# Congress of the United States

# House of Representatives

***One Hundred Eleventh Congress***

## Committee on Oversight and Government

## Subcommittee on Information Policy, Census and National Archives

***"The National Archives' Ability to Safeguard the Nation's Electronic Records"***

**Thursday, November 5, 2009 at 2:00 P.M.**

**2154 Rayburn House Office Building**

**Testimony of**

**Alan E. Brill**

Chairman Clay, Ranking Member McHenry, and Members of the Subcommittee. My name is Alan Brill. I am currently a Senior Managing Director at Kroll Ontrack, but I am here not here today as a representative of Kroll Ontrack, but as an individual, to share whatever knowledge and experience I have in the fields of information security, data protection and data recovery to assist the Subcommittee with the vital work it performs I am grateful for the opportunity to speak with you.

The reality is that in today's environment, a substantial proportion of the information that is being created within our government is generated, exchanged and stored digitally. It is produced and stored on computers, be they the desktop or laptop computers of individuals or the massive processing arrays and networks of large agencies. It is also a simple fact that most of the data that is created, and which

may have import for extended periods will never in the course of normal use be printed. How do we safely and efficiently preserve electronic records when the technology involved in producing and storing those records changes at what certainly seems to me to be accelerating and certainly a breathtaking rate. Consider that the first computer I used at the Pentagon in 1968 had a total memory size of two thousand characters. Today, my wristwatch has exponentially more than that. Storage has evolved from being measured in kilobytes, went through megabytes pretty quickly, got to gigabytes, and is now moving on to terabytes. In my firm's data center, we measure our storage capacity in petabytes. One petabyte is equal to one million gigabytes.

I've been involved in the security and recovery of data from computers for more than 40 years. My recent experience has involved working with private sector organizations to safeguard sensitive data and to help those organizations respond to data security incidents. I've learned a few lessons that I hope will be helpful to the Subcommittee when it considers how best to carry out its oversight role in assuring the preservation of records which are a vital part of our national heritage.

1. Don't assume that the devices currently used to store data will be commonly used – or even reasonably available – into the future. I could name a wide range of storage media ranging from 8-inch diskettes I to 7-track magnetic tapes to Magnetic Card Selectric Typewriter cards, to dozens of other formats that are no longer with us. It is very easy to confuse the storage of information with the storage of media containing information. This is not a new concept of course. Paper records have to be stored in a manner that protects the ability to read the information they contain. Magnetic and optical media also have environmental requirements. I've seen tapes stored in tropical climates that actually have moss growing on the reels. Above all else, we must ensure that we can access the information stored on the media we use to preserve important information. This means that we either have to preserve the reading mechanisms (and be prepared to develop interfaces from what will be essentially antique devices to the computers of the future) or periodically transfer the data to contemporary media, as new storage technology obsoletes the old. If we don't pay heed to this, the information may be in our warehouses, but it will be as unreadable as if it were in an ancient language that cannot be translated. Put another way, you might have a great collection of 8-track audio tapes, but you're going to have a problem playing them unless you've preserved player hardware as well, or transferred the data to some other format.

2. Don't assume that data cannot be restored, even if the storage medium appears to be damaged or beyond repair. The technology of data and media recovery has advanced quickly. Take a quick example. Following the tragic loss of the Space Shuttle Columbia in 2003, NASA located 3 hard drives in the debris field. The Glenn Research Center sent them to my firm for examination. Two were beyond hope. The surfaces containing the data had been heated to the point that, in fact, no data remained. On the third drive, plastic had melted onto the drive surfaces. We rebuilt the mechanical components, cleaned the disks, and were able to recover over 99% of the data, which turned out to be vital for completing a long-term physics experiment. With today's technology, unless the media containing the data is utterly destroyed, the data is at least potentially recoverable, potentially readable. And this can be true even for disks that are part of large storage arrays. There are many variations of such arrays, and how they store data. I fully understand that because some storage arrays distribute data across many disk volumes, so that if one disk fails it can be replaced and the data automatically restored to it by the computer using copies on other disks, there is sometimes the belief that individual disks can't be read. That without the whole of the array, one disk is useless. But in many cases, that is not true. It is quite often possible to read the disk and to see at least some of the data that it may contain. Does this mean that it is impossible to completely erase data from a disk drive? No. There are a number of ways to wipe data from a disk very effectively. I know that when I am moving to a new laptop computer, for example, after I have transferred the data that I need, I use software to completely wipe out the information on the drive. Until I do that, I try to protect it with whole disk encryption software, and a number of other safeguards. I believe that best practice is that when a device contains sensitive data that is even potentially recoverable, it must be handled appropriately, and that before the device is decommissioned or discarded, the data must be destroyed through physical or other means. Disks can be cut or smashed. CDs or DVDs can be destroyed with a few seconds of microwave energy. Degaussers can quickly and irrevocably destroy data. But as the disk from the space shuttle showed, data can be tough to destroy. If it's being done, it has to be done right, and such destruction should be documented.
3. What you see is often not all that you can get. Computer programs don't just contain the data that we think about. We all use word processors. And we know that they create files that contain the words we write. But they contain more. There are a number of data fields that are automatically created and maintained by the program. Some are obvious – the date and time the file was originally written, how many times it was edited, when it was last opened. But it can

contain more. For example, it may contain a record of changes made in the course of revision and review. Other information is maintained by the computer's operating system. When you see a list of files, you know that you often see the creation date and size. This specialized information is called metadata, and it is important to the understanding of the underlying data. This is not a new issue. When we look at Abraham Lincoln's handwritten manuscript of the Gettysburg Address, we can see how he edited it, what it looked like before he made the changes. The same can be seen through examination of metadata, but only if it is preserved. Unfortunately, unless care is taken in regard to the processes by which data is preserved, metadata can be inadvertently changed or lost. Our courts recognize this. They have held that merely printing and storing a document may not be enough to properly preserve its value. The metadata can be vital in establishing the authenticity of an electronic document. A will purportedly dated July 1, 2003 might be questioned, for example, if examination of the digital file showed that the file wasn't created until 2005. So data preservation must also take into consideration how to best preserve not only the basic document – the words in an email or the numbers in a spreadsheet, but the metadata as well. To ignore metadata is to contstrain our understanding of the file. Preserving this metadata is not particularly difficult, but it does require a detailed technical understanding of how various copying or preservation processes affect metadata so that the proper methodology can be selected.

4. Ensuring data security must be more than an afterthought. There is no question that there is a cost to data protection. Planned effectively, these costs can be controlled. There is always a trade-off between cost and protection. Identifying the level of protection that is reasonable and appropriate to the data being protected is not necessarily easy. Protective measures that are sufficient today may be insufficient tomorrow as threats mature and evolve. Perhaps the best way to summarize it is to say that if you are complacent about information security, assuming that whatever you're doing today is sufficient and appropriate, and will stay that way, you're setting yourself up for an unpleasant surprise. This is a lesson that has been very publically and painfully learned by organizations across the globe in recent years. While I am not an expert in the various security standards that are used by federal agencies, I have found that there are a number of centers of knowledge which can be of immense value in understanding the risks and alternatives. The work of the professionals at NIST come to mind. I have no doubt that this Subcommittee is aware of the ongoing work there to identify risks, protective measures and to provide publications that can help professionals and managers in both the public and private

sector to do a better job of securing sensitive data. The other reality is that the cost of not protecting data appropriately can be very high. What is the cost of compromising millions of credit card records? Or sensitive medical information? What is the cost to future knowledge if electronic records of today's decisions and activities are lost through security failures, or through permitting security needs to change while protective measures stagnate?

5. Finally, I believe that the expertise exists to assist and advise our government on this complex and continually changing issue. There are many specialists like myself who recognize that service on advisory councils and other appropriate mechanisms is part of our civic and professional duty. Why not call on this pool of knowledge. The reality is this: If we don't collect data and collect it properly, if we don't maintain it in a usable and complete form, and if we don't safeguard it appropriately, it won't be there for the benefit of future generations. Technology is making it possible to not only collect vast amounts of data, but to index it and make it more accessible and useful than ever before. I believe this can be done without undue risk to our privacy and security, if the risks are recognized and there is a commitment to protecting that privacy and taking the right steps to have reasonable security. Can we guarantee 100% security? Of course not, but we can minimize the incidents through the use of encryption, access controls and logging, making sure that users have access to only the information they need, and other techniques. Equally important, we must assure that both public and private sector organizations have a plan for what they will do if there is a data protection incident. Trying to develop a crisis management plan in the middle of a crisis is difficult at best. Recognizing that incidents can occur, and preparing for them is far more effective.

I want to thank the Subcommittee for inviting me here today. I'm fortunate to have had the opportunity to work with information security colleagues in federal service, including the FBI, Secret Service, Inspector General offices and Department of Defense, among other agencies, and I hope you appreciate their service as much as I do. They are fine professionals who could probably earn more in the private sector, but who recognize the value of public service. The subject of today's hearing is important, and the public is well-served by the Subcommittee's interest and focus on this area.

Thank you.