

Five Things You Need to Know about Your Data

Pitney Bowes Legal Solutions

On-site Management | Litigation and Document Services

The Top Five Things You Need to Know About All Your Data

eDiscovery Inside and Outside the Comfort Zone

It is impossible to overstate the impact advances in digital technology have made on Electronically Stored Information (ESI) in the last ten years. In 1999, dot-com businesses were exploding onto the scene as millions of people began to depend on emails and the Internet for both business and personal information. Today, smart phones, iPods and thumb drives transport information to the far corners of the earth. Through texting, people use their phones as data delivery systems. Phone conversations become data packets on Voice over Internet Protocol (VoIP) networks. For the legal professional, these developments have made tracking down ESI during eDiscovery an entirely different, much more complex process.

Litigation Becomes Bit-Mapped

Today, 90 percent of legal documents are electronic. This means that for litigation, compliance and evidentiary chain of custody, law firms, government agencies and large corporations need to efficiently manage massive amounts of ESI. Just one lost or overlooked email or file can have a huge impact on a case.

In 2006, there were 944 million Instant Messaging accounts. By 2010, there will be 1.4 billion. In 2007, 3.5 exabytes of email data were sent. An exabyte is the equivalent of 600 trillion books, each 500 pages in length.¹

On top of this, the Federal Rules of Civil Procedure (FRCP) amendments implemented on December 1, 2006, established a very broad definition of ESI as "... data compilations stored in any medium from which information can be obtained—translated, if necessary, by the

respondent into reasonably usable form...."²The amendments also include ESI as part of initial disclosure, now defined as: "... a copy of, or a description by category and location of, all documents, ESI, and tangible things that are in the possession, custody, or control of the party..."³

Clients need to avoid the situation where they are held accountable by the courts for innocently organizing their data but not knowing where it all is. In a \$1.5 billion settlement against Morgan Stanley, the court recognized that innocently organized data is no longer acceptable.

The world of eDiscovery has grown from 55 vendors in 2002 to over 600 in 2007, serving a \$2 billion market.⁴ In addition to the massive increase in ESI volume and the impact of the FRCP amendments, this growth in eDiscovery has been driven by billion dollar sanctions and the convergence of compliance and regulatory oversight.

The variety of formats and locations for ESI keeps growing, while there is little perceived burden regarding the job of discovering all this data. The fact remains, attorneys need to know everything. The fear of sanctions mounts and clients assume that everyone in their law firm is an expert in eDiscovery.

The Top Five Things to Know About Data— Inside and Outside the Comfort Zone

Let's look at the key things they need to find out in eDiscovery. This brings us back to the basics: the who, what, when, where and how of the data. It's important to realize there is ESI that we're all familiar with – data that's inside the

eDiscovery comfort zone -- but there's also "exotic" eDiscovery, where we are dealing with new ESI formats and locations, which lie outside the comfort zone.

The ESI comfort zone includes four categories of easily recognizable files. These are the types of data we pursue in "standard" eDiscovery:

- Email (pst, nsf)
- Microsoft Office files (doc, xls, ppt)
- System files (sys, com, exe)
- PDF files

There is no mystery to these file locations, all of which are accessible – servers, backup tapes and drives, desktops, laptops, CDs, DVDs and home computers.

Once we go outside the comfort zone, however, we enter a world where ESI has a wide variety of sources and formats for how the data is packaged, transported and presented -- and these formats are constantly changing. iPods and thumb drives are now data transport devices. In fact, most electronic gadgets today hold discoverable data. Here's a list of data types and storage locations that need to be searched as part of exotic eDiscovery:

- Cell phones
- Business application databases
- Building security systems
- Debit and credit card databases
- Audio evidence
- VoIP telephone networks
- Websites
- Text Message servers Instant Message servers, chat rooms and bulletin boards
- Blogging
- PDAs and handhelds
- Digital cameras and camcorders
- Jump and thumb drives
- iPods and MP3 players
- Legacy data

Now let's review those top five things we need to know about all types of data and observe the

differences between standard and exotic eDiscovery:

1 WHO:

Standard eDiscovery: For these files, you need to identify custodians, authors and creators.

Exotic eDiscovery: The "who" of these data formats also includes system administrators and web developers, as well as IM discussion threads and social networks.

2 WHAT:

Standard eDiscovery: All you basically need to know is the file type, size and whether any compression was used.

Exotic eDiscovery: Here, you could be also dealing with legacy applications, encryption, password protection, formulas, database specifics, VoIP networks, xml and docx, audio files and data embedded in RAM.

3 WHEN:

Standard eDiscovery: You just need dates created and modified, which are typically easy to find.

Exotic eDiscovery: For this data, you may need to check time zones and the timings on any network metadata. Other issues include retention policies, legal holds and litigation readiness.

4 WHERE:

Standard eDiscovery: With these types of ESI, you look for network servers, hard drives, laptops and backup devices, both local and networked.

Exotic eDiscovery: Here, you have many more device types and a huge number of each type. This covers everything from PDAs, cell phones, iPods and MP3 players, to collection servers, internet servers and portable media such as thumbdrives, CDs and DVDs.

5 HOW:

Standard eDiscovery: You need to know about the office suite or other applications, special formatting and email client/server information.

Exotic eDiscovery: Here you also may be required to know the operating system, legacy platform, and what collaboration tools, hidden

text, formulas, databases and web pages may be employed.

Roadblocks to Exotic eDiscovery

When sleuthing for information outside of standard eDiscovery, here are some of the difficulties typically encountered:

AUDIO/VIDEO EVIDENCE

- This may be located in voicemail, on VoIP networks, in web conferences, as part of depositions and in call centers. Issues to deal with include: the expense of audio review in real time; the risk of unknown content; the need to work under short deadlines and overcome reasonable accessibility because of FRCP Amendments; and finally, eternal retention requirements, which do not make analysis feasible.
- Searching challenges include: the quality of the recording; semantics, such as slang and code words; foreign languages; accents, and the need for digital conversion from tape.
- Search methodologies are either expensive or non-existent. You can: sit and listen; transcribe, then search and review; or use phonetic technology such as Nexidia or Kroll. Video evidence is discoverable, although nothing presently exists to index based on visual cues.

LEGACY APPLICATIONS AND DATABASES

- These may contain valuable information for the 26(f) Meet and Confer rule of the FRCP Amendments. But problems begin with the fact that legacy applications and data points are often unusable outside the native environment. Finding legacy hardware, typically mainframes, is challenging.
- You can exercise the option of converting to current application software and/or hiring experienced legacy programmers with Y2K experience.

- Just getting the data points isn't enough.

WEBSITES AND WEB DATA

- The first challenge here is determining the relevancy of the data. You need to determine the scope of the site data and whether this web content constitutes the best evidence available or you need to explore other sources.
- The second challenge is dealing with dynamic web pages. Static pages can be saved locally or captured by PDF as well as "spidered" for all pages, but dynamic pages may have multiple data sources and preservation issues.

TEXTING AND INSTANT MESSAGING

- These communication streams are more convenient to use than email but harder to capture. Conversations are stored at the service provider and on local machines. Corporate IM systems may regularly be locked down or flushed.
- Both these technologies make discoverable personal communications that deal with company issues. Like conversation, these communications are often more damaging than email.
- Characterization is difficult (what is a document in this context?) and production can be problematic (what is an IM unit?).
- Cut and Paste are not acceptable in these formats and there can be authentication issues.

CELL PHONE DATA STREAMS

- Exotic eDiscovery needs to go beyond LUDs (Local Usage Details). SIM cards specify both internal and external information.
- The service provider also has data points, logs, text messages and other information.
- PDAs offer similar recoverability of an even wider range of both current and deleted data.

RAM AND MORE

- New precedents were set in the landmark case, *Columbia Pictures v. Bunnell*, 2007 WL2080419 (C.D. Cal. May 29, 2007). These include the fact that IP log files must be tracked and captured and are discoverable.
- In addition, RAM is now considered ESI and comes under the FRCP Amendments.
- All this eDiscovery should be localized to the facts of a particular case, but it could start an avalanche of temporary data cache discovery issues.

MICROSOFT® OFFICE 2007

- This popular suite of office applications was re-written for PCs running Microsoft's latest Windows® Vista operating system. Many of its features make eDiscovery both more important and more challenging.
- BitLocker is now hardware based, so you may need the whole computer, not just the hard drive, to access data. In addition, there are now multiple layers of security during the booting up process.
- Aiding eDiscovery, point-in-time shadow copies are created as the user works, along with an automatic version trail.
- There are also new file formats and a save to xml.

Concluding Thoughts: Getting Ahead of It All

Given the size and complexity of Exotic eDiscovery, many companies and law firms regularly enlist the help of eDiscovery experts such as Pitney Bowes Legal Solutions to assist

in the work. The real goal however, should be to stop performing a costly "fire drill" response to every eDiscovery need. This best approach requires taking control of your data early on in the process and combining Records and Information Management (RIM) with litigation readiness. Happily, there are tools to put this in place without making massive changes to the corporate information infrastructure.

Most companies don't know where all their information is. Data comes from many sources and is not easily captured. The data silos are international. Outside the silos, information is collected and transported literally everywhere. The RIM approach maps these various silos and sites. The goal of RIM is to stay ahead of the situation – to be ready for the next eDiscovery project before it happens. RIM effectively converges the information residing in email servers, document management servers and file servers. It provides a proactive understanding of both how the data is collected and what it contains.

When records management, knowledge management and compliance all converge, you work with an overall organization of all documents in all forms in all locations. As such, when litigation happens, you are ready for it.

The whole idea of RIM is to know the "who-what-when-where-and-how" of your data – to be able to identify all collection points and all documents. Courts are holding companies accountable for innocently organizing their data and not knowing where it all is. The big question to ask: is this information important to us? If the answer is yes, you need to build a structured approach to records retention, records management and litigation readiness.

¹The Radicati Group, Inc.

²FRCP Rule 34 (a)(1)

³FRCP Rule 26 (a)(1)(B)

⁴Socha-Gelbmann 5th Annual Electronic Discovery Survey