

GUIDELINES FOR THE DISCOVERY OF ELECTRONIC DOCUMENTS IN ONTARIO

A. Introduction: Purpose of E-Discovery Guidelines

In its Report, the Task Force on the Discovery Process in Ontario recommended the development of a “best practices” manual to address the discovery of electronic documents. These Guidelines respond to that recommendation.¹

The preservation, retrieval, exchange and production of documents from electronic sources in electronic form are together referred to as “e-discovery.” In these Guidelines, that term also includes the use of automated tools to produce documents in electronic form, whether they originate in hard copy or electronic sources. While documents from hard copy sources can be produced in electronic form, and paper copies of electronic documents can be printed out for production in litigation, these activities would not, in themselves, constitute “e-discovery” as the term is used, generally or in these Guidelines.

The development of best practices for e-discovery is not unique to Ontario. A number of other organizations and jurisdictions have implemented or published similar guidelines that have been instructional in the development of these Guidelines. These are referred to as appropriate in the commentary.

The premise of these Guidelines is that existing Rules already provide a legal foundation for the requirement that parties address issues relating to e-discovery, because the definition of “document” in applicable civil Rules already includes “data and information in electronic form.”² However, those Rules and the case law to date provide little clear guidance to parties and their counsel on *how* to fulfill that requirement. The suggestions in these Guidelines have been developed to address this issue with respect to production of documents in civil litigation.

E-discovery is already widely used as an integral part of the discovery process in complex cases and, increasingly, in many types of litigation that are less complex. In part, this is because of the inclusive definition of “document” referred to above. In addition, however, as the available technology matures, lawyers have begun to recognize its capacity, in some cases, to manage document production more efficiently, and to support the discovery process more effectively, than traditional paper-based methods permit.

However, many lawyers have yet to fully recognize the impact of this technology on the discovery process. The overall orientation of the profession remains towards printed documents. This, combined with the absence of clear guidelines on the scope and manner of e-discovery, means that many lawyers remain unfamiliar with their clients’ obligations to preserve and produce electronic documents, and with the technology available to retrieve, search and produce them in a cost-effective manner.

Accordingly, Section C below sets out a number of principles that are intended to guide lawyers, clients and the judiciary in the e-discovery process. It is hoped that these Guidelines will provide an appropriate framework to address *how* to conduct e-discovery, based on norms that the bench and bar can adopt and develop over time as a matter of practice. They are not intended to be enforceable directly, as are the *Rules*

¹ The Discovery Task Force wishes to thank the members of the e-Discovery Sub-Committee for their excellent work: Sara Blake, Peg Duncan, Martin Felsky, Michael Fraleigh, Derek Freeman, Karen Groulx, Christopher Leafloor, Daniel Pinnington, Mohan Sharma, Glenn A. Smith and Phil Tunley.

² Rules of Civil Procedure, Rule 1.03

of Civil Procedure, although they may support the enforcement of agreements between parties or provide the basis for court orders. Mandating how e-discovery is conducted through the enactment of detailed rules, at this stage, could be counterproductive, and risk imposing a “one-size fits all” approach that may not be appropriate in different types of litigation or responsive to new technologies as they emerge. It could also add unnecessary complexity to the Rules, and lead to more disputes and related motions.

Rather, the objective of these Guidelines is to educate the legal profession, including the judiciary and the practicing bar, on issues relating to e-discovery and how those issues can be addressed in practice. They are intended to provide practical suggestions for the profession, both on how to fulfill parties’ existing obligations respecting the preservation and production of relevant documents from electronic sources, and how to improve the cost effectiveness of the discovery process. They suggest how to reach early agreements in the e-discovery process, in order to minimize the potential for undue cost and delay.

These Guidelines also include some suggestions to take advantage of electronic tools, in order to minimize unnecessary cost and delay. Despite the apparent complexity of some e-discovery issues, technology increasingly offers improved methods of retrieving, reviewing and producing documents electronically. In many circumstances, this can offer significant savings of cost and time compared to paper-based methods.

In order to serve as an educational guide for the profession, it may be necessary for some readers to review the basic concepts and terminology relating to e-discovery. For those readers, Section B following provides this review in a practical context. It outlines the stages in the process of discovery of electronic documents, and some key terminology and concepts that lawyers and judges need to master at each stage.³ Those readers who are already familiar with this terminology and the e-discovery process may prefer to go directly to Section C.

B. Key Issues and Terminology in the E-Discovery Process

At every stage of the e-discovery process, lawyers are asked to give advice to clients about issues that involve new concepts, and new terminology, that highlight key differences between the discovery of electronic documents and traditional paper-based files. At each stage, disputes may arise about those issues that require court resolution. As a result, to deal effectively and consistently with these issues, both lawyers and the judiciary need to become familiar with new concepts and related terminology in the area of e-discovery.

This section introduces some of the most important ones that arise at each stage of the e-discovery process.

The stages of the e-discovery process do not themselves differ from those involved in traditional hard copy discovery. They are:

- (a) **LOCATION** of potential document sources;
- (b) **PRESERVATION** of potentially relevant materials;
- (c) **REVIEW** of documents for relevance, privilege and other issues; and
- (d) **PRODUCTION** to other parties, for use in court proceedings.

³ For a detailed glossary of frequently used terms, see The Sedona Conference® Glossary For E-Discovery And Digital Information Management. A Project of The Sedona Conference® Working Group on Electronic Document Retention and Production (WG1) RFP+ Group May, 2005; available on The Sedona Conference website (www.thesedonaconference.org).

Only by understanding the new concepts and terminology that come into play at each of these stages in the case of e-discovery, can lawyers and judges make informed decisions, avoid potential disputes in this area, or resolve them in a manner consistent with the Rules. This includes when and why it may make sense to seek or order production of electronic documents, and how to do so in a manner that remains cost effective to the parties.

(i) The Location of Electronic Documents

The first question that arises is what must be located, within the existing Rules definition of “data and information in electronic form”?

Generally speaking, documents are referred to as “electronic” if they exist in a medium that can only be read through the use of computers, as distinct from documents that can be read without the aid of such devices. It is also generally accepted that this definition includes many familiar types of electronic “documents,” such as e-mail, web pages, word processing files, and databases that are stored on computer.⁴ However, both the definition and case law suggest that a broader range of electronic “data and information” may also be covered in some cases. The limitations on what may be covered are not to be found so much in technical distinctions, as they are in the familiar criteria of relevance.

The next obvious question is what computer systems the client has, or had at the relevant time, that may contain relevant data or information. Again, depending on the nature of the case, the answer may include enterprise systems or networks, as well as personal computers (desktops, laptops, and even hand-held devices), and even individual components and media relating to them, such as memory chips, magnetic disks (such as computer hard drives or floppy disks), optical disks (such as DVDs or CDs), and magnetic tapes.

The variety of hardware and media involved can pose problems for lawyers, clients and the courts. For example:

- some items may be in use by individual witnesses, others in storage in different areas or departments, and the documents may be in a wide variety of different electronic formats;
- copies of the same document may be stored in multiple locations in the course of normal operations: for example, an e-mail sent from one person to another on a networked system may be saved by each of the sender and recipient on their own computers, and further copies retained by the system for a variety of purposes;
- relevant electronic documents, even those created using systems that were once commonplace, may have become unreadable over time because of the unavailability or obsolescence of key software or hardware components;
- in some cases, the sheer volume of data can be enormous, both because of the expanding use of computer systems and their increasing storage capacity, and also because of the way they affect the behavior of people and organizations: for example, e-mail is not only replacing traditional paper-based communications such as letters and memoranda in many circumstances, it is also replacing many

⁴ THE SEDONA PRINCIPLES: Best Practices Recommendations & Principles for Addressing Electronic Document Production. A Project of The Sedona Conference[®] Working Group on Best Practices for Electronic Document Retention & Production, published January 2004.

informal exchanges that in the past were not documented fully or at all, such as telephone calls and even casual conversations.

These factors can all make the process of locating and assembling electronic documents for litigation purposes more difficult than for traditional paper-based materials. The involvement of clients' IT staff is often essential to ensure that the assembly process is complete and problem-free.

In order to ensure the completeness of searches, lawyers also need to understand some of the different sources of documents that may exist within a given organization's computer systems, and their different purposes. Here, discussion with IT staff or consultants is essential, and the use of correct terminology can anticipate problems and avoid mistakes. For example, electronic documents familiar both in personal and business usage - such as word processing, spreadsheet, database and e-mail documents - may be found in several different electronic locations and formats. A complete search should consider the following possible sources:

- “**Active data**” is data that is currently used by the parties in their day-to-day operations. This type of data is normally straightforward to identify and access using the current systems. However, because this data is in active use, significant issues may arise for lawyers and courts concerning the need to preserve the integrity of this data for litigation, to design and manage searches to avoid business disruption, and to separate relevant from irrelevant information.
- “**Archival data,**” on the other hand, is data organized and maintained for long-term storage and record keeping purposes. Some systems allow users to retrieve archival data directly, but others require special equipment or software, and the involvement of IT staff.⁵
- “**Backup data**” is similar to archival data, except that this term refers to an exact copy of system data, which serves as a source for recovery in the event of a system problem or disaster. Backup data is generally stored separately from active data, and is distinct from archival data both in the method and structure of storage that reflect its intended uses. It is generally not accessible to ordinary system users, and requires special (and sometimes expensive) intervention before it is “readable.”

Archival and backup data both constitute a set of electronic data and information collected for a particular purpose, and perhaps as at a moment in time. That purpose and timeframe may or may not be related to the litigation, and their relevance and completeness need to be assessed in that light.

Lawyers and the judiciary should also be aware that certain electronic sources, such as internet web-pages or database applications, may be under constant revision as new information is published on the site or added to the system. Unless these documents are located promptly, the available active copy may not reflect what the data actually looked like at the point in the past that is relevant to the litigation. Lawyers should be prepared to question their clients, to confirm which of the available versions are the best evidence for litigation purposes.

The documents most commonly requested and produced in litigation are those created by word processors, databases, spreadsheets, e-mail, and other familiar programs. These documents are routinely used and exchanged in business and private dealings. As noted above, these documents are normally quite easy to identify and locate. However, in discussions with IT staff involved, lawyers also need to be aware that many other, different kinds of “information and data” can exist in computer systems, in order to assess how and when they may be relevant. These may include less familiar kinds of documents, such as web-pages,

⁵ The Sedona Conference® Glossary For E-Discovery And Digital Information Management. A Project of The Sedona Conference® Working Group on Electronic Document Retention and Production (WG1) RFP+ Group May, 2005.

browser history files that track a user's movements between web-sites and pages on the internet, cell-phone logs, and many other kinds of information stored on computer-based devices in their day-to-day operations. Most users may be completely unaware these documents even exist.

In addition, there may be hidden data or information associated or related to electronic documents that should be considered, particularly if there are issues of authorship or authenticity raised with respect to a document. Case law suggests that any data or information that can be readily compiled into viewable form, whether presented on the screen or printed on paper, is potentially within the definition of "document" under Rule 30.01 of the *Rules of Civil Procedure*. Again, some understanding of the concepts, as well as the terminology involved, is essential.

- "**Meta-data**" refers to electronic information that is recorded by the system about a particular document, concerning its format, and how, when, and by whom it was created, saved, assessed, or modified. For example, most word processing software records who created or modified a document, as well as the dates and times of document revisions. Most e-mail software records the dates and times e-mails are created, sent, opened, and saved as well as the names of the originator and all recipients, including those "blind copied." This information may not be seen by users or appear in a print-out of the document in the ordinary course of business. However, meta-data is generally readily available, and can be extracted in searchable or printable form if it is relevant to litigation. Meta-data may be relevant directly to the litigation or it may be relevant to the authenticity and admissibility in evidence of the electronic documents with which it is associated, where this is disputed. Accordingly, its importance should not be underestimated.⁶
- "**Residual data**" refers to any information that remains stored on a computer system after a document has been deleted. The computer does not necessarily "wipe clean" the disk or memory space in which the file was stored, but merely "tags" it as re-usable by the system. The "deleted" data may not become truly unavailable until this space is re-used. Hence, deleted files or fragments of deleted files are often retrievable for some period of time after "deletion." This can provide information about a document, and sometimes about changes made in successive revisions of a document, that would not otherwise be available. This kind of information is only recoverable using special "forensic" methods, and is unlikely to have significance in most litigation.
- "**Replicant data**" is created when a software program, such as a word processor, makes periodic back-up files of an open file (e.g. at five minute intervals) to facilitate retrieval of the document where there is a computer malfunction. Each time the program creates a new back-up file, the previous back-up file is deleted, or tagged for reuse.

Lawyers must understand the different kinds of electronic documents that may exist, and their characteristics, in order to assess whether and how they may be relevant, and where they may be found in a given case. Without some guidance from their lawyers on these issues, parties involved in litigation are unlikely to be able even to identify and locate the various electronic information and data that may have key relevance to their dispute.

(ii) Preservation of Electronically Stored Documents

A party's duty to preserve electronically stored documents that are relevant to contemplated or threatened litigation arises in the same way as for paper documents.

⁶ The Sedona Conference® Glossary For E-Discovery And Digital Information Management. A Project of The Sedona Conference® Working Group on Electronic Document Retention and Production (WG1) RFP+ Group May, 2005.

However, the discussion and terminology reviewed above highlights some special problems that can arise in the preservation of electronic documents, and also suggests how they can be addressed. Specific guidance is offered in Section C below, but the following are some examples of practical problems that arise from the lack of such understanding, and of the solutions that may often be available.

- Electronic documents or media containing them may be considered obsolete by the client in terms of its current business systems, but may nevertheless be recoverable to a readable form by specialized forensic methods. The costs involved, at least for many of the most commonly used methods, have declined to a point that may be cost effective in an increasing range of litigation.
- Relevant meta-data may exist at the time an electronic document or source is located, but may be altered or lost simply in the process of making a copy of the relevant electronic files for litigation purposes. This again is avoidable, as relatively affordable techniques exist, either to make “forensic copies” or “mirror images” that are specifically designed to preserve the integrity of the meta-data, or to capture the relevant meta-data from the original source documents before they are copied.
- Preserving web-site files in electronic form, rather than simply printing them up at a point in time, may enable a party, at minimal cost, to recreate the website electronically in a courtroom, in order to demonstrate dynamically any relevant links, relationships, and special features that characterized the site at the time the litigation arose.
- Formal document retention policies are a relatively recent development, and even today may not be standard except in the very largest and most sophisticated organizations. Moreover, sound business reasons may exist for practices that result in the destruction of relevant electronic documents: for example, routine deletion or omission to back-up e-mail to maintain storage space. For these reasons, early discussion with IT staff is often necessary to prevent continued deletion after litigation is threatened or commenced.

These examples illustrate the point that, in order to understand how to comply with or enforce the obligation to preserve electronic data and information for litigation, parties, lawyers and the courts first need to understand the characteristics of electronic documents and the concepts and terminology of e-discovery discussed above.

(iii) Electronic Document Review

The preceding discussion of the ways electronic documents differ from paper also affects the approaches to the review of available electronic materials for litigation purposes.

Review of electronic documents is essential, first, to separate relevant materials, which should be produced, from irrelevant material, which should not. Over-production of irrelevant electronic documents may be just as damaging to clients’ interests and the litigation process as incomplete production.

However, the sheer volume and particular characteristics of electronic documents may be a significant barrier to effective review, for a number of reasons:

- Many institutions and businesses save a copy of their entire system onto back-up tapes periodically, and some retain them for long periods of time. Computer back-up tapes can store huge amounts of data, which may be organized for purposes of disaster recovery, rather than normal usage. It often needs to be converted back to readable form, before it can be searched or printed out to determine relevance. The

volume and organization of archive and backup data, and costs of conversion, can be significant barriers to production, especially as restoration may require processing a complete set of back-up tapes together.

- Depending upon the institution's retention policies, the resulting set of documents (although complete and accurate for the purposes for which they were stored) may be incomplete or may not fully reflect the status of the same documents at the time relevant to the litigation.
- The document set may also contain multiple duplicates. Electronic documents are easily duplicated and, as noted above, copies of the same document may be stored in multiple locations in the course of normal operations. Consequently, although a user may have deleted his/her own copy, others persist in other locations, often without the user's knowledge.
- Earlier versions (including drafts) or later versions may still be retained. Unless clearly marked – or better yet, unless the relevant meta-data has been preserved - it may be impossible to know which version is earlier or later, and which version is relevant to the timeframes and issues raised in the litigation.
- Since even meta-data could, in certain cases, contain or reveal privileged, secret, or other sensitive information, an organization may determine that it too must be separately reviewed before the documents are produced.

Once the files are collected in readable form, manually searching for and retrieving specific files may be cumbersome, time-consuming and prohibitively expensive. Depending on the documents and the technology used, however, automated search tools may offer solutions. E-discovery has been greatly facilitated by new technologies that permit some kinds of electronically created documents to be converted from one digital form into another, in large volumes, often at minimal cost. This means that in some cases the practicing lawyer and client may no longer face prohibitive cost and technology barriers to the review and searching of electronic documents, particularly with respect to many common forms of electronic documents, such as e-mail.

In some cases, however, even the available electronic tools may not permit complete review for production in litigation on a cost-effective or timely basis. Lawyers and the judiciary in such cases need to seek agreements, or arrive at terms for court orders, that target the most relevant data and information.

(iv) Production of Documents in Electronic Form

The question lawyers are increasingly asked to advise on (and courts may be asked to adjudicate) is whether parties may simply print out electronic data such as e-mails, or whether they are obliged to produce them to the opposing party in electronic form. The answer in any given case may involve a balance of competing considerations.⁷

In order to maximize the benefits of e-discovery, the courts and the profession need to gain experience with respect to such issues as: what circumstances call for electronic production as opposed to paper production;

⁷ For example, many electronic documents involve more than mere printable text. In a database application, individual pieces of information may be meaningless, unless they are produced within their context or environment, and the ability to manipulate relevant information using the original software application in which it was created may bring added benefits. However, a database may often contain irrelevant, confidential, and even privileged information, together with the relevant information, or the software application may not be available commercially, or at all, to third parties. In such cases, standard or custom "reports" displaying the relevant information with the context in a readable form might be generated, without producing the entire system, and may be sufficient.

how the cost of production should be fairly allocated; how to ensure that electronically produced documents are compatible with courtroom technology to facilitate production at trial; how to provide for the redaction of privileged and irrelevant material in electronic form; and how to ensure appropriate retention of electronic records.

These issues are very much affected by the availability of new technology, and its increasing use by lawyers and courts. Most litigation support software provides for exporting production sets in formats that allow them to be imported by a recipient party into the litigation support tool of their choice. Many of these tools are designed to produce properly redacted versions of documents⁸, to permit the creation of special fields for production of relevant meta-data, and to allow the user to select which fields will be exported.

Similarly, large volumes of hard copy documents can be scanned as image files, and exchanged on CDs or via web-based software, often at less cost than would be involved in producing a similar number of photocopy sets. This is especially important in multi-party litigation, and where parties have the opportunity to share the costs of scanning. With the assistance of available software tools, electronically scanned documents can be much easier and more efficient to store, organize, manage and search, than equivalent volumes of paper documents. These developments are rapidly reducing cost and technological barriers to high-volume document cases, even where the client's source documents exist in paper form.

However, the use of these new tools and methods is still limited, and sometimes inconsistent, among lawyers and the judiciary. These Guidelines are intended to promote the efficient use of technology in the discovery process. The control of escalating costs, together with increased effectiveness for lawyers and parties advancing their case through the discovery process, is an important part of the rationale behind these Guidelines.

C. Principles that should Guide the E-Discovery Process

(i) Discovery of Electronic Documents (“E-Discovery”)

Principle 1: ***Electronic documents containing relevant data and information are discoverable pursuant to Rule 30.***

Commentary: As soon as litigation is contemplated or threatened, it is essential for parties and their counsel to go beyond paper file searching, and consider what electronic data and information exists that they may need to produce. Parties must take reasonable steps to locate and preserve electronic documents containing data and information that can reasonably be expected to be relevant to litigation. Further, parties should consider what relevant electronic documents other parties may have, that they may want to request be preserved for production in the course of the litigation.

Principle 2: ***The obligations of the parties with respect to e-discovery are subject to balancing, and may vary with (i) the cost, burden and delay that may be imposed on parties; (ii) the nature and scope of the litigation, the importance of the issues, and the amounts at stake; and (iii) the***

⁸ Counsel using such tools should ensure that redactions are permanently embedded in the production copy of the document, and cannot be electronically “undone”. Counsel should also ensure that, if a full-text or OCR version of the documents is also being produced, this version, as well as the image, should be redacted.

relevance of the available electronic documents, and their importance to the court's adjudication in a given case.

Commentary: This principle is consistent with Rule 1.04(1), and the objective of securing the just, most expeditious, and least expensive disposition of litigation on its merits.

Even where there has been complete production in paper form, electronic versions of the same documents may contain relevant meta-data that may not appear in a printout or scanned version of the document. Meta-data may be directly relevant in the litigation, or it may be relevant where there is an issue as to the authorship or authenticity of a document. In such situations, it may also be necessary to produce the relevant meta-data in some form. Parties should consider whether it may be preferable to produce the entire document, including the meta-data, in electronic form.⁹

The questions to be considered in determining whether to require the use of forensic techniques to recover back-up or obsolete sources include not only the costs involved, and the potential amount, usability, reliability and relevance of the information to be obtained, but also:

- whether the party believes that the materials available from active electronic and paper sources are reasonably complete;
- whether the party has rules for printing up or retaining important documents in electronic form, and whether they are monitored for compliance; and
- the availability and completeness of the back-up or obsolete sources.

Parties should use the most cost-effective methods to locate, preserve, review and produce electronic documents. Electronic documents may be easier to search than printed or scanned copies, and therefore more effective in litigation, and production of documents in electronic form may be more cost-effective than print production.

The costs to be considered may, where appropriate, include the costs of counsel and any necessary consultants, hardware, software or other facilities or services required (i) to recover or make electronic documents available in a readable form; (ii) to search documents in various formats to identify relevant material, and separate irrelevant material; (iii) to

⁹ An example of a case where resort to back-up tapes was ordered by the court is in the U.S. decision of *Zubulake v. UBS Warburg LLC*, 2003 W.L. 21087884 (S.D.N.Y. May 13, 2003), an action claiming gender discrimination and illegal retaliation, where a request for an order compelling UBS to produce various e-mails now existing only on back-up tapes and other archived media was before the court. Despite the fact that UBS had already produced approximately 100 pages of e-mails, Zubulake believed it had more based on the fact that she herself had produced approximately 450 pages of e-mails. The court determined that UBS should provide tangible evidence of what the backup tapes might have to offer in the form of a sample. UBS was therefore ordered to produce responsive e-mails from any five back-up tapes selected by the plaintiff. UBS was also required to prepare an affidavit detailing the results of its search, as well as the time and money spent. Following the production of relevant e-mails taken from the sample back-up tapes, UBS was ordered to restore its back-up tapes and produce responsive e-mails from these tapes. The case suggests that, where a party on proper evidence convinces a court that documents have not been produced and that such documents are likely stored on a computer hard drive or other electronic storage medium, such as back-up tapes, but the party in possession of the computer asserts it has printed or produced all that it has, then the only solution would be to allow inspection of the storage medium itself or restoration of the documents from back-up tapes.

