Why electronic documents are different

For litigators and non-litigators, discovery and document production are familiar parts of the litigation process. Formal court rules, the law of evidence and informal practices – all of which evolved in a paper-based world – direct and govern how a discovery proceeds.

But, the world has changed.

Dan Pinnington

In just the last decade, personal computers, e-mail and the Internet have transformed how the world functions and communicates. By some estimates, up to 95 per cent of all new information is created in an electronic format, the majority of which is never printed. Most people are now very comfortable with creating and sharing electronic documents, and use e-mail as a primary communications tool.

These electronic documents and e-mails are sometimes relevant in legal disputes, and in many cases, are pivotal evidence. The courts recognize this, and electronic data is producible on discovery, both under the *Ontario Rules of Civil Procedure* and the Federal Court Rules.

What are "electronic documents"?

In the *Ontario Rules*, a "document" is defined broadly as including "data and information recorded or stored by means of any device." Thus, a document includes wordprocessing files, e-mail messages, Web pages, information stored in databases, and virtually anything else stored in any electronic device, including cell phones, Blackberries, PDAs, voicemail systems, instant messaging clients, iPods, TiVos, digital copiers, and even printers (some multi-function printer/scanner/fax machines have hard drives in them).

Electronic information can also be found on a variety of media, including magnetic disks (such as computer or laptop hard drives, floppy disks, or zip drives), optical disks (such as DVDs or CDs), magnetic tapes (most commonly backup tapes) and USB memory sticks.

In a number of obvious, and sometimes not so obvious ways, electronic documents are profoundly different from their paper counterparts: Members of the legal profession, including lawyers and judges, need to understand how electronic documents are different, and how these differences affect the preservation, collection and disclosure of information from electronic sources, now commonly known as "electronic discovery."

Vast accumulations of electronic data

In today's world, electronic documents vastly outnumber paper documents, and are being created at rates far greater than paper documents ever were. As a result, the amount of information potentially available for discovery has increased exponentially.

Consider the example of e-mail. In all of 2004, Canada Post handled 10.9 billion pieces of mail. In North America today, billions of e-mails are sent every day.

Hard drives are much larger than they were just a few years ago, and can hold massive amounts of information. Today a typical desktop computer has a 40 gigabyte hard drive, which can potentially hold about four million pages of information. A typical network server might have an 80GB hard drive, which could potentially hold eight million pages of information, and the backup tapes for that server would likely be the same size and hold a similar number of pages.

With typical e-mail usage and large hard drives, even a small business with only a dozen computers can have significant amounts of electronic information. Medium or large businesses have unimaginable amounts as they may have dozens of servers, hundreds or thousands of desktop computers, and boxes full of backup tapes. Try to fathom the amount of electronic information generated by the Ontario government.

Just using a computer creates more electronic information. In various ways, Windows will track and keep a list of the documents you opened or accessed, the network servers you connected to, the Web sites you visited, and more. Some of this information is lost each time you turn your computer off, but some of it will remain indefinitely in various files on your computer hard drive.

Data can easily replicate itself

One reason electronic documents are so widespread is because they are more easily replicated than paper documents. On a large scale, electronic information is replicated by users, and in various automated ways, all without any degradation of the data.

Consider the e-mail example again. E-mail users frequently send the same e-mail to multiple recipients, often with attachments. Some of those recipients may reply to that message, or forward it to others, and so on. Many e-mail systems automatically keep information about sent and received messages, sometimes including actual copies of messages. In many business environments, copies of data on servers, including e-mail logs and actual messages, are backed up on a daily basis. Most businesses don't keep all their backup tapes because they rotate them, but many keep month-end, quarter-end or year-end tapes going back several years. Thus, identical copies of a widely distributed e-mail could be found in many different places.

The auto-recover or auto-save feature found in many software programs, including Word and WordPerfect, can also replicate data. This feature is designed to prevent data loss by automatically creating a complete and identical backup copy of any currently open document at a regular specified interval, often every few minutes. This data – called replicant data – is stored on the hard drive as separate documents, which are supposed to be deleted when programs are closed. Because they are often not deleted, they can provide copies of documents long since changed or deleted.

Deleted does not mean deleted

Electronic documents tend to be much more difficult to dispose of than paper documents. If you delete a file from your hard drive, and take the extra step of deleting it from your Recycle Bin, the common assumption is that the file is gone.

Nothing could be further from the truth.

In fact, when you delete a document on a hard drive, you are only erasing the pointers to the location of the file data on the hard drive. The actual data itself remains on the hard drive, completely untouched.

This data, called residual data, is invisible to Windows and the computer user. Yet often one-third or more of the physical space on a hard drive contains information from deleted files; this information will be "deleted" only when the computer recycles the space by placing new information in it – weeks, months or even years after it was first created. Deleted files (or at least portions of them) therefore can be recovered long after they supposedly have been deleted – although to do so requires specialized software tools and assistance from an IT person or forensics expert.

To completely erase deleted files you must use specialized software that "scrubs" the hard drive.

And remember, you may not have to go to the trouble of trying to recover deleted documents if they were on a server hard drive, as you may be able to find them on one or more backup tapes or as attachments to e-mail messages in Sent e-mail folders.

Increasingly, e-mail messages are the pivotal evidence in a matter, and they are almost impossible to delete after they have been sent. Let's say you send an e-mail, delete it from your Sent folder, and ask the recipient to delete it when they receive it. It's gone, right? Absolutely not.

Electronic footprints from that message can exist in several places, and actual copies of the message are potentially available in at least a few locations. The moral of the story: Don't say anything in an e-mail that you would ever want your mother, children or boss to read on the front page of the newspaper. E-mails are often out there somewhere – all you have to do is find them.

Metadata – friend or enemy?

Metadata can be simply described as "data about data." Think of it as a hidden level of extra information that is automatically created and embedded in a computer file. Most software programs, including Word and WordPerfect, have metadata in their files. Metadata can be necessary for the operation of the software, or in some cases, is simply intended to be helpful to computer users – even though most don't even know it is there.

Parties exchanging documents electronically as part of a discovery (and lawyers sending e-mail attachments to clients or opposing counsel) need to appreciate that electronic document files include both the information you see on the screen, and metadata, which you don't see. This metadata is often sensitive or confidential information that can be damaging or embarrassing if seen by the wrong eyes. It can make or break your case.

Among many other things, metadata can include the following types of information:

- the date the file was created;
- the name of the person who created the file;
- the names of people who edited a file, and the date and time they did so;
- document revisions, including deleted text that is no longer visible on the screen;
- the name of the computer the file was created on; or
- the name of the hard drive the document was saved on.

E-mail messages can be divided into two parts: the body, and header – which is metadata. The body contains the part of the message you see on the screen (To, From, the subject line, and the contents of the message). The header, which you never see, has a large amount of metadata which details, step-by-step, the entire route the e-mail took as it crossed the Internet. This information can be useful in verifying who saw an e-mail or where it was sent.

Although some metadata can be viewed within the program that has created a file (in Word or WordPerfect documents click on File, then Properties, and review the information in the Properties dialog box), in most circumstances it can only be seen with specialized software.

For more information on metadata see the **Dangers of metadata** article from the June 2004 issue of LawPRO Magazine (www.practicepro.ca/metadata).

In most cases, metadata will have no material or evidentiary value, as it will not matter when a document was printed, or who typed revisions, or when edits were made before the document was circulated. However, in some cases metadata may help authenticate a document, or establish facts material to a dispute such as when a file was created or accessed, or when an e-mail message was sent. Understanding when metadata needs to be preserved and produced represents one of the biggest challenges in electronic discovery and document production. It can be very expensive and time-consuming to capture and preserve metadata; as well, at an early stage of a matter, it is often not clear if metadata will be relevant, and if so, what steps a lawyer and client should take to prevent metadata from being lost or destroyed.

Dynamic and changeable content

Electronic documents and data (and their associated metadata) are dynamic and can change over time, even without human intervention. Consider for example, information in a database, or Web pages that are built with information from a database. Unlike paper documents, many electronic documents and databases are never in a fixed and final form, and there isn't always a copy that shows exactly what the data looked like at any given point in the past.

Moreover, the act of merely accessing or moving electronic data can change and even destroy it. For example, several hundred files are accessed and changed when you simply boot up a computer, and potentially relevant files may be over-written. Moving a wordprocessing file from one hard drive to another can change a number of attributes in the file. Opening and reading an e-mail message can change metadata information for that message, and in some cases, could potentially over-write metadata that might be helpful on the matter in question. It is not uncommon for internal IT staff to destroy potentially helpful metadata when they are dispatched to find information relevant to a potential or pending litigation matter. For this reason, it can be helpful to have a forensics expert assist in preserving and collecting electronic data. They will use special tools to make a true image of all data on hard drives, including deleted data. This will properly preserve all available data for a thorough forensic analysis, if required.

The dynamic nature of electronic documents also makes them much easier to change than paper documents. It is easy to "spoof" or fake the sender's name on an e-mail. Spammers do it all the time. Documents in electronic form can be modified in numerous ways that are sometimes difficult to detect, even with computer forensic techniques. In some cases metadata can assist in verifying the authenticity of an electronic document.

Environment dependence

In many cases electronic data will be meaningless when separated from its original or native software environment. For example, consider data within a database that includes custom reports to organize and present summaries of the data. If the raw data is produced, it might appear as a long list of undefined information and numbers. To make sense of the data, the viewer needs the native software to access and manipulate the data. In some cases,

E-Discovery resources

PracticePRO resources

practicePRO has posted a number of additional resources to help you learn more about electronic discovery. At www.practicepro.ca/ediscovery you'll find:

Electronic Discovery – A Reading List
Peg Duncan, Director, Business Opportunities and Emerging Technologies in the Information Management Branch of the Federal Department of Justice has prepared a list of some of the best Web sites and online articles on various ED issues.

ED request letter (sample)
A sample letter based on the annotated discovery request letter provided by Martin Felsky and Peg Duncan and featured in the September 2005 issue of LawPRO Magazine.

Additional resources

Electronic Discovery Guidelines
A first draft of the Electronic Discovery Guidelines prepared by the Electronic Discovery Sub-committee of the Discovery Task Force will be posted on the Ontario Courts Web site in mid-fall (www.ontariocourts.on.ca/)

Electronic Discovery and The New ED Guidelines – A Roadmap for Dealing with Electronic Information
An ED CLE program, jointly put on by the Ontario Bar Association and The Advocates Society on Monday, November 28, 2005, from 9 a.m. to 4:30 p.m. at the OBA Conference Centre in Toronto. For program information and to register, go to www.softconference.com/oba

the courts have recognized a duty to produce electronic evidence in a form and manner that is usable by the party receiving it.

Technology obsolescence

The frequent obsolescence of computer systems due to changing technology can create many issues for recovering electronic documents that are no longer in active data sources, that is, data sources that are in regular use every day and easy to access. Over the course of many years, a business may use different e-mail systems or different backup hardware and software. Organizations often find themselves with boxes of backup tapes they can't read, or data on backup tapes that can't be opened as there is no software or hardware available to access this old or legacy data. Keep in mind that if you keep data, you face the potential obligation to produce it, regardless of the time or expense required to do so. A good document retention policy that ensures the destruction of legacy data can help to reduce the exposure to crippling and costly productions in the event of litigation.

Who really had access to the document?

It can be more difficult to determine the providence of electronic documents than paper documents. Electronic files are often stored in shared network folders that multiple users can access. As well, the increased use of collaborative software allows the group editing of electronic documents, which makes it more difficult to determine authorship.

Searching & finding needles in many haystacks

While an employee's paper documents will often be consolidated in a handful of boxes or filing cabinets, the same employee's electronic documents can reside in numerous obvious locations, such as the work desktop, laptop computers, network servers, floppy disks, and backup tapes. They may also be found in not-so-obvious locations such as home and cottage computers, and personal or browser-based e-mail accounts. To some degree, lawyers will need to assist clients in identifying what must be preserved, collected and produced, and they will have to question the opposing party to ensure everything that should have been produced was produced.

On the plus side, some forms of electronic data and electronic media can be searched far more quickly and accurately than paper versions. With a well-developed search strategy, you can narrow the scope of your search and find the small amount of relevant data within vast collections of electronic data. Search strategies involve identifying specific search terms that will target relevant data, and setting other parameters that will limit and filter search results.

For example, you might want to look at e-mail messages sent to or received by a particular person in a narrow time frame which contain a certain term.

A good search strategy will comb through large amounts of data and give you a collection of documents that is smaller and more manageable in size. Specialized electronic evidence tools can also de-dupe search results to remove extra, identical copies of documents or e-mail messages.

e-Discovery in an electronic world

The differences between electronic documents and paper documents make it clear that discovery can be different in the electronic and paper worlds. To meet their obligations to assist clients in preserving, collecting and producing all relevant data, and to have the ability to ask appropriate questions to ferret out and find relevant electronic data from the opposing side, lawyers need to better understand what electronic documents and data are, and where they can be found.

Dan Pinnington is director of practicePRO, LawPRO's risk and practice management program. Dan can be reached at dan.pinnington@lawpro.ca