

Understanding Media

(10 Useful things to know about IT and the people who work there)

Peg Duncan
IT and eDiscovery
Ottawa, Ontario

April, 2008

Understanding Media

(10 Useful things to know about IT and the people who work there)

1. IT builds systems and applications

People in the IT unit develop, maintain, operate and upgrade the technical components that make up the data processing environment. This may seem like an obvious statement, but there are also corollaries:

- Content belongs to the business application owner
- Until recently, email was left entirely in the hands of users
- Decisions regarding what to keep and what to delete are made by users

There's a very clear division of responsibility between the owners of the business applications and IT. IT people will build or implement commercial-off-the-shelf applications to corral, organize and manage this information as long as the owner of the system provides the business requirements, including the business rules for information workflow, retention, archive and destruction.

Illustration i: A medium-sized market research company has been operating a groupware product (e.g. Lotus, Groupwise, Groove etc.) for a number of years. Users in the company communicate and collaborate in the groupware product, often creating and controlling access to their own document libraries and special sites. There is no overall business owner for the groupware product, no standards for naming conventions and access management, and no central inventory of the libraries and sites. The software allows for documents to be “replicated” to other sites and users have created personal copies on CDs. The IT group has been responsible for installing, updating and troubleshooting the software, but has otherwise not been involved. The servers housing the libraries have been backed up to tape, but there is no process to archive or cull the libraries. When the company sues one of its clients for lack of payment on a contract going back several years, it cannot find the library containing the related working papers and the employee who worked on the contract has since moved on.

The problem with this state of affairs is that even the business owners themselves expect IT to know about the information in the corporation.

IT *can* tell you whether or not there are email management/archiving systems in place, how often systems are backed up, which servers have which data *stores* on them, etc.

2. Email accounts and user logon names are in the names of the user, not the organization.

The email for an organizational unit is not necessarily all stored in the same place either, especially if the employee has moved around. Accounts get assigned to servers depending on load balancing to ensure optimal performance.

Illustration ii: An innocent, but key individual in an alleged fraud investigation had left the corporation about 6 months ago. His workstation was wiped and re-assigned to someone else, and his email stores were deleted from the Exchange server. IT still has the backup tapes, but does not know which server the email store was on. The tape backup system only stores the indexes for the current backup. All tapes have to be restored and sampled to extract the email store.

Illustration iii: Outside counsel in a litigation case have asked for the email stores of all custodians who had worked for the CEO and CFO. However, staff turnover has been particularly high and there have been a number of executive assistants and administrators. The researchers decided to restore the executives’ email stores and mine the social network to identify former incumbents of the positions. Although they did track down all the names, they learned that several had been in the habit of printing and deleting the emails.

3. The Properties information inside the document may be misleading.

Often Properties contains neither the name of the author NOR the name of the work unit unless the user has made an effort to reset the default author field in Properties. If the document was modified from a version created by another user, the Properties will contain the original user’s name unless it has been changed. Most users are totally unaware of this “metadata” unless they are trained, and even then they may not think to change it.

Illustration iv: A company uses a shared drive to store documents created and used by a number of different people in a work unit. Everyone in the real estate management group has full “rights” to the shared drive, including read, write, change and delete. The author field in the document Properties contains the name “Bob Smith” on almost all the documents; however, no one named Bob Smith has ever worked in real estate management group. It turns out that “Bob Smith” was the name of the IT person who created the workstation image for the last operating system and office software upgrade.

4. IT worries about data availability, integrity and confidentiality

The networks a corporation uses are under relentless threat from hackers, spammers and viruses; constant vigilance is required to close down holes. Software vendors issue patches on a weekly basis to combat the attacks, but sometimes only an operating system upgrade will solve a serious security vulnerability.

Illustration v: An investment firm is aware of a pending investigation by the Securities Commission and puts out a notice to all employees not to delete any information. Shortly after the litigation hold is issued, the company is hit with a viral attack and the IT department realizes that the best option is to wipe the disks on the desktops and install the upgraded operating system, which offers better control over user accounts. The upgrade must be installed quickly to avoid the ongoing threats of information destruction and loss of system availability. Users are asked to copy important files to the shared drives on the servers. Did anyone ask Legal if this approach was appropriate and defensible while the firm was under a litigation hold?

5. New tools are cool, but come with their own management problems.

Wikis, blogs, Facebook and other new tools offer users a rich environment for sharing information, both for the good and for the less helpful. There may or may not be an audit trail, and the systems offered on the web are outside the control of the corporation.

Illustration vi. A well-respected industry leader has criticized the actions and policies of a large company on her personal blog, to which many industry watchers subscribe. Under threat of litigation, the author erases the offending entries from the blog but it is not clear if the entries were erased by subscribers to the RSS feed. The remarks may still be on other websites.

6. Backups aren't archives

Backups are designed for disaster recovery. Although the technology is changing, servers in many cases are still backed up to tape, a medium that deteriorates over time. When the disk becomes unreadable or the entire server is lost, the most recent set of backup tapes is used to recreate the disk. If those tapes are unreadable, an earlier backup version may be restored. Some IT units keep the last backup of the month “just in case,” but all other tapes are returned to a pool for re-use.

The point to keep in mind is that backup happens every night, but recovery happens extremely rarely. For this reason, the design is focused on improving backup rather than on ease of recovery/retrieval. As storage capacity increases, the technology changes to increase the speed and efficiency of backup; tapes created at one point may be unreadable by the newer technology five years down the road.

By contrast, archives are designed for long term storage of information that may be required for business, legislative or historical reasons in the future. Where active data is stored and frequently accessed using desktop applications available in the ordinary course of business, archive data is stored in a stable long-term format that will still be understandable 10 years hence. The archive may be kept in “near off-line” storage and migrated over time to other storage devices as long-term storage technology advances.

Illustration vii: A government archive agency stores database data in comma delimited format with documentation describing the fields. This is the lowest common denominator that can be understood by any software. Because the documentation describes the fields and the relationships, the relational database can be reconstructed in the future with the software of the time.

Illustration viii: During the course of the collection of information for the discovery phase of litigation, the company becomes aware that the only source of email that was exchanged by a custodian 5 years ago is on the backup tapes. The tapes are found in a storage box in the corner of the server room. The box indicates that it contains the month end tapes for 2003 from the servers and each tape is labelled with the server identifier. The IT department no longer has the equipment to read these tapes. All tapes must be turned over to a data recovery company in order to retrieve the custodian’s email.

7. Storage is cheap

Each year the new desktops, laptops and servers have larger disks than the year before. What would have been an acceptable configuration two years ago is seen by the consumer as wholly inadequate. The cost per gigabyte of storage has plummeted over the past few years.

Rather than solve the problem of information management, many IT units have been content to simply add more disk. While adding disk may solve the immediate problem, in the long run there are consequences:

- Multiple copies of the same email, document, etc.
- Large numbers of documents returned in a search query (needle in a haystack problem)

- Increased time to back up the systems
- Increased time (possibly days!) to restore the information in the event of a disaster
- Larger volumes to process in the collection phase of discovery in the event of litigation

8. IT specialties

As with other occupations, Information Technology breaks down into several sub-specialties, each with its own education, skills and unique experience requirements. A software engineer is not the right person to ask about a high-speed network link to the Yukon.

Because of the nature of information systems in the 21st century, it is no longer safe to assume that any one person has a complete picture of what is happening in IT in a corporation. Moreover, the IT unit does not necessarily have the right people to take on and manage the complex preservation, collection and processing steps involved in, say, a fraud investigation. This doesn't imply that you have to hire an e-discovery service bureau to take over the process but it is prudent to consider bringing in a consultant with e-discovery or forensic expertise to guide the IT group through the process.

A corporation that is particularly prone to litigation may well want to develop this kind of litigation support skill set inside.

9. Privacy...

The introduction of the IBM personal computer introduced – well – the notion of “personal computing,” which evolved to be anything done on a desktop or laptop on an employee's desk. Until recently, employees still felt they had an expectation of privacy and that any monitoring was possibly an intrusion into that privacy. The culture of email and the use of local disk drives and peripherals can reinforce the sense that the desktop is a personal tool and the information stored there belongs to the individual rather than to the corporation.

As a consequence, a lot of “personal computing” does go on in the office. Jokes, eBay solicitations, on-line purchases and banking are not that rare. They leave a footprint on the desktop in the form of internet temp files, cookies, passwords and other objects that users seem blissfully unaware of.

An Anton Piller order can go even further and open a party's private and personal Yahoo account – and even their home computer – if there's an allegation that commercially sensitive information has been stolen. Home computers contain everything from Quicken

databases to illegal Torrent downloads. In certain cases, the full contents of the computer have been given directly to the plaintiffs, barring all.

10. Applications evolve over time

Back in the late-1980s, Wordstar, Multiplan and dBaseIII were the most popular desktop applications used in the office setting. While it is possible to find Wordstar to Word converters on eBay, fully restoring the document, spreadsheet or database created using any of these tools is going to be difficult.

Not only is software obsolescence a challenge, but the evolution of a database application over time can also cause problems.

Illustration ix: For a number of years a company used Q&A for contact and event management. From time to time the administrator would clone the database and add or change a field. Because the application was only used by a few people, no one made any effort to define what should be in the fields in a coding manual. The administrator would merge the databases for certain reports, recognizing that some of the fields were not present in all the records. Later, after he had moved to another company, the Q&A software was no longer supported by the vendor and the new administrator had to convert the databases to the recommended replacement software. Because the new software required a more disciplined approach to data definition and management, the database was considered to be corrupted. An outside consultant had to be brought in to analyze and rebuild the database at some expense.

Systems are designed to meet current and anticipated requirements not to reconstruct how things were at a point in time in the past.

Illustration x: An aircraft company publishes its assembly manuals and parts inventories in an application accessed over a private wireless network using a handheld device within the manufacturing plant. A newly designed transport aircraft had just started rolling through production; unbeknownst to the company, the assembly manual instructed the technician to install a part that wouldn't withstand the stress of landing in heavy wind conditions. The error in the manual was eventually identified and corrected so that the current version refers to the correct part, but not before a transport plane was damaged along with its cargo during its maiden voyage. Does the online application allow the investigation team to reconstruct the manual to the point in time before the correction?

For more information about Electronic Discovery in Canada, please see:

ELECTRONIC DISCOVERY - A Reading List, published on the Lawyers Professional Indemnity Company *practicePRO* website:

http://www.practicepro.ca/practice/eDiscovery_Rlist.asp

E-Discovery Canada – the Canadian E-Discovery Portal – published on the LexUM website hosted by the Faculty of Law at the University of Montreal:

<http://www.lexum.umontreal.ca/e-discovery/>

© Peg Duncan 2008