William Mitchell Law Review

Volume 40 | Issue 2 Article 6

2014

IT Technologies and How to Preserve ESI Cost Effectively

Mary T. Novacheck

Molly B. Thornton

Jeffrey J. Beard

Mark Burns

IT TECHNOLOGIES AND HOW TO PRESERVE ESI COST EFFECTIVELY

Minnesota E-Discovery Working Group 3†

† The Minnesota E-Discovery Working Group is a grassroots organization that was founded in 2011 with the goal of writing five separate papers that address various aspects of e-discovery best practices from a Minnesota perspective and could be used as a resource by both judges and lawyers in Minnesota. Its members consist of members of the Minnesota judiciary, in-house attorneys, attorneys practicing with law firms across Minnesota, and e-discovery experts. The Working Group and the *William Mitchell Law Review* thank Briggs & Morgan, P.A. and Fredrikson & Byron, P.A. for their financial contribution to this joint project.

About the Authors[††]

Mary T. Novacheck (Co-Chair) is a partner practicing in product liability defense and leads Bowman and Brooke LLP's E-Discovery and Discovery Coordination Group. She is a member of the Sedona Conference Working Group on ESI, has briefed and argued ESI issues in state and federal courts throughout the country, and is a frequent lecturer on ESI issues.

Molly B. Thornton (Co-Chair) is a generalist attorney at Cargill, Inc., where she advises clients on both transactional and litigation matters. Prior to joining Cargill in 2011, Thornton was a shareholder at Briggs and Morgan, P.A., and a judicial law clerk to U.S. District Court Judge Paul Magnuson.

Shari Aberle is a partner in the Trial Group at Dorsey & Whitney LLP, where she represents clients in all types of complex commercial litigation, including finance industry disputes, health industry disputes, consumer fraud and deceptive trade practice act claims, contract disputes, copyright and trademark infringement suits, class actions, and appellate advocacy. Ms. Aberle is also a member of Dorsey's Electronic Discovery and Privacy and Social Media Practice Groups, the Sedona Conference Working Group on Electronic Document Retention and Production, and the Minnesota State Bar Association's E-Discovery Working Group.

Jeffrey J. Beard is an Information Governance Consultant with IBM, where he assists organizations faced with challenging e-discovery and data retention issues. He works with corporate legal, records, and IT departments to identify, evaluate, and implement in-house e-discovery and information governance solutions, including policy and process considerations.

Mark Burns is an E-Discovery Manager for the Boston Scientific Chief Litigation Office. Mark is responsible for the identification, preservation, collection, processing, and review of data for product

†† The opinions expressed in this publication do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong. Working Group 3 wishes to acknowledge committee members Roy Ginsburg, attorney at Dorsey & Whitney LLP; Dean LeDoux, attorney at Gray, Plant, Mooty, Mooty & Bennett, P.A.; Sonya Seidl, attorney at Target Corporation; and Jason White, of White Light Solutions, Inc., for their contributions to the Working Group.

liability, IP, and employment litigation; government investigations; and other corporate litigation disputes.

Megan I. Brennan received her BA from the University of Minnesota in 2003 and her JD from Hamline University School of Law in 2006, both *summa cum laude*. As an attorney at Nichols Kaster, PLLP, she has been involved in a wide variety of individual and class action cases representing employees.

John Carney is Chief Technology Officer and a practicing digital forensic examiner at Carney Forensics where he has been assisting clients with electronic and mobile device evidence for the past five years as an expert witness. He graduated from the Massachusetts Institute of Technology Media Lab with a Bachelor of Science degree and is a federally and state licensed attorney in Minnesota.

Mary Frantz is the managing partner of Enterprise Knowledge Partners, LLC (EKP). EKP has three practice areas: Security and Compliance, E-Discovery, and Technology Strategy. EKP provides advisory services to counsel and information technology professionals in security, enterprise risk management and e-discovery technology, processes, and data mapping, and are vetted expert witnesses.

Sean Hora is a manager with KPMG's Forensic Technology Services practice and has an MS in Software Systems from the University of Saint Thomas (2006). Sean works with corporate clients on e-discovery projects and technology implementations.

Anna Horning Nygren is an attorney with the law firm Lockridge Grindal Nauen in Minneapolis. Her practice includes employment law, antitrust law, and real estate law. Ms. Horning Nygren has presented on social media discovery and proportionality in e-discovery, and has experience with electronic discovery both from the plaintiff's and the defendant's perspective.

Renee Jackson is the General Counsel of The Dolan Company and is a regular speaker on electronic discovery and other legal and strategic issues.

William Lutz is Assistant General Counsel at PBH Marine Group, LLC, where he manages litigation in state, federal, and international jurisdictions. He previously served as Vice President–Litigation at Genmar Holdings, Inc., and as Acting Assistant General Counsel for Litigation at Honeywell International. He also litigated state and federal cases as a complex litigation attorney at Robins, Kaplan, Miller & Ciresi, LLP.

Daniel Prokott is a partner in the Labor and Employment Group of Faegre Baker Daniels LLP and is past co-chair of Faegre Baker Daniels' E-Discovery Working Group. Dan specializes in advising businesses regarding complex workplace matters, which includes advising clients regarding records management practices and the preservation of information in connection with legal disputes.

Amey Schnabel is a paralegal at Lindquist & Vennum LLP. She has extensive experience managing cases involving e-discovery in law firms, and has previously advised law firm clients with their data forensics and e-discovery needs as a technology vendor.

Nforsi Moutchia is in litigation support management at Bowman and Brooke LLP, consulting on e-discovery as well as data reporting and analytics. He has extensive experience in the litigation support field and is a Certified E-Discovery Specialist with a graduate degree in management of information systems.

I. INTRODUCTION

The duty to preserve requires a party to identify, locate, and maintain electronically stored information (ESI) that is relevant to specific and identifiable or reasonably anticipated litigation.[1]

A "litigation hold" is the act of holding from destruction those records that are the subject of pending or potential litigation.[2] It requires a corporate defendant to suspend its routine document retention/destruction policy and to put in place a "litigation hold" to ensure the preservation of documents.[3]

Preservation of all client information sources is not required. Rather, client data sources must be evaluated to determine whether they likely contain *relevant* or *discoverable* evidence. The goal of this publication is to educate Minnesota attorneys on specific types of data sources to consider when identifying where potentially

1. *See* Multifeeder Tech., Inc. v. British Confectionery Co., No. 09-1090 (JRT/TNL), 2012 WL 4135848 (D. Minn. Sept. 18, 2012); *see also* MINN. R. CIV. P. 37.05 advisory committee's comment (2007 amendment) ("The good-faith part of this test is important and is not met if a party fails to take appropriate steps to preserve data once a duty to preserve arises.").

2. *See In re* Napster, Inc. Copyright Litig., 462 F. Supp. 2d 1060, 1070 (N.D. Cal. 2006) (citing Zubulake v. UBS Warburg L.L.C., 220 F.R.D. 212, 218 (S.D.N.Y. 2003)).

3. *Id.*

relevant information exists and cost-effective options currently being utilized by Minnesota litigants to preserve relevant data.

Proportionality in electronic discovery is important to this analysis as well.[4] Are available preservation options and methods disproportionately expensive or burdensome? In 2010, the Sedona Conference issued an important Commentary on Proportionality in Electronic Discovery.[5] The Conference issued six guiding "Principles" to assist with the decision whether to preserve and discover electronically stored information:

1. The burdens and costs of preservation of potentially relevant information should be weighed against the potential value and uniqueness of the information when determining the appropriate scope of preservation.
2. Discovery should generally be obtained from the most convenient, least burdensome, and least expensive sources.
3. Undue burden, expense, or delay resulting from a party's action or inaction should be weighed against that party.
4. Extrinsic information and sampling may assist in the analysis of whether requested discovery is sufficiently important to warrant the potential burden or expense of its production.
5. Nonmonetary factors should be considered when evaluating the burdens and benefits of discovery.
6. Technologies to reduce cost and burden should be considered in the proportionality analysis.[6]

On February 4, 2013, the Minnesota Supreme Court issued an Order Adopting Amendments to the Rules of Civil Procedure and General Rules of Practice Relating to the Civil Justice Reform Task Force.[7] These rules, effective July 1, 2013,[8] include amendments to

4. *See* FED. R. CIV. P. 26(b)(2)(B) (limiting discovery of ESI "from sources that the party identifies as not reasonably accessible because of undue burden or cost"); MINN. R. CIV. P. 26.02(b) (stating a general mandate that discovery "must comport with the factors of proportionality").

5. The Sedona Conference, *The Sedona Conference Commentary on Proportionality in Electronic Discovery*, 14 SEDONA CONF. J. 155 (2013).

6. *Id.* at 157.

7. Order Adopting Amendments to the Rules of Civil Procedure

Minnesota Rules of Civil Procedure 1 and 26.02(b),[9] and specifically require a showing of proportionality prior to conducting discovery:

Rule 1:

> It is the responsibility of the court and the parties to examine each civil action to assure that the process and the costs are proportionate to the amount in controversy and the complexity and importance of the issues. The factors to be considered by the court in making a proportionality assessment include, without limitation: needs of the case, amount in controversy, parties' resources, and complexity and importance of the issues at stake in the litigation.[10]

Rule 26.02:

> **(b) ~~In General~~ Scope and Limits.** Discovery must be limited to matters that would enable a party to prove or disprove a claim or defense or to impeach a witness and must comport with the factors of proportionality, including without limitation, the burden or expense of the proposed discovery weighed against its likely benefit, considering the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the action, and the importance of the discovery in resolving the issues. Subject to these limitations, ~~P~~parties may obtain discovery regarding any matter, not privileged, that is relevant to a claim or defense of any party, including the existence, description, nature, custody, condition and location of any books, documents, or other tangible things and the identity and location of persons having knowledge of any discoverable matter. Upon a showing of ~~For~~ good cause and proportionality, the court may order discovery of any matter relevant to the subject matter involved in the action. Relevant information sought need not be admissible at the trial if the discovery appears reasonably

(Minn. Feb. 4, 2013), *available at* http://mn.gov/lawlib//archive/supct/1302/ORADM108051-020413.pdf.

8. *Id.* at 1.
9. *Id.* at 2, 7.
10. *Id.* at 2.

> calculated to lead to the discovery of admissible evidence.[11]

The Court further amended Rule 26.02(b) to require the trial court's analysis to consider "proportionality" when determining whether to allow discovery of electronically stored information identified by a party as not reasonably accessible because of undue burden or cost:

> (2) Limits on Electronically Stored Evidence for Undue Burden or Cost. A party need not provide discovery of electronically stored information from sources that the party identifies as not reasonably accessible because of undue burden or cost. On motion to compel discovery or for a protective order, the party from whom discovery is sought must show that the information is not reasonably accessible because of undue burden or cost. If that showing is made, the court may nonetheless order discovery from such sources if the requesting party shows good cause and proportionality, considering the limitations of Rule 26.02(b)(3). The court may specify conditions for the discovery.[12]

Proportionality must be considered during the preservation stage.[13] Therefore, *for the analysis presented below, assume that the needs of the case, the amount in controversy, the parties' resources, and complexity and importance of the issues at stake in the litigation outweigh the cost and burden of preserving the ESI located in the locations we discuss.*

So, when confronted with new litigation and a technology-dependent client, what are the data sources where discoverable ESI may lie, and what steps should be taken to preserve it?

II. FORENSICALLY SOUND COLLECTION AND ADMISSIBILITY

When litigators need to preserve and collect ESI for analysis to be used as evidence in litigation, care must be taken to ensure that an accurate copy of the information is collected.[14] The method by

11. *Id.* at 7.

12. *Id.* at 8.

13. *See* The Sedona Conference, *The Sedona Conference Commentary on Preservation, Management and Identification of Sources of Information That Are Not Reasonably Accessible*, 10 SEDONA CONF. J. 281, 285 fig.1 (July 2008).

14. *See* The Sedona Conference, *The Sedona Conference Commentary on Achieving Quality in the E-Discovery Process*, 10 SEDONA CONF. J. 299, 302 (2009).

which the collection is performed will determine whether an admissible copy of the information has been obtained.

There are some methods of copying information whereby metadata may be altered in the collection (copying) process. Metadata is information about the document itself, and is like a fingerprint—it is different for every document (or version of a document) created. If collection methods are used that alter the types of metadata that are relevant to the case or necessary for proper analysis and use of the ESI, the collected information is no longer considered to be an *accurate copy* of the original, and may be inadmissible as evidence. *Forensically*[15] *sound procedures* are defined as those "used for acquiring electronic information in a manner that ensures it is 'as originally discovered' and is reliable enough to be admitted into evidence."[16] This does not, however, necessarily mean that entire computer hard drives or other media need to be forensically imaged[17] in every case—perhaps only

What is Metadata?

"Meta" is a term used for "about." "Metadata" is data about the information in an electronic file.

Metadata typically is generated by the person creating electronically stored information and also by the computer application that is being used to create the ESI.

While various forms of ESI can contain dozens or even hundreds of metadata values, some examples of common metadata value include: to, from, cc, bcc, subject, author, title, date created, date last modified, etc.

15. *Forensic* is defined as "relating to or dealing with the application of scientific knowledge to legal problems." MERRIAM-WEBSTER'S COLLEGIATE DICTIONARY 456 (11th ed. 2003). As it pertains to litigation involving ESI, *computer forensics* is defined as "the discipline that combines elements of law and computer science to collect and analyze data from computer systems, networks, wireless communications, and storage devices in a way that is admissible as evidence in a court of law." U.S. COMPUTER EMERGENCY READINESS TEAM, COMPUTER FORENSICS 1 (2008), *available at* http://www.us-cert.gov/sites/default/files/publications/forensics.pdf; *see also EDRM Collection Standards*, EDRM, http://www.edrm.net/resources/standards/collection (last updated Oct. 31, 2012).

16. *Forensically Sound Procedures,* EDRM, http://www.edrm.net/resources/glossary/f/forensically-sound-procedures (last visited Nov. 16, 2013).

17. *See* Madihah Mohd. Saudi, *An Overview of Disk Imaging Tool in Computer Forensics*, SANS § 3 (2001), http://www.sans.org/reading-room/whitepapers/incident/overview-disk-imaging-tool-computer-forensics-643 (defining "imaging"

portions of them will be copied, but the copying process (i.e., collection) will need to be handled using forensically sound procedures.[18]

The level at which ESI is stored is the main criterion used to determine the proper method for acquiring a forensically sound version of the information. Servers, network shares, cloud storage areas, desktop and laptop computers, tablets and other mobile devices, as well as countless other places where ESI is located each have discrete procedures of collection that will produce a forensically sound version of the information that they store.

III. DEFENDING THE ESI COLLECTION PROCESS

Each step in the process of acquiring and collecting applicable ESI should be documented to prevent and defend against potential future challenges to its authenticity. As the Sedona Conference suggests, "As with all evidentiary material, when working with ESI as tangible evidence, it should be handled with care to ensure that there is a defensible chain of custody."[19]

Proportionality is a factor when determining the level of documentation used for any given ESI collection. While some matters will involve the use of a chain of custody form for every piece of evidence, other matters may use a documentation approach that outlines the collection and evidence handling processes used universally for all evidence in a matter. The latter documentation approach would include, for example, the process used to locate, collect, move, and check the integrity of e-mail data.

as a "[t]erm given to creating [a] physical sector copy of a disk and compressing this image in the form of a file"). This file is said to be a *forensic copy* when it is "an exact copy of an entire physical storage media, including all active and residual data and unallocated or slack space on the media." *The Sedona Conference Glossary: E-Discovery & Digital Information Management*, SEDONA CONF. 23 (Sept. 2010) [hereinafter *Sedona Conference Glossary*], https://thesedonaconference.org/download-pub/471.

18. Collection of ESI in a manner that does not alter relevant metadata creates a "forensically sound" copy of the information. For collected data to be described as forensically sound, the use of collection methods that are precise, repeatable, observable, and testable must be employed to prevent the altering of relevant metadata or the data themselves. *See generally* The Sedona Conference, *supra* note 14.

19. The Sedona Conference, *The Sedona Conference Commentary on ESI Evidence & Admissibility*, 9 SEDONA CONF. J. 217, 232 (2008).

This level of documentation should also include a section on how any individually collected ESI can be determined to be a forensically sound copy and is typically reviewed by counsel at the onset of a matter so that those representing a client can understand and ask the right questions during the collection process.

What Should Be Included in a Chain of Custody Form?

Chain of Custody forms should, at a minimum, include:

- Description of the evidence
- Unique identifier of the evidence
- Who has controlled the evidence
- Date and times of the handoffs

Evidence Description	*Evidence Serial Number*
Laptop	1234567abc

Date	*Time*	*From*	*To*
1/1/12	8:00am	Tech 1	Attorney 1
1/2/12	8:00am	Attorney 1	Outside Counsel 1

Other items that are typically found on Chain of Custody forms include bar label IDs, shipping vendor, shipment tracking numbers, and signatures for those people giving and taking possession of the evidence.

Is "Forensically Sound" Necessary?

Litigation may not require that all metadata for each of the relevant documents remain perfectly intact. According to the Department of Justice, "The standard for authenticating computer records is the same as for authenticating other records."[20] Witness testimony, comparisons, reasonable analysis of existing metadata fields, or other criteria may be sufficient—and sometimes preferable—to more stringent and scientific means of authenticating documents collected.[21] Indeed, the Sedona Conference has declared:

> Civil litigation should not be approached as if information systems were crime scenes that justify forensic investigation at every opportunity to identify and preserve every detail. Theoretically, a party could preserve the contents of waste baskets and trash bins for evidence of statements or conduct. Yet, the burdens and costs of those

20. COMPUTER CRIME & INTELLECTUAL PROP. SECTION, CRIMINAL DIV., U.S. DEP'T OF JUSTICE, SEARCHING AND SEIZING COMPUTERS AND OBTAINING ELECTRONIC EVIDENCE IN CRIMINAL INVESTIGATIONS 198 (3d ed. 2009), *available at* http://www.justice.gov/criminal/cybercrime/docs/ssmanual2009.pdf.

21. *See* FED. R. EVID. 901(b); MINN. R. EVID. 901(b).

> acts are apparent and no one would typically argue that this is required. There should be a similar application of reasonableness to preservation of electronic documents and data.[22]

Accordingly, the cooperation of counsel and a reasonable application of proportionality of the cost and extent of the collection of relevant ESI to the case at hand are vital for the fair application of justice to all involved parties. The sections that follow discuss the various types of ESI, objectives to consider in the course of preservation, and various technical aspects associated with such preservation efforts.

IV. COST-EFFECTIVE PRESERVATION TECHNOLOGIES

A. Custodian Data on Computers, Laptops, and Servers

The most effective method for preserving data for a particular custodian is to apply a "custodian-based view to the data."[23] Counsel must consider these questions: What role does this witness have in the litigation? Is this person's role "key"? What relevant or discoverable documents or files did the witness create and where are they kept?

In other words, look for the potential locations of data from the view of the person who created the information. Determine the case's need for preservation of the witness's data in any or all locations, and ask the witness where files are stored. This should provide the basis of a written plan for locating, preserving (imaging), and collecting the witness's data. Conferring with the witness will indicate whether the files sought are stored on the witness's computer workstation, a laptop, another location such as a server on an employer network, or e-mail accounts.

22. THE SEDONA CONFERENCE, THE SEDONA PRINCIPLES ADDRESSING ELECTRONIC DOCUMENT PRODUCTION 34 (Jonathan M. Redgrave et al. eds., 2d ed. 2007).

23. The Sedona Conference, *supra* note 14, at 317.

1. *Internal and External Hard Drives, Media Storage Devices, and Personal Drives on Servers*[24]

If files are stored on the individual's computer, the employer's IT department or an outside computer forensics technologist can forensically copy only those files that are relevant to the case or image the custodian's entire hard drive to preserve its contents. It may be discovered that the custodian utilizes an external hard drive or other external media storage devices that likely have relevant information that also should be imaged. Employees frequently utilize "personal" drives on corporate servers, and it may be discovered that relevant files are kept there. The employer's IT department or the outside technologist can and may need to create images of portions of these potential locations of relevant data.

Objectives to Consider

Active Files and "Unallocated Space"

- Technologists are often asked to image the entire contents of a custodian's personal computer's hard drive. This sounds like a simple task and a clear instruction, but the attorney needs to understand options available to the technologist, and the attorney must give clear direction on precisely *what* to image.
- A technologist can make a complete mirror image of the custodian's hard drive or apply filters to collect only active files that match the specified criteria. "Active files" are those files that can be readily accessed by the computer's operating system, such as program files, operating system files, and user-created files.[25] "Unallocated space" is the unused area on a hard drive where deleted files reside until the space they occupy is needed to store "active" files.[26]
- A potential pitfall concerns the possible failure to collect deleted files housed in "unallocated space." It must be determined whether it will be acceptable to the opposing party

24. For sample preservation and collection technologies relating to hard drives see Appendix A, *infra*.

25. *See* THE SEDONA CONFERENCE, *supra* note 22, at 45 (discussing the different methods by which data can be removed from active files).

26. *Appendix A: Technical Primer*, TECHNET (Sept. 16, 2005), http://technet.microsoft.com/en-us/library/bb457138.aspx.

and the court to image only specific user-created files. If the litigation demands the preservation of deleted (in addition to active) files, these reside on the computer unknown to the user. In that case, it is important to capture inactive files residing in the unallocated space on the hard drive during the collection process.

- The technologist makes choices during the imaging process that impact whether files in the "unallocated space" are collected. The technologist should be instructed whether to make a forensic collection of only active files from a custodian's hard drive, or to include, in addition, the unallocated space.
- The risk is this: it might be impossible later to go back and collect the deleted files available in the unallocated space because they have been overwritten by new active files saved to the custodian's hard drive. To avoid spoliation claims and potential sanctions, the written preservation and collection plan needs to clearly define the collection methodology to avoid this potential pitfall.

Corporate Server Backups of Custodian Files

- Frequently, electronic documents and corporate e-mail are automatically backed up to an easily accessible archive as long as the custodian's computer is connected to the company's network. This enables the technologist to collect a custodian's files without the custodian's knowledge, and without the need to retrieve and image all of the custodian's actual computer.[27]
- However, it is not always true that the archives exist or that they have backed up the custodian's relevant files. It depends on the company's policies, its technology infrastructure, and the custodian's habits. Inquiries about the structure of the company's electronic support system are essential when creating the written preservation and collection plan and determining whether the custodian maintains relevant files in locations or on devices that are not backed up.

27. *See* The Sedona Conference, *supra* note 13, at 19. However, be mindful of the "unallocated space" issue discussed in the preceding section. *See supra* text following note 26.

Technologist Competence

- When creating a written plan for collecting custodian data, be aware of and appreciate the potential pitfalls that can occur during the preservation phase and discuss them with the technologist imaging the data as outlined in the plan.
- Engage a competent technologist who is aware of methods to preserve data without changing metadata, understands the potential issues outlined above, and can serve as a competent witness to assist throughout the process.
  - This may or may not be an employee working within the client's IT department.
  - Counsel must judge the competence of corporate IT employees and determine whether they are familiar with preservation and imaging methods that do not change the files or their metadata.
  - Moreover, even if competent, the IT employee may be required to testify as to the methods employed and work completed. The client may choose to avoid burdening its in-house IT staff by retaining an outside vendor technologist to image the data outlined in the plan.
  - Counsel must determine what steps will be taken to preserve and collect data and who will do that work while remaining within the confines of the electronic discovery budget.

2. E-Mail

Electronic Mail, commonly referred to as e-mail or email, involves various technological systems for sending and receiving written electronic messages, attachments, and other content across internal and external networks. Although newer forms of electronic communication have proliferated with the rise of social media and mobile devices, e-mail remains ubiquitous and is an important source of ESI to address. Like other forms of ESI, e-mail messages and related content are discoverable and subject to the duty to preserve.

For example, it is common to find mass quantities of e-mail in enterprise, personal, and cloud-based computer systems. Content typically includes messages, calendaring, contacts, tasks, and a wide range of file attachments including documents, spreadsheets, presentations, scanned items, photos, and more. Due to the ease of

including multiple recipients as well as various storage and backup options, e-mail can often be located in multiple storage locations and contain a substantial amount of duplication. Common storage locations include e-mail servers and network storage, local hard drives, portable storage devices (e.g., external flash drives and hard drives), mobile devices (such as smartphones and tablets), backup media, e-mail archive systems, document management systems, intranets, extranets, third-party enterprise and personal systems, and cloud-based systems. In addition, "legacy" computer systems that are no longer used by the organization may contain relevant e-mail content.

Adding to this complexity, a number of e-mail systems are configurable regarding content retention.[28] For example, e-mail systems can often be set to purge or automatically delete messages based on age. This can introduce significant preservation challenges throughout and across matters. In addition, user account size quotas or limits may be in place to limit the volume of e-mail retained.

It is important to understand the normal operation of these systems. Once the duty to preserve has arisen, relevant and responsive ESI may be at risk of destruction if appropriate steps are not taken in a timely manner to identify and preserve it. Courts have issued sanctions for various failures relating to e-mail, including monetary sanctions, adverse inferences, and dismissal.[29]

Objectives to Consider

- Proper and timely preservation of e-mail is a significant concern, with courts finding spoliation or mandating preservation by organizations and individual account owners.

28. *See* The Sedona Conference, *The Sedona Conference Commentary on Email Management: Guidelines for the Selection of Retention Policy*, 8 SEDONA CONF. J. 239, 241 (2007).

29. *See* Merck Eprova AG v. Gnosis S.P.A., 901 F. Supp. 2d 436, 443 (S.D.N.Y. 2012); Voom HD Holdings L.L.C. v. Echostar Satellite L.L.C., 939 N.Y.S.2d 321, 331–33 (App. Div. 2012); 915 Broadway Assocs. L.L.C. v. Paul, Hastings, Janofsky & Walker, L.L.P., 950 N.Y.S.2d 724, 2012 WL 593075, at *13 (Sup. Ct. 2012); *see also* Disability Rights Council of Greater Wash. v. Wash. Metro. Area Transit Auth., 242 F.R.D. 139 (D.D.C. 2007).

- Counsel should work with the relevant persons to understand who is responsible for and knowledgeable of these systems and their operation. The inquiry can often include IT professionals as well as the relevant custodians.[30] Seek out technical assistance, as necessary, to understand the applicable systems and locations, their operation, e-mail retention periods, and other information retrieval details.[31]
- It is critically important to identify whether auto-delete or purge mechanisms are in effect and to disable or suspend them for the relevant custodians.[32] Consider informing the appropriate IT personnel to disable or suspend any auto-deleting functions. It may also be necessary to adjust or suspend size quotas and limits for the relevant custodian accounts. Inform IT staff with clear directions and confirm their understanding, including follow-ups. Consider establishing communication and escalation channels.
- Likewise, it is important to follow up and monitor such suspension throughout the matter. This can impact decisions

30. *See* Green v. Blitz U.S.A., Inc., No. 2:07-CV-372 (TJW), 2011 WL 806011, at *3–4 (E.D. Tex. Mar. 1, 2011) (stating employee was "solely responsible for searching for and collecting documents relevant to ongoing litigation" and employee "did not institute a litigation-hold of documents, do any electronic word searches for emails, or talk with the IT department regarding how to search for electronic documents"). In *Green*, the court ordered monetary sanctions of $250,000 and required that defendant provide said order to plaintiffs "in every lawsuit it has had proceeding against it, or is currently proceeding against it, for the past two years." *Id.* at *10–11. The court also ordered the defendant to file the order in every case for five years. *Id.* at *11.

31. *See generally* BARBARA J. ROTHSTEIN ET AL., MANAGING DISCOVERY OF ELECTRONIC INFORMATION (2d ed. 2012), *available at* http://www.fjc.gov/public/pdf.nsf/lookup/eldscpkt2d_eb.pdf/$file/eldscpkt2d_eb.pdf (providing a question and answer guide for judges on managing e-discovery).

32. *See* Apple Inc. v. Samsung Elecs. Co., 881 F. Supp. 2d 1132, 1150–51 (N.D. Cal. 2012) (granting adverse inference instruction for defendant's failure to disable e-mail system's auto-delete function and to appropriately follow up with litigation hold recipients). Interestingly, as both parties failed to preserve ESI, they and the court subsequently agreed not to include the adverse inference instruction for the jury. *See* Apple Inc. v. Samsung Elecs. Co., 888 F. Supp. 2d 976, 1000 (N.D. Cal. 2012).

regarding "preservation in place"[33] versus "collection to preserve"[34] the ESI.

- Because custodians may be spread across multiple e-mail servers and locations with differing retention policies and settings, it is important to identify custodians early.
- Contemplate including former employees' e-mail in the inquiry and preservation efforts. Due to layoffs, reductions in force, and the like, it is not uncommon for former employees' ESI to receive less attention.
- Prioritize and triage the various storage and preservation risks. Since e-mail content may be dispersed across multiple systems and locations, it may be difficult and even impractical to attempt to address every conceivable location concurrently.
- Consider the example of a key custodian's e-mail content stored on a local hard drive. It may contain unique ESI if the custodian archived e-mail content to the drive. Thus it presents a single point of failure should the hard drive fail or suffer a mishap after the duty to preserve has arisen, but before the ESI has been collected from the drive.[35]
- Consider engaging digital forensics experts or mobile forensics examiners to ensure proper preservation and retrieval of e-mail message evidence in those situations warranting it—for example, where unique e-mail content resides on mobile devices or where foul play is suspected.[36]

33. Maintaining files on their original media, but suspending deletion.

34. Copying or exporting the data from their original media or data source to preserve it from potential modification or deletion.

35. *See* Multifeeder Tech., Inc. v. British Confectionery Co., No. 09-1090 (JRT/TNL), 2012 WL 4135848, at *4–6, *10 (D. Minn. Sept. 18, 2012) (holding the defendant in contempt due to use of wiping software and deletion of a .PST file by key employees, as well as the failure to reveal the existence of encrypted data, and ordering sanctions of $600,000 to be paid to the plaintiff and $25,000 to be paid to the court).

36. *See, e.g.*, Multifeeder Tech., Inc. v. British Confectionery Co., No. 09-1090 (JRT/TNL), 2012 WL 4128385, at *6 (D. Minn. Apr. 26, 2012), *adopted in part, rejected in part*, 2012 WL 4135848, at *1 (noting that the court ordered the exclusive court-appointed, third-party expert to prepare a detailed log of all relevant documents after the expert discovered in the course of imaging and searching the defendant's computers that certain files had been purged or deleted shortly before the expert arrived at the defendant's facility).

- Determine who has possession, custody, or control of third-party-hosted systems, such as outsourced, personal, or cloud-based e-mail systems.
- Keep in mind that e-mail content may be stored in a compressed format, such as in .PST files.[37] The data often

37. There are two main types of data files used by Microsoft Outlook with their filenames ending in .PST and .OST. A .PST file is the Microsoft Outlook data file that contains a user's local Outlook data, such as e-mail and attachments, calendar items, contacts, tasks, and notes. The .PST file is also referenced by different descriptions, such as "Personal Storage Table," "Personal Folder File," "Personal Folders," or "Personal Archive." In normal operation, there can be one or more .PST files on a user's local drive, as it can be used as the primary storage file (typically for non-Exchange accounts) and also as separate archive files.

In contrast, an .OST file is utilized when the e-mail account is hosted on a Microsoft Exchange server and the user wants to be able to work off-line (i.e., when his or her computing device is disconnected from the e-mail server). An .OST file is also referred to as an "Off-line Storage Table," "Off-line Storage Folder," "Off-line Storage," or "Off-line Folders." The contents of the .OST file are periodically synchronized with the user's mailbox account on the Exchange server when Outlook is connected to the server. By design, this results in duplication of data between the Exchange server and the user's system as the .OST file holds off-line copies of the synchronized content from the Exchange server.

.PST and .OST files should not be confused. Both .PST and .OST files can coexist on the same local drive, and they store information in different formats. Indeed, there are numerous third-party .OST-to-.PST conversion software programs available.

In environments where IBM Notes (formerly IBM Lotus Notes) is used, the main e-mail data files typically include one or more files ending in .NSF (Notes Storage Facility) and an .ID file. The .NSF file is roughly analogous to the .PST/.OST files, while the user ID file (ending in .ID) enables access to the contents of the .NSF file when security is enabled. Thus, it's often very important to preserve the user's .ID file along with the .NSF file(s). If the user's .ID file is not available, consider consulting with the appropriate Notes IT professional to inquire whether the .ID file can be recovered via other methods.

It's also important to note that these Outlook and Notes data files can be present in locations other than a computer's local drive, such as when they are copied to external or network storage.

For further information on .PST and .OST files applicable to Outlook 2010, including potential file locations, see *Introduction to Outlook Data Files (.PST and .OST)*, MICROSOFT, http://office.microsoft.com/en-us/outlook-help/introduction-to-outlook-data-files-pst-and-ost-HA010354876.aspx (last visited Dec. 23, 2013). For further information on .PST and .OST files applicable to Outlook 2013, including potential file locations see *Introduction to Outlook Data Files (.PST and .OST)*, MICROSOFT, http://office.microsoft.com/en-us/outlook-help/introduction-to-outlook-data-files-pst-and-ost-HA102749465.aspx (last visited Dec. 23, 2013).

expand when processed or extracted into individual messages and files. This can result in significantly underestimating the volume, time, and cost when planning and budgeting various ESI and e-discovery-related tasks, such as processing, indexing, culling, loading, reviewing, and producing the data.

- Where e-mail archiving systems are utilized, a trap for the unwary is the collection of message "stubs" instead of the actual message. The practice of "stubbing" involves moving e-mails from the custodian's electronic mailbox to a new location (e.g., an e-mail archive), while replacing the original e-mail in the custodian's mailbox with a small placeholder message. This small placeholder, or stub, points to the new location of the e-mail. It is helpful to think of a stub as a shortcut or link that contains the information to point to the actual content. It may also present searching challenges as the full content of the e-mail is no longer present in the custodian's mailbox, but rather, in the e-mail archive system.
- Consider potential challenges in addressing some content types, such as voicemail attached to e-mail. This may occur when unified messaging solutions are used to send voicemail audio messages to users via e-mail attachments. As audio files, their content cannot be searched or found by text-based keyword searches, for example. Similarly, scanned document attachments without OCR text[38] are not searchable as they are essentially images.

38. Optical character recognition (OCR) is "the conversion of a scanned document into searchable text." *OCR*, EDRM, http://www.edrm.net/resources/glossaries/glossary/o/ocr (last visited Nov. 3, 2013).

> When a paper document is scanned into a computer, an image is created. The computer does not recognize the characters of the document as text until OCR software converts the image into text. OCR systems vary widely in the accuracy of their conversion. Even seemingly high accuracy rates can, however, still result in significant numbers of words being misrepresented. A 99% accuracy, for example, would still result in one word out of 20 being misspelled.

Id. In the case of PDF file attachments, it is common to have some PDFs with OCR text included, but also PDFs that contain only the scanned image without any corresponding text. While the former are generally searchable, the latter are not, at least not without going through an OCR process.

- Consider the impact that various data privacy and data protection laws may have on accessing and collecting e-mail in various countries.[39]

B. Cell Phones, Smart Phones, and Tablets

Very few people live or work today without cell phones or smart phones and, increasingly, tablets. In fact, wireless industry data show more wireless subscribers in the United States than people—the number of subscribers is equal to 102% of the population.[40] And 56% of all cell phones in the United States are now estimated to be smart phones.[41]

Cell phones, smart phones, and tablets are rich with ESI, including:

- Subscriber and equipment identifiers, device characteristics;
- Phone address book (contacts);
- Appointments and calendar;
- Dialed, received, and missed call logs;
- Text messages (SMS);
- Electronic mail (e-mail);
- Photographs;
- Audio and video recordings;
- Voice memos;
- Multi-media messages (MMS);
- Instant messaging and web browsing activities;
- Electronic documents;
- Date and time stamps, language, and other settings;
- Geolocation information (Geotags);
- EXIF data from onboard camera snapshots and video recordings;

39. *See The Sedona Conference International Principles on Discovery, Disclosure & Data Protection*, SEDONA CONF., at i–ii (public comment version, Dec. 2011), https://thesedona conference.org/download-pub/495 (developing a model protective order and a model data process and transfer protocol).

40. *Wireless Quick Facts*, CTIA, http://www.ctia.org/your-wireless-life/how-wireless-works/wireless-quick-facts (last visited Nov. 17, 2013).

41. Aaron Smith, *Smartphone Ownership 2013*, PEW INTERNET (June 5, 2013), http://www.pewinternet.org/Reports/2013/Smartphone-Ownership-2013/Findings.aspx.

- Access point data from wi-fi logins and activity;
- Application identity, usage, logs, and user data.

It is therefore no surprise that cell phone and smart phone data are clearly within the definition of ESI, and recognized by the Seventh Circuit Electronic Discovery Pilot Program as a standard category and focus of e-discovery.[42] Cell and smart phone ESI is often a high-profile feature in litigation, both civil and criminal, and the Sedona Conference has specifically addressed admissibility issues for text messages, which are a primary source of ESI from cell and smart phones.[43]

Gathering and preserving mobile phone (cell and smart phone) data has been generally done by one or two approaches: (1) subpoenaed through the mobile service provider's business records and (2) data mined from the device itself.

1. Business Records of the Service Provider

Under the traditional approach, business records are sought from the mobile phone provider by subpoena. It is a cumbersome

42. *See E-Discovery Practical Guide: What Everyone Should Know About the Mechanics of eDiscovery*, SEVENTH CIRCUIT ELECTRONIC DISCOVERY PILOT PROGRAM 8 (Apr. 6, 2011), http://www.discoverypilot.com/sites/default/files/MLS_7Circuit_Slides.pdf.

> The Seventh Circuit Electronic Discovery Pilot Program Committee ("Committee") was formed in May 2009 to conduct a multi-year, multi-phase process to develop, implement, evaluate, and improve pretrial litigation procedures that would provide fairness and justice to all parties while seeking to reduce the cost and burden of electronic discovery consistent with Rule 1 of the Federal Rules of Civil Procedure. To that end we brought together the most talented experts in the Seventh Circuit from all sectors of the bar, including government lawyers, plaintiffs' lawyers, defense lawyers, and in house lawyers from companies with large information systems, as well as experts in relevant fields of technology. These experts developed Principles Relating to the Discovery of Electronically Stored Information ("Principles"), and a Standing Order by which participating judges implement the Principles in the Pilot Program test cases.

About the Committee, SEVENTH CIRCUIT ELECTRONIC DISCOVERY PILOT PROGRAM, http:// www.discoverypilot.com/about-us (last visited Nov. 17, 2013).

43. *See* The Sedona Conference, *supra* note 19, at 222.

process usually requiring a prompt letter of preservation, then a subpoena, court order, or search warrant.

Time is critical when using this approach, because data stored by cell phone providers are fleeting and can evaporate through normal business policies in just a few days. In urgent situations, subpoenas can be sought and issued for provider-held ESI on cell phones or smart phones.[44]

With this business records approach, there are privacy hurdles that may protect a provider from this discovery under the Stored Communications Act (SCA).[45] Case law has increasingly restricted the ability of civil litigants to obtain information this way on the grounds that Congress intended only law enforcement to make use of the SCA to obtain evidence from cell or smart phones.[46]

Data received through this method can also be incomplete. The ESI content one receives from a service provider is usually limited in scope and similar to what is provided in a cellular phone bill. For example, subscriber data, call detail records (date/time of call, who called/received, duration, and optionally originating cell phone tower and terminating tower), and text message logs (date/time of message, but not its content) are generally provided. Key information that may be missing includes:

- Phone address book;
- Photos;
- Videos;
- Audio clips and voice memos;
- Ring tones;
- Text messages;
- Anything deleted.

44. Humphrey v. Sallie Mae, Inc., No. 3:10-cv-01505-JFA, 2010 WL 2522743, at *2 (D.S.C. June 17, 2010) (ordering discovery of cell phone records from Verizon on expedited basis).

45. 18 U.S.C. §§ 2701–2712 (2012).

46. *See* O'Grady v. Superior Court**,** 44 Cal. Rptr. 3d 72, 89 (Ct. App. 2006) (holding only law enforcement can use the SCA to obtain ISP information; civil litigants must get information through civil subpoena to sender of e-mails); *see also In re* Subpoena Duces Tecum to AOL, L.L.C., 550 F. Supp. 2d 606, 611 (E.D. Va. 2008) (relying on *O'Grady*).

2. *Data Mining from the Phone's Handset and Memory Card*[47]

"Data mining" presents the new frontier for mobile phone data retrieval and analysis. Unlike the traditional business records approach, data mining may be more effective because it can bypass the need to go through the service provider and eliminates red tape, delay, and significant limitations.

Instead, data mining applies mobile forensic tools to the mobile phone itself, and case law demonstrates that forensic examination of a phone can be ordered so that a range of ESI including photos, texts, and e-mails can be extracted.[48]

Minnesota cases on mobile phone ESI range from orders to produce phones for examination by an expert in state district court,[49] to reversal of a homicide conviction by the Minnesota Supreme Court on grounds that the trial court erred in refusing to admit evidence from a cell phone that identified a legitimate alternative perpetrator.[50]

3. *Text Messages*[51]

Text messages, more formally referred to as "Short Message Service" (SMS), are "message[s] of up to 160 characters used to communicate with text over mobile networks."[52] "Multimedia Messaging Service" (MMS) is a "message that includes multimedia content such as pictures, video, or sounds used to communicate

47. For sample preservation and collection technologies relating to data mining phone memory cards, see Appendix B, *infra*.

48. *See* Robinson v. Jones Lang LaSalle Americas, Inc., No. 3:12-cv-00127-PK, 2012 WL 3763545, at *1 (D. Or. Aug. 29, 2012) (holding there was "no principled reason to articulate different standards for the discoverability of communications through text messages, e-mails, or social media," and ordered ESI to be produced); *Humphrey*, 2010 WL 2522743, at *2 (ordering discovery of cell phone records from Verizon on expedited basis); Smith v. Cafe Asia, 246 F.R.D. 19, 22 (D.D.C. 2007) (ordering plaintiff to produce images on phone); Moreno v. Ostly, No. A127780, 2011 WL 598931, at *8 (Cal. Ct. App. Feb. 22, 2011) (affirming sanctions for violating a court order by failing to produce a cell phone for inspection of text messages and e-mails).

49. *See* State v. Powers, No. 27-CR-11-9265 (Minn. Dist. Ct. Nov. 15, 2011).

50. *See* State v. Ferguson, 804 N.W.2d 586, 592–93 (Minn. 2011).

51. For sample preservation and collection technologies relating to text messages, see Appendix C, *infra*.

52. *Sedona Conference Glossary*, *supra* note 17, at 48.

over mobile networks."[53] SMS and MMS are globally accepted communications protocols that allow the transmission of messages between mobile subscribers and external systems like e-mail, paging, and voicemail systems.

Text messages present a relatively new area of applicable discovery rules. Like other methods of ESI, text messages are discoverable and subject to the duty to preserve. However, emerging technologies, like self-deleting texting applications, continue to challenge the accessibility and availability of such electronic information.

Objectives to Consider

- SMS evidence is particularly prevalent in criminal cases,[54] including automobile accidents,[55] as well as employment disputes.[56] It has also been relevant in various civil actions including negligence, insurance fraud, insider trading, consumer protection, contract, family law disputes, and debt collection.[57]
- Attorneys should work with the relevant persons to identify the custodians who use SMS technologies and determine which devices contain potentially relevant SMS evidence. The inquiry should extend to both employer-owned and personal devices.
  - Identify the applicable service providers (which can include common wireless carriers like Verizon, AT&T, and Sprint as well as other entities that provide corporate text messaging services, mass text messaging services, and marketing messaging services, among others).

53. *Id.* at 34.

54. *See, e.g.*, United States v. Finley, 477 F.3d 250, 260 (5th Cir. 2007) (upholding a search of a cell phone, which revealed text messages related to drug use and trafficking).

55. *See, e.g.*, State v. Small, No. 09AP-1175, 2010 WL 4323032, at *1 (Ohio Ct. App. Nov. 2, 2010) (noting that cell phone records indicated the defendant was texting when she hit a person with her car).

56. *See, e.g.*, City of Ontario v. Quon, 560 U.S. 746 (2010).

57. *See, e.g.,* The Sedona Conference, *supra* note 19, at 222 (discussing the admissibility of SMS evidence under the Rules of Evidence and two example cases (arbitration award dispute and bankruptcy dispute) where such evidence had been admissible).

- Identify other messaging applications in use. Examples include:
  - Secure messaging services like TigerTextPRO, which allow users to send HIPPA compliant text messages;
  - Self-deleting messaging services like TigerText or SnapChat, which allow users to send text messages that will be deleted from all servers and users' devices after a period of time selected by the user; and
  - Universal texting applications, which may provide unlimited text messaging, free text messaging, and instant messaging. Examples include:
    - Text Plus and Text Free, which provide free unlimited text messaging;
    - PingChat and KiK Messenger, which allow users to send instant messages;
    - Voxer, a Walkie Talkie application, which allows users to send instant audio, text, and photo messages; and
    - WhatsApp, which is a cross-platform mobile messaging service that works through the device's internet data plan.

Determine the scope of the duty to preserve, considering the service providers and applications in use. Identify any potential backup resources like iTunes backups for iPhones, Blackberry Desktop Messenger for Blackberries, or even iCloud for iPhones. Determine which nonparties, if any, may be within the party's control and thus subject to the duty to preserve. Contractual privity may be sufficient to establish "control."[58]

- Seek out technical assistance, as necessary, to understand the database structure, text message retention periods, and other information retrieval details. Engage digital forensic experts

58. *See* Flagg v. City of Detroit, 252 F.R.D. 346, 352–54 (E.D. Mich. 2008) (holding that Rule 34 of the Federal Rules of Civil Procedure requires disclosure of information within a party's "control," and the city's ability to withhold consent to disclose also meant that the city exercised sufficient "control" over the evidence to permit production, even though the evidence was in the possession of a third-party vendor).

and mobile forensic examiners to ensure proper preservation and retrieval of text message evidence when appropriate.

- Contemplate issuing both internal and external litigation holds. Internal holds should be issued to clients advising them to preserve and cease using any self-deleting applications. Other deletion scenarios include cell phone text message quotas or limits, user deletion, and improper SIM card usage. An external document hold may be necessary for applicable service providers.
- Beware of statutory hurdles that may interfere with the discoverability of text messages. The SCA[59] prohibits service providers from releasing contents of electronic communications unless certain consent requirements are met.[60] The SCA also divides electronic communications, like text messages, into two types: electronic communications systems (ECS) and remote computing services (RCS).[61] If a communication is an ECS, then lawful consent of either the sender or recipient is required. If the communication is an RCS (i.e., the service provider simply stores the communication), then the provider can release the information with the consent of the sender, recipients, or subscriber. There are then several exceptions to the SCA that allow for broader release of information.[62] Other federal statutes that may come into play include the Computer Fraud and Abuse Act,[63] the Telephone Consumer Protection Act,[64] Whistleblower Protection Act of 1989,[65] and the Communications Assistance for Law Enforcement Act.[66]
- Employee privacy issues may arise if employees have a reasonable expectation of privacy in communications that are sent on personally owned devices. Employers have a competing interest in preserving the confidentiality of

59. 18 U.S.C. §§ 2701–2712 (2012).
60. *See id.* § 2702.
61. *See id.*
62. *See id.* §§ 2701–2712.
63. 18 U.S.C. § 1030.
64. 47 U.S.C. § 227.
65. Pub. L. No. 101-12, 103 Stat. 16 (codified as amended in scattered sections of 5 U.S.C.).
66. 47 U.S.C. §§ 1001–1010.

corporate assets and other sensitive information stored on employee-owned devices.[67] Case law on this point continues to evolve.

- Admissibility of SMS evidence is within the sound discretion of the trial court and, like all other evidence, must be authenticated.[68] SMS evidence can be authenticated by direct proof or circumstantial evidence.[69] Although text messages include the telephone numbers of the originator and recipient, the devices (smartphone) or applications (e-mail accounts or other applications) that transmit the messages are not always used exclusively by the telephone number owner.[70] At least one court has declared that "more than mere confirmation that the number or address belonged to a particular person" is needed to authenticate SMS evidence and that circumstantial evidence "which tends to corroborate the identity of the sender" is required.[71]

C. Social Media

Social media involves various online technology tools that enable people to communicate easily via the Internet in almost real time to share resources and information.[72] Social media formats may include text, video, audio, and graphics.[73] Given its wide variety of forms and formats, there is not a consistent definition for social media, and it is therefore helpful to consider the typical features of social media content.[74] Specifically, social media content is commonly shared with others, interactive, accessed via the Internet,

67. *See* Flagg v. City of Detroit, 252 F.R.D. 346, 351–52 (E.D. Mich. 2008) (discussing the City of Detroit's interest as an employer in protecting information stored on city-issued text messaging devices).

68. *See* Pennsylvania v. Koch, 39 A.3d 996, 1005 (Pa. 2011) (stating that detective's description of how text messages were transcribed was not sufficient to establish the identity of the sender, which was essential for admissibility).

69. *Id.*

70. *See id.* at 1004–05.

71. *Id.* at 1005.

72. *See* The Sedona Conference, *The Sedona Conference Primer on Social Media*, 14 SEDONA CONF. J. 191, 196 (2013).

73. *Id.* at 194.

74. *Id.*

informal in tone, and involves personal commentary.[75] Popular examples of social media websites include Facebook, LinkedIn, Twitter, YouTube, and MySpace.[76] Social media is essentially a form of cloud computing, with social media data largely residing on various websites.[77] Moreover, because it provides a more instantaneous communication platform, typically with a larger audience as compared with e-mail, use of social media has grown exponentially.[78]

There is no question that social media evidence may be relevant to any number of legal disputes including, for example: misappropriation of trade secrets, unfair labor practices, employer monitoring of or ownership of employee social media accounts, consumer advertising, use of social media posts for making hiring decisions, and criminal matters. Parties increasingly seek social media evidence to support and defend litigation. Courts have made clear that the failure to properly preserve such evidence may subject the party in control to sanctions for spoliation.[79] It is important to note that this is a rapidly developing area of law and to the extent this outline provides guidance to Minnesota's practitioners, care should be taken to confirm such guidance with current legal practices, principles, and legal authority.

A significant issue is the use of social media by employees and their employers' efforts to monitor or otherwise control employee

75. *Id.*

76. *Id.*

77. *Id.* at 223.

78. *See* Alfonso Serrano, *The Social Media Explosion: By the Numbers*, FISCAL TIMES (Sept. 12, 2011), http://www.thefiscaltimes.com/Articles/2011/09/12/The-Social-Media-Explosion-By-the-Numbers.aspx#page.

79. *See* Gatto v. United Air Lines, Inc., No. 10-cv-1090-ES-SCM, 2013 WL 1285285, at *4 (D.N.J. Mar. 25, 2013) (concluding that a party's deletion of his Facebook profile raised an inference of spoliation); Lester v. Allied Concrete Co., Nos. CL08-150, CL09-223, 2011 WL 9688369, ¶ 1 (Va. Cir. Ct. Oct. 21, 2011) (ordering plaintiff and plaintiff's counsel to pay defendants over $700,000 in fees and expenses); Lester v. Allied Concrete Co., Nos. CL.08-150, CL09-223 2011 WL 8956003, ¶¶ 100, 115 (Va. Cir. Ct. Sept. 1, 2011) (reducing jury award by over $4 million because plaintiff "deliberately delete[d] Facebook photos that were responsive to a pending discovery request" at counsel's direction), *aff'd in part, rev'd in part*, 285 S.E.2d 699 (Va. 2013); *see also* The Sedona Conference, *The Sedona Conference Commentary on Legal Holds: The Trigger & the Process*, 11 SEDONA CONF. J. 265, 269–70 (2010) (discussing spoliation and "providing a framework an organization can use to create its own preservation procedures").

social media posts. The National Labor Relations Board recently released a series of memos focused on social media issues in the labor and employment field.[80] As one memo notes, while companies are increasingly putting in place social media policies to address concerns related to employee social media use at work, many such policies include overbroad and potentially unlawful provisions that may violate federal law.[81]

Another area that has generated a great deal of attention is the use of social media in the hiring process as employers have started asking applicants to provide access to social media accounts as a precondition of employment. Such practices could have First Amendment and privacy rights implications. Some states have responded by prohibiting employers from requiring employees to provide access to their social media accounts. Companies should have well-defined social media policies in place to address these various issues.

Objectives to Consider

- As with other forms of ESI, relevant social media evidence must be preserved at the point that litigation is reasonably anticipated.[82] Proper and timely preservation of social media is a significant concern, with courts finding spoliation or mandating preservation by individual account owners.[83]
- To the extent applicable, the rules governing preservation and collection of ESI apply equally to social media; however, given the interactive nature of social media, the text and imagery preservation and collection standards created for other ESI

80. *The NLRB and Social Media*, N.L.R.B., http://www.nlrb.gov/node/5078 (last visited Oct. 17, 2013).

81. *Id.* (citing Memorandum from Anne Purcell, NLRB Office of the Gen. Counsel, to All Reg'l Dirs., Officers-in-Charge, & Resident Officers (Jan. 24, 2012), http://mynlrb.nlrb.gov/link/document.aspx/09031d45807d6567).

82. *See, e.g.*, Offenback v. L.M. Bowman, Inc., No. 1:10–CV–1789, 2011 WL 2491371 (M.D. Pa. June 22, 2011); EEOC v. Simply Storage Mgmt., 270 F.R.D. 430 (S.D. Ind. 2010); Bass *ex rel.* Bass v. Miss Porter's Sch., No. 3:08cv1807, 2009 WL 3724968 (D. Conn. Oct. 27, 2009); Patterson v. Turner Constr. Co., 931 N.Y.S.2d 311 (App. Div. 2011); Romano v. Steelcase, Inc., 907 N.Y.S.2d 650 (Sup. Ct. 2010).

83. *See, e.g.*, Katiroll Co. v. Kati Roll & Platters, Inc., No. 10-3620, 2011 WL 3583408 (D.N.J. Aug. 3, 2011).

often do not fit. Traditional methods of collection and preservation may not capture metadata and logging data or reflect how easily social media data can be navigated.[84]

- Complex data management ownership issues must be carefully considered as they may present preservation challenges to ensure the ability to properly authenticate social media evidence.
- Given the transient and cloud-based nature of social media, traditional preservation and collection methods are often inadequate. Social media evidence should be collected with a proper chain of custody that preserves associated metadata, thus ensuring a greater likelihood of being able to establish authenticity at trial.
- Whether a party has possession, custody, or control of social media content is a threshold issue in determining that party's preservation obligations. This issue can be complicated by the fact that social media content is generated and stored in a variety of ways. Typically, the user is considered to have control of his or her own social media content and must therefore be responsible for proper preservation of any content he or she can access on demand.[85]
- Whether an organization has possession, custody, or control of social media posted by an employee on an external social media site is less straightforward. It has been suggested that where an employer has informed its employees that information created, stored, or exchanged from the employer's computers belongs to the employer, the employer may then be held to "control" such information even if it was created for personal use and stored physically on a third party's servers.[86]
- Data on social media sites are often protected by various permissions and privacy settings and therefore public searches are unlikely to yield much useful information. To access and collect social media evidence, it must be accessed directly by logging on to the account or by using specialized software.[87]

84. *See* The Sedona Conference, *supra* note 72, at 227.

85. *Id.* at 224.

86. *Id.*

87. *See* Holter v. Wells Fargo & Co., 281 F.R.D. 340, 343–45 (D. Minn. 2011) (limiting access to social media to relevant portions thereof).

- Tools to facilitate the preservation and collection of social media evidence are being developed by various vendors and are constantly evolving. Companies should take care to research and utilize the appropriate vendor or the particular requirements related to the social media evidence at issue.
- Vendors are developing solutions that go beyond capturing static, single-point-in-time images and instead allow content to be collected in a way to better preserve metadata fields associated with the social media, which significantly decrease concerns related to authenticity.
- There are important ethical issues that must be considered when accessing certain social media evidence. For example, it is unethical for an attorney to covertly (either directly or through another) "friend" a person to gain access to postings with heightened privacy settings.[88] Additionally, the terms of use of the social media site may preclude an attorney from collecting without consent an account holder's public-facing social media information that is not protected by privacy settings.[89] Moreover, such methods of collection place the data

88. *See* MODEL RULES OF PROF'L CONDUCT R. 4.2, 8.4(c) (2012); Craig D. Klausing, *It's the Steak, Not the Sizzle That Counts*, MINN. LAW., Dec. 7, 2009, *available at* http://lprb.mncourts.gov/articles/Articles/It%E2%80%99s%20the%20steak,%20not%20the%20sizzle%20that%20counts.pdf; *see also* PHILADELPHIA BAR ASS'N PROF'L GUIDANCE COMM., OP. 2009-02 (2009), *available at* http://www.philadelphiabar.org/WebObjects/PBAReadOnly.woa/Contents/WebServerResources/CMSResources/Opinion_2009-2.pdf (explaining that a lawyer directing a third person to "friend" a non-party witness in order to view the witness's private Facebook and Myspace posts would violate Pennsylvania Rule of Professional Conduct 5.3 on responsibilities regarding nonlawyer assistants, Rule 8.4 on misconduct in engaging in fraud or deceit or inducing others to violate the rules, and Rule 4.1 on truthfulness in statements to others).

89. *See, e.g.*, *Statement of Rights and Responsibilities*, FACEBOOK, https://www.facebook.com/legal/terms (last updated Nov. 15, 2013).

> This Statement of Rights and Responsibilities . . . is our terms of service that governs our relationship with users and others who interact with Facebook. By using or accessing Facebook, you agree to this Statement . . . .
>
> . . . .
>
> 5. Protecting Other People's Rights
>
> . . . .
>
> 7. If you collect information from users, you will: obtain their consent, make it clear you (and not Facebook) are the one

collector in the chain of custody and could thus require his or her testimony at trial in order to authenticate the data.[90]

- Direct access to a social media site for the opposing party should only be given as a last resort and only where it is critical for the opposing party to have interactive use of the content.
- Evidentiary issues of authentication, hearsay, preliminary matters of admissibility, and conditional relevance must all be fully considered when seeking to offer social media evidence.[91] Given the transient and cloud-based nature of social media data, traditional collection methods are often inadequate.
- It is important to document and verify the process and results of social media collection to ensure the evidence will have a greater likelihood of admissibility at trial. Communication and cooperation with the adverse party regarding the data to be preserved and methods of production can greatly reduce the likelihood of a spoliation claim.
- The SCA[92] has been broadly construed to encompass various social media content and imposes different levels of protection on data requested directly from the service provider depending on how the provider's services are categorized (i.e., is it "electronic communication service" or "remote

> collecting their information, and post a privacy policy explaining what information you collect and how you will use it.

Id.

90. *See, e.g.*, Cook v. J & J Snack Foods Corp., No. 2:09–CV–02297–GED–EFB, 2010 WL 3910478, at *5 (E.D. Cal. Jan. 28, 2010) (stating that printouts of a website attached to defendant's declaration were inadmissible absent evidence that the declarant had personal knowledge as to their accuracy); Toytrackerz L.L.C. v. Koehler, No. 08–2297–GLR, 2009 WL 2591329, at *6 (D. Kan. Aug. 21, 2009) ("To authenticate printouts from a website, the proponent must present evidence from a witness with personal knowledge of the website at issue stating that the printout accurately reflects the content of the website and the image of the page on the computer at which the printout was made.").

91. *See, e.g.*, Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 538 (D. Md. 2007). *See generally* Paul W. Grimm et al., *Back to the Future:* Lorraine v. Markel American Insurance Co. *and New Findings on the Admissibility of Electronically Stored Information*, 42 AKRON L. REV. 357 (2009) (detailing the considerations to be made in applying Federal Rules of Evidence 104(a), 104(b), 901, 902, 803, and 804 to ESI).

92. 18 U.S.C. §§ 2701–2712 (2006).

computing service"?).[93] The answer will largely depend on the type of data at issue and their current state. It is important to understand this body of law and how it may be applied when seeking to obtain social media evidence directly from service providers.[94]

- The same considerations that apply to selection of a review platform apply to production issues, which will turn on the importance of being able to review the social media data interactively and as they appeared on the social media site. Where this is not as important, it may be sufficient to produce the social media data in searchable format, with or without metadata.

D. Data on Websites[95]

A web page (or webpage) is a document on the World Wide Web consisting of an HTML (hypertext markup language) file and any related files for scripts and graphics.[96] It may provide navigation

93. *See* Crispin v. Christian Audigier, Inc., 717 F. Supp. 2d 965, 973 (C.D. Cal. 2010).

94. *See, e.g.*, Konop v. Hawaiian Airlines, Inc., 302 F.3d 868, 879–80 (9th Cir. 2002) (holding that even though Wong and Garnder consented to Davis' use of the website, the district court did not make any findings on whether Wong and Gardner actually used Konop's website, and thus they were not users under the SCA); Facebook, Inc. v. Power Ventures, Inc., No. C 08-05780 JW, 2010 U.S. Dist. LEXIS 93517, at *35–36 (N.D. Cal. July 20, 2010) (finding "that Power did not act 'without permission' within the meaning of Section 502 when Facebook account holders utilized the Power website to access and manipulate their user content on the Facebook website, even if such action violated Facebook's Terms of Use"); *Crispin*, 717 F. Supp. 2d at 972–76 (concluding that the plaintiff has standing to move to quash the subpoenas that were issued under the SCA because an "individual has a personal right to information in his or her profile and inbox on a social networking site" and his or her webmail inbox); Viacom Int'l Inc. v. Youtube Inc., 253 F.R.D. 256, 264–65 (S.D.N.Y. 2008) ("ECPA does not bar disclosure of non-content data about the private videos (*e.g.,* the number of times each video has been viewed on YouTube.com or made accessible on a third-party website through an 'embedded' link to the video).").

95. For sample preservation and collection technologies relating to web pages, see Appendix D, *infra*.

96. *Web Page*, TECHTERMS.COM, http://www.techterms.com/definition/webpage (last visited Nov. 10, 2013); *Web Page*, THEFREEDICTIONARY, http://www.thefreedictionary.com/web+page (last visited Nov. 10, 2013); *Web Page*, WIKIPEDIA, http://en.wikipedia.org/wiki/Web_page (last visited Nov. 10, 2013).

to other web pages through hypertext links.[97] It may provide not only static content, but also dynamic content, such as so-called "inline links," which is content hosted by another web server or even another website.[98]

A website (also written as website or site) is a group of connected web pages generally located on the same server and controlled by a person, group, company, educational institution, government, or organization.[99] Generally, the pages of a website can be accessed from a Uniform Resource Locator (URL) called the web address.[100] A typical site structure includes a home page with links to the website's supplemental pages, such as the "about" and "contact" pages.[101]

Websites convey information to the public or to a more limited target viewing audience. Certain websites require a subscription to access all or a portion of their content, such as many business sites, message boards, web-based e-mail, and social media networking websites.[102]

Websites can be rich sources of ESI, because they often contain representations about a party that are authored by or for the party itself. Although a website's content may change often,

97. TECHTERMS.COM, *supra* note 96; THEFREEDICTIONARY, *supra* note 96; WIKIPEDIA, *supra* note 96.

98. *See* Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1156 (9th Cir. 2007) ("The process by which the webpage directs a user's browser to incorporate content from different computers into a single window is referred to as 'in-line linking.'"); *see also* Righthaven L.L.C. v. Choudhry, No. 2:10-CV-2155 JCM (PAL), 2011 U.S. Dist. LEXIS 48290, at *5 (D. Nev. May 3, 2011) (defining in-line linked images as "HTML instructions that direct a user's browser to a website publisher's computer that stores the full-size photographic image"); *Inline Linking*, WIKIPEDIA, http://en.wikipedia.org/wiki/Inline_linking (last visited Nov. 10, 2013).

99. TECHTERMS.COM, *supra* note 96; THEFREEDICTIONARY, *supra* note 96; WIKIPEDIA, *supra* note 96.

100. TECHTERMS.COM, *supra* note 96; THEFREEDICTIONARY, *supra* note 96; WIKIPEDIA, *supra* note 96.

101. TECHTERMS.COM, *supra* note 96; THEFREEDICTIONARY, *supra* note 96; WIKIPEDIA, *supra* note 96.

102. *See Sedona Conference Glossary*, *supra* note 17, at 54; *see also* MERRIAM-WEBSTER'S COLLEGIATE DICTIONARY, *supra note* 15, at 1418; *Website*, TECHTERMS.COM, http://www.techterms.com/definition/website (last visited Nov. 10, 2013); *Website*, THEFREEDICTIONARY, http://www.thefreedictionary.com/web+site (last visited Nov. 10, 2013); *Website*, WIKIPEDIA, http://en.wikipedia.org/wiki/Web_site (last visited Nov. 10, 2013).

and may be under the control of a third party, it has been held that there is "no reason to treat websites differently than other electronic files."[103]

Objectives to Consider

- Attorneys should work with the relevant persons (potentially IT and in-house counsel) to determine who is responsible for website content, who maintains the data, and where the data are stored.[104]
- If a party has "control" of a website, it has a duty to preserve. This is true whether the website is maintained on the party's server or that of a third-party administrator. The question is whether the party has the ultimate authority to add to, delete, or modify the website's content.[105]
- If the party does use a third-party administrator for the site, counsel should determine the time span over which the third party has stored client information. Moreover, if the party reasonably anticipates litigation, there should be a control in place to notify the third party of the need for retention of the data.[106]
- Whether a party needs to preserve the entire website or a web page depends on whether it is legally relevant or will lead to admissible evidence. Generally, it is best to grab the whole site. However, if, for example, the client's site is Ebay.com, that may not be practical. And many tools that can be used for preserving a single web page cannot preserve the entire site. Moreover, certain pages of the site may be inaccessible to

103. Arteria Prop. Pty Ltd. v. Universal Funding V.T.O., Inc., No. 05-4896 (PGS), 2008 WL 4513696, at *5 (D.N.J. Oct. 1, 2008).

104. *See* Allison Stanton & Virginia Vance, *Retain and Preserve Electronic Information to Minimize Risk and Cost*, HOGAN LOVELLS, http://www.hoganlovells.com/files/Publication/b8a71a9b-25d5-48ef-af91-9d5b2e20f537/Presentation/PublicationAttachment/3de1c9f1-94fd-4fb1-8825-5e61636bf442/RetainandPreserve.pdf (last visited Jan. 25, 2014).

105. *See Arteria*, 2008 WL 4513696, at *5 (concluding that defendant's control over content posted on its website also meant that "it had the power to delete such content"; defendant's "*ultimate* authority, and thus control, to add, delete, or modify the website's content" gave rise to spoliation inference).

106. *See* Stanton & Vance, *supra* note 104, at 3–4.

persons other than the site administrator. Therefore, attempts to preserve an opposing party's site may fall short.

- A producing party should preserve the contents of a website once litigation is reasonably anticipated to preserve existing evidence.[107]
- There are various ways to preserve websites. It will be important from the outset to determine the most appropriate means of preservation. This will vary significantly based on whether the information preserved is on the client's site or the site of another party.
- One simple way to preserve a web page is to use the built-in "web capture" tool in Internet Explorer or in Adobe Acrobat. One advantage of saving static images to capture a web page is that the resulting PDF or document can be Bates stamped and produced. However, this method does not preserve the web page in its native form. Thus, one would need a more comprehensive capture if, for example, metadata or underlying formatting are at issue.[108]
    - To save a copy of the website to a computer, go to the web page in Internet Explorer and select "Save As."[109] (Note, however, attorneys should never engage in self-collection.)[110]

107. *See* JAY E. GRENIG ET AL., ELECTRONIC DISCOVERY AND RECORDS MANAGEMENT GUIDE § 7:1 (2009).

108. *See* Ira P. Rothken, *Web Spiders and Crawlers in E-Discovery*, MOREDATA (May 22, 2007, 8:03 PM), http://www.moredata.com/home/web-spiders-and-crawlers-in-e-discovery.html; *Web Crawler*, WIKIPEDIA, http://en.wikipedia.org/wiki/Web-crawler (last visited Nov. 10, 2013).

109. Susan Ardisson, *Website Content: Duty to Preserve*, QUBIT, Nov. 2008, at 1, *available at* http://www.bit-x-bit.com/sites/default/files/userfiles/files/articles/qubit-nov-08-website-content-duty-to-preserve.pdf.

110. *See* Leonard Deutchman, *Steer Clear of the Perils of Self Collection*, L. TECH. NEWS (Apr. 16, 2008, 12:00 AM), http://www.lawtechnologynews.com/id=900005508773; Mark Kerzner, *Technology for Lawyers and Paralegals: Evidence Authentication—Web Site Content,* SHMSOFT BLOG (Sept. 1, 2008, 7:34 PM), http://shmsoft.blogspot.com/2008/09/technology-for-lawyers-and-paralegals.html; *see also Proving Publication in Cyberspace*, WILLIAM R. WOHLSIFER, PA, http://infringement-attorney.com/provingpublicationincyberspaceecommercelawaspenlawandbusinessnewyorknyoct2001.html ("Third-party authentication services help overcome multiple objections to admissibility and always increase the weight of the proffered evidence.").

  - To save a static image of the web page, press the *Print Screen* key and then paste it into a document.[111]
  - Adobe Acrobat also allows the user to choose *Print to Adobe PDF* to save a static image of the web page.[112]
- A web crawler (also referred to as ants, automatic indexers, bots, web spiders, web robots, or web scutters) is another way to preserve websites. A web crawler is a computer program or script that browses the World Wide Web in a methodical or automated fashion. Web crawler software programs can be used to "mirror" web pages and entire websites as of a certain point in time. Web crawlers can be used to make a copy of all the visited pages, which may include a single page or the entire website, depending on how many levels deep the program was instructed to capture. Consider using a more complex spidering tool if many web pages will need to be preserved, there is a long time frame at issue, or periodic reviews are needed to capture any changes to the web page. Grab-a-Site, Web Copier, and WebWhacker are popular programs; because they each have advantages and disadvantages, it may be worthwhile to use more than one for the same matter.[113]
- Online cloud-based services also provide web page or website preservation services, such as Iterasi.[114] Some services are designed specifically for attorneys for e-discovery purposes.[115]
- Additionally, screen-recording programs can assist with preservation. They allow the user to record the video and audio of a website, which can be played at a later date for the court or jury. One popular program is Camtasia Studio.[116]

111. *See* Ardisson, *supra* note 109, at 1. But note, however, this will not accurately collect information that is dynamic, such as certain types of advertisements and inline links. *Id.*

112. *Convert Web Pages to PDF in Internet Explorer and Firefox (Windows)*, ADOBE, http://help.adobe.com/en_US/acrobat/X/pro/using/WS58a04a822e3e50102bd615109794195ff-7f60.w.html (last visited Oct. 7, 2013).

113. *See* Rothken, *supra* note 108.

114. ITERASI, http://www.iterasi.com (last visited Nov. 18, 2013).

115. *See* NEXTPOINT, http://www.nextpoint.com/#&panel1-1 (last visited Oct. 7, 2013).

116. *See* Ira P. Rothken, *Snagit and Camtasia Studio*, MOREDATA (Mar. 17, 2007, 11:20 PM), http://www.moredata.com/home/snagit-and-camtasia-studio.html.

- Older versions of a web page may be found by searching the Google cache or using the Wayback Machine.
  - Google takes a snapshot of each page it examines and stores ("caches") that version as a backup.[117] This is useful if, for example, the original page is unavailable because the owner recently removed the page from the World Wide Web.[118] Most search results are accompanied by a cached link.[119] Alternatively, the user can search the Google cache by typing in "cache": and then the web address into Google's search box.[120]
  - The Wayback Machine is a service created by the Internet Archive that enables users to see archived versions of web pages.[121] By typing in a URL and selecting a date range, users can browse old versions of a particular site.[122]
    - There are certain ways to block Internet Archive from copying files from a website and disabling access to all of the Wayback Machine's archives of those files. Blocking is achieved through the use of a file called robots.txt on the web server, which crawlers will find in the root-context and will respect.[123] However, a web page cannot be created from the past that was not cached by archive.org (because of a robots.txt).[124]
- Printouts of web pages are not self-authenticating documents; thus, be prepared to properly authenticate the contents of any printed web pages. This can generally be accomplished by a

117. *Cached Pages*, GOOGLEGUIDE, http://www.googleguide.com/cached_pages.html (last updated Dec. 28, 2011).

118. *Id.*

119. *Id.*

120. *Id.*

121. *Wayback Machine*, PC MAG., http://www.pcmag.com /encyclopedia/term /56218/wayback-machine (last visited Nov. 18, 2013).

122. *Internet Archive Wayback Machine,* INTERNET ARCHIVE, http://archive.org /web/web.php (last visited Nov. 18, 2013).

123. *See Robots Exclusion Standard*, WIKIPEDIA, http://en.wikipedia.org /wiki/Robots_exclusion_standard (last visited Oct. 7, 2013).

124. For a decision discussing this technology, see *Netbula, L.L.C. v. Chordiant Software, Inc.*, No. C08-00019 JW (HRL), 2009 WL 3352588, at *1–2 (N.D. Cal. Oct. 15, 2009) (concluding that plaintiffs had "control" over archived web pages and were obligated to produce such evidence).

webmaster or someone with personal knowledge of the web pages.[125] Likewise, be prepared to authenticate evidence derived from web crawlers or screen recording programs.

- If the web content in question is from a website that is under control of the client, their web publishing solution may have tools that allow preservation of web content and recreation of web pages at a certain point in time.

E. Databases

A database is an organized collection of ESI.

> Traditional databases are organized by fields, records, and files. A field is a single piece of information; a record is one complete set of fields; and a file is a collection of records. For example, a telephone book is analogous to a file. It contains a list of records, each of which consists of three fields: name, address, and telephone number.[126]

There are many types of database models, including a flat file, hierarchical, network, relational, dimensional, and object. The goal of each model is the same: to organize data. However, depending on the type of data, each model will provide different opportunities for control over structure, development, and performance.[127]

In addition to the many types of database models, there are also many systems to manage databases. The database management system is a software package with computer programs that control the creation, maintenance, and use of a database. It allows organizations to conveniently develop databases for various applications by database administrators (DBAs) and other specialists. These include Oracle, IBM DB2, Microsoft SQL Server, Microsoft Access, PostgreSQL, MySQL, and SQLite.[128]

125. *See* Kristen L. Mix, *Discovery of Social Media*, 5 FED. CTS. L. REV. 119, 134 (2011); *see also* INTERNET ARCHIVE, http://archive.org/about/faqs.php#274 (last visited Oct. 8, 2013) (addressing the questions: "How can I get pages authenticated from the Wayback Machine? How can [I] use the pages in court?").

126. *Database*, WEBOPEDIA, http://www.webopedia.com/TERM/D/database.html (last visited Nov. 19, 2013).

127. *See* JEFFREY D. ULLMAN & JENNIFER WIDOM, A FIRST COURSE IN DATABASE SYSTEMS 1–2 (2d ed. 2002).

128. *Id.*

Finally, a database will also have a view for the end-user to interact with the data. The view will likely vary depending on the type of database and the database management system. This may include printed reports, reports generated on a computer, table views, and record views generated by the database management system.

Due to the complexity of working with data maintained in databases, the Sedona Conference issued six "Database Principles":

1. Absent a specific showing of need or relevance, a requesting party is entitled only to database fields that contain relevant information, not the entire database in which the information resides or the underlying database application or database engine.
2. Due to differences in the way that information is stored or programmed into a database, not all information in a database may be equally accessible, and a party's request for such information must be analyzed for relevance and proportionality.
3. Requesting and responding parties should use empirical information, such as that generated from test queries and pilot projects, to ascertain the burden to produce information stored in databases and to reach consensus on the scope of discovery.
4. A responding party must use reasonable measures to validate ESI collected from database systems to ensure completeness and accuracy of the data acquisition.
5. Verifying information that has been correctly exported from a larger database or repository is a separate analysis from establishing the accuracy, authenticity, or admissibility of the substantive information contained within the data.
6. The way in which a requesting party intends to use database information is an important factor in determining an appropriate format of production.[129]

129. *The Sedona Conference Database Principles Addressing the Preservation & Production of Databases & Database Information in Civil Litigation*, SEDONA CONF., at iii (public comment version Mar. 2011), https://thesedonaconference.org/download-pub/426.

Objectives to Consider

- It is important to understand what type of database has been identified. This will help in the process to identify the structure and scope of the database. For example, records can span over many tables or even databases. Understanding the structure of a database will help in the process of defining and satisfying hold obligations.
- Identify the database management system (DBMS) used to manage the database. The DBMS should provide insight into the organization and storage of the data. The DBMS may also provide capabilities for preservation and retention. In addition, if a backup of the entire database is required, it may require specific versions of DBMSs to restore the data.
- Identify the database administrator. The database may be hosted in-house or by a vendor. However, there may be in-house support for both solutions. The database support should have knowledge of backup schedules and retention schedules. In addition, the database support may have the ability to export data, modify retention schedules, and hold backups.
- The location of the stored database data may be different than the server where the DBMS is located. It is important to identify the location of the storage. This could be in-house, at a vendor, or in the cloud.[130] There may be a difference in the process to extract data in-house versus the cloud. It is important to understand the storage capability.
- Does the database have retention schedules in place? The DBMS may have tools to apply retention schedules. If applicable, identify the process to remove any auto-deletion schedules based on litigation hold requirements. They also may have logging functions that collect and purge activity data, which may need to be preserved.
- Identify the data that are relevant to the matter. It may be possible that the relevant data are a small percentage of the larger database. Understanding the volume may support arguments of undue burden for holding an entire database.

130. *See infra* Part IV.F (defining "cloud").

- Databases are dynamic systems, meaning that records may be updated and added frequently. Identify if there needs to be an ongoing hold on the database data. This will determine if there needs to be point in time or ongoing backups in lieu of a record retention system.
- It is important to identify what kind of data are stored (fields with data, documents, or other information). This may determine the best method for preservation and extraction.
- In addition, it is important to understand how the data are viewed and identify a format that is reviewable. This will depend on the data. Document management systems may have extracted files with a metadata load file and both are reviewable items. In addition, there may be raw records that need to be reviewed in a form or report format. This may require development of a form to review the record and associated files.

Sample Technologies for Preservation and Collection

Prior to preservation or collection from a database, collaboration with the IT resources responsible for the database administration is essential in understanding the database structure, design, and impact preservation activities may have. Preservation of database ESI typically is accomplished in either an online or off-line manner. Preserving database ESI online requires suspending auto-delete functions or modifying end user permissions to prevent the editing or deletion of ESI. Preserving database ESI off-line requires point-in-time copies of the database ESI that are preserved separate from the online system. Disaster recovery backups should not be relied upon for these copies.[131] When choosing the off-line option, there must be a method to restore or image the database ESI without modification of metadata and place it into the same database structure as the online system.

Prior to collecting data from a database, a determination needs to be made on reasonable options for data extraction and a format suitable for review by all parties. Coordination is required between the IT professionals, legal parties, and vendors to facilitate the collection, processing, review, and production of database ESI.

131. See Backup Systems overview, *infra* Part IV.G, for more information.

Collection from databases can be time consuming and expensive due to the technical skill sets required to facilitate extracting data from the database and converting it into a reviewable format.

F. Cloud Computing

Cloud computing is defined as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[132]

More simply, cloud computing stores client data and applications in shared data centers operated by third parties (often around the world) so clients can access their data or run applications from any location with an Internet connection. The essential characteristics of cloud computing include: (1) on demand self-service, (2) broad network access, (3) resource pooling, (4) rapid elasticity, and (5) measured service.[133] Examples of cloud computing services include Google Gmail, Google Apps, iCloud, Dropbox, SkyDrive (Microsoft), Amazon cloud drive, Shutterfly, Snapfish, and Netflix. Figures 1 and 2 below illustrate how data is stored on the cloud and how clients interact with the cloud.

132. PETER MELL & TIMOTHY GRANCE, NAT'L INST. OF STANDARDS & TECH., THE NIST DEFINITION OF CLOUD COMPUTING § 2 (2011), *available at* http://www.nist.gov/customcf/get_pdf.cfm?pub_id=909616; *Sedona Conference Glossary*, *supra* note 17, at 9. *See generally* VIVEK KUNDRA, EXEC. OFFICE OF THE PRESIDENT OF THE U.S., FEDERAL CLOUD COMPUTING STRATEGY (2011), *available at* http://www.whitehouse.gov/sites/default/files/omb/assets/egov_docs/vivek-kundra-federal-cloud-computing-strategy-02142011.pdf (discussing use of the cloud).

133. MELL & GRANCE, *supra* note 132, § 2.

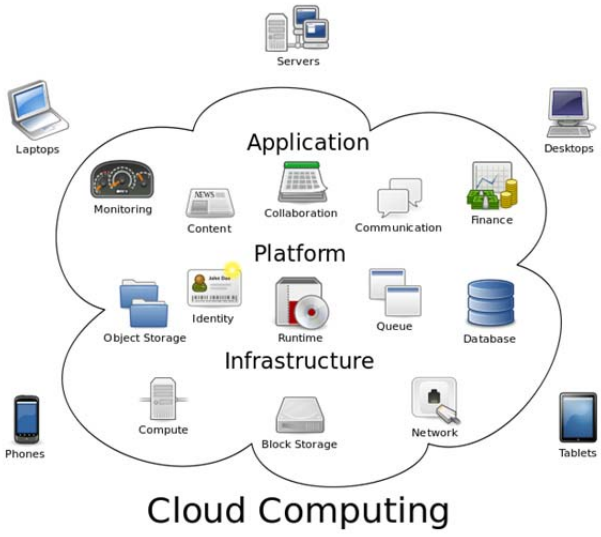


Figure 1: Illustration of Cloud Computing[134]

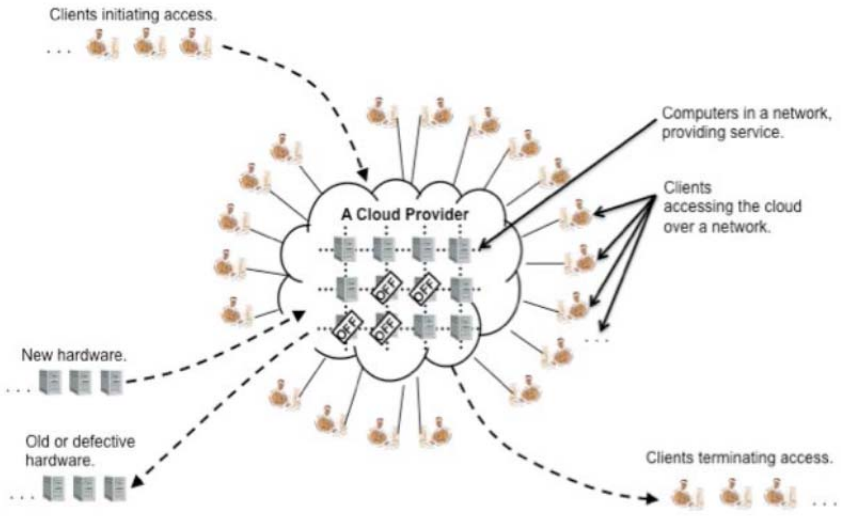


Figure 2: General Cloud and Subscriber View[135]

134. Sam Johnston, *Cloud Computing*, WIKIPEDIA (Mar. 3, 2009), http://en.wikipedia.org/wiki/File:Cloud_computing.svg.

135. LEE BADGER ET AL., NAT'L INST. OF STANDARDS AND TECH., CLOUD COMPUTING SYNOPSIS AND RECOMMENDATIONS 4-1 fig.1 (2012), *available at* http://csrc.nist.gov/publications/nistpubs/800-146/sp800-146.pdf.

There are four types of cloud models:

- *Private Cloud*: The cloud is used exclusively by a single organization. The private cloud may be managed by the organization or a third-party provider and may be either on-site or off-site.[136]
- *Community Cloud*: The cloud is used by a specific community of consumers that have shared concerns. It may be managed by the organization or a third-party provider and may be either on-site or off-site.[137]
- *Public Cloud*: The cloud is open for use by the general public.[138]
- *Hybrid Cloud*: The hybrid approach stores the majority of data on a public cloud and sensitive data on a private cloud.[139]

Cloud computing also has three types of service models:

- *Software as a Service (SaaS)*: Provides cloud-based software for consumers.[140] Computer applications are run on the cloud infrastructure and are accessible through an interface such as a web browser.[141]
- *Platform as a Service (PaaS)*: Primarily used by developers to create software for consumers.[142] The consumer does not manage or control the underlying cloud infrastructure (i.e., network, servers, operating systems, or storage); the consumer has control over the applications and possibly configuration settings for the application-hosting environment.[143]
- *Infrastructure as a Service (IaaS)*: Primarily used by developers to access cloud-based or virtual hardware.[144] The consumer does not manage or control the underlying cloud infrastructure, but has control over operating systems, storage, and

136. MELL & GRANCE, *supra* note 132, § 2.
137. *Id.*
138. *Id.*
139. *Id.*
140. NICOLE BLACK, CLOUD COMPUTING FOR LAWYERS 4 (2012).
141. MELL & GRANCE, *supra* note 132, § 2.
142. BLACK, *supra* note 140, at 4.
143. MELL & GRANCE, *supra* note 132, § 2.
144. BLACK, *supra* note 140, at 4.

applications, and may have limited control of select networking components (e.g., host firewalls).[145]

Objectives to Consider

Location of Data:

- In traditional data storage, data are stored in a data center and the data owner controls where the data are stored. In contrast, cloud service providers may store data in a variety of locations and, as a result, the data may be stored in multiple data centers all over the world.[146] This is true even for a "private cloud" that is run by a single entity. Data in the cloud can be transferred across multiple borders, which (as discussed below) may have significant legal implications.[147]

Possession, Custody, or Control:

- Rule 34 of the Federal Rules of Civil Procedure calls for the production of documents in the responding party's "possession, custody, or control."[148] Unlike traditional data storage, the cloud consumer has the right to access the data stored in the cloud, but the data are in the possession of a third-party provider.[149] Because a third-party has possession of the data, production of data stored on the cloud may be subject to unique discovery problems, such as whether the data are actually in the responding party's "possession, custody, or control."[150]

145. MELL & GRANCE, *supra* note 132, § 2.

146. *See* BADGER ET AL., *supra* note 135, § 8.4.2.

147. *See id.* §§ 4, 8.4.2–8.4.3, 9.2; Laurin H. Mills, *Legal Issues Associated with Cloud Computing*, SECUREIT (May 13, 2009), http://www.secureit.com/resources/Cloud%20Computing%20Mills%20Nixon%20Peabody%205-09.pdf.

148. FED. R. CIV. P. 34(a)(1).

149. BLACK, *supra* note 140, at 88–89; *see also* BADGER ET AL., *supra* note 135, § 8.4.

150. *See* Mark L. Austrian & Martin Krolewski, *Basic Steps in E-Discovery Continued: Legal Hold Policies Where Information Is Within the Company, in a Cloud or on a Social Media Site*, 19 METROPOLITAN CORP. COUNS., Apr. 2011, at 1, 1–2, *available at* http://www.kelleydrye.com/publications/articles/1470/_res/id=Files/index=0/1470.pdf; Alberto G. Araiza, Note, *Electronic Discovery in the Cloud*, 2011 DUKE L. & TECH. REV. 8, ¶ 5 (2011); Christine Soares, *Applying E-Discovery Best Practices to Cloud Computing*, FOX ROTHSCHILD LLP (Jan. 31, 2012), http://

- When contracting with a third-party cloud provider, a consumer should ensure it has the practical or contractual ability to access its ESI stored on the cloud.[151] To ensure this, the contract with the provider should state that the consumer owns the data, has the authority to manage its data, has the ability to access its data, and the data are protected from inappropriate disclosure.[152]

Data Privacy Issues Arising with Cloud Storage:

- Because data is stored with other consumers' information, consumers should consider possible cloud security risks, such as the risk of inadvertently producing data from other consumers.[153] While cloud service providers generally provide safeguards to ensure security and privacy, consumers should inquire about a provider's security management systems prior to choosing a provider.[154]

Privacy:

- Further, consumers should remember that information stored on the cloud is subject to different protections than information stored in a traditional data center. Data stored in cloud storage may be subject to compelled disclosure to the government under the Electronic Communications Privacy Act, the SCA, the USA Patriot Act (including National Security Letters and FISA warrants), or warrants and subpoenas.[155] Similarly, the physical location of the server storing the data

www.foxrothschild.com/newspubs/newspubsArticle.aspx?id=4294970830; *see also* Josiah Dykstra & Damien Riehl, *Forensic Collection of Electronic Evidence from Infrastructure-as-a-Service Cloud Computing*, 19 RICH. J.L. & TECH. 1, ¶ 20 (2012); Sean L. Harrington, *Collaborating with a Digital Forensics Expert: Ultimate Tag Team or Disastrous Duo?*, 38 WM. MITCHELL L. REV. 353, 392 (2011) ("[T]raditional computer forensics approaches are likely to be stymied by both the data storage architecture and the data delivery infrastructure.").

151. *See* BLACK, *supra* note 140, at 84–92; Soares, *supra* note 150.

152. *See* BLACK, *supra* note 140, at 84–92; Soares, *supra* note 150.

153. BADGER ET AL., *supra* note 135, §§ 8.5, 8.5.4; Christopher Wolf, *Privacy and Data Security Issues in Cloud Computing*, *in* PRACTICING LAW INST., CLOUD COMPUTING 2012: CUT THROUGH THE FLUFF & TACKLE THE CRITICAL STUFF 113, 125 (2012).

154. *See* Araiza, *supra* note 150, ¶ 6; Mills, *supra* note 147.

155. *See* Mills, *supra* note 147.

may have legal implications.[156] In particular, U.S. and foreign privacy laws may be in conflict.[157]

Hybrid Storage:

- When thinking about how to maintain privacy of data stored in cloud storage, consumers may want to think about using the hybrid storage approach—storing sensitive data in a private cloud or simply keeping particularly sensitive data off the cloud entirely.[158] Additionally, consumers may consider encrypting all ESI and metadata with consumer-specific keys so it remains indecipherable to other consumers.[159]

Preservation and Collection Considerations

Because the services are being provided by a third-party provider, simply issuing a litigation hold notice may not be sufficient for preservation of data.[160] To ensure that data can be properly preserved at the time a cloud computing agreement is entered into, terms should be negotiated to address:

- The preservation of data for purposes of e-discovery and the timeframe within which the preservation process can be implemented;[161]
- The provider's ability to stop end users from deleting relevant data;[162]
- The preservation of data and metadata as part of the normal course of business;[163]
- Ownership of the data stored on the cloud;[164]

156. *See* BADGER ET AL., *supra* note 135, § 8.4.2; Janine Anthony Bowen, *Overview of Cloud Computing*, *in* PRACTICING LAW INST., *supra* note 153, at 51, 57.

157. BLACK, *supra* note 140, at 75–85; Bowen, *supra* note 156, at 57.

158. Rachel Beth Evans, *Cloud Computing Due Diligence*, *in* PRACTICING LAW INST., *supra* note 153, at 79, 87; Araiza, *supra* note 150, ¶ 39; *see also* BADGER ET AL., *supra* note 135, § 8.5.1.

159. *See* Araiza, *supra* note 150, ¶¶ 16–17 (defining metadata and stating that the information is not readily available to the reader).

160. Soares, *supra* note 150; *see also* Araiza, *supra* note 150, ¶¶ 10–12.

161. BADGER ET AL., *supra* note 135, § 9.1.

162. Soares, *supra* note 150.

163. *Id.*

164. BLACK, *supra* note 140, at 91.

- Recovery procedures in the event of a disaster;[165]
- Limits on storage locations (by state or country, for example).

In addition, consider:

- Charges that may result from ongoing preservation;
- Whether agreements with the cloud provider cover the legal issues raised by the duty to preserve;[166]
- The provider's policies and capabilities regarding ESI retention, retrieval, and collection—including whether self-collection is possible;[167]
- How the provider will cooperate in responding to e-discovery requests;
- How the provider will apply and release legal holds.[168]

G. Backup Systems

Federal Rule of Civil Procedure 26(b)(2)(B) and its Minnesota counterpart, Minnesota Rule of Civil Procedure 26.02(b)(2), state that a party need not provide data from sources that are "not reasonably accessible" because of "undue burden or cost."[169] Disaster recovery backup data are typically considered inaccessible because they are not maintained and accessible within the active data of a company. However, if backups contain unique data or are used as an "archive" this may not apply.

A backup system makes copies of ESI to be used if the original ESI is lost due to a data loss event. As a result, most such systems are referred to as "Disaster Recovery Systems" or "DR Systems."

There is no typical DR system, but rather many different approaches and technologies for backing up ESI. The following is a list of the most common utilized methods for backing up ESI:

- *Online*: Used to backup ESI to another online source (i.e. network attached storage (NAS), server attached storage (SAS)). Although this provides the ability to restore ESI very

165. Mills, *supra* note 147.

166. *Id.*

167. *Id.*; Katie Taylor, *Cloud Considerations: E-Discovery*, LEGAL CLOUD CENTRAL (Oct. 29, 2012), http://www.legalcloudcentral.com/2012/10/articles/in-the-courts/cloud-considerations-ediscovery/.

168. *See* Austrian & Krolewski, *supra* note 150.

169. FED. R. CIV. P. 26(b)(2)(B); MINN. R. CIV. P. 26.02(b)(2); *see also* The Sedona Conference, supra note 13, at 283.

quickly in the event of a data loss, it is relatively expensive to implement. It is also more vulnerable to data corruption and deletion than other backup methods. Due to the high cost to implement and run these systems, the amount of historical data available is also limited.

- *Near-line*: Used to backup data to a device attached to the computer network, but not as readily available as online systems. The typical implementation of a near-line system would include a tape library, which is a mechanical device used to read and write ESI on magnetic tapes. When the tapes are not being read or written to, they are stored within the tape library, similar to a jukebox. These systems take longer to read and write to than an online system, but provide an advantage over other methods due to reduced time to restore data. Near-line systems provide less advantage, though, over off-line and off-site methods because they are still exposed to online system failures.
- *Off-line*: Used to backup data to a device that is only connected to the original ESI source during the backup. This could include portable disk drives or even a tape library, but the difference from a near-line system is the storage media is disconnected from the location of the original ESI, providing an advantage to online and near-line systems in that the backup media is not susceptible to issues with the computing environment. These systems are also less expensive to implement than other options.
- *Off-site*: Identical to off-line storage with the exception that the backup media is sent to a separate physical location for storage. This provides the advantage of mitigating risks to the original data's computing and physical environment. The disadvantage is that it takes the longest amount of time to restore data in the event of a data loss of all the backup options.

Objectives to Consider

- When a company uses a system of tapes to store its backup data, those tapes are maintained many miles off-site and are not connected to the company's server. The tapes are typically kept in atmospherically controlled conditions but are not

connected to any electronic device. They are tapes stored in a box.

- Data on backup tapes are typically compressed, so they need to be "restored" before they can be used.
- Data on backup tapes may lose integrity with time. Backup tapes lose their business value as a "backup restoration" resource when subsequent tapes are created.
- The backup tapes are typically not indexed—the data are much like a collection of records without a table of contents. They need to be "indexed" by an IT professional to show what records are located where on the tapes before it can be determined where relevant records exist on the tapes.
- Backup tapes are often a set of tapes made at specific intervals (daily, weekly, monthly). The most recent set provides the owner of the data with a ready resource for restoring its data system in the event of a catastrophe that causes a total loss of the business's data (hurricane, tornado, fire)—this is why the tapes are kept at a secure, atmospherically-controlled location far from the site of the company.
- Organizations typically design their backup processes to make the most use of the backup media (i.e., backup tapes). In most cases, this results in the mixing of data types (e-mail, share data, application data) on individual or sets of backup tapes. As a result, it is increasingly difficult for organizations to place preservation holds on subsets of ESI; entire backup tape populations generally need to be preserved to ensure that the relevant data are preserved. Costs to the organization increase in two ways: (1) by placing populations of tapes on hold, the organization cannot reuse tapes with older data for ongoing disaster recovery backups and must purchase new media for the ongoing need; and (2) because of the increasing number of backup tapes used, ongoing off-site storage costs continue to increase.
- Due to employee and organizational habits, often times "old" data that would seem to only exist on backup media are also still available in online systems. Examples include PST/ Personal Archives for e-mail, personal share drives, and computer hard drives. Indexes of both online and backup media can provide metadata such as filenames, created dates, folder path, and last modified dates. These can then be

compared to demonstrate how much or how little difference there is between the backup media and the online storage.

Preservation and Collection

Preservation of backup media typically is done through the suspension of two practices: reuse and destruction. Organizations typically reuse backup media that contains stored ESI that is no longer needed for recovering from a disaster, thereby minimizing the cost of purchasing new backup media for ongoing needs. Another practice that is generally part of a DR system is to destroy the media (render the media unreadable or eliminate stored ESI) through mechanical means, typically through degaussing (de-magnetizing) or shredding. This practice is used when the backup media is no longer compatible with the systems used to read and write to it or when more efficient and cost-effective media replaces it. Counsel needs to advise clients about the likelihood that relevant information exists on existing backup tapes and to suspend these typical practices.

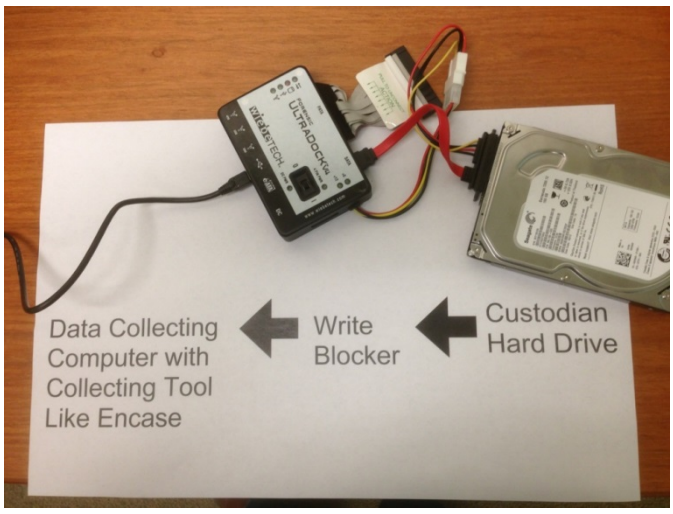
Collection from backup tapes can be a time-consuming and expensive process. Generally, subsets of data stored on backup tapes cannot be restored without first restoring all of the data on the backup tapes by reading from magnetic media and writing to hard drives, a generally slow process.

APPENDIX A: SAMPLE TECHNOLOGIES FOR PRESERVATION AND COLLECTION—HARD DRIVES

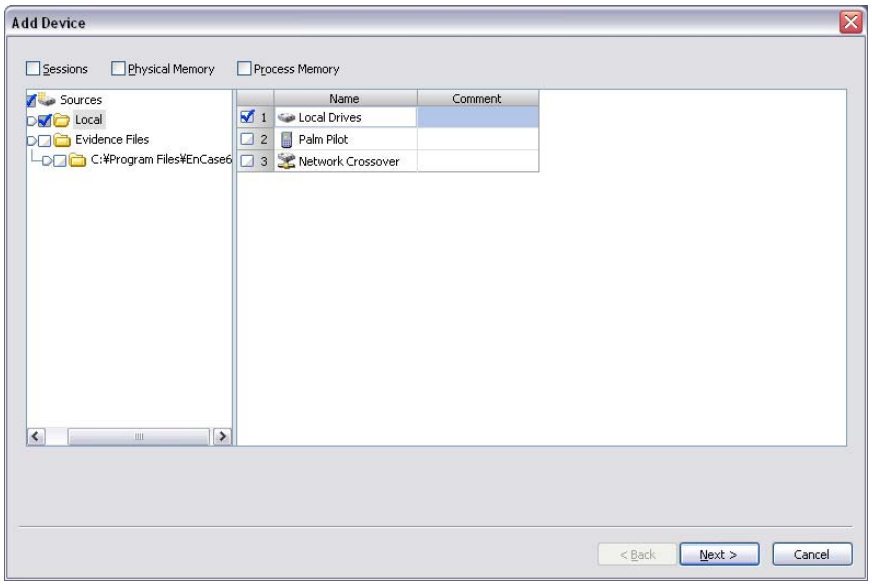
Step 1: The major steps taken during a data preservation and collection process are defined in the checklist below and it is important for the person collecting the data to have a similar checklist in order to log and account for all the actions taken in the process.

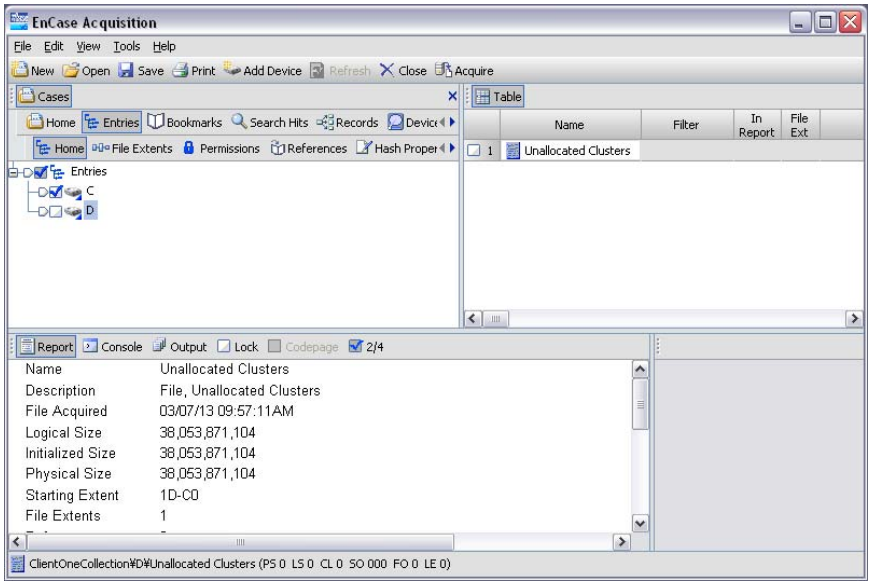
Practice Support Forensic Collection Checklist					
Item Name/ Description					
Date	Action	Check If Completed	By	Comments/Details	Initials
	Update Media's Chain of Custody				
	Use of Write Blocker				
	Forensic Collection Tool Used				
	Target Collection Performed				
	Entire Disk Collection Performed				
	Verification Performed				
	Protected				
	Protected				
	Image File Created				

Step 2: Extract the custodian's hard drive from his or her computer and plug it into a write blocker to prevent data alteration during the preservation/collection process, as shown below.



Step 3: Connect the write blocker to the computer with a data-collecting tool like EnCase (forensically sound) and then add the device to EnCase (see illustration below). Also add the hard drive or device that will store the collected data. After adding the devices, if targeted collection is required, burrow down into the custodian drive (labeled by a letter) and select or check the appropriate files or folder. If a full collection is required, then select the device and click "Acquire" to begin the process.





Step 4: Define the location and properties of the output similar to below. Ideally, the output should be placed in a password protected hard drive. The output data are an E01 forensic image, which is a secure way of storing the data, and with Acquisition MD5 selected (see illustration below), the file will be automatically verified once the process completes to ensure integrity.

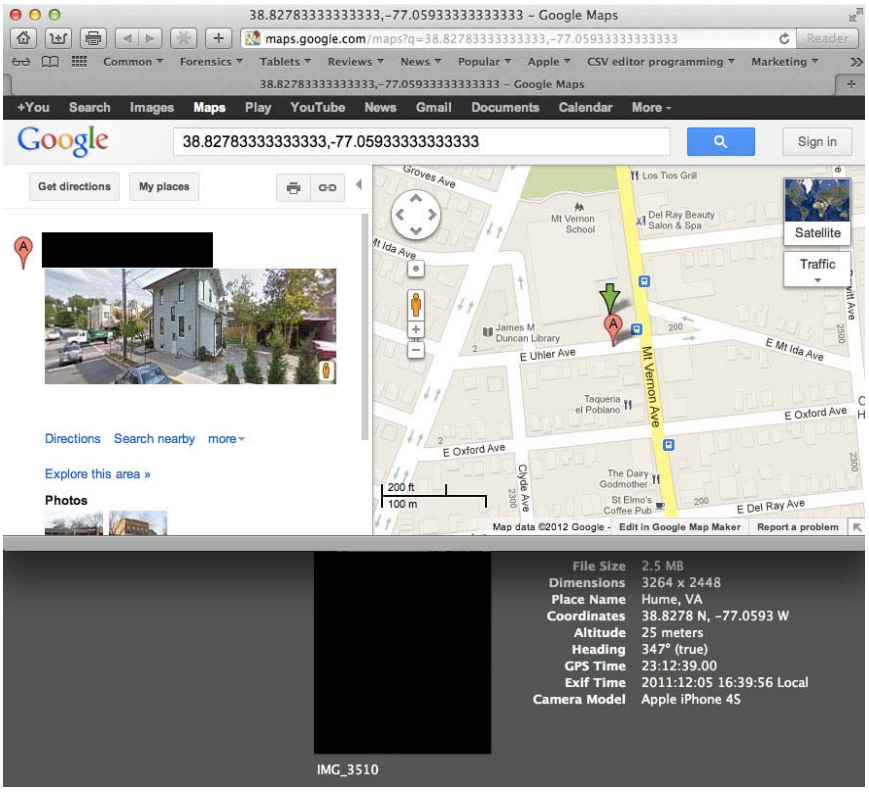
Step 5: When the verification process completes, safely unplug the hard drive with the collected data from the collecting computer and make sure it is physically secure and safe. Unplug the write blocker from the collecting computer, unhook the custodian hard drive, and put the hard drive back in his/her computer. Complete the chain of custody process and make sure all the requirements in the checklist are met.

APPENDIX B: SAMPLE TECHNOLOGIES FOR PRESERVATION AND COLLECTION—DATA MINING PHONE MEMORY CARDS

Photographs and Geotags

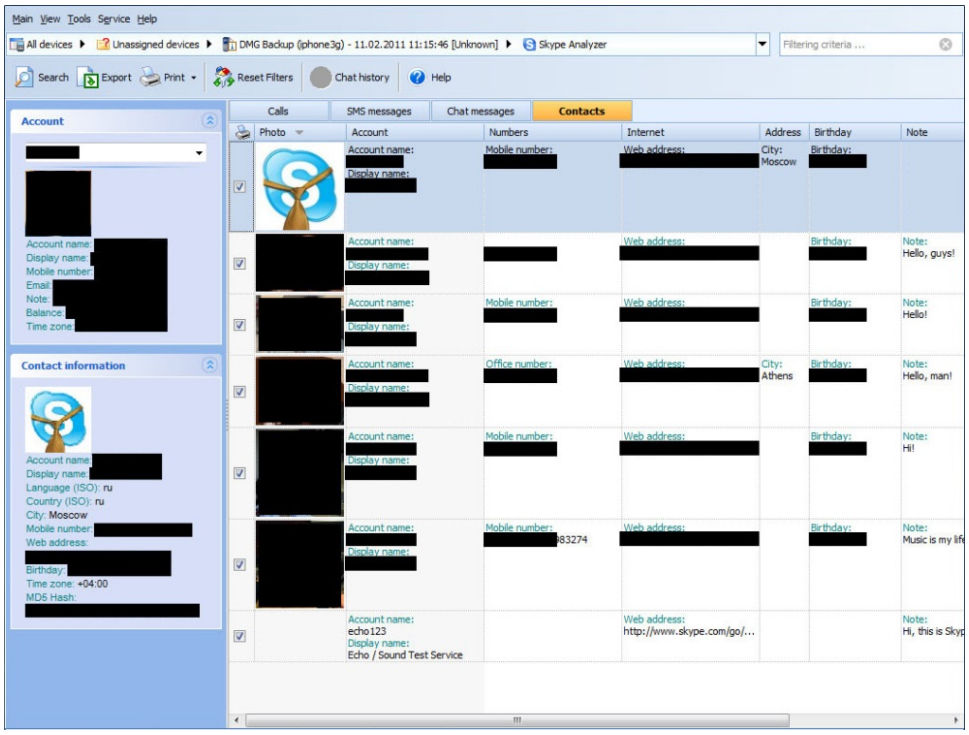
Mobile forensic reports on photographs stored on a smart phone can provide more information than just the photograph alone.

The following visual is an example of a photograph taken by the onboard camera on a smart phone. It also shows photograph metadata with camera model, date and time, place name and coordinates, and a Google Map of the location in which the photo was taken—including the Google Street View picture of the vicinity.



Skype Video

The following visual shows a mobile messaging smart phone app called Skype with its user profile and several of the phone user's address book of contacts. Each contact features a thumbnail photo, mobile phone number, web address, street address, birthday, note, and more.



APPENDIX C: SAMPLE TECHNOLOGIES FOR PRESERVATION AND COLLECTION—TEXT MESSAGES

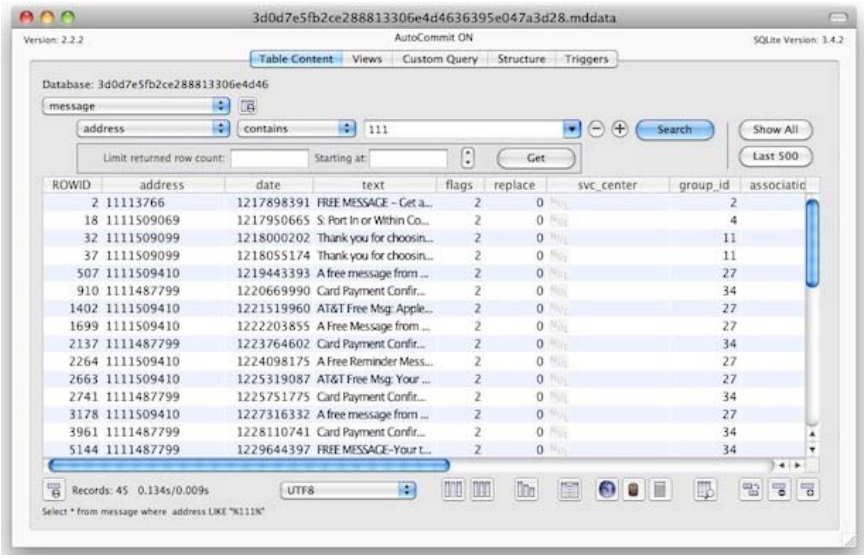
Step 1: All SMS/text messages available are stored within the phone. For preservation and collection purposes, create a backup of the phone. This ensures the process does not affect or alter active data.

This example outlines the process for an iPhone, but the process is very similar for most phones. The SMS/text messages are backed up when the phone is backed up and stored within the standard iPhone backup location located at: ~/Library/Application Support/MobileSync/Backup/. When backed up to the computer, SMS/text messages can found at the following locations:

- Mac iPhone backup file: %APPDATA%\Apple Computer \MobileSync\Backup\
- Windows XP: C:\Documents and Settings\[your username] \Application Data\
- Windows Vista: %APPDATA% = C:\Users\[your username] \AppData\Roaming
- Windows 7: C:\Users\user\AppData\Roaming\Apple Computer \MobileSync\Backup

The messages are stored in randomly generated hexadecimal filenames such as: 7182649a9879a8798c798e98794798f9279877c98 7984. This file is a small database called SQLite and can be read by any application that reads a SQLite database. There are plenty of SQLite and free applications online for Windows and Mac. For example, MesaSQLite for Mac OS X.

Step 2: Download a SQLite application and open the SMS/text file. It should look like:



Step 3: Query the information just like any other database. For example, the text number, the date of the text, or keywords in the text message can be queried. The text messages can also be dragged to a text editor like notepad or textpad.

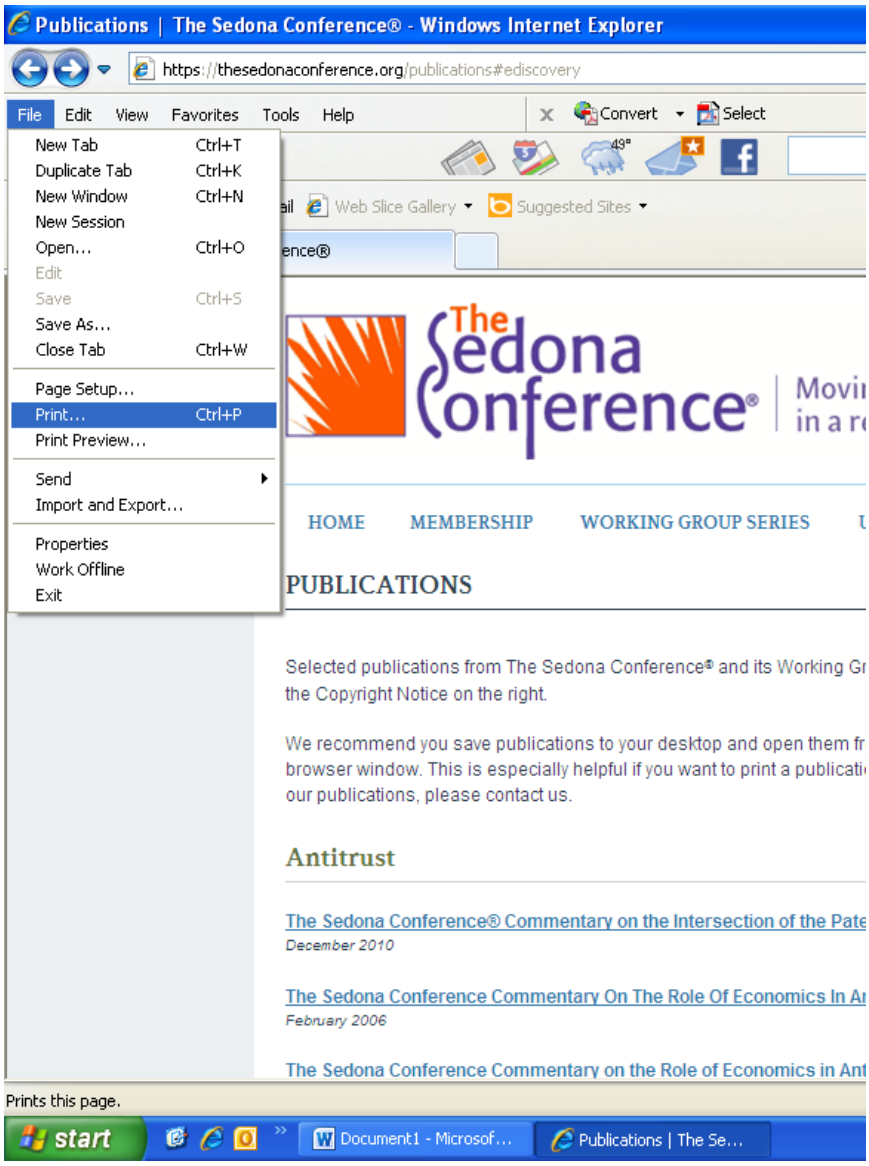
Step 4: After a review of the file and confirmation that the text message can be viewed, close out of the application and copy the SMS/text file to an encrypted hard drive with a password for storage.[170]

170. William Pearson, *How to Access and Read the iPhone SMS Text Message Backup Files*, OS X DAILY (July 8, 2010), http://osxdaily.com/2010/07/08/read-iphone-sms-backup/.

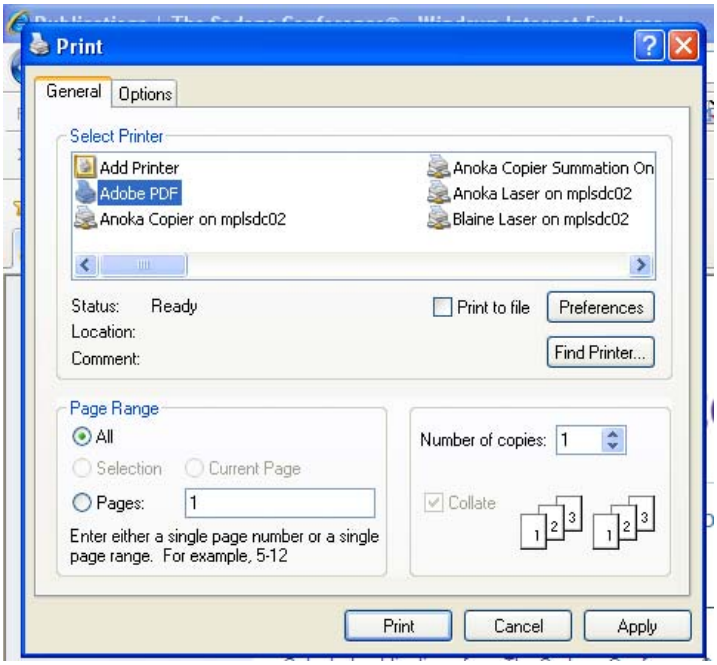
APPENDIX D: SAMPLE TECHNOLOGIES FOR PRESERVATION AND COLLECTION: PDFS AND THE WAYBACK MACHINE

Print to PDF Using Adobe Acrobat

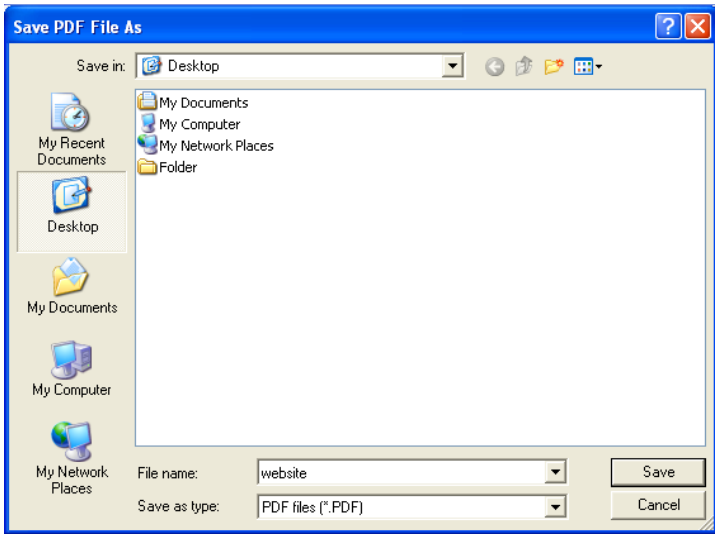
Step 1: Go to desired web page. Choose *File* > *Print* in the application. The Print dialog box will open.



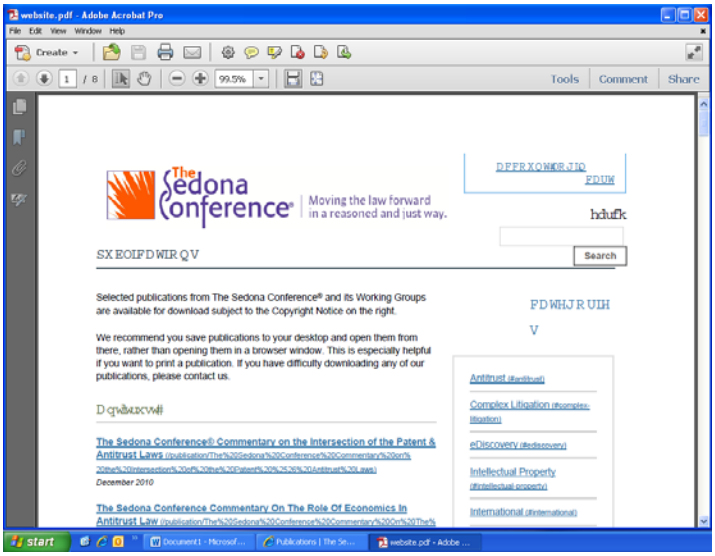
Step 2: Under *Print* menu, choose *Adobe PDF* as the printer selection. Click *Print*.



Step 3: When the standard *Save* dialog box opens, type a name for the file. Next, select the location that the user would like to save the file (for example, the desktop). Then, click *Save* or *OK*.

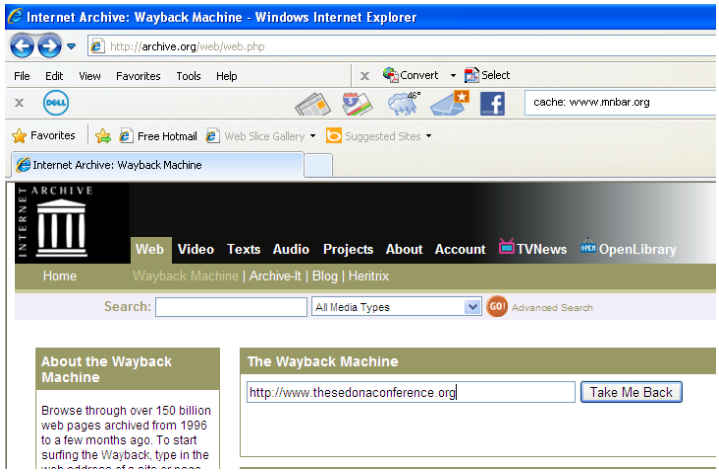


Step 4: A newly created PDF file should then appear (in this example, it was saved to the desktop).



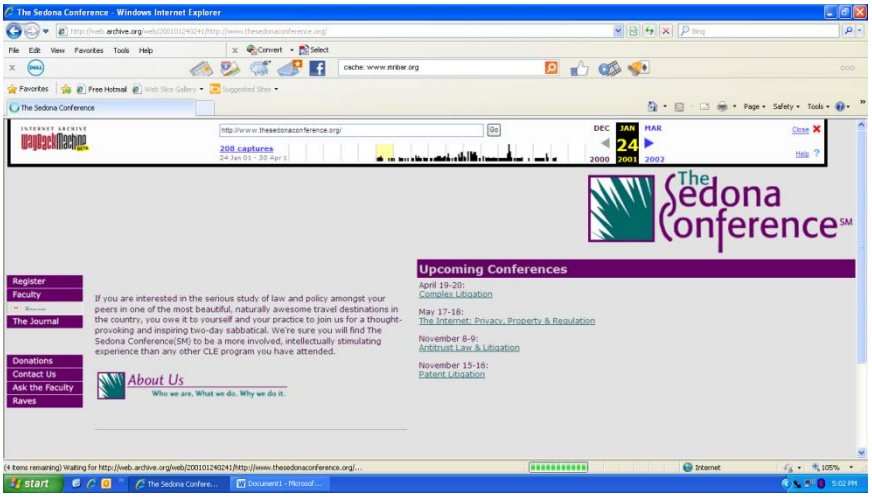
Using the Wayback Machine

Step 1: Go to http://archive.org/web/web.php. Then, enter the desired web page in the search box under the Wayback Machine. Next, click *Take Me Back*.



Step 2: The Wayback Machine will pull up a calendar listing the available dates, if any, that contain an archive of the web page.

Step 3: Click on the desired date and the past web page will appear. (In the example below, January 24, 2001 was selected.)



Using Print Screen

Each of the above-referenced static images of Web pages was created by:
Step 1: Pressing the Print Screen (*Prt Scr*) key; and then
Step 2: Pasting it into this document.