a claim or defense. Requesting the production of "all versions" of all

hyperlinked documents, however, may not actually be needed or appropriate in many circumstances due to burden and proportionality considerations. Parties should avoid setting arbitrary volume requests for multiple versions of hyperlinked documents and instead demonstrate a need for the multiple versions.

Understanding the capabilities and potential limitations of a collaboration platform can help evaluate potential burden issues for the producing party. For example, matching the as-sent version of a hyperlinked document with the communication containing the hyperlink can be difficult to accomplish depending on how the data is stored. Furthermore, it may not even be possible to identify the contemporaneous version of a hyperlinked document because the version may have been deleted before the party's duty to preserve was triggered. Thus, one of the very reasons organizations utilize hyperlinks, i.e., to foster collaboration on documents, also presents one of the greatest challenges in discovery.

d. Collaboration Platform Functionalities and Hyperlinks

Electronic discovery and records retention demands have resulted in the development of functionalities within collaboration platforms to help address discovery- and preservation-related challenges. The licensing level an organization purchases may dictate the extent to which helpful solutions are available. Thus, an organization should understand the capabilities and level of access available when utilizing a collaboration platform. For example, the Business Standard license of early versions of Microsoft 365, which can include Teams, did not provide any ability to automatically preserve the as-sent versions of hyperlinked documents, nor to identify or preserve

hyperlinked content even in locations within the Microsoft 365 environment, such as SharePoint or OneDrive.

An early Business Premium license for Microsoft 365 provided more electronic discovery-related functionality than other license levels but still may not have had all the capabilities ideally needed for electronic discovery purposes. For example, upon implementation of a legal hold, it would not be possible to use that early version of Purview, Microsoft's data governance tool, to associate the contemporaneous version of a hyperlinked document with the message containing the hyperlink. It would, however, facilitate more generally identifying and putting on hold sources within the Microsoft 365 environment such as OneDrive and SharePoint.

Nevertheless, with any level of Microsoft 365 license, an organization would have needed to choose among various imperfect preservation strategies involving risks of either over-preservation or under-preservation, such as preserving the entire storage source where the hyperlinked document is stored (e.g., folders in SharePoint or OneDrive), or to use date ranges, search terms, or other methods to preserve a subset of documents stored in those sources. For review and production purposes, the organization might be faced with the task of manually trying to align a close to as-sent version of the hyperlinked document—possibly a burdensome and resource-intensive process. An open dialogue with the requesting party can avoid overcollection, as in many cases they have no need for an entire folder or drive pointed to by a hyperlink and will cooperate to make sure that productions are tailored toward only responsive material.

Subsequent versions of the Microsoft 365 software addressed some of these problems. A feature was added for Premium subscribers allowing organizations to retain a copy of

the as-sent version of a hyperlinked document (within the Microsoft 365 environment) at the time it was sent. The system would then treat this copy of the hyperlinked document and the message containing the hyperlink as a family. This feature needs to be set in advance, before the message with the hyperlink is sent, and cannot be implemented retroactively. Indexing versions of documents in a source within the environment may also facilitate searching for the as-sent version of a hyperlinked document, but conducting a search could still be a burdensome process.

e. Electronic Discovery Software Tools and Hyperlinks

Software providers have been developing tools that can, in some circumstances, collect and associate hyperlinked documents with messages containing the hyperlinks. One example at the time of publication of this *Commentary* is a computer forensics software that claims to automatically detect and acquire hyperlinked Google Drive documents during Gmail and Google Workspace collections, presenting them in a package that maintains relationship between the email and hyperlinked document. Another application claims to identify and collect hyperlinked documents from wherever they may exist in the SharePoint or OneDrive environments and to associate them with Teams messages containing the hyperlinks. The tool also compares the communication date of the Teams message to the version history of each file to associate the version shared at that point in time. Caution should be exercised regarding the actual capabilities of such tools, however, as their capabilities may be limited to certain environments or software configurations.²⁵

25. See, e.g., *In re Uber Techs., Inc. Passenger Sexual Assault Litig.*, 2024 WL 1772832, at *2 (third-party tool could “retrieve active Google Email and

Regardless of the tool, for an organization to adapt to preserve, collect, and produce hyperlinked documents, it may require coordination with IT regarding the functionalities of the organization's systems and applications. The collection, review, and production of hyperlinked documents may require detailed and tracked workflows that can capture these documents while retaining the metadata required for production.

As the use of hyperlinks continues to grow, additional tools will likely be developed to address the challenges associated with discovery of hyperlinked documents. The tools described above that exist as of the time of publication of this *Commentary*, however, may have significant limitations in that they may only be able to associate hyperlinks in collaboration platform messages with hyperlinked documents in the collaboration application's environment. They may not be able to make these associations where the hyperlinked content is outside of the collaboration platform's environment, e.g., a Microsoft 365 email

contemporaneous versions of linked Google Drive documents, but it does not have the ability to do the same with Google Email and Drive documents archived using Google Vault."'). *See also In re Meta Pixel Healthcare Litig.*, No. 22-CV-03580-WHO (VKD), 2023 WL 4361131, at *1 (N.D. Cal. Jun. 2, 2023) ("The Court is persuaded that the commercially available tools plaintiffs suggest may be used for automatically collecting links to non-public documents have no or very limited utility in Meta's data environments or systems, and even that limited utility (i.e. using the Microsoft Purview electronic discovery (Premium) tool to collect links to SharePoint and OneDrive cloud attachments in Microsoft Exchange environments) would disrupt Meta's standardized workflow for ESI-related discovery processing across all of its platforms and systems. Accordingly, the ESI protocol should make clear that hyperlinked documents are not treated as conventional attachments for purposes of preserving a 'family' relationship in production."').

with Google Drive links or Google Mail links to OneDrive or SharePoint.²⁶

In addition to being knowledgeable about the availability and functionalities of tools to address these issues, organizations will also need to be cognizant of the costs involved. Indeed, organizations may be well served by considering these issues in configuring their information systems and in their information governance, to avoid complex and expensive problems in fulfilling electronic discovery obligations when litigation comes. Courts may analyze these issues within the framework of burden and proportionality.²⁷ While courts may

26. See, e.g., *In re Uber Techs., Inc., Passenger Sexual Assault Litig.*, 2024 WL 1772832, at *1–5, in which the defendant introduced evidence that although certain tools can automatically collect contemporaneous versions of documents within the Google Workspace environment—i.e., in Google Drive from hyperlinks in Gmail and Google Chat—they could not do so from Google Vault, where unique versions of the documents had been archived for preservation and electronic discovery process purposes. Recognizing that tools with improved functionalities could later become available, and that existing tools may work outside of the Google Vault environment, the court ordered that language be included in the ESI Protocol requiring the production of contemporaneous versions of hyperlinked documents “to the extent feasible on an automated, scalable basis with existing technology[.]” Because the defendant established that, at the time of the decision, it was not able in an automated fashion to associate with the hyperlinks contemporary versions of the hyperlinked documents in Google Vault, the court ordered that the defendant “is not required to produce the contemporaneous document version at the time the email or message was sent, as this is not possible through an automated process with existing technology.” Nevertheless, the court ordered that plaintiffs could identify up to 200 hyperlinked documents for which the defendant would be required to manually search and produce.

27. See, e.g., *In re Meta Pixel Healthcare Litig.*, 2023 WL 4361131, at *1 (finding that “the commercially available tools plaintiffs suggest may be used for automatically collecting links to non-public documents have no or

very limited utility in Meta's data environments or systems"); *In re Insulin Pricing Litig.*, 2024 WL 2808083, at *7–8 (adopting defendants' proposed language regarding hyperlinks and family relationships in ESI protocol because defendants submitted evidence sufficient to show that the commercial tools whose use plaintiffs proposed "are either not feasible whatsoever or unduly burdensome to apply to their respective data environments"); *In re Uber Techs., Inc. Passenger Sexual Assault Litig.*, 2024 WL 1772832, at *3–4 (declining to order the defendant to build a program based upon a "proof of concept" proposed by plaintiffs for collection of contemporaneous versions of hyperlinks because the effectiveness of the program was not assured); *Nichols v. Noom*, 2021 WL 948646, at *4–5 (finding based on a proportionality analysis, that the defendant did not need to re-produce hyperlinked documents that were previously produced separately and associate them with emails containing hyperlinks. However, holding that plaintiffs could request "an additional targeted pull or production or clarifying information about a hyperlinked document's identity or Bates number." The court also rejected plaintiffs' proposal that the defendant should be required to have its software programmers write an application that would extract hyperlinked documents in an automated fashion.). *See also* *Shenwick v. Twitter, Inc.*, 2018 WL 5735176, at *1 (granting motion to compel production of hyperlinked G Suite documents along with the emails containing associated hyperlinks over defendant's burden objections); *Iqvia, Inc. v. Veeva Systems, Inc.*, 2019 WL 3069203 at *5 (D.N.J. July 11, 2019) (granting motion to compel re-association of separately produce hyperlinked documents to the extent possible because the request was neither unduly burdensome nor disproportionate considering proportionality factors and that the producing party had agreed they were relevant.).

decline to order a party to develop functionalities to enable such collection,²⁸ a party's failure to adopt an existing technology could factor in a court's assessment of burden claims.²⁹

f. Practical Solutions for Addressing Hyperlinks

Due to the technical challenges and burdens involved, it may be beneficial to both sides in litigation to share early in the litigation (for example at the Rule 26(f) early meeting of counsel) information sufficient to understand the technical capabilities of how hyperlinks are associated with a collaboration platform. This can allow the parties to have an informed discussion regarding the scope of production of hyperlinked documents and the burdens associated with producing relevant hyperlinked information.

There are no bright-line rules from statutes or case law governing the production of hyperlinked documents. The parties should engage in an informed dialogue about the nature of their respective data and hyperlinks. A productive dialogue will most likely be accomplished if both sides come prepared with appropriate technical knowledge regarding how hyperlinks work and specific information applicable to the particular platform(s) at issue. A robust understanding and communication of the policies, protocols, and tools in place for preserving, collecting,

28. See, e.g., *In re Uber Techs., Inc. Passenger Sexual Assault Litig.*, 2024 WL 1772832, at *4 ("the Court will not order Uber to expend potentially significant time and resources to develop such a program in order to produce discovery in this MDL, as the program's effectiveness is not assured."). See also *Nichols v. Noom, Inc.*, 2024 WL 948646, at *4 (rejecting plaintiffs' proposal that the defendant's programmers create a program to automatically extract hyperlinked documents).

29. See *In re Uber Techs., Inc. Passenger Sexual Assault Litig.*, 2024 WL 1772832, at *4 ("The Court is mindful of the burdens to Uber but also recognizes that Uber has chosen Google Vault as its storage method.").

processing, reviewing, and producing relevant hyperlinked documents can be helpful for counsel to resolve issues of burden and proportionality.

4. Collection Challenges

Collaboration platforms may present various collection challenges for purposes of electronic discovery. One of the primary issues is that some collaboration platform communications and documents may not be stored with the collaboration platform itself, but rather with other applications both within its immediate environment (e.g., certain types of communications and files in Microsoft Teams stored in SharePoint or OneDrive) and without (e.g., communications in Slack with hyperlinks to files in Google Drive or Box). Accordingly, for purposes of collection, it is important to understand where relevant and responsive information is likely to be located.

The best practice for collection is usually to engage a data collection professional with expertise in collecting data for litigation purposes. Such professionals may be found either within an organization's information technology department or through external providers. They should have knowledge of the options available for collection, including the functionalities available in the collaboration platform application and other software tools or approaches available to collect data associated with collaboration applications.

It can also be helpful for in-house and outside counsel to understand collection options and the extent to which there are limitations associated with them. One option, perhaps the most basic, is collecting data through download functionality in the

platform itself. Known as DYI (download your information),³⁰ these tools usually limit the scope, and accordingly the volume, of information that can be extracted from the platform. Therefore, it is important to understand what these limitations are before relying on this technique for electronic discovery collection purposes.

Additionally, DYI downloads may not be directly loadable into an electronic discovery-review platform or in a format that can be readily ingested by an electronic discovery-processing program.³¹ Commercial software tools exist, however, that can process some DYI downloads into RSMF (Relativity Short Message Format) or review platform-ready load files. Unless DYI downloads are conducted carefully, however, metadata can be altered or deleted. Those conducting DYI downloads should be alert to this potential issue and be prepared to address it, including through detailed instructions and appropriate supervision.

Data may also be extracted from a collaboration platform through the platform's API (application programming interface). APIs are interfaces for programs and programmers to access a platform programmatically. Counsel should determine license levels early in the preservation planning process, both for their own clients and for third-party sources. For example, Google provides an API to enable its clients and commercial software developers to interface directly with Google Drive documents

30. See *Download a Copy of Your Information on Facebook*, FACEBOOK, <https://www.facebook.com/help/212802592074644> (accessed March 25, 2025); See *How Do I Download My Data from Snapchat*, SNAPCHAT, <https://help.snapchat.com/hc/en-us/articles/7012305371156-How-do-I-download-my-data-from-Snapchat> (accessed March 25, 2025)

31. For example, Facebook DYI downloads come in HTML or JSON format.

and folders by creating programs that retrieve files in Google Drive.³² Similarly, Microsoft provides an API to retrieve SharePoint documents using Microsoft's SharePoint REST API.³³

Commercially available tools and those used by electronic discovery service providers may use multiple APIs from different platforms to address some heterogeneous environments—for example, to collect Slack threads and messages and then to collect the hyperlinked documents resident on Google Drive. There may not be commercial programs that can collect all the information together or in an associated fashion for heterogeneous environments—for example, using Microsoft Teams but storing hyperlinked documents outside of the Teams environment.

APIs, moreover, can present challenges in data collection. As stated in *The Sedona Conference Commentary on ESI Evidence and Admissibility*, “[a]n 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.”³⁴

Various commercial tools may be available on the market at any given time for the collection of data from collaboration platforms and the applications within those platforms. There also may be specialized tools that focus on specific aspects of collaboration application data, such as collecting versions of

32. See, e.g., *Introduction to Google Drive API*, available at <https://developers.google.com/drive/api/v3/about-sdk> (accessed on June 5, 2023).

33. See, e.g., *Get to know the SharePoint REST service*, available at <https://learn.microsoft.com/en-us/sharepoint/dev/sp-add-ins/get-to-know-the-sharepoint-rest-service> (accessed on June 4, 2023).

34. See *The Sedona Conference, Commentary on ESI Evidence and Admissibility*, 22 SEDONA CONF. J. at 139–140 (2021).

hyperlinked documents. It also has been asserted that custom applications may be developed to collect collaboration platform data. Courts, parties, and their counsel should be careful, however, to assess whether the commercial or custom applications can perform as advertised in the collaboration platform environment at issue.³⁵

Finally, the ability to collect collaboration platform data may depend on the type of license or subscription the producing party has for the platform. For example, the Enterprise version of Slack has more options for preservation and collection than the free version.³⁶ Similarly, Microsoft's Purview electronic discovery Premium (E5 license) has more advanced functionality than its Standard (E3) offering.

5. Culling, Search, Review, and Production Challenges

a. Culling

Once collaboration platform data is collected, parties must decide whether additional culling should occur prior to search, review, and production. Additionally, decisions may need to be made at the processing stage that will affect the form in which the data is reviewed and produced. For example, document unitization—e.g., dividing continuous communication streams such as chat or instant messages into more

35. See *supra* note 31.

36. For example, in *Calendar Research LLC v. Stubhub Inc.*, No. CV 17-4062 SVW (SSx), 2019 WL 1581406, at *3-4 (C.D. Cal. Mar. 14, 2019), the plaintiff requested Slack messages from an individual defendant's employer. Because the employer had a free Slack account, certain Slack folders were not retrievable. The defendants paid for an upgraded account, but Slack denied full access because not all the parties on the account had consented. Slack provided a "utility tool," however, to target the channels used by the individual defendants.

manageable units, such as a 24-hour period—may need to take place at the processing stage. Different types of documents and communications on collaboration applications also may involve different types of culling, search, review, and production issues.

Culling collaboration platform data may involve examining the data for content, context, metadata, and other attributes, and applying various techniques to filter, categorize, and prioritize the data. For example, some common techniques for initial culling of chat and instant-message data are keyword search, date range, concept search, and near-duplicate detection.

Deduplication is an additional culling issue that should be considered. For chat and instant messages, one must discern whether they will be deduplicated across custodians or whether there is value in producing the same messages multiple times. If deduplication is preferred, the producing party should verify that the tools used can perform this type of deduplication.

Similar to deduplication decisions, the producing party should determine whether all versions of a collaboration document may be relevant and proportional to the claims or defenses, or whether the production can be limited in some way. Collaboration tools may store each modification as a separate version, resulting in thousands of versions, many of which may have only immaterial differences.

For chat and instant-message data, the producing party should consider whether the collection should be limited to certain date ranges. Unitizing (i.e., breaking up) continuous message streams into smaller pieces after collection is becoming a common practice. Options include breaking up streams into certain time periods (e.g., 24 hours) or into a certain

number of messages.³⁷ If the relevant participants are all in the same time zone, it may be best to process in that local time zone and unitize at, for example, midnight. If the participants are in different time zones, then processing by UTC zone and unitizing either by 24-hour periods or by breaks in discussion may be preferred. The potential downside of unitization by time period or number of messages is that it can artificially break up continuing conversations into separate documents, requiring the parties to reassociate them for use in the litigation.

Another potential solution may be organizing the collection by subject matter, although there can still be issues where related conversations span multiple days. An early version of Microsoft's Purview Premium, for example, created transcript files in 24-hour windows. Messages could be converted to Relativity's Short Message Format (RSMF) in 24-hour or other time periods for review and production. Other technologies allow parties to create different groupings of messages.

Producing parties should consider how comments within documents will be treated during review and production. Producing parties should also determine whether comments can be extracted during processing and loaded to a metadata field. Also, producing parties should determine whether a review platform will show the comments in the viewer and extracted text and, when the documents are imaged, whether the comments will show on the image. The answers to these questions will determine what can be done with them from a production standpoint. Producing parties should set the expectations of

37. See *Lubrizol Corp. v. IBM Corp.*, No. 1:21-CV-00870-DAR, 2023 WL 3453643, at *4 (N.D. Ohio May 15, 2023) (ordering producing party to produce entire Slack and Microsoft Teams conversations if the conversations had twenty or fewer messages, and the ten prior messages and ten subsequent messages if the conversation had more than twenty messages).

requesting parties regarding how these will be handled during the review and production process. If the comments can be reviewed and produced, consistent with proportionality considerations, then they should be produced to the extent they are not privileged or otherwise protected from disclosure.

b. Search and Review

Search methodologies in electronic discovery include using keyword searches (also referred to as search terms), technology-assisted review (TAR), and artificial intelligence (AI). AI may also be able to assist in grouping topically related chat and instant messages together (“smart grouping”).

Despite the availability of TAR and other AI methodologies, keyword searches are still commonly used for search and review. Producing parties should carefully consider whether search terms are a viable option for chat and instant messages. If used, search terms should be tailored to the nature of both chat and instant messages in general and the specifics of the matter, but flexible enough to collect communications where abbreviations are commonly used, and misspellings can occur. Thus, it is necessary to understand what abbreviations, acronyms, shorthand, and slang are likely to be used by the communicants in the specific matter. Doing so can also be important for identifying privileged communications, as attorneys may just be referred to by an abbreviation or a nickname in chat and instant messages. Once the data is collected, producing parties can analyze samples to get an understanding of how users are communicating within these platforms.

In analyzing the data, pay special attention to the use of emojis and GIFs. Depending on the nature of the matter, these may be important to decoding the communications. Verify

what capabilities the platform has for searching and analyzing emojis and GIFs and apply them accordingly.

Part of the analysis can also include evaluating whether different terms may be necessary for the chat and instant-message data compared to emails and other forms of ESI. Producing parties should consider whether search terms should be used at all, and whether other search methodologies would be more effective. Other possible approaches including conducting a more detailed review of a person's message threads within a certain time period, or in exchanges with certain people. Users often communicate much more informally on collaboration platforms than they do in emails, for example. Therefore, the same search terms may not yield the intended results.

As with all searches, the efficacy of keyword searches may be tested by running proposed search terms and sampling the hits and nonhits, and using other analytics tools to look for additional terms. Such sampling may also help determine whether search terms may or may not be appropriate. If the parties have negotiated an ESI protocol, it can be helpful if that protocol contains provisions that allow for such an iterative approach, with analysis and sampling, and parties should not agree to specific search terms prior to testing them.

TAR and other AI-based tools may also be used for search-and-review of collaboration-platform data, although parties should undertake appropriate measures to ensure the process is providing reasonable results, including validation of the results.³⁸

38. For a discussion of judicial decisions regarding validation of TAR results, see The Sedona Conference, *TAR Case Law Primer, Second Edition*, 24 SEDONA CONF. J. 1, 55–62 (2023).

c. Production

(1) Content of Production

There are several considerations when determining what information from a collaboration platform should be produced. Parties should discuss and attempt to agree upon the content of production to avoid potential disputes.³⁹

Producing parties reviewing chat messages on collaboration platforms may need context to determine relevance. One method to organize responsive messages is to produce a certain number of messages before and after a responsive chat message. Another method is to identify and organize communications from a particular time range or date range.⁴⁰ The potential downsides to these approaches include that there may be messages outside the selected message range that would also provide helpful context, and those messages would not be produced as a single conversation if the parties are following such a protocol.

Not breaking up the message stream into different units, however, may result in both irrelevant and relevant messages being produced together. The production of such irrelevant messages may increase the risk that personal and sensitive information is produced. Communication within chat and instant-messaging applications may also jump from one subject to another in quick succession, and then back to earlier subjects.

39. For example, as discussed above, it can be problematic to include hyperlinked documents within the definition of a document “family.” See *In re StubHub Refund Litig.*, No. 20-md-02951-HSG (TSH), 2024 WL 2305604, at *1-*5 (N.D. Cal. May 20, 2024).

40. See *supra* Section II.B.5.a., discussing advantages and disadvantages of different approaches to unitization.

Breaking up the conversation into different units may therefore require the parties to reassociate the related messages.

With respect to privilege, depending on the parameters of the review and the platform, parties may redact the privileged messages in a message stream or, if the messages are being produced as separate units, to withhold and log the separate privileged messages. Some platforms allow users to code at the message level and allow users to select messages that are responsive and create a separate record for production. Users can also select messages that are privileged and create a separate record for logging purposes. Regardless of the approach, the fact of withholding privileged content should be disclosed and producing parties should keep in mind that the context (e.g., surrounding messages) of responsive messages can also be relevant.

Identifying privileged communications may be more complicated when dealing with chats and instant messages. Attorneys may be referred to by first names only or by initials, for example, or using some other shorthand, which reviewers may miss. Another complication is that people may be communicating with an attorney by email, then jump platforms to Slack or Teams and start discussing the legal advice provided in the email, but not include the attorney in the chat or instant-message communication. It may be difficult for reviewers to recognize that privileged content without the benefit of the context of the other platforms.

(2) Form of Production

As with other forms of ESI, if a request does not specify a form for producing ESI, a party may produce it in a form in which it is ordinarily maintained or in a reasonably usable

form.⁴¹ Generally, a reasonably useable form will be one that is viewable and searchable in most document review platforms. To avoid later problems, parties should discuss the proposed form of production early in the case and, if they are unable to reach agreement, seek the assistance of the court.

Parties should memorialize—for example, in an ESI protocol—their agreements regarding the form of production with respect to collaboration platform data, just as they would with other sources. Note that such an agreement may also be technology-dependent and may change as the technology advances. Standard .tiff or .pdf files, plus text, metadata, and load files are usually sufficient forms of production for chat and instant messages. There is no “native format” production option for chat and instant messages. Parties should evaluate the technological options, keeping in mind how the data will be used in depositions and at trial.

Parties should discuss what metadata should be produced from chat and instant messages. If parties are reviewing in 24-hour daily periods, they may only have the metadata from the first message of that day, or they may not be able to provide metadata for all the individual messages. Or they may be able to provide all the participants and a date range, but not an entry for each message. It is important to understand the technical limitations before agreeing to a format that is not readily available. Parties should discuss what metadata is readily available, and they should also discuss any burdens associated with production of such information.

Some courts’ decisions reflect that information contained in continuous message streams may need context, and therefore even messages that by themselves may be considered

41. See FED. R. CIV. P. 34(b)(2)(E).

irrelevant may nevertheless be discoverable because they provide needed context to other relevant communications.⁴² Other courts, however, have held that the producing party can unilaterally withhold portions of a text message chain that are not relevant to the case.⁴³

42. See, e.g., *Bidprime, LLC v. Smartprocure, Inc.*, No. 1:18-CV-478-RP, 2018 WL 6588574, at *3 (W.D. Tex. Nov. 13, 2018) (ordering producing party to produce full chat logs including portion that the producing party had withheld as allegedly irrelevant). See also *Lubrizol Corp. v. IBM Corp.*, No. 1:21-CV-00870-DAR, 2023 WL 3453643, at *4 (N.D. Ohio May 15, 2023) (ordering producing party to produce entire Slack and Microsoft Teams conversations if the conversations had twenty or fewer messages, and the ten prior messages and ten subsequent messages if the conversation had more than twenty messages); *Sandoz, Inc. v. United Therapeutics Corp.*, No. 19-CV-10170, 2021 WL 2453142, at *2 (D.N.J. June 16, 2021) (“Accordingly, the Special Master holds that RareGen shall produce context-related messages for each of the messages UTC has identified in Exhibit C to the Motion, including the preceding text message or responding text message, if they exist. If they do not exist, RareGen must so state, and provide an explanation based on information available to it why they do not exist.”); *Advanced Magnesium Alloys Corp. v. Dery*, No. 1:20-CV-02247-RLY-MJD, 2022 WL 3139391, at *4 (S.D. Ind. Aug. 5, 2022) (court ordered a party to produce all text messages between two individuals during a particular period, but also held that the producing party could redact any message or portion of a message that was purely personal in nature or related to business matters other than those at issue in the case).

43. See, e.g., *Marksman Sec. Corp. v. P.G. Sec., Inc.*, No. 19-CV-62467, 2021 WL 4990442, at *2 (S.D. Fla. Mar. 19, 2021) (“The undersigned finds that Defendants are in the best position to determine which text message conversations require additional context and which do not, and therefore which texts require the surrounding conversation to be produced.”); *In re Pork Antitrust Litig.*, No. 18-CV-1776, 2022 WL 972401, at *14-15 (D. Minn. Mar. 31, 2022) (holding that party need only produce relevant text messages, and stating “[j]ust because there may be some relevant texts within a data set does not make all texts within that set presumptively relevant”); *Baillet v. Colorado Springs*, No. 20-CV-01600,-WJM-KMT, 2021 WL 2912921, at *2 (D.

6. Evidentiary, Privilege, and Privacy Issues

Documents and communications exchanged and stored on collaboration platforms present evidentiary and privacy issues, often without analogues to traditional ESI.

When documents are shared and simultaneously edited by multiple employees in a workspace that overwrites the document with every edit, issues affecting admissibility may arise. If multiple people work on a document, it may be difficult to identify the best person to authenticate or lay a foundation to overcome hearsay issues. If a document is shared with recipients by hyperlink in an email or chat communication, it may be difficult to determine whether the version of the document produced in litigation is the same version that was shared. Consideration also should be given to privilege and privacy issues. For example, when multiple individuals in different jurisdictions are working on a document, it may impact which jurisdiction's privilege or privacy laws apply.

Parties should consider and discuss early in the process evidentiary and privacy issues that may arise from the discovery of information from collaboration platforms.⁴⁴

Colo. July 12, 2021) (holding that nonresponsive, withheld text messages did not require any objection, redaction, or notation in a privilege log).

44. See The Sedona Conference, *Commentary on ESI Evidence & Admissibility*, Second Edition, 22 SEDONA CONF. J. 83 173–74, 179 (2021) (Hereinafter *Commentary on Admissibility*) (parties should consider ESI evidentiary issues early in the case and ensure they have defensible preservation and collection protocols).

d. Evidentiary Concerns

(1) Authenticity

Collaboration platforms may present unique issues pertaining to the authenticity of a document. Given that collaboration platforms allow multiple users to view and edit a document simultaneously, authentication may be more complex.⁴⁵ It may be necessary to rely upon other evidentiary rules to authenticate documents from a collaboration platform.⁴⁶ Another avenue to consider is whether documents from collaboration platforms can be authenticated by a qualified person pursuant to Federal Rule of Evidence 902(14).⁴⁷

(2) Hearsay

As with authentication, testimony often lays the foundation for a hearsay exception—for example, that an out-of-court statement is a statement by a party opponent or that it meets

45. See FED. R. EVID. 901(b)(1) (providing that a witness with knowledge may be proffered to authenticate an item of evidence); *Commentary on Admissibility supra* note 47 at 156 (2021) (discussing difficulties in determining the owner or creator of some ESI).

46. See *Commentary on Admissibility supra* note 47 at 139–140 (discussing collaboration tools and the various ways to authenticate them); see also *id.* at 194–209, 218 (identifying methods of authentication for various types of evidence and providing supporting case citations).

47. See *id.* at 98–104 (discussing new subsections 902(13) (certified records generated by an electronic process or system) and 902(14) (certified data copied from an electronic device, storage medium, or file) to Federal Rule of Evidence 902, which streamline the authentication of electronic evidence); *id.* at 156–168 (discussing digital identification methods); *id.* at 149–151 (discussing the challenges of Rule 902(14)).

the criteria for a business record.⁴⁸ Where multiple people are collaborating on a document, however, identifying the author of a particular statement may be difficult.

Additionally, employees may use the chat features of collaboration platforms to discuss mixed business and personal matters, which may raise questions as to whether specific chat records are properly considered to be business records. Furthermore, chat platforms may be used by employees but not officially sanctioned by the organization for use. Thus, identifying the use of a collaboration platform is necessary to understand how the information may be admissible as a business record or other exception to the hearsay rule.

e. Privilege and Work-Product Concerns

(1) Work-Product Doctrine

Sometimes, collaboration platforms may not specifically maintain records of the author, recipients, participants, or contributors of non-communication information, which may complicate the analysis of determining whether the work-product doctrine applies. With traditional documents, identifying work product was more likely to be straightforward. A memorandum prepared in anticipation of litigation or for trial by counsel providing legal advice is often recognizable for what it is. The informal nature of messaging in collaboration platforms can make it difficult to determine whether content or materials were being gathered or prepared for litigation purposes. Establishing a clear record at the front end that information is work product and limiting access to the communication from its initial inception forward can save time when reviewing

48. See generally FED. R. EVID. 803 (discussing exceptions to the rule against hearsay).

information and may prevent the inadvertent waiver of the work-product protection.

(2) Attorney-Client Privilege

The attorney-client privilege may apply if a confidential communication is made to render or receive legal advice or to enable the provision of legal services. Accordingly, for privilege purposes, it is important to understand the context of the communication. Communications in collaboration tools such as Slack, however, often involve a long stream of different conversations over time, which can make identifying and separating out privileged communications more difficult. Additionally, ongoing chats may contain mixed-purpose communications in one long chain of text. This characteristic can make it more difficult to identify the context of the communication, whether a primary purpose of the communication is for legal advice, and whether the communications for purposes of legal advice can be separated from the remainder of the communications.

When attorney-client communications are made using a collaboration platform, consideration should be given to expressly stating in the communication that it is for the purpose of seeking or providing legal advice, or labeling the communication as such (for example, “Confidential Attorney-Client Privileged”). Consideration also should be given to limiting permission rights and access to such communications to avoid the potential risk of waiver. Taking steps at the time of the communication—such as explicitly identifying the communication as one with an attorney, seeking legal advice, or facilitating rendering legal advice—can help ensure that the attorney-client privilege will be properly identified and asserted, and will help to prevent inadvertent waiver of the privilege. Invoking privilege where it does not apply, however, can undermine

proper invocations of privilege, rendering such labeling less effective.

f. Privacy Considerations with the Use of
Collaborative Platforms Within the Borders of the
United States

(1) U.S. Privacy Considerations

Those adopting collaboration platforms must be aware of privacy issues involved in using such platforms for communications that cross state and international borders. While the United States has a mix of federal laws to protect specific types of data, “[it] has no overarching and preemptive national ‘privacy law’ or ‘data security law’ in place.”⁴⁹ Rather, privacy in the U.S. is mostly addressed by state-specific laws (among which there is a notable lack of uniformity).⁵⁰

Communication and collaboration tools are used by multi-state and global entities, so the users may be in different states and in different countries. When there are users based in different jurisdictions, consideration should be given to choice of law and whether and how privacy laws will govern the information and its retention.⁵¹

49. See The Sedona Conference, *Commentary on Quantifying Violations Under U.S. Privacy Laws*, 22 SEDONA CONF. J. 489, 497 (2021) (“Commentary on U.S. Privacy Laws”).

50. See generally, *id.* at 497–505. As of March 2024, there were 15 U.S. state laws on Privacy. See, e.g., California Privacy Rights Act (CPRA), Cal. Civ. Code § 1798.100; Illinois Personal Information Privacy Act, (PIPA), 815 ILCS 530; Connecticut Data Privacy Act, (CTDPA) Conn. Gen Stat § 42-522.

51. See e.g., *Lieberman v. Unum Group*, Case No. 5:20-CV-1798-JGB (SPx), 2021 WL 4807643, at *2 (C.D. Cal. Oct. 14, 2021) (rejecting arguments that other state’s privacy laws apply, deciding that the law governing the substantive claims and defenses governs).

Information created in collaboration platforms may raise additional privacy concerns for personal information—which could be identifying information regarding users, protected health information, or other sensitive personal information. While the Health Insurance Portability and Accountability Act (HIPAA) protects against the disclosure of protected health information, the HIPAA privacy rule only applies to covered entities—namely health plans, health plan clearing houses, and health care providers that electronically transmit any health information in connection with transactions for which the U.S. Department of Health and Human Services has adopted standards.⁵² To the extent a covered entity possesses protected health information within a communication or collaboration tool and such information is sought in litigation, then HIPAA may apply.⁵³ In such a scenario, it would be wise for the parties to agree on a HIPAA-compliant protective order.⁵⁴

(2) Foreign Privacy Considerations

The use of collaboration platforms may also raise specific foreign data privacy laws and tangentially other foreign laws, such as labor laws, which restrict retention of information about employees. Consideration should thus be given to the identification of the location of information and users of the platforms. The EU has a comprehensive data privacy law, the General Data Protection Regulation (GDPR), which governs the

52. See 45 CFR § 160.102.

53. See e.g., *Lillard v. Univ. of Louisville*, 2014 WL 12725816, at *18 (W.D. Ky. Apr. 7, 2014) (denying request for “all relevant” Slack messages related to the University of Louisville School of Medicine because, in addition to being “vastly overbroad” and “unduly burdensome,” it would be a potential HIPAA violation in the event the messages contains patient information).

54. See 45 C.F.R. § 164.512.

rights of personal information of its citizens and addresses how data is processed and controlled within the European Union.⁵⁵ The GDPR applies to protect its citizens and certain information, such as personal identifying information or sensitive personal information, in any documents, data, or information collected or processed.

Notably, the GDPR can apply to businesses based in the United States if the business has an establishment in the European Union or if the business targets individuals in the European Union for offering goods and services or monitors their activities.⁵⁶ Where an entity has employees located in the European Union or does business in the European Union, and discovery is sought from a collaboration platform that the entity's employees use, there may be certain information stored within the collaboration platform that is subject to the GDPR provisions affecting production and, furthermore, may limit retention of such information only as necessary to the purpose for which it was created.

Understanding the specific collaboration tool, whether user-specific or other personal identifying information is collected, how data is shared, where it is maintained, and whether user-specific information is collected is necessary to determine whether foreign data privacy laws will apply. Additionally, many labor councils in the EU limit the retention period of an enterprise in keeping such communications that

55. Regulation (EU) 2016/679 of the European Parliament and of the Council 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>

56. See generally Commentary on U.S. Privacy Laws at 284-343.

contain personal information. Compliance is often insisted upon by the governing Council to prevent violation of the Labor Codes, and that may be a basis to explain in discovery why certain information is missing for certain contributors to collaborative platforms.

C. Information Governance Considerations

Collaboration tools have created great efficiency, accessibility, and opportunities for sophisticated workflow management and information exchange within an organization. Collaboration platforms have also enabled greater social engagement and information sharing by employees and an increase in collaborative work productivity. Use of collaboration platforms, however, should involve careful consideration for the implementation of appropriate data and information-governance principles applicable to their use.⁵⁷

An organization should understand collaboration tools before their deployment, including whether and how such tools fit within the organization's business processes and information systems (e.g., requirements for records retention, audit, and discovery in litigation). This understanding is commonly accomplished, for example, through the creation of a team that

57. ARMA International sets forth as its highest level of Records Management Principles (Level 5—Transformational) as follows: "This level describes an organization that has integrated records management and its infrastructure and business processes such that compliance with the organizations' policies and legal/regulatory responsibilities is routine." Available at <https://www.arma.org/page/PrinciplesMaturityModel> (last visited June 22, 2024).

works together to vet the collaboration platform and to understand its capabilities.⁵⁸

An organization should have a representative from the legal department either on the team or to provide input to the team to ensure consideration of potential issues that may arise from the collaboration platform's functionalities, including the electronic discovery capabilities and limitations at different license levels. Information or data governance programs should align with the objectives of the organization, while maintaining enough flexibility to respond to changes and opportunities.⁵⁹

If an organization is considering a collaboration platform, the team vetting the platform should understand the needs of the users and business units and understand the business, legal, data-security, and regulatory requirements related to the platform. The vetting team should consider whether the collaboration platform has certain functionalities, including information retention, preservation, and legal hold. When vetting any collaboration platform, the organization should ensure that its data security features are sufficient and will integrate with the organization's information system.

Consideration should also be given to the potential need for third-party services that can provide technological capabilities needed to satisfy data preservation, extraction, and review

58. See generally The Sedona Conference, *Commentary on Information Governance, Second Edition*, 20 SEDONA CONF. J. 95, 102 (2019) ("Organizations should consider implementing an Information Governance program to make coordinated, proactive requirements and manage risks while optimizing value").

59. See *id.* at 116 (discussing the business case for information governance).

from the platform.⁶⁰ The administrator of the platform within an organization, as well as the legal electronic discovery team and any other relevant business unit authorized to administer the platform, should understand where the collaboration platform's data is stored, who has access to it, how it is accessed, and how the data may be preserved, collected, reviewed, and ultimately produced.

A collaboration platform's records retention settings should be clear regarding information and content that may be created on the platform, and the settings should align with the organization's record retention policies. As with all retention policies, the policy should establish a records retention period for created content and different types of messaging components (e.g., email, chat). The records management owner should also consider legal, business, privacy, and regulatory requirements for the retention of content or communications.

An organization should determine if the collaboration platform has the capability to preserve information subject to legal hold, and the manner in which the tool can implement such preservation.⁶¹ The stakeholders should create a strategy to implement and release legal holds on information in the platform.⁶² The collaboration platform should have the ability of auto-deletion or to purge data when the retention period is

60. See, e.g., *Red Wolf Energy Trading*, 626 F. Supp. 3d at 505 (noting that defendant provided "changing, unconvincing explanations for why [defendant] did not employ an experienced vendor to search the Slack messages.").

61. See, e.g., *Commentary on Legal Holds*, *supra* note 6 at 398 (discussing the difference between data collection and data preservation).

62. See *id.* at 408 (discussing the release of legal holds and termination of the duty to preserve).

expired and when a legal hold is lifted.⁶³ To ensure the ability to comply with electronic discovery obligations, data about usage and ability to access data, file sizes, searching capabilities across the collaboration platform or its components, and exportability should be available and readily communicable as a prerequisite for implementing such platforms.

An organization should establish best practices for collecting and extracting content from collaboration platforms, as these tasks are an important part of a company's well-crafted information governance plan and will be necessary when discovery arises. It is important that an organization understands which collaboration tools are being used, which parts of the organization are using these tools, what information is being created in them, and in what form and locations it is being maintained.

Once a collaboration platform is selected and implemented, there should be clearly communicated acceptable use standards, retention standards, and education on the version of the platform being used, its capabilities, and the approved scope of its use.⁶⁴

Finally, creating a long-term retention strategy appropriate to the value and type of information created or used on the collaboration platform involves considering a broad range of factors pertaining to the digital assets and the circumstances of the organization itself. These factors should include: business value; regulatory importance; legal requirements; global and local privacy concerns; retention schedule; legal hold status; file

63. See *id.* at 409 (discussing automated software for the efficient management of legal holds).

64. Paul Kirvan, *Acceptable Use Policy (AUP)*, TECHTARGET, <https://www.techtarget.com/whatis/definition/acceptable-use-policy-AUP> (last visited June 22, 2024) (explaining elements of an acceptable use policy).

format; continued availability of the technologies required to access and read the information; the likely failure rate of the storage medium as it is configured; the available budget and resources of the organization for the tool and—for third-party services such as cloud storage, software as a service, etc.—the contractual agreements between the customer and provider.

D. GenAI Considerations

Many collaboration platforms have built-in generative AI (GenAI) capabilities, and at the time of publication of this *Commentary*, many organizations have begun to license generative AI solutions as plug-ins to, or features of collaboration platforms. Organizations should consider the use of these solutions and how they may create additional relevant, responsive data for discovery purposes.

Microsoft Teams and OneDrive are integrated with Microsoft Copilot, which currently uses GenAI to draft emails and to create reports. Copilot also currently allows users to create a GenAI summary of meetings and a list of action items if such meetings were recorded and/or transcribed. Copilot can also assist users in drafting or editing documents shared via OneDrive.

Companies should consider that GenAI outputs within collaboration platforms may be based on communications or work product from counsel, highly confidential business information, or protected personally identifiable information (PII). Because GenAI tools can aggregate content from multiple sources, organizations should account for these risks in their information governance policies and practices. GenAI inputs and outputs may also be subject to preservation and production obligations, provided that they are relevant and their

preservation and production is proportional to the needs of the case.

Accordingly, it is important to understand whether the collaboration platform and related tools have GenAI functionalities, and if so, whether any of those functionalities have been active during the relevant time period, as well as and whether they have generated relevant content. Organizations will also need to understand where and how GenAI information is stored and the technical capabilities (if any) for collecting that data.

In summary, effective records retention and information governance for collaboration platforms require clear policies, technical capabilities to manage legal and regulatory obligations, and ongoing user education. With the rise of generative AI features, organizations must carefully assess new sources of data, associated risks, and ensure their strategies and controls adapt to evolving technological and compliance landscapes.

III. CONCLUSION

Organizations must be prepared to identify, preserve, collect, and produce relevant ESI from collaboration platforms in a defensible manner, especially given the complexities associated with hyperlinked documents, message streams, and varied data locations. Policies and procedures should anticipate the unique challenges of e-discovery involving collaboration platforms.

Success in these endeavors is facilitated by sound information governance and records retention strategies, including methodologies for identifying, retaining, collecting, and producing relevant content. Hyperlinked documents may require particular attention, as they may be stored separately and

involve multiple versions not contemporaneous with the message containing the hyperlink.

By keeping policies and technical capabilities up to date and ensuring readiness for discovery across a rapidly evolving technological environment, organizations can fulfill legal and regulatory obligations with confidence. As collaboration platforms and e-discovery-related technologies advance, organizations will be well served by adapting governance frameworks and technical tools as needed to comply with discovery requirements. Proactive planning, regular policy review, and ongoing user education can help organizations navigate the evolving demands of modern e-discovery as they apply to collaboration platforms.