The 2008 Analysis of Information Remaining on Disks Offered for Sale on the Second Hand Market

Dr. Andy Jones[1,2,] **Dr Glenn S. Dardick**[2,3] **Mr. Gareth Davies**[4]
Dr. Iain Sutherland[2,4] **Dr. Craig Valli** [2]

1 Security Research Centre, BT
2 Edith Cowan University
3 Longwood University
4 University of Glamorgan

andrew.28.jones@bt.com; Phone: +44 1473 646133

Abstract. The use of computers that contain hard disks to process and store information has been ubiquitous across organisations in both the public and private sector for more than two decades and is being ever more widely used by individuals in the home. During that time, the processing capability of the computers has increased enormously. At the same time the storage capacity of the computers has increased from tens of Megabytes to hundreds of Gigabytes and the use of Terabyte storage devices in both commercial and private locations is now becoming increasingly common. In recent years, because of social change and alterations in the way in which organisations work, there has also been an increasing trend in the use of the same computer to process and store both the organisation's and the individuals personal information. It is clear that the majority of organisations and private individuals still remain ignorant or misinformed of the potential volume and type of information that is stored on the hard disks contained within these computer systems. As a result, they have not considered, or are unaware of, the potential impact of this information becoming available to an unintended third party.

This is the fourth study in an ongoing research effort that is being conducted into the volume and type of information that remains on computer hard disks offered for sale on the second hand market. The research has been undertaken to gain an understanding of the level and types of information that remains on these disks and to determine the damage that could, potentially be caused, if the information was misused. These studies have examined a large number of disks that have been purchased in a number of countries. The rationale for this was to determine whether there are any national or regional differences in the way that computer disks are disposed of and to compare the results for any regional or temporal trends.

The first study was carried out in 2005 and has been repeated annually with the scope extended to include additional research partners and countries during each of the subsequent years. The studies were carried out by British Telecommunications and the University of Glamorgan in the UK, Edith Cowan University in Australia and Longwood University in the USA.

The core methodology of the research has remained the same over the duration of the study: to acquire a number of second hand computer disks from a range of sources and then to determine whether they still contained information relating to a previous owner or if the device had been effectively erased. If the disks still contained information, the research examined whether it was in a sufficient volume and of enough sensitivity to the original owner to represent a risk if unintentionally exposed to a third party. One of the results of the research has been that for a very large proportion of the disks that have been examined, there was significant information present and both organisations and individuals were potentially exposed to the possibility of a compromise of sensitive information. Potential impacts of this might include embarrassment to individuals and organisations, fraud, blackmail and identity theft. It is noted that where the disks had originally been owned by organisations, they had, in most cases, failed to meet their statutory, regulatory and legal obligations.

In the 2008 study, the fourth in the series, the research methodology that had been followed in the previous studies was repeated, but in addition the scope was again broadened geographically to include disks sourced from within France.

Keywords: Computer forensics, disk analysis, data recovery, data disposal, data destruction, data leakage, privacy.

1. Introduction

The first in the series of annual studies was carried out in January 2005, (Jones et al, 2005) and revealed that a significant proportion of the disks that were examined still contained large amounts of information, much of which would have been considered sensitive by the previous owner. Prior to the publication of the report of this first study, there had been limited investigation in this area, with the most significant findings being reported by Garfinkel and Shelat (2003). At that time, there had also been a small number of commercially sponsored investigations and newspaper reports on the subject of personal data being found on incorrectly disposed disks. The 2005 report identified that the majority of the random sample of disks that were obtained still contained significant amounts of sensitive information that the researchers considered had the potential to cause embarrassment or financial harm to either the organisation or the individual.

In the 2006 study, the research effort included contributions from the University of Glamorgan in Wales and Edith Cowan University in Australia and the Security Research Centre of British Telecommunications. The research of 2005 was repeated and the scope was expanded to include a number of additional countries. The aim of the 2006 research was to determine whether there had been any change in the level or potential sensitivity of information that remained on the disks during the intervening period and also to gain an understanding of how the results compared between the countries that had previously been surveyed and the additional countries. The report of the study (Jones et al, 2006) revealed that for countries surveyed in both years, there had been no significant changes in the results and that the amount and sensitivity of information that could be recovered remained at a similar level. The results from the disk drives obtained from countries that had not been included in the 2005 survey were broadly similar to those from countries that were included in both the 2006 and 2005 surveys. The research undertaken in 2006 was sponsored by British Telecommunications (BT) and Life Cycle Services (LCS) (now Sims Lifecycle Services) who funded the purchase of the disks.

In the 2007 study the research effort included contributions from the University of Glamorgan in Wales and Edith Cowan University in Australia, Longwood University in the USA and the Security Research Centre of British Telecommunications. The research of the previous two years was repeated and the scope was further expanded to include additional countries. The report of the study (Jones et al, 2006) revealed that for countries surveyed in all three years there had been no significant changes in the results and that the amount and sensitivity of information that could be recovered had remained at a similar level. The aim of the 2007 study reflected that of 2006, to determine whether there had been any change in the level or potential sensitivity of the information that remained on the disks during the intervening period and to gain an understanding of how the results compared between the countries that had previously been surveyed and the additional countries. The results from the disk drives obtained from countries that had not been included in the 2006 study were broadly similar to those from countries that were included in all three studies. The research undertaken in 2007 was again sponsored by British Telecommunications (BT) and Sims Lifecycle Services (SLS) who funded the purchase of the disks.

All of the research had been conducted under the same conditions (using commonly and easily available tools that had similar capabilities) and the results then compared. The outcome of the research found that a number of conclusions and recommendations were made on ways in which the destruction or removal of data from disks that were being disposed of could be improved.

This paper, the report on the fourth and latest survey, contains the results of the 2008 research which has again extended the scope of the countries included in the survey and had the same objectives as the research in the two previous years. The research was again sponsored by British Telecommunications (BT) and Sims Lifecycle Services.

2. The Research

The same basic objectives, processes and procedures that were used in the previous studies have been followed throughout the period of the research. All the disks used in the research were purchased at computer auctions, computer fairs or through eBay in the respective regions. The disks were acquired discretely by a number of purchasers so that the sellers would not have any indication of the reason for purchase. For the most part the disks and computers were obtained either singly or in small batches in order to minimise any influence the disposal practices of one seller may have had on the overall result. For example, if a large number of disks were obtained from one seller and they obtained the disks from a particular source or wiped all of the disks that they resold, then this may have an effect on the results by affecting the proportion that were from one sector of the market or that contained no data.

In 2008, in line with the research that had been carried out in each of the previous years, the disks were supplied ‘blind’ to the researchers so that they had no external visual indicators of the potential source of the disks. The only markings on the disks that were provided to the researchers were sequential serial numbers so that each disk could be uniquely identified throughout the process. By supplying them in this way, any information that is recovered by the researchers can be clearly identified as having been the result of the data that was available on the disk.

The research methodology remained unchanged from the earlier researches (Jones 2005, 2006, 2008), with each disk being forensically imaged using verified software and then placed in secure storage. The analysis was undertaken on the forensic images of the original disks. The rationale for this time consuming step was that it met two requirements.

The first was that it was considered that there was a need to preserve the original media in an unaltered state and store it in a secure area in case there was a requirement to pass the disks on to the police. This would be necessary in the event that notifiable criminal activity was discovered and enable a chain of custody to be preserved for an investigation by law enforcement.

The second was to allow the research to be carried out in a non-intrusive manner that did not affect or change the original data in case any anomalies were detected with the image and it was necessary to validate the data against a second image created from the original. As with the previous researches, this proved to be a sensible precaution as four of the disks was found to contain material that necessitated them being passed to law enforcement for further investigation.

The tools used in the 2008 study were fundamentally the same as those used in the previous years (although the versions of the tools may have changed). The tools performed similar functions to the Windows Unformat and Undelete commands and that of a hex editor (which was used to view any information that exists in the unallocated portions of the disk). Tools that perform this type of functionality are freely available: examples include the Linux based Autopsy (Version 2.08) and Sleuthkit software. These types of tools do not require significant levels of skill or knowledge to effect the recovery of remnant data from storage media and there are now numerous online tutorials for operation of these tools for the purposes of recovery.

The objectives remained the same as in previous years: firstly, to determine if the disks had been effectively cleansed of data or whether they still contained information that was either visible or easily recoverable with the tools identified above. The second objective of the research was to determine whether the information available on the disk would allow for the identification of the organisation or individual(s) that had used the disk's host computer.

The results of the 2008 survey once again indicate that there has been very little change over time in the amount or sensitivity of the organisational information that remained on disks that were made available in the second-hand market. The level of sensitive personal data that has been recovered has shown a small but consistent reduction over the period. Before detailing the results of the 2008 survey, the results of the studies in the preceding years are briefly described below.

3. Summary of the Previous Research Results

The results of the previous studies highlighted a number of issues that have been identified throughout the period. These included the fact that, to date, nearly half of the second hand disks that were obtained and could be accessed, had had some attempt made to remove the data, but the majority of those attempts were unsuccessful. In fact, the vast majority of those second hand disks that could be accessed contained data that could easily be recovered. Of those disks that could be accessed, approximately half of them still contained sufficient data to allow the previous owner, whether an organisation or an individual, to be identified. Around one in five of them contained financial information relating to organisations, including staff salary details, sales receipts and profit and loss reports. There were also a significant number of disks that had come from computer systems that had been used in the organisations of critical infrastructure providers such as power generation, water and telecommunications utilities.

There is an increasing awareness of the impacts of data breaches, data losses and identity theft that has come about as a result of regular publicity of incidents together with the publication of the results of a number of surveys and reports (Synovate, 2003; Price Waterhouse Cooper, 2006; Johannes, 2006; Verizon, 2008; ITRC,2008). In addition to this, there is an increasing level of regulation in the area of data storage and disposal. Despite all this, there was no indication from the research that organisations have modified their procedures to ensure the effective removal of data before the disposal of computer hardware. In fact the 2008 survey indicated that there had been an increased use of ineffective practices such as the formatting of disks.

4. The 2008 Research Results

This section details the results for the study carried out during 2008, covering the UK, the USA, Germany, France and Australia. As in previous years, the results of the study are broken down into the individual countries to enable comparison.

For the 160 disks obtained in the UK:

- 73 (46% of the disks) were physically damaged and could not be accessed.
- 27 (31% of the readable disks) had been wiped and contained no data.
  - Of the remaining 60 (69 % of the readable disks)[1],
- 29 (48%) contained sufficient information for the organisation that they had come from to be identified.
- 32 (53%) contained sufficient information for individuals to be identified.
- 47 (78%) indicated that attempts had been made to remove data from the disks by deletion, formatting or reinstallation of an operating system.
- 4 (5% of the readable disks) contained information that was considered to be illicit.

For the 63 disks that were obtained from North America:

- 7 (11% of the disks) were physically damaged and could not be accessed
- 3 (5% of the readable disks) had been wiped and contained no data.
  - Of the remaining 53 (95% of the readable disks),
- 11 (21%) contained sufficient information for the organisation that they had come from to be identified.
- 9 (17%) contained sufficient information for individuals to be identified.
- 14 (26%) indicated that attempts had been made to remove data from the disks by deletion, formatting or reinstallation of an operating system.
- 8 (13% of the readable disks) contained information that was considered to be illicit.

For the 28 disks that were obtained from Germany:

- 11 (39% of the disks) were not in working order and could not be accessed.
- 9 (53% of the readable disks) had been wiped and contained no data.
- Of the remaining 8 (47% of the readable disks),
  - 2 (25%) contained sufficient information for the organisation that they had come from to be identified.
  - 5 (62%) appeared to be from individuals.
  - 6 (75%) indicated that attempts had been made to remove data from the disks by deletion, formatting or reinstallation of an operating system.
- None of the readable disks contained information that was considered to be illicit.

For the 44 disks that were obtained from France:

- 22 (50%) of the disks were physically damaged and could not be accessed
- 9 (41% of the readable disks) had been wiped and contained no data.
- Of the remaining 13 (59%),
  - 5 (38%) contained sufficient information for the organisation that they had come from to be identified.
  - 5 (38%) contained sufficient information for individuals to be identified.
  - 9 (69%) indicated that attempts had been made to remove data from the disks by deletion, formatting or reinstallation of an operating system.
- 4 (18% of the readable disks) contained information that was considered to be illicit.

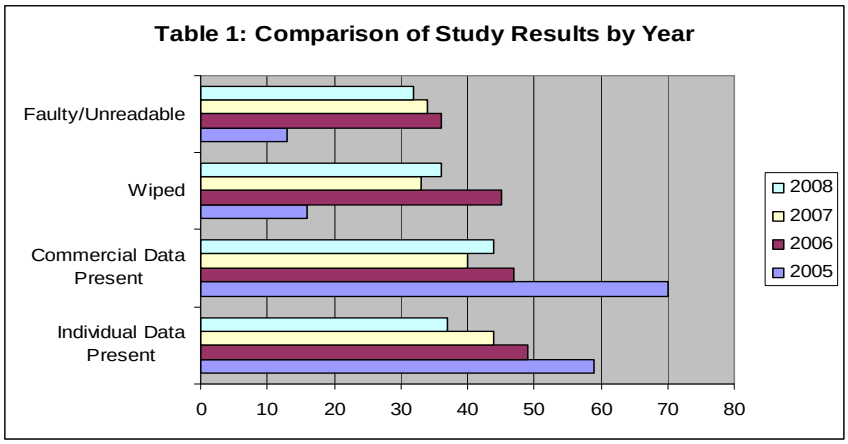
For the 43 disks that were obtained from Australia:

- 6 (14%) of the disks were physically damaged and could not be accessed
- 13 (34% of the readable disks) had been wiped and contained no data.
- Of the remaining 25 (66%),
  - 15 (60%) contained sufficient information for the organisation that they had come from to be identified.
  - 4 (16%) contained sufficient information for individuals to be identified.

[1] The three categories in this section are independent of each other. A disk may contain information on just an individual or an organization or both and any of them may also have had attempts made to remove the data.

  - 7 (28%) indicated that attempts had been made to remove data from the disks by deletion, formatting or reinstallation of an operating system.
- 3 (12% of the readable disks) contained information that was considered to be illicit.

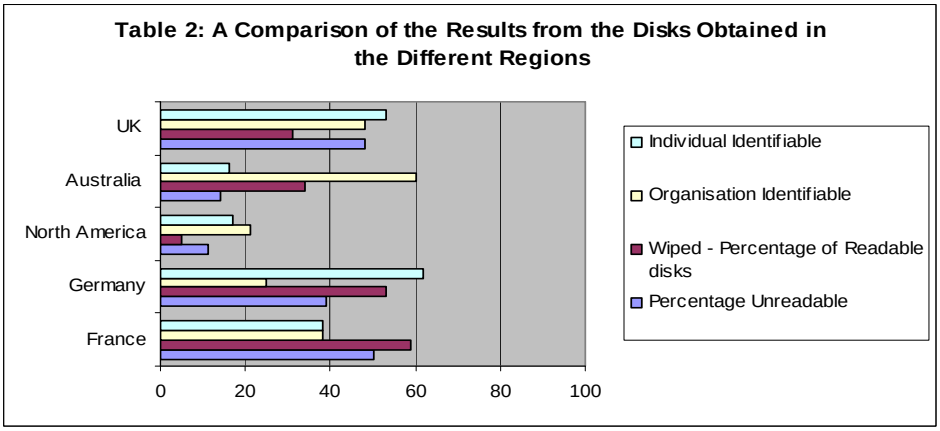
Table 1 shows a comparison of the results of the 2005, 2006, 2007 and 2008 disk surveys.



Note: The results for the wiped disks are given as a percentage of the disks that were readable.
Note: The results for data present are given as a percentage of readable disks that had not been wiped.

The results of the 2008 survey appear to indicate that there is a slow improvement in the majority of the results over the period. With the exception of the 2005 results, the proportion of the obtained disks in the three years that followed that were faulty/unreadable has slowly declined from 36% to 32%. The percentage of disks that had been successfully wiped rose sharply from 2005 to 2006 but now appear to have stabilised in the area of 33% to 36%. The percentage of disks on which commercial data was found dropped significantly between 2005 and 2006, but then appears to have stabilised in between 40% and 46%, with a slight rise in the number detected this year over last. The percentage of disks on which personal data that could be identified to an individual has shown a consistent decline over the whole period from a high of 59% in 2005 to 37% in the latest study. One noticeable change that has taken place over the period has been an increase in the number of disks that have been found to contain significant volumes of both personal and corporate information. This may be the result of organisations implementing more liberal policies with regard to the personal use of corporate computers or may indicate the increased use of home computers for business purposes.

The table below shows a comparison of the results obtained during the 2008 study from the disks acquired in the different regions.



Note: The results for the wiped disks are given as a percentage of the disks that were readable.
Note: The results for data present are given as a percentage of readable disks that had not been wiped.

The results of the 2008 survey that were obtained from the individual regions have again revealed what appears to be a number of significant disparities. The first is that the number of disks that were unreadable varied significantly. The numbers this year ranged from 50 percent in France to 11 percent in the USA. This is a change

from previous years when the highest proportion of unreadable disks came from Germany and the lowest came from Australia (this year the results for those regions were 39 percent and 14 percent respectively).

The second disparity that was noted was the proportion of the disks that had been effectively wiped. The results of the study showed that in Germany, the number of disks that could be accessed had increased from 29 percent to 61 percent over the period of the studies and the percentage of the readable disks that had been wiped also increased from 42 percent to 53 percent. Over the same period, the results from the disks from the USA show that the proportion of unreadable disks has dropped from 50 percent to 11 percent over the period but that the proportion of disks that had been effectively wiped has declined from 8 percent to 5 percent this year with a peak of 19 percent in 2007.

While it is still too early to draw any conclusions on possible trends, there are a number of indications that are of significance. Once again, a surprisingly large range and quantity of information that could have a potentially commercially damaging impact or pose a threat to the identity and privacy of the individuals involved was recovered as a result of the survey. An indication of the quantity and type of material that was recovered in 2008 from the disks that originated in central government, local government, the healthcare sector, commercial and academic establishments included:

- A number of disks were recovered which contained data belonging to a well known UK based fashion company. These disks contained a range of corporate data included information marked private and confidential, information relating to trading performance, budget, etc. stock control and discount codes, internal email and email addresses, Customer web orders, including the customer names and addresses. The disks also contained some network configuration information and traces of personal usage including private images and Facebook details.

- One disk was recovered in the UK that had originated in a large metropolitan local council. The information on the disk mostly related to a number of planning application details for the local area but also included some financial information and internal emails. This disk also included some network information and firewall configuration information.

- One disk recovered in the UK contained corporate data which appears to have originated from one of the major motor manufacturing companies. The disk contains references to design / engineering that appears to relate to the Fixture Design and Evaluation System for what appeared to be vehicle interiors The disk also contained information that included Corporate Data Protection warnings dated from 1996 to 2001 – "Confidential XXXXXX Motor Company, files that contained the names of their authors and some system support network configuration information

- Two disks were recovered in the UK which appear to have data belonging to a water management service provider and which also showed a high degree of personal use (including pornographic material). It is probable these are disks from personal computers, but which also contain a significant volume of corporate information. One of the disks contains email addresses and emails passed from water management service provider and the other contains email addresses. One of the disks again appears to be from a personal computer but contains details of the water management service provider's projects. The disks contained:
  - A number of images of rivers / watercourses
  - Scanned university degree certificate
  - A report on Manhole covers
  - Extensive hardcore pornography

- Two disks recovered in the UK appeared to have originated from within the regional National Health Service Trust. One contains a number of references to two hospitals. One disk appears to contain patient data relating to radiology / x-rays. The disks also contained details of medical staff shifts, sensitive and confidential staff letters, information from a medical system that appears to refer to patients and thumbnail images of x-rays.

- One disk recovered in the UK originated from a Scottish legal firm that contained a number of pieces of information which enables individuals to be identified. The material recovered included a file relating to a successful injury claims and compensation amount, client correspondence on legal matters, details of solicitors weekly meeting schedules and User's account login information for corporate websites.

- Six disks recovered in the UK had originated in a number of academic institutes. These ranged from universities to colleges and a secondary school. The disk from the University appears to have belonged to a postgraduate student with access to a range of material including copies of university forms (not completed), sensitive / personal material from research interviews / survey data with subjects including their names and addresses and a range of hardcore pornographic material. The material recovered from four disks from a college included the names of teachers, a student CV with home address and telephone number, information on the users Internet surfing habits which included: music, news and very personal sex / abortion advice and a small amount of very soft core pornographic material. One disk that was recovered had originated in a Church sponsored primary (elementary) school and contained data, such as the student name and information on the school network.

- One disk that was recovered in the UK appears to have originated from an international corporation with operations around the globe. The material on the disk included a large volume of information relating to rail transport such as images of railway tracks, cabling and bridges, track / rail / safety condition reports and references to Railtrack. Also on the disk were staff CVs and some soft-core pornography.

- One disk recovered in the UK contained data from a small firm in the carpet fitting trade. The material recovered included company emails, orders / jobs, spreadsheets and accountancy data and an archive of invoices.

- One disk acquired in the UK appears to contain data relating to a major mobile (cell) phone manufacturer including some material marked as company confidential. The disk contained information marked 'company confidential', images of cell-phone circuitry, minutes of meetings, personnel details, lists of chemical substances, what appears to be device test data and manuals for disassembly and troubleshooting phones. Also found were internal contact information for employees such as telephone, fax and mobile numbers, secretary's names, business group and projects, personal ID numbers and job descriptions. A significant amount of material present included data on the recycling of old cellular phone components. The information was relatively dated but would still have been potentially embarrassing.

- A disk was recovered from France that appears to have originated from the Embassy of another European Country. This disk came from a Linux system and contained corporate information relating to 'Cidale'. *Centre for Information and Documentation* belonging to the Embassy which is located in Paris. This disk contains a range of material including network and security data, internal IP addresses, security logs, a domain key for Dotnetflux and the minutes of internal meetings.

- One disk recovered from France appears to have originated from small French 'Junior' company that was formed to give web design experience to 10 students. The disk contains corporate information and a range of data including emails, spreadsheets, some pornographic material, scanned images of passports / identity cards and images of a highly personal nature of some employees. Some of the material recovered could potentially have been used for blackmail.

- From a disk obtained in the USA were Test Launch Procedures documents from a Government Contactor. Also found on the disk were design documents, documents relating to the subcontractor, security policies, blueprints of facilities and personal information on employees and SSN.

- On a disk obtained in the USA was information relating to $50 Billion currency exchange proposals with reference to 15% transaction fees. The currency exchange was between US dollar and Spanish Euros, $1Billion secured line of credit and million dollar bank transfer docs. Also found on the disk were bank account numbers and details of business dealings with a number of countries including Venezuela, Tunisia, and Nigeria. Correspondence from someone who appeared to be a member of the Federal Reserve Board suggested that the acquisition of a Bank Guarantee might not be forthcoming because of some questionable circumstances.

- A disk acquired in the USA that originated from a computer belonging to a part-time legal assistant. On the disk were documents that contained client information and court documents.

- One RAID array that was acquired as a single group purchase from an Australian chemical supply company contained the entire catalogue and ordering systems, together with a complete website for the company. The dates on the drive were less than 3 weeks old. Also on the array were business documents, spreadsheets and other data you would expect to find in a medium sized enterprise.

- Another RAID array from Australia was recovered from a Sun Server. The server was sold as being unable to boot, however when the system was switched on, the server booted to a Solaris login prompt. The subsequent analysis revealed that the disks had originated in a superannuation board. This was supported by the asset stickers that were on the external case of the server. This server yielded administrator passwords, configuration files and also network details.

- Four disks that were acquired as a set in Australia were all unformatted and appeared to have originally belonged to a mining company. Analysis of the drives indicated that they were separate SCADA masters. The disks contained information including, the Historian database, iFix and other SCADA components relevant to the companies process control. In addition, passwords and usernames, including that of the administrator, were still present on the drives. There were also photos of the site and personnel present.

- Three disks that were acquired as a set in Australia had belonged to an organisation that provided IT support for Australian and international banks, as well as network services for other critical infrastructure providers. The drives had been formatted, but the information that was recovered included: Remote access policies and procedures for staff to access the banks networks and network diagrams. Names of staff and what type of equipment they had been issued were also present.

- One disk that was acquired in Australia had been formatted, but again, data was recoverable. The drive appeared to be a Windows server domain controller from a community health centre. It contained minutes from meetings and some other documentation. It also contained illegal software keys.

- Another disk that was acquired in Australia had been formatted but the data was easily recoverable. The drive had originated from a Nursing home and contained an extensive range of information. This included patient information, letters from medical doctors, drug information, pictures of patients and their wounds, menus, accounts and minutes of meetings.

The type of material that was recovered from the disks that originated from home PCs in the UK included:

- A disk from one individual contained the name, address, phone number and a range of other material including school notes, photographs of the of family and house, solicitors letters and letters of complaint.

- One disk that came from a family contained a large number of mp3's and photographs and included some very personal data. The material recovered included an invitation to a double wedding and details of a visit to Pakistan, possibly for a wedding that included a number of pictures. Also recovered were pictures of a young man posing with a pistol, a scanned image of passport and a scanned image of a credit card.

In the report on the survey for the previous year, the issue of the increase in the availability, use and reliance on computers and the services that they enable was discussed. Also addressed was the resultant ability and requirement to store, process and transmit increasing volumes of information, a proportion of which will inevitably be of a sensitive nature. It is reiterated that the technology has been available for some time that addresses the requirement to protect this information to an appropriate level and to destroy it effectively. The failures that are shown by these studies would indicate that the failure to properly dispose of the information would, within organisations, appear to be attributable to a lack of corporate policies and procedures for the disposal of obsolete equipment. A second contributory factor appears to be the lack of awareness of individual

users, whether on a business system or their home computer, of the impact of storing such information or of effective methods for ensuring that it is destroyed when it is no longer required.

Data security breaches have now reached an unprecedented level and are reported in the press on a very regular basis The (ITRC, 2008) report stated that '"Reports of data breaches increased dramatically in 2008." The Identity Theft Resource Center's 2008 breach report reached 656 reported breaches at the end of 2008, reflecting an increase of 47% over last year's total of 446'. While it is accepted that this report address all reported data security breaches, it supports the argument that there is increasing awareness of the issues relating to the storage, handling and disposal of data.

The subject of the disposal of disks on which the data has not been effectively destroyed first made the news in the Canadian Globe and Mail (Canadian Globe and Mail, 1993) and has been reported with increasing frequency in the intervening years to date. The reports that appear in the press tend to be those relating to high profile individuals and significant organisations such as government departments and financial institutions. One of the early reports (Calvert and Warren, 2000) related to the disposal, by the Morgan Grenfell Asset Management merchant bank, of a disk that contained details of Sir Paul McCartney's bank account. More recent reports have included:

- From the UK:

- A (BBC, 2007) report on the loss by HM Revenue and Customs department (HMRC) of 2 computer disks holding the personal details of all families in the UK with a child under the age of 16. The Child Benefit data on them includes name, address, date of birth, National Insurance number and, where relevant, bank details of 25 million people.

- An article on a Ministry of Defence investigation into what was claimed to be its worst information security breach after a portable hard drive with details of about 100,000 servicemen and women and 800,000 applicants to join the Armed Services was found to be missing (Kerbaj, 2008)

- A report in the Register (Oates, 2008) that a computer hard disk containing more than one million sets of bank details was purchased on eBay for just £35.The disk contained details of American Express, NatWest and Royal Bank of Scotland customers The information in the files included names, addresses, sort codes, account numbers, credit card numbers, mobile phone numbers, mothers' maiden names and even scans of signatures.

- A report in the Register (Leyden 2008) on the loss of a sensitive portable hard drive by the private contractor EDS which contained the details of an estimated 5,000 prison officer and admin staff. The report went on to comment about the loss of a computer memory stick containing the details of 84,000 prison inmates by a different consultancy in the previous month.

From the USA:

- A report on the loss of an unencrypted backup tape from the Bank of New York Mellon, the loss of which had potentially exposed information on 4.5 million customers of that bank and of People's United Bank of Bridgeport, CT(BankInfo Security, 2008). The missing tape contained social security numbers and bank account information on the 4.5 million customers and several hundred thousand depositors and investors of People's United Bank.

- An article (InfoWorld, 2007) in which TJX Companies, a large retailer that operates over 2,000 stores under brands such as Bob's Stores, HomeGoods, Marshalls, T.J. Maxx and A.J. Wright had admitted that it suffered a computer breach. This had occurred in its network that handled credit card, debit card, check, and merchandise transactions in the U.S. and overseas. Later reports put the number of accounts affected at 45.6 Million.

- An article in Newsvine (Kerr, 2008) reported that sensitive information on about 1,000 patients at Walter Reed Army Medical Center and other military hospitals had been exposed in a security breach. According to hospital officials, the information that was released included names, social security numbers, birth dates and other information

From Germany:

- A report in the Register (Oates, 2008a) that T-Mobile had admitted losing 17 million German customer records including names, addresses, phone numbers, dates of birth and email addresses. The records were stolen in 2006 and included 'secret' addresses of politicians, an ex-federal president, celebrities and others likely to be at risk from having their contact details released. No bank details were included in the stolen data. The records were offered for sale online, but the report stated that no one had bought them.

- An article in (Robertson, 2008) examined whether a computer intrusion at a Best Western hotel in Germany had allowed a hacker to steal the records of 8 million customers.

- An article in (Heise Online, 2007) reported that the Hamburg Kartenhaus ticket sales office had informed its customers of the theft of credit card numbers and billing addresses. Some 66,000 customers who had made purchases with a credit card from the Kartenhaus.de website between October 24, 2006 and September 30, 2007 were affected.

From Australia:

- A article by (Deane, 2008) on Symantec's first Australian data loss survey, that reported almost 80 per cent of local organisations have experienced a data breach in the past five years, with a further 40 per cent reporting between six and twenty known breaches during the period.

- An Article by (Sharma, 2008) gave details of the loss of an unencrypted disk containing the name, address and super fund tax file numbers for 3,122 trustees, when a courier failed to deliver it to the tax office.

It is notable that, as has been reported in earlier studies, until 2005 the subject of data losses only made the mainline news once every two or three years. In the USA alone in 2008 , there were 656 reports of data losses. It is also worth noting that these are only the cases where the loss or the data was discovered or the media was recovered and reported. The problem of the effective removal of data was first raised in an academic conference paper (Gutmann, 1996) and then through the intervening period (Gutmann, 2001) and (Garfinkel and Shelat, 2003).

Since 2004 there have been a number of reports and articles (Leyden, 2004), (Valli, 2004), (TechWeb News, 2005), (Herald Tribune, 2007), (Channel 4, 2008), (BBC News, 2008) all on the topic of data that had been recovered from disks that had been obtained on the second hand market. Over the same period, there had been a huge increase in the level of publicity on the problem of identity theft.

The ever increasing and regular exposure that the issue of the proper protection of personal information has received over recent years has so far failed to make any significant impact on the results that are being observed. Together with the publicity regarding the issue has been an increasing availability of suitable tools2 that can ensure the effective removal of information from disks. With the level of publicity, the increasing legislative pressure to take action and the increased availability of effective and affordable tools, it is strange that the one area of apparent improvement is that of data from which an individual private user can be identified. If validated, this would indicate that the home user has taken more notice and changed the way in which they deal with old computers and disks in a more effective manner than organisations. It would appear that the general level of awareness of the potential problems related to the disposal of computer equipment remains poor.

It was noted in the report for the 2007 study that changes in the way in which people work and the way in which they interact with technological devices will impact on the volume and type of information that they store on computers. An increasingly mobile workforce and a greater number of people either working from home or working from a home base, is likely to result in a greater mix of organisational and personal information stored on any disk. This problem is exacerbated by the relaxation of the rules on the actual use of computers in the workplace, with the acceptance by many organisations that some level of personal use is reasonable. These changes may result in data being stored on personal systems over which an organisation has no control and also in personal information being stored on organisational computers of which the managers have no knowledge. The results of the study this year again identified a significant number of disks that contained a mix of corporate and personal data. In a small number of cases, it was clear that the owner was a one-person or very small organisation where the computer had been used for both personal and business purposes.

Independent of this study, the University of Glamorgan is carrying out research on the disks that could not be accessed. They will produce a separate report into the level of difficulty and the complexity involved in restoring these disks. While not directly relevant to this study, this related work will provide an insight into the

[2] Such as the Blancco tool which is approved for use by the UK Government.

potential risk to organisations if more sophisticated tools and techniques are applied to the disks. It is worthy of note that the proportion of disks that do not work is between 30 and 40 percent of the total and that, because they do not work, it is probable that in many cases, all of the data that was stored on them at the time that they failed will be available.

The study this year has reinforced the findings of those in the preceding years, with a significant number of disks found that contained sensitive information on both organisations and individuals. The potential effect of the failure of organisations and individuals to destroy sensitive information on disks that are disposed of is that this information will continue to be accessible to the purchasers of the disks and cause embarrassment, potential financial losses and leave the original owner susceptible to blackmail or provide the opportunity for identity theft.

5. Conclusions

While the improvements in the results that were found in 2006 and 2007 have largely been confirmed, it was disappointing to note that the percentage of disks that have been effectively wiped remains at only 36 percent during the 2008 study. It is also difficult to understand why the only consistent improvement has been in the reduction in the level of information from which an individual could be identified, while the percentage of disks from which an organisation could be identified had got slightly worse.

One of the notable changes detected in the results of the study was the percentage of disks from Germany that were unreadable. These results have shown a significant improvement over those in previous years with only 39 percent compared to 71 percent in the 2007 study. The reason for this dramatic change is not known.

It remains clear from the results of this research that there is an ongoing requirement for significant improvements in the awareness of staff within organisations of the potential risks and the actions to be taken for the safe disposal of information stored on computer disks. This can be achieved through awareness campaigns, education and training programmes for appropriate staff. Without this effort and suitable policies, procedures and the availability of appropriate tools, staff will continue to fail to fulfil their organisational responsibilities with regard to the safe disposal of computer systems.

There remains a strong argument for leadership by government in this area in the form of public awareness campaigns and suitable legislation that is supportive of a range of measures to improve the whole area of information security.

6. Recommendations

Throughout the period that this research has been taking place, a number of recommendations have been made regarding the measures that can be taken by both individuals and organisations to reduce to reduce the level of sensitive information that is exposed when computer devices and the hard disks that they contain are disposed of. These are re-iterated below:

- User Education - A public awareness campaign by Government, the media, commerce and/or academia. An example of this is the information provided by the disk manufacturer Seagate in their 'What is Data Erasure?[3]' and 'Drive Disposal Best Practices[4]' documents or documents containing advice such as 'How To Permanently Erase Data from a Hard Disk'[5].

- Organisational Risk Assessments – Carry out organisational risk assessments to determine the sensitivity of the information on disks.

- Best Practice - The introduction into organisations of procedures to ensure that computer systems and computer hard disks are disposed of in an appropriate manner.

- Physical Destruction - Where appropriate, the physical destruction of the disks using services such as the Ultratec Secure Data Erasure service[6] or that offered by DataTerminators[7].

[3] Seagate – what is data erasure? - http://www.seagatedatarecovery.com/assets/resources/(CB-20-0508-e)-What-is-Data-Erasure-Screen.pdf (Accessed 05 Mar 2009)
[4] Seagate - Drive Disposal Best Practices, http://www.seagate.com/docs/pdf/whitepaper/Disposal_TP582-1-0710US.pdf (Accessed 05 Mar 2009)
[5] Sedory, DB (2008), How To Permanently Erase Data from a Hard Disk, http://mirror.href.com/thestarman/asm/mbr/WIPE.html
[6] Ultratec Limited - http://www.ultratec.co.uk/
[7] DataTerminators - http://www.data-terminators.co.uk/

- Data Erasure – The development of and access to the tools such as Blancco data erasure tool[8] and facilities to enable individuals to effectively remove the information from their computers.

- Encryption - The full encryption of hard disks to ensure that information could not be easily recovered using software such as TrueCrypt[9] or PGP whole disk encryption[10] or hardware encryption devices such as the Secure Data Vault[11].

- Asset Tracking - It is also suggested that organisations may more effectively secure their data if asset tracking is conducted at a storage device level. This would require that asset tags are placed on individual disks rather than the computer system unit to ensure safe disposal as increasingly systems are offered with more than one physical storage device.

- Legal – Assign responsibility to those charged with receiving discarded or damaged hard disks. Disks considered dead or faulty should have the same disposal practices applied to them as disks removed from a working system.

From the results of the studies that have taken place over the last four years there appears to have been a consistent but marginal improvement in the area of the level of personal information remaining on the disks, while the level of organisational data has slightly risen this year after consistently falling over the last three years.

While it is increasingly difficult to believe that either organisations or individuals are not aware of the potential problems that failing to take appropriate measures to protect data can cause, it is likely that the results or the study are a result of:

- The time it takes to organisations to develop, validate, implement and promulgate new processes and procedures.
- An ongoing lack of commitment or the structure to implement education and awareness campaigns at the levels that are required.
- A failure to understand the problem at an organisational and governmental level and a failure to prioritize the issue.

Acknowledgements

In addition to the individuals named as authors for this paper, we would like to acknowledge the significant efforts of the people who assisted in the imaging and analysis of the large number of disks required for the research and the preparation of this report. The people involved were:
Andrew Woodward, Gogoba Daa Dabibi, Thomas Martin, Konstantinos Xynos, and Simon Harries.

Contributing organisations

British Telecommunications (BT). BT is one of the world's leading providers of communications solutions serving customers in Europe, the Americas and Asia Pacific. Its principal activities include networked IT services, local, national and international telecommunications services, and higher-value broadband and internet products and services. In the UK, BT serves more than 20 million business and residential customers with more than 30 million exchange lines, as well as providing network services to other licensed operators.

Edith Cowan University (ECU). The SECAU - Security Research Centre at ECU conducts research into all aspects of Computer and Information Security from the technological aspects of computer forensics and network security to the 'softer' side involving issues such as perception management and information policy. The group has numerous doctoral, masters and honours candidates. Its main areas of interest are information operations, computer/network forensics, RFID security, mobile computing security, honeypots and the use of deception in security

[8] Blancco - http://www.blancco.com/en/frontpage/
[9] TrueCrypt - http://www.truecrypt.org/downloads.php
[10] PGP Corporation, Whole disk ensryption - http://www.pgp.com/products/wholediskencryption/index.html
[11] Secure Systems Secure Data Vault - http://www.securesystems.com.au/

Longwood University (LU). The University offers interdisciplinary programs in Homeland Security and in Digital Forensics, Security and Law. Longwood University's program in Homeland Security offers students an interdisciplinary exposure to the global economic, cultural and political issues relevant to homeland security. Its Digital Forensics, Security and Law program brings together students from Information Systems, Criminal Justice and Computer Science.

University of Glamorgan (UoG). The Information Security Research Group from the Faculty of Advanced Technology at the UoG has a strong and well established theme in the areas of Computer forensics, Computer Network Management and Computer Network Defence. The Information Security Research Group is focused on the issues associated with the design and development of early warning systems that are capable of detecting and responding to a variety of cyber based attacks, and on the issues associated with computer forensic science. The research is conducted mainly in the two specialised laboratories of the group, the Network Security Laboratory and the Computer Forensics Laboratory. The research feeds into the undergraduate and postgraduate degree schemes in forensics and computer systems security offered at the university.

References

1. BankInfo Security (2008), Bank of New York Mellon Investigated for Lost Data Tape, http://www.bankinfosecurity.com/articles.php?art_id=862, (accessed 03 Mar 2009)
2. BBC News (2007), UK's families put on fraud alert, BBC, 20 November 2007, http://news.bbc.co.uk/1/hi/uk_politics/7103566.stm, (accessed 03 Mar 2009)
3. BBC News (2008), HSBC loses customers' data disc, http://news.bbc.co.uk/1/hi/business/7334249.stm (Accessed 08 Mar 2009)
4. Calvert, J, Warren, P (2000), Secrets of McCartney Bank Cash Are Leaked, *Daily Express*, 9 February 2000, pp 1–2. (Accessed 22 Feb 2009)
5. Canadian Globe and Mail (1993), Disk Slipped Into Wrong Hands, *Canadian Globe and Mail*, 2 Aug 1993. (Accessed 25 Feb 2009)
6. Channel 4, Science and technology (2008), 'Melted' disk drive data recovered, http://www.channel4.com/news/articles/science_technology/melted+disk+drive+data+recovered/2166472 (Accessed 09 Mar 2009).
7. Deane, K., (2008). Data breach hits 80% of local companies: survey, The Australian, http://www.australianit.news.com.au/story/0,,24530567-15306,00.html (Accessed 01 Mar 2009)
8. Garfinkel S.L, Shelat A, (2003), Remembrance of Data Passed: A Study of Disk Sanitization Practices. *IEEE Security & Privacy*, Vol. 1, No. 1, 2003.
9. Gutmann, P. (1996), Secure Deletion of Data from Magnetic and Solid-State Memory, *Sixth USENIX Security Symposium Proceedings*, San Jose, California, 22-25 Jul 1996.
10. Gutmann, P. (2001), Data Remanence in Semiconductor Devices, *10th USENIX Security Symposium*, Washington, D.C., 13-17 Aug 2001.
11. Heise Online (2007), Theft of credit card data affects tens of thousands of Kartenhaus customers, http://www.heise.de/english/newsticker/news/96992 (Accessed 04 Mar 2009)
12. Herald Tribune (2007), UK government lost personal data on driving students in US, http://www.iht.com/articles/ap/2007/12/17/europe/EU-GEN-Britain-Disappearing-Data.php (Accessed 04 Mar 2009)
13. Identity Theft Resource Centre (ITRC). (2008), ITRC Breach Report 2008 Final, http://www.idtheftcenter.org/BreachPDF/ITRC_Breach_Report_2008_final.pdf (Accessed 03 Mar 2009)
14. Infoworld. (2007), Retailer TJX reports massive data breach, http://www.infoworld.com/article/07/01/17/HNtjxbreach_1.html (Accessed 04 Mar 2009)
15. Johannes, R. (2006), The Demographics of Identity Fraud: Through education and vigilance, banks can prepare and protect those most vulnerable, Javelin Research, http://www.javelinstrategy.com/uploads/607.R_2006_IDF_Demographics.pdf, Aug 2006. (Accessed 04 Mar 2009)
16. Jones, A., Mee, V., Meyler, C., and Gooch, J,(2005), Analysis of Data Recovered From Computer Disks released for sale by organisations, *Journal of Information Warfare*, (2005) 4 (2), 45-53.
17. Jones, A., Valli, C., Sutherland, I., and Thomas, P,(2006), The 2006 Analysis of Information Remaining on Disks Offered for Sale on the Second Hand Market, *Journal of Digital Forensics, Security and Law*, (2006) 1 (3), 23-36.
18. Jones, A., Valli, C., Sutherland, I., and Dardick, G., (2008), The 2007 Analysis of Information Remaining on Disks Offered for Sale on the Second Hand Market, *International Journal of Liability and Scientific Enquiry 2009 - Vol. 2, No.1 pp. 53 – 68.*

19. Kerbaj, R., (2008), MoD investigates biggest ever data loss, The Times , 10oct 2008, http://www.timesonline.co.uk/tol/news/uk/article4918986.ece, (Accessed 03 March 2009).
20. Kerr, J.C., Newsvine. (2008), Walter Reed says patient data may be compromised, http://www.newsvine.com/_news/2008/06/02/1531942-walter-reed-says-patient-data-may-be-compromised. (Accessed 07 Mar 2009)
21. Leyden, J. (2004), Oops! Firm accidentally eBays customer database, The Register, 07 June 2004. (Accessed 07 Mar 2009)
22. Leyden, J., (2008), Prison officers slam EDS data loss, The Register, 08 Sept 2008, http://www.theregister.co.uk/2008/09/08/prison_officer_lost_disc_debacle/, (Accessed 03 March 2009).
23. Oates, J., (2008), Million bank details sold on eBay, The Register , 26 Aug 2008, http://www.theregister.co.uk/2008/08/26/more_details_lost/
24. Oates, J., (2008a), T-Mobile joins data breach elite, The Register , 6th Oct 2008, http://www.theregister.co.uk/2008/10/06/t_mobile_records_lost/
25. Price Waterhouse Cooper (2006), DTI Information security breaches survey 2006, http://www.dti.gov.uk/industries/information_security Sept 2006. (Accessed 07 Mar 2009)
26. Robertson, J., Newsvine. (2008a), Best Western rebuts claims of massive data breach, http://www.newsvine.com/_news/2008/08/26/1789209-best-western-rebuts-claims-of-massive-data-breach (Accessed 07 Mar 2009).
27. Sharma, M., (2008), Tax data unencrypted, The Australian, http://www.australianit.news.com.au/story/0,,24597034-15306,00.html (accessed 04 Feb 2009)
28. Slashdot (2008), Computer With UK Bank Customer Data Sold On eBay, http://it.slashdot.org/article.pl?sid=08/08/27/0055251&from=rss . (Accessed 08 Mar 2009)
29. Synovate, (2003), *Federal Trade Commission – Identity Theft Survey Report*, Federal Trade Commission, June 2006.
30. Techweb, (2005), Seven-In-Ten Second-hand Hard Drives Still Have Data, Bank Systems and Technology, 01 Jul 2005, http://www.banktech.com/risk-management/showArticle.jhtml?articleID=165600008. (Accessed 07 Mar 2009)
31. Valli, C. (2004), Throwing out the Enterprise with the Hard Disk, In *2nd Australian Computer, Information and Network Forensics Conference*, We-BCentre.COM, Fremantle Western Australia.
32. Verizon Business Risk Team, (2008), Data Breach Investigations Report, http://www.verizonbusiness.com/resources/security/databreachreport.pdf , (Accessed 12 Mar 2009).