THE VANCOUVER
courier.com

Your source for Vancouver News and Entertainment

Digital extinction

By Chris Miller
Staff writer

Lane McDonald pauses outside the Inner Sanctum and taps his pass card against a black alarm box. When the box's light switches from red to green, McDonald—a calm, balding, middle-aged man—opens the door and walks into a room dominated by a machine 15 feet long and five feet high.

A computer terminal rests on top of the smooth, rectangular surface and a sheet of microfilm sits in a small opening in its side. Connected to a bank of computers on the wall, the machine takes online information—such as hospital patients' records and daily bank transactions—then writes it onto microfilm.

But thanks to a trend towards preserving records on laser discs instead of microfilm, the $500,000 machine—and a large duplicator across from it—may soon be collecting dust. "I fully expect within two to three years, these things will be shut off," says McDonald, western regional manager for Lason: The Information Management Co. "They won't be running anymore."

In addition to putting records on microfilm, Lason copies them onto digital devices such as CD-ROMs. Right now, demand for the two techniques is about 50-50, but McDonald says that won't be the case for long—digital storage is the wave of the future.

Unfortunately, as government departments and local municipalities latch onto the trend, they risk its pitfalls as well. Neither tapes nor disks nor software programs last long before breaking down or becoming obsolete. NASA discovered one day it couldn't read data from its Mars probes because they were created in an outdated computer format.

The threat of losing records is significant. Accurate records hold governments accountable—through documents such as minutes from a council meeting—and protect individual rights, providing proof of permits for a house renovation, for example, or payment of auto insurance. Health files help doctors determine the best way to treat patients and criminal records mean dangerous people can be screened out of jobs at daycare facilities or schools.

Loss of such records would be devastating, yet, as more government functions happen on-line, experts are struggling with the question of how to ensure that doesn't happen.

How long will information on your Zip drive be readable?

Computer science professor Richard Rosenberg hopes it's a long time—he stores stuff there regularly—but he knows better than to count on it. "We forget how quickly things change, and I'm sure that everyone has things on computer that now can't be read."

He's reminded of this fact each time he sees the computer tape on his desk. The tapes, used in 1960s-era tape drives, are now collected by computer scientists like fossils. Floppy disks, which store little information by today's standards, may be the next in a long list of storage devices to go the way of the Dodo.

Even commonly used storage devices seem to have Achilles Heels. Microfilm lasts for 500 to 1,000 years, but trying to find specific information on it can be a nightmare. Good-quality paper lasts just as long, but is costly, difficult to reorganize and takes up lots of space. Video and audiotapes last as little as 10 years before the magnetic layer, which stores pictures and sounds, begins flaking away. With CD-ROMs, the stored information spreads over the disk, rendering it next to useless in 10 to 20 years.

"You can put a lot of stuff on CD-ROMs, but the question is, can we read that in the future and will they physically last?" asks Rosenberg, who maintains industry can make digital preservation easier for everybody. Microsoft, for example, has kept successive versions of Word compatible with each other because it's so widely used. Other experts believe private companies should save every software program and operating system available.

Apple is doing this in the lowest-tech way imaginable. In the 1990s, the computer megacorporation concluded the best way to archive software was to print out the computer code—and the code for the operating system—on good-quality paper.

At the government level, however, old files and records are likely being lost. "I'm sure there are some data they've lost because they haven't updated it far enough," says Rosenberg. "Are they storing information now that they don't realize is going to break down? I would guess so."

Given the problems of digital storage—the cycle of software obsolescence and lack of durable storage devices—trouble looms. At the same time, the Canadian government is pushing a program called Government Online that aims to shift activities from phone, mail and in-person visits to digital "offices" by the year 2004. Papers will be scanned or copied into a computer file, then destroyed, to reduce paper use. With such a system, records must be filed correctly or chaos will reign.

Though the federal government is leading the electronic revolution, the province and some municipalities—Richmond and Coquitlam lead the way in the Lower Mainland—aren't far behind.

The City of Vancouver is already edging into electronic business by allowing people to order archival photos and renew dog licences on-line, although the transactions must be completed in person and leave a paper trail. In future, city business conducted on-line—including e-mails—will have to be stored digitally. Andrew Power, the city's corporate information analyst, thinks the city may even have to look at recording copies of its web site on a back-up tape, thanks to a lawsuit south of the border.

The U.S. archive and records administration used to advise government departments to print out records of e-mails. But this was challenged successfully in court by a plaintiff group that pointed out print isn't the same as an electronic document, which may have hypertext links and interactive features that can't be preserved on paper.

"It has legal implications," says Power. "If you don't preserve it in the way it was used, there's some question of legality."

Most city records that need to be kept in perpetuity—like council minutes and legal descriptions of city lots—are preserved on paper. And the city's 1993 policy states that employees should also print out copies of records they make, though Power suspects this isn't always happening.

Unlike some institutions—such as CNN and BBC London—the city archives isn't looking at converting its stored materials to electronic formats, then doing away with the originals. The move is seen as counter-productive, since some 150 to 160-year-old archived papers and black-and-white photos of slightly newer vintage still look as good as new. And so far, there's still enough space to store everything. Electronic records, though easier to sort and search, require migrating into new formats or storage devices.

Due to public demand, the archives scanned 15,000 photographs onto the Internet for public perusal. The originals are still around, but the archives have to periodically copy the digital pictures onto new CD-ROMs before the old disks start losing data.

That's small potatoes compared to the digital flood that could eventually emanate from city hall. Archivist Sue Baptie says the city's archives are already preparing for electronic financial records and engineering maps from the city. Baptie and conservator Sue Bigelow are studying digital preservation so they're ready once the data starts flowing.

A few years ago, the City of Richmond's records-keepers and information technology staff were experimenting with a new computer system that automatically creates and stores new records. Now, all the city's 900 staff are on it. "I think if we were to take it away from staff, they wouldn't be able to do their jobs," muses Dovelle Buie, Richmond's records and information management analyst.

When a clerk, engineer or planner saves a new document, it's automatically sent to a common database. Staff use coded titles, incorporating important details such as rezoning numbers so files are easy to find. Documents can't slip through the cracks by being buried on desks or saved to the wrong computer file, and coding makes it easier for archivists to comb through records, deciding what needs to be saved and what can be let go, Buie said.

Though the system is among the best in the industry, Buie admits the city may have to migrate information onto new applications or storage devices in future, a process that isn't easy. Richmond ran into problems opening

documents in Word '97 after switching from WordPerfect a few years ago.

The City of Richmond offers a few services on-line, some of which require use of a credit card. For instance, people can pay parking tickets or buy garbage tags—for exceeding the two-bag limit—by computer. Once a tag is purchased, the system automatically sends an e-mail to a city department that puts the tags in an envelope and sends them out.

Other services are information-based. City residents can look up their tax and utility accounts on-line, and beginning this fall, users will be able to register for community centre courses.

The city began piloting remote services in 1997, first on telephone, then on computer terminals inside the eight-storey city hall. Though people can't pay their property taxes over the Internet, they can at the terminals, which include a keypad where you can swipe your bank card.

Colleen Smith, Richmond's web site co-ordinator, says the city is always looking for opportunities to expand its list of on-line services. "Opportunities will probably come to use just through the evolution of the software industry," Smith says, noting web-based software allows for more interactive features.

The Richmond Library has also been an information technology leader. On-line users can apply for library cards, pay late fees, book library computers, register for library programs, search catalogues and renew and put holds on books.

"There's so much that we can do to build a virtual city hall," says Buie. "And I think there's a lot of interest in the community in doing that."

A file that arrived at the provincial archives one day stumped archivists at the B.C. Archives in Victoria. They knew it contained a database of film and video collections, but couldn't read it because it had been constructed by a once-popular brand of software from the 1970s.

Brant Bady, manager of emerging and applied information technology group for the B.C. Archives, knew he had applications that would help him extract information from the file, but it would be a tough slog. Eventually, he had to reduce the file to binary code —the Os and Is that are the basis of all computer programming— then reconstruct it.

"I was surprised we were able to get anything out of it," says Bady.

Almost all the material stored at the archives, located in the Royal B.C. Museum complex, is either in its original format or stored on microfilm, but that won't always be the case, Bady says. Recently a royal commission report was sent to the archives in digital form only, leaving Bady struggling over how to store it. Though these types of questions are rare now, they'll be commonplace when the great data deluge hits. "We know that blip is coming, or some say it's a tidal wave."

Bady says the government is looking at limiting the number of word processing programs it uses and choosing software that lends itself to digital storage, but hasn't enacted any policies yet.

With the flood of information flowing onto computer, Luciana Duranti, a digital preservation expert at UBC, worries saving records will be dependent on government budgets. If there isn't money to save them all, how do you choose which records will be preserved and which lost?

The government can still pull back from its headlong rush into electronic agony, she says. It simply needs to evaluate which records should be kept on computer, and which should be left in old

formats such as microfilm and even—gasp!—paper.

"Lots of records are good for electronic use—if there's hypertext or you need to do random searches," Duranti says. "But if the information was first generated on paper, leave it on paper."

Lively and slim, with greying blonde hair and a strong Italian accent, Duranti has been leading a crusade against loss of computer data. A professor in the university's School of Library, Archival and Information Studies, Duranti embarked on an ambitious program in 1997 to permanently preserve the electronic records created each day by governments, businesses and private citizens. The project, called InterPARES, which stands for International Research on Permanent Authentic Records in Electronic Systems, is vital, she says. "We live constantly in the present, and it's very difficult to see how we got there, from three months in the past to six months in the past, let alone years ago."

This year, Duranti and 16 grad students have been working on the problem with the help of federal funding, partner agencies and academics in more than 15 countries, including the United States, United Kingdom, France, Australia, Italy, Sweden, China and the Netherlands. Ten national archives are participating, in addition to experts at the San Diego Supercomputer Centre, the Smithsonian Institution and top universities.

Though Duranti is convinced digital information is constantly being lost now, it's difficult to know how much. That's partly because the government doesn't broadcast its embarrassing mistakes, and partly because you might not know information is lost until you try to open a file.

Files that have been transferred to ASCII, a text-only format, often lose their meaning because the format can't reproduce simple layout features, like column headers and points for point-form notes. Yet ASCII remains a popular format because it's one of the few sure-fire ways of converting electronic records. "That's really not acceptable," Duranti says. "Records are more than data."

Most records that are no longer "active"—meaning they've been moved to storage—are in a format that makes them difficult to read. Wax audio cylinders from Edison's time can't be played on anything but a phonograph, and the same principle holds true for outdated software programs, operating systems and computers. The Vancouver School Board, which is automating its library collection, was forced to scrap computer servers it set up just a few years ago to make way for an Internet-compatible system. Because the board couldn't "migrate" entries from the old system, employees had to manually re-type the book information they'd typed in a few years before.

In some cases, other details, such as colour, might be lost. That's not a big deal if the record's an e-mail, but if it's a colour-coded map, you're in big trouble, Duranti notes.

After a couple of years of prep work, the InterPARES project was officially launched Jan. 1, 1999 and has been gaining steam since, with new experts and archivists joining the fold. Within the next several years, Duranti hopes InterPARES experts will develop international policies and standards to ensure electronic records can be preserved forever. In the meantime, she remains a voice of caution.

"[Computers] are good for many things, but when it comes to preserving for future generations evidence of who we were and what we have done, the safest way is not the digital way."

back to top