A Case of Mistaken Identity? News Accounts of Hacker, Consumer, and Organizational Responsibility for Compromised Digital Records

Kris Erickson

Department of Geography
University of Washington

Philip N. Howard

Department of Communication
University of Washington

The computer hacker is one of the most vilified figures in the digital era, but to what degree are organizations actually responsible for compromised personal records? To examine the role of organizational behavior in privacy violations, we analyze 589 incidents of compromised data between 1980 and 2006. There were more reported incidents in 2005 and 2006 than in the previous 25 years combined. Excluding a particularly large security breach at Acxiom, hackers account for the largest volume of compromised records, some 45%, while 27% of the volume is attributed to organizational mismanagement and 28% remains unattributed. In terms of incidents, 9% were an unspecified type of breach, 31% of the incidents involved hackers, and 60% of the incidents involved organizational mismanagement: personally identifiable information accidentally placed online, missing equipment, lost backup tapes, or other administrative errors. Options for public policy oversight are discussed.

doi:10.1111/j.1083-6101.2007.00371.x

Introduction

Recently, electronic personal records have become the subject of a great deal of public interest. Their ubiquity has spurred debates about the nature of democracy, the potential for electronic panopticism, and the erosion of personal privacy in an era of increased police surveillance. Attention has been leveled at the various aspects of data collection, data management (or mismanagement), and the potential for unwanted disclosure of private records through loss or theft. In early 2005, a series of high-profile cases culminating in the loss of more than 140,000 customer credit records by the data mining firm ChoicePoint helped generate significant public interest in the dangers associated with digital records of personal information. Then that summer, another 40,000 credit card records appeared on the black market for personal data, and in the summer of 2006, the U.S. Department of Veterans' Affairs admitted that some 18,000 personal records had been compromised. Data security is never perfect, and credit card companies, universities, and government agencies cannot predict security lapses. But the growing number of news stories about compromised personal records reveals a wide range of organizational mismanagement and internal security breaches: lost hard drives and backup tapes, employee theft, and other kinds of administrative errors.

So far, blame has been directed at all parties involved: at the state, for being lackadaisical in regulating institutions and businesses that deal with electronic records; at the private sector, which is accused of de-prioritizing personal privacy and information security; and finally at the end-users themselves, who are enjoined by a variety of authorities and experts to take better care of managing their online identities in order to mitigate the risk of fraud. A significant amount of the information in these records concerns health and credit records, but such data are often combined so that businesses, lobbyists, and politicians can generate a convincing electronic portrait of an individual, thus effectively reconstituting their identity (Howard, Carr, & Milstein, 2005). These stolen identities can also be used fraudulently to deceive government agencies and credit institutions.

The threat of electronic data theft also has serious implications for societies that increasingly rely on the security of data networks for day-to-day life. For example, as more of our political system becomes computerized, there is a stronger possibility that electronic data could contain information about an individual's political beliefs or voting records, which are now both easier to access and highly detailed (Howard, 2002, 2006). Yet most U.S. citizens report being uninterested in learning how to manage their personal data better or in learning about how organizations mine for data (Fox, 2000; Milne & Culnan, 2004). Today, however, both policy makers and computer software and hardware companies are aggressive in enrolling individual consumers in the task of securing their own data against loss or theft.

At the center of these privacy breaches is often the hacker archetype. Corporate and government leaders have reframed the meaning of computer hacking, using intellectual property law, court challenges, and amicus briefs, from a character working for freedom of access to technology and information to one who is deviant and criminal (Nissenbaum, 2004). However, the actual role of hackers in the computer security sector is considerably more complex. Many hackers not only enjoy technical challenges, but are sometimes enlisted by corporations and governments for their specific skills (Samuelson, 2003; Universal City Studios, 2000). Even though the campaign against hackers has successfully cast them as the primary culprits to blame for insecurity in cyberspace, it is not clear that constructing this target for blame has improved the security of personal digital records.

This article explores how responsibility for protecting electronic data is currently attributed and examines legislation designed to manage the problem of compromised personal records. This makes it possible to compare the aims of legislation with an analysis of reported incidents of data loss for the period of 1980-2006. A discrepancy between legislative responses to electronic data loss and the actual damages incurred reveals that responsibility for maintaining the security of electronic personal records has been misplaced and should be re-examined. We conclude with a brief discussion of the options for public policy oversight.

U.S. Legislation to Secure Electronic Records

Legal scholars often point out that new information technologies consistently present legislators with the challenge of regulating issues for which there are no readily apparent legal precedents. Lawmakers are frequently cast as lagging behind technological innovation, as they struggle to catch up with new forms of behavior enabled by rapidly evolving technology. Traditional legal concepts such as private property and trespass often become problematic when applied in online contexts enabled by information and communication technologies.

For example, Cavazos and Morin (1996) have argued that in the case of defamatory, libelous, and obscene speech, the law has struggled to account adequately for the nuances of computer-mediated communication. Offline publishers and republishers of defamatory statements can be held liable, because it is expected that they possess considerable editorial control over their own published content. However, when publication moves into an online setting, the distribution of liability becomes less clear. Not all Internet publishers maintain strict editorial control, and some media outlets function more like "conduits" through which news is automatically updated. Other websites allow users to generate content, with limited moderation provided by the system administrator. In both of these cases, it becomes more difficult to assign responsibility for defamatory material.

The decentralized nature of computer networks poses other challenges for regulators. In cases involving obscenity, lawmakers in the United States have employed a method known as the "community standards test" to determine whether published material can be considered obscene. Material is deemed to lie outside the protections afforded by the First Amendment of the U.S. Constitution when it is found to be offensive to the norms and standards of the community in which it is located. While this method has functioned adequately in offline settings, it is less effective when individuals from diverse communities can transmit information to one another, often across state and national boundaries (Cavazos & Morin, 1996; Zook, 2003). Early applications of the community standards test to online publishers proved unworkable. In the case of *United States v. Thomas* (6th Cir., 1996) a website operator located in California was tried and convicted in Tennessee for violating the obscenity laws in the jurisdiction where the material was accessed, rather than where the material was stored ("United States v. Thomas," 1996). This case is often cited as evidence that current legislation is anachronistic and lags behind the requirements of communication technologies that bypass traditional jurisdictional boundaries.

These jurisdictional conflicts become even more apparent in cases where lawmakers have attempted to regulate behavior across several legal jurisdictions, such as in the cases of music piracy and online gambling. Faced with an overwhelming number of users, along with the relative anonymity provided by computer-mediated communication, prosecutors in the U.S. have tended to focus efforts on website operators rather than on end users. The jurisdictional challenges posed by computer networks continue to hamper their efforts in this regard, however, since offending websites can be operated offshore in areas with less stringent regulation. The U.S. has pursued this strategy in regard to online gambling, with limited success. Charges brought by New York State against 22 online gambling websites in 1999 yielded only one arrest, when the operator visited the U.S. on vacation (Wilson, 2003).

An additional problem facing legislation aimed at controlling online behavior is its questionable effectiveness as a deterrent. The Computer Fraud and Abuse Act (CFAA) was passed in 1984 in response to growing political and media attention surrounding the dangers of computer crime. The act criminalized unauthorized access to private computer systems, making it a felony offense when trespass leads to damages over a certain monetary threshold. The CFAA underwent major revisions in 1986 and 1996, and it was further strengthened by the passage of the USA Patriot Act in 2002. Overall, these revisions have served to make the act more broadly applicable to various kinds of computer crime, while also increasing the punitive response to these offenses.

For example, the revisions in 2002 were tailored to make it easier to surpass the $5,000 felony threshold. The threshold was waived in cases where the computer systems involved are used for national security or law enforcement purposes. In cases not involving national security, the definition of "damage" was broadened to include costs relating to damage assessment and lost revenue during an interruption of service. The $5,000 threshold is also cumulative over multiple machines if more than one system is involved in an attack.[1] Additionally, the maximum sentence for felony computer trespass was raised from five to 10 years for first-time convictions, and from 10 to 20 years for repeat offenders (Skibell, 2003).

Given the relatively harsh penalties for computer trespass compared with other crimes where victims suffer personal physical harm, it is surprising that the CFAA has not been more effective as a deterrent. The apparent surge in computer-related offenses, including the theft of online personal records, suggests that the punitive nature of this legislation is not having the desired effect. Skibell (2002) argues that not all computer crime is committed by self-interested or malicious criminals. The belief that all hackers are malicious is essentially a myth—many members of the computer hacker subculture do not condone destructive behavior and do not consider their activities to be particularly malicious. Criminals who make use of hacker techniques to access private data are rarely members of hacker communities, and often less sophisticated in their hacker skill-set (Skibell, 2002). More legitimate computer hackers appear to be motivated by codes of conduct internal to their community and are therefore less likely to be deterred by legal sanctions. According to Jordan and Taylor (1998), these legitimate computer hackers are motivated by a variety of concerns that make comparisons with other types of criminal behavior problematic:

> How often does a burglar leave behind an exact copy of the video recorder they have stolen? [...] What bank robbers ring up a bank to complain of lax security? The simple analogy of theft breaks down when it is examined and must be complicated to begin to make sense of what hackers do. (p. 772)

These scholars argue that hacker projects are shaped by an ethical framework formed by a strong sense of imagined community. Many hackers are interested in the intellectual challenge and sense of mastery provided by computer networks, rather than monetary rewards that could be gained from accessing sensitive information. They seek to differentiate themselves from other computer criminals who use computer networks for destructive, rather than creative, processes.

Furthermore, in recent years the core hacker community has been somewhat successful at contesting the malicious meaning attached to the term "hacker." While the press often continues to report hackers as those responsible for most forms of computer crime, more legitimate hackers have worked hard to distance themselves from the sensationalist definition used by the news media. Many of them divide their community into "white hat" and "black hat" constituencies to help distinguish those who use their computer skills with malicious intent from those who do not. The term "cracker," which now denotes individuals who destroy rather than improve computer systems, indicates a deliberate rhetorical strategy on the part of some hackers to bring nuanced understanding of the different aspects of computer hacking, particularly among scholars interested in computer subcultures (Jordan & Taylor, 2004; Thomas, 2002). In contrast, "gray hats" are those who publicly expose security flaws, without concern for whether the act of exposure allows administrators to patch the flaw or allows others to exploit the flaw. Moreover, mainstream computer security experts have co-opted the term "blue hat" to further distinguish the community of skilled computer users who hack in the service, and often employment, of Microsoft.

Arguably, the most significant threat posed by computer criminals comes not from the core group of white, blue, or grey hat hackers but from individuals who make use of hacker techniques to invade systems for monetary gain. Since knowledge and tools developed by more experienced hackers can easily be obtained on the Internet, the capability to penetrate insecure networks has propagated outside of the legitimate hacker community to other groups, ranging from inexperienced teenagers to international crime syndicates.[2] These individuals may feel protected from the law due to the relative anonymity of computer-mediated communication, or they may be located in jurisdictions where harsh criminal penalties for computer fraud do not apply.

While the CFAA aids in the prosecution of criminals who engage in electronic data theft and trespass, individual states have recently taken additional legal steps to regulate the management of electronic records. In 2003, the state of California introduced a new provision to the Information Practices Act, termed the "Notice of Security Breach." This addition to the California Civil Code obliges any business or agency that has been the victim of a security breach to notify any parties whose personal information may have been compromised. The California legislation defines "personal information" as an individual's full name, in combination with one of the following types of data:

(1) Social security number
(2) Driver's license number or California Identification Card number
(3) Account number, credit or debit card number, in combination with any required security code, access code, or password that would permit access to an individual's financial account.

The company or institution responsible for handling the compromised data must notify potential victims individually, unless the cost of notification exceeds a threshold amount of $250,000, or if the total number of individuals affected is greater than 500,000. In these cases, substitute notification can be made using a combination of e-mail notification and disclosure to major media outlets. Notification must be carried out:

> in the most expedient time possible and without unreasonable delay, consistent with the legitimate needs of law enforcement [...] or any measures necessary to determine the scope of the breach and restore the reasonable integrity of the data system. (California Civil Code 1798.29)

Following California's footsteps, 34 additional states have enacted similar legislation as of 2007.[3] For the most part, individual state legislatures have maintained the spirit of the California provision, including the extension of liability to both businesses and agencies, as well as the notification threshold.

The aim of the "notification of breach" legislation is significantly different from that of the CFAA. By making corporations and institutions liable for damages potentially incurred by customers and clients, this legislation to some extent seeks to discipline offenders who engage in poor record-keeping practices. Both the indirect threat of future litigation and the potential for public embarrassment are intended to improve data security in both the public and private sector. Unlike the CFAA, however, this legislation does not directly address the issue of network security. It does not formalize standards or rules for information security, nor does it make businesses and institutions accountable for poor security practices that may make them vulnerable to attack. The legislation punishes businesses only for failing to notify the public, rather than for negligence in securing electronic records.

Since adequately securing a computer network from intrusion is an expensive prospect, this legislation essentially lets businesses off the hook, by making them liable for damages only when they fail to notify affected individuals that their data have been compromised. Interestingly, by failing to assign responsibility for data loss to those agencies that manage electronic personal information, this legislation serves in part to shift that responsibility to the individual user, since it is he or she who must take steps to protect their identity once notified of a breach.

This sentiment is supported by the California Department of Consumer Affairs, which maintains a website devoted to online privacy protection. The agency has also distributed a flyer listing the "top 10 tips for identity theft prevention." This list enjoins consumers to take active steps to avoid becoming the victims of electronic fraud, by shredding personal documents, installing up-to-date computer virus and firewall software, and becoming vigilant about which sites they visit and how they use their credit cards. Consumers are also urged to take a more proactive role in monitoring their personal credit rating, in order to detect potential fraud. The Department of Consumer Affairs recommends that individuals apply for free credit reports at least three times per year in order to prevent misuse of their electronic identity.

So far, the legal responses to electronic identity theft in the U.S. have sought to minimize the direct involvement by the state, instead relying on a partnership between the interests of private institutions and the consumers of those services. The two major forms of legislation governing the security of computer records in the U.S.—the CFAA and the California Notification of Breach laws—closely resemble offline governmental strategies that seek to place responsibility on individual consumer-citizens while disciplining those who do not adequately protect themselves (Burchell, 1996; Peck & Tickell, 2002). Moreover, the reticence of public agencies in the U.S. to draft legislation that would directly influence the terrain of data security is consistent with the overall trend of regulatory devolution, a shift that began before the information sector occupied such a primary position in the national economy.

The legislative choices that policymakers in the U.S. have made to combat the problem of data insecurity have been shaped by the tenet that governments should interfere only minimally with markets. Thus, legislative initiatives have eroded public policy oversight of corporate behavior. In the arena of data security for private information, this erosion has meant de-emphasizing the role of government and public policy oversight in data security, encouraging industry self-regulation among the firms benefiting in the retention of personal data, and increasing individual responsibility for managing one's personal data.

Analysis of Compromised Electronic Records, 1980–2006

We conducted a search of incidents of electronic data loss reported in major U.S. news media from 1980 to 2006. These included print publications with national circulation such as the *New York Times*, the *L.A. Times*, and *USA Today*, along with major broadcast news media. Because some news reports contained references to more than one incident, we employed a snowball methodology to expand our analysis by including additional security breaches mentioned in the same article. Duplicate entries were eliminated by comparing news stories on the basis of organizations involved, dates, and other incident details. In cases where newspapers reported different quantities of lost records, we chose the most conservative report.

We also consulted lists of electronic data breaches compiled by third party computer security advisories, such as the Identity Theft Resource Center (www.idtheftcenter.org) and Attrition.org. Our method yielded 589 incidents, 550 of which were successfully cross-checked with LexisNexis and Proquest to ensure accuracy, and 39 of which we discarded for involving citizens of other countries or for being unverifiable in major news media reports.[4]

There are interesting advantages and disadvantages to using printed news sources to construct the history of computer hacking and breached private records. As stated above, the mainstream media often equate hackers with any crime involving a computer and use the misnomer "hacker" without a nuanced understanding of the history of more legitimate computer hacking. We continue to use the term in this analysis, because it is the most commonly used term in media reports where an intruder was deemed responsible for compromised data. Instead of media reports, criminal records would certainly provide details about the prevalence of malicious intrusions. Unfortunatley, such records are extremely difficult to collect nation-wide. Moreover, a survey of incidents composed through criminal records would significantly *over-sample* incidents where an individual hacker was at fault, and significantly under-sample incidents where an organization was culpable but not deemed criminally negligent. Over the decade, journalists would not have discovered all incidents, and even though current California law requires that a person whose data had been compromised be so informed, such a breach is not necessarily noted in news archives. However, journalists do their best to report the facts, and in the absence of a public agency that might maintain comprehensive incident records on privacy violations, news accounts provide a good accessible resource.

Our list of reported incidents is limited to cases where one or more electronic personal records were compromised through negligence or theft. We acknowledge that there may be occasions where end-users consider their personal information compromised when the data are sold among third parties for marketing purposes without their informed consent. For this study, we look only at incidents of compromised records that are almost certainly illegal or negligent acts.

For the purposes of this study, we define electronic personal records as data containing privileged information about an individual that cannot be readily obtained through other public means. Rather than become involved in the broader debate about the virtues and dangers of online anonymity, we have chosen to focus only on data that are more sensitive than the information that we regularly volunteer in the course of surfing the web (such as one's name or IP address). We define "personal data" to be information that should reasonably be known only to the individual concerned or be held by an organization under the terms of a confidentiality agreement (such as between a patient and a care provider). Electronic personal records therefore could include individuals' personal credit histories, banking information such as credit card numbers or account numbers, medical records, social security numbers, and grades earned at school. We focused only on incidents where compromised personal records were kept for a legitimate purpose by a company, institution, or government agency. Consequently, "phishing" or spoofing scams where victims are deceived into volunteering their own personal information are not included in our analysis. All of the incidents in our analysis deal with data that were maintained as digital files, although in some cases compromised data were contained on a lost or stolen laptop computer.

Between 1980 and 2006, some 1.9 billion records were reported compromised by government agencies, firms, hospitals, universities, and the military. This is the sum of compromised records from 529 cases in which some estimate of the volume of lost records was offered, although in 60 of these incidents the impact of the security breach was unknown. In a sense, this number of lost records is larger than we might expect because a few landmark incidents account for large portions of the total number of records compromised. On the other hand, the number of confirmed incidents—550 in all—may seem smaller than expected given the 26-year time frame of our search. Some articles report multiple incidents, and of course many incidents were covered by journalists on multiple occasions.

In 2004, the Census Bureau estimated that there were 217 million adults living in the United States. We can conservatively estimate that for every U.S. adult, in the aggregate, nine private records have been compromised. Unfortunately we cannot know how many of these compromised private records have actually been used for identity theft, or how many were sold to marketing companies.

Table 1 reveals the number of reported incidents and volume of compromised records between 1980 and 2006, along with their distribution by sector. The majority of incidents involved commercial actors, less than one-third of the incidents involved colleges and universities, and the remainder involved government, hospitals, and the military. When the exceptional loss of 1.6 billion personal records by Acxiom Corporation is removed, the commercial sector still accounted for approximately 252 million individual compromised records, four times that of the next-highest contributor, the government sector.[5]

The education sector accounted for a small percentage of the overall quantity of lost records, but accounted for 30% of all reported incidents, suggesting that educational institutions suffer from a higher rate of computer insecurity than might be anticipated. This could be explained by the fact that colleges and universities generally maintain large electronic databases on current and past students, staff, faculty, and alumni, and have an organizational culture geared towards information sharing. However, medical institutions—which presumably also maintain large quantities of electronic data—reported a significantly lower number of incidents of data loss. These differences may be the result of strong privacy legislation in the arena of medical information, but comparatively weak privacy legislation in the arena of educational and commercial information.

Although Table 1 has aggregated 26 years' worth of incidents, the bulk of the reports occur in 2005 and 2006, after legislation in California, Washington, and other states took effect. There were three times as many incidents in the period between 2005 and 2006 as there were in the previous 25 years. Interestingly, the mandatory

Table 1 Reported incidents and volume of compromised records by sector, 1980–2006

Time Period			Commercial	Educational	Government	Medical	Military	Total
1980–1989	Records	N	90,000,002	0	0	0	4,190,000	94,190,002
		%	96	0	0	0	4	100
	Incidents	N	3	0	1	0	3	7
		%	43	0	14	0	43	100
1990–1999	Records	N	53,369,339	0	20	3,010	461	53,372,830
		%	100	0	0	0	0	100
	Incidents	N	16	0	1	2	3	22
		%	73	0	5	9	14	100
2000–2006	Records	N	1,708,156,069	8,121,234	63,543,351	4,640,097	90,601	1,784,551,352
		%	96	0	4	0	0	100
	Incidents	N	190	166	107	51	7	521
		%	36	32	21	10	1	100
Total	Records	N	1,851,525,740	8,121,234	63,543,392	4,643,118	4,281,129	1,932,114,613
		%	96	0	3	0	0	100
	Incidents	N	209	166	109	53	13	550
		%	38	30	20	10	2	100

Note: A zero value in sectors with no incidents indicates that no records were compromised. A zero value in sectors with incidents indicates that the volume of compromised records was not reported.

reporting legislation seems to have exposed educational institutions as a major source of leakage of private data. In total, 38% of the incidents involved commercial firms, but specifically in 2005 and 2006, 35% of the incidents involved educational institutions. These kinds of organizations may have been the least equipped to protect the data of their students, staff, faculty, and alumni.

For the majority of incidents, the news article reports some information about how the records were compromised. A closer reading of each of the incidents, however, reveals that most incidents involve combinations of mismanagement, criminal intent, and, occasionally, bad luck. The hacker label is often used, even when the theft is perpetrated by an insider, such as a student or employee. Moreover, company public relations experts often posit that personal records were only “exposed,” not compromised, when employees post private records to a website or lose a laptop and the company cannot be sure that anyone has taken specific advantage of the security breach.

Table 2 reveals that the legislation has also seemed to have the effect of forcing the reporting organizations to reveal more detail about the ways these private records get compromised. In the early reports, most incidents were described as an unspecified breach or as the general result of hacker activity. However, for the period between 2000 and 2006, only 31% of the incidents were about a breach caused by a hacker, 8% of the incidents involve an unspecified breach, and 61% of the incidents involved different kinds of organizational culpability. For example, sometimes management accidentally exposed private records online, administrative error resulted in leaked data, or employees were caught using the data for activities not related to the work of the organization. On some occasions, staff simply misplaced backup tapes, while on others, computer equipment such as laptops were stolen.[6]

A single incident, involving 1.6 billion compromised records at Axciom, accounts for a large portion of the volume of records lost in the period 2000-2006.[7] If this event is removed from this period, then 32% of the compromised volume and 30% of the incidents are related to hackers, 48% of the compromised volume and 62% of the incidents involve organizational behavior, and 20% of the compromised volume and 8% of the incidents remain unattributed. If this event is removed from the volume of compromised records for the whole study period—between 1980 and 2006—then 45% of the total volume of compromised records related to hackers, 27% of the volume was attributed to the organization, and 28% remained unattributed. If this event is removed from the total number of incidents for the whole study period, then 31% of the incidents involved hackers, 60% involved organizational management, and 9% remain unattributed. Regardless of how the data are broken down, hackers never account for even half of the incidents or the volume of compromised records.

If we distinguish the reported incidents that clearly identify a hacker from those concerning some other form of breach, the organizational role in these privacy violations moves into sharp relief. Figure 1 separates the count of stories in which a hacker was clearly identified as the culprit from those stories where the cause of the

Table 2 Reported incidents and volume of compromised records by type of breach, 1980–2006

Time Period			Administrative Error	Exposed Online	Insider Abuse or Theft	Missing or Stolen Hardware	Stolen – Hacked	Unspecified Breach	Total
1980–1989	Records	N	0	0	0	0	90,000,002	4,190,000	94,190,002
		%	0	0	0	0	96	4	100
	Incidents	N	0	0	1	0	3	3	7
		%	0	0	14	0	43	43	100
1990–1999	Records	N	0	3,030	20	20,000	33,430	53,316,350	53,372,830
		%	0	0	0	0	0	100	100
	Incidents	N	0	3	1	1	10	7	22
		%	0	14	5	5	45	32	100
2000–2006	Records	N	33,281,120	4,605,967	6,844,162	44,397,886	1,659,391,166	36,031,051	1,784,551,352
		%	2	0	0	2	93	2	100
	Incidents	N	18	81	24	198	159	41	521
		%	3	16	5	38	31	8	100
Total	Records	N	33,281,120	4,609,014	6,844,203	44,417,892	1,749,424,795	93,537,590	1,932,114,614
		%	2	0	0	2	91	5	100
	Incidents	N	18	84	26	199	172	51	550
		%	3	15	5	36	31	9	100

Note: A zero value in a type of breach with no incidents indicates that no records were compromised. A zero value in sectors with incidents indicates that the volume of compromised records was not reported.

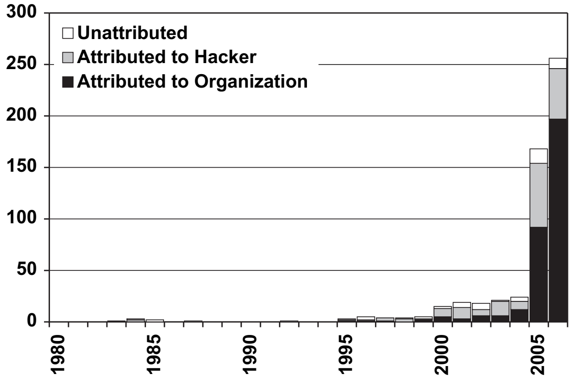

Figure 1 Hacker and organizational culpability in reported incidents of compromised records, 1980–2006

breach was unspecified and from those stories where the cause of the breach was related to organizational action or inaction. In this latter category, we consider organizational behavior to include four types of security breach: accidental exposure of personal records online, insider abuse or theft, missing or stolen hardware, or other administrative error.

First, it is noticeable that as more states require organizations to report compromised digital records, the overall volume of annual news stories on the topic increases significantly. In fact, there were more reported incidents in 2005 and 2006 than in the previous 25 years combined. We found 126 incidents of compromised records between 1980 and 2004, and 424 incidents between 2005 and 2006. Just summing the incidents from 2005 and 2006, when mandatory reporting legislation was in place in many states, we find that 68% of the stories concern data that were accidentally placed online or exposed through administrative errors, stolen equipment, or other security breaches such as employee loss of equipment or backup tapes.

Several factors might explain the pattern of increasing incidents and volume of compromised data over time. First, there is the possibility that the results are skewed due to the relative growth of new, fresh news stories devoted to this issue, and the loss of older stories that disappeared from news archives as time passed. Perhaps there have always been hundreds of incidents every year, but only in recent years has the severity of the problem been reported in the news. If this were the case, we would expect to see a gradually decaying pattern with greater numbers of reported cases in 2006 than in 2005, 2004, and so on. However, the dramatic difference in reported incidents between later years and early years suggests that this effect does not adequately explain our observations.

A second possibility is that increased media attention or sensational reporting in 2005 and 2006 led to a relative over-reporting of incidents compared with previous years. Literature on media responses to perceived crises or "moral panics" would suggest that a similar effect commonly accompanies issues that are granted a disproportionate amount of public attention, such as with the case of the mugging scare in Great Britain in the 1970s or the crackdown on the rave subculture in the 1990s (Critcher, 2003; Hall, Critcher, Jefferson, Clarke, & Robert, 1978). While it is unlikely that media outlets have exaggerated the amount of electronic personal record loss, it is possible that in previous years a certain number of events went unreported in the media due to lack of awareness or interest in the issue of identity theft.

A third possibility is that there were a greater number of reported incidents of data loss in 2005 and 2006 because institutions are maintaining and losing a larger quantity of electronic data, and because a changing legislative environment in many states is obliging institutions to report events publicly that may have gone unreported in previous years. The fourth possibility, and the most plausible one, is that mandatory reporting legislation has exposed both the severity of the problem and the common circumstances of organizational mismanagement.

It is likely that a combination of factors explain our observations. The Notification of Breach legislation that requires the prompt reporting of lost records in California came into effect in 2003. However, the legislation was not widely adopted and implemented by other states until 2005, which might help to explain the dramatic increase in reported cases. The Notification of Breach legislation in California, as in many other states, requires notification when a state resident has been a victim of data loss, regardless of where the offending institution resides. Therefore, institutions located in states without Notification of Breach laws, such as Oregon, are still required to report cases to victims who live in states that have enacted this type of legislation, such as New York. The nature and complexity of many databases means that in many cases, compromised databases are likely to contain information about residents who are protected by notification of breach legislation, thus increasing the total number of reported cases.

Conclusion

Surveying news reports of incidents of compromised personal records helps expose the diverse situations in which electronic personal records are stolen, lost, or mismanaged. More important, it allows us to separate incidents in which personal records have been compromised by outside hackers from incidents in which breaches were the result of an organizational lapse. Of course, organizations should be expected to perform due diligence and safeguard the digital records holding personal information from attack by malicious intruders. But often organizations are both the unwilling and unwitting victims of a malicious hacker. Through this study of reported incidents of compromised data, we found that two-fifths of the incidents over the last quarter century involve malicious hackers with criminal intent.

Surprisingly, however, the proportion of incident reports involving hackers is smaller than the proportion of incidents involving organizational action or inaction. While 31% of the incidents reported clearly identify a hacker as the culprit, 60% of the incidents involve missing or stolen hardware, insider abuse or theft, administrative error, or accidentally exposing data online. The remainder of the news stories record too little information about the breach to determine the cause: Either organizations or individual hackers might be to blame for some of these incidents.

Organizations probably can be blamed for the management practices that result in administrative errors, lost backup tapes, and data exposed online. And even though an organization can be the victim of theft by its employees, organizations might still be expected to develop suitable safeguards to ensure the safety of client, customer, or member data. Even using the news media's expansive definition of hacker as a basis for coding stories, we find that a large portion of the security breaches in the U.S. are due to various forms of organizational malfeasance.

One important outcome of the mandatory reporting legislation is improved information about the types of security breaches. Many of the news stories between 1980 and 2004 report paltry details, with sources being off the record and vague estimates of the severity of the security breach. Since reporting legislation became mandatory in many states, most news coverage provides more substantive details. In 2006, only 10 of the 257 news stories were unable to make some attribution of responsibility for a security breach.

Legislators at the federal and state level have adopted two main strategies to address the problem of electronic record management. On one hand, they have directly targeted those individuals (computer hackers) whose actions potentially threaten the security of private electronic data. The CFAA has been repeatedly strengthened in response to a perception that electronic data theft represents a material and growing concern. The fact that punishments for digital trespass now surpass those for many other more violent forms of crime suggests that federal legislators consider computer crime to constitute a serious threat to our personal and collective security. However, the data in this study suggest that malicious intrusion by hackers makes up only a portion of all reported cases, while other factors, including poor management practices by organizations themselves, contribute more to the problem.

The second strategy employed by regulators might be thought of as an indirect or "disciplinary" strategy. Notification of Breach legislation obliges institutions that manage electronic data to report any loss of that data to the individuals concerned. While this directly addresses the problem of consumer protection by empowering individuals to protect themselves in case of lost or stolen data, it has probably been intended to produce secondary effects. Companies and institutions, wary of both the negative publicity and the financial costs generated by an incident of data loss, are encouraged to adopt more responsible network administration practices. Similarly, end-users are urged to weigh both the risk of doing business electronically and the costs associated with taking action once they are notified of a potential breach. The practice of using a risk/reward calculus to achieve policy objectives through legislation has been termed governing "in the shadow of the law" by some authors in the critical legal studies and governmentality literature (Mnookin & Kornhauser, 1979; Rose, 1999).

One potential problem with this strategy is that the risks and rewards will be unequally distributed among various individual, state, and corporate actors. While a large corporation might possess the resources and technical skill necessary to encrypt data, secure networks, and hire external auditors, other institutions in the private or public sector may not find the risk of potential record loss worth the expenditure necessary to secure those data. Governing through this type of market discipline is likely to result in a wide spectrum of responses from differentially situated actors.

There are a number of alternatives open to lawmakers and policy advisors that could materially strengthen the security of electronic personal records in this country. Alternatives include setting stricter standards for information management, levying fines against institutions that violate information security standards, and mandating the encryption of all computerized personal data. However, the introduction of legislation to directly regulate institutions that handle electronic information would certainly be controversial. A wide variety of agencies, companies, and organizations manage personal records on a daily basis. This complexity would hinder the imposition of standardized practices such as encryption protocols. Corporations would probably balk at the prospect of having to pay fines or introduce expensive security measures and would accuse the government of heavy-handed interference. Others might argue that the imperatives of free-market capitalism demand that the government refrain from adopting punitive legislation, especially in order to maximize competitiveness. Identity theft can have a significant impact on individuals whose identity is stolen, and can taint the reputation of the organization that was compromised. But in the incidents studied here, the security breach is most often attributed to large corporations, and increasingly universities, rather than individual hackers.

Although computer hacking has been widely reframed as a criminal activity and has received increasingly harsh punishments, the legal response has obfuscated the responsibility of commercial, educational, government, medical, and military organizations for data security. The scale and scope of electronic record loss over the past decade would suggest that organizational self-regulation or self-monitoring is failing to keep people's personal records secure, and that the state has a more direct role to play in protecting personal information. State-level initiatives have helped expose the problem by making it possible to collect better data on the types of security breaches that are occurring and to make some judgments about who is responsible for the breaches. If public policy can be used to create incentives for organizations to manage personally identifiable information better and punish organizations for mismanagement, such initiatives would probably have to come at the state level. Electronically stored data might very well be weightless, but the organizations that retain personally identifiable information must shoulder more of the heavy burden for keeping such data secure.

Notes

[1] In practice, the monetary felony threshold has proved somewhat meaningless, since the value of computer code compromised during intrusion is often quoted well in excess of $5,000. In the case of *United States v. Mitnick* (9th Cir. 1998), Sun Microsystems claimed $80 million in damages related to the cost of research and

development of the source code that Mitnick copied during his intrusion ("United States v. Kevin Mitnick," 1998).

2 In our survey, some incidents involving U.S.-based organizations or U.S. citizens were reportedly carried out by individuals working outside the United States. For example, the 2001 theft of customer account information from Bloomberg Financial was carried out by a Kazak citizen named Oleg Zezov, who threatened to expose the information unless the company paid him $250,000.

3 The National Conference of State Legislatures maintains a website to track this kind of legislation, whether enacted or pending: http://www.ncsl.org/programs/lis/cip/priv/breach.htm.

4 We retained incidents that had been reported in multiple sources, even if no exact number of compromised records was reported. To be conservative, we recorded these incidents as having 0 compromised records. In news stories where it was only reported that "hundreds" or "thousands" of personal records were compromised, we recorded 100 or 1,000 compromised records.

5 The records lost by Acxiom Corporation consisted of credit card numbers, purchasing histories, and marital status of individuals.

6 We believe it is more likely that computer equipment is stolen for personal use or resale value, rather than for the data that thieves might suspect is on the hard drives of the equipment they steal.

7 This single case is illustrative of the challenge of compiling and comparing incidents of compromised personal records. For example, the Axciom incident involved an employee of Snipermail.com, who removed 8.2 gigabytes of personal data in 137 separate incidents between April 2002 and August 2003. To be consistent with our sampling, we record this as a single incident occurring in 2004, because the news coverage and his arrest did not occur until 2004. Axciom, the company that was entrusted with personal records, and even justice officials commenting on the case, describe the culprit as a hacker. However, there was actually a client relationship between the two firms, and Snipermail.com staff legitimately had the correct password to upload data to Axciom servers. Someone at the Snipermail.com firm guessed that the same password might also be used to download data, though Snipermail.com was not legitimately allowed to do so. Some might argue that this is an example of a poor security choice on the part of Axciom, rather than an example of an ingenious technical exploitation by a rogue outsider with a hacker's skills. However, the majority of cases we label as "insider abuse" involve employees. The culprit in this case did legitimately have some insider information about Axciom's security. To be conservative, and since we are interested in how the news media frames issues of data security, we code this incident as involving data stolen by a hacker, because that was the language used in news coverage; we do not code it as insider abuse, because the culprit was not an employee.

References

Burchell, G. (1996). Liberal government and techniques of the self. In A. Barry, T. Osbourne, & N. Rose (Eds.), *Foucault and Political Reason: Liberalism, Neoliberalism, and Rationalities of Government* (pp. 19–36). Chicago: University of Chicago Press.

Cavazos, E. A., & Morin, D. (1996). A new legal paradigm from cyberspace? The effect of the information age on the law. *Technology in Society*, **18**(3), 357–371.

Critcher, C. (2003). *Moral Panics and the Media*. Buckingham, UK: Open University Press.

Fox, S. (2000). *Trust and Privacy Online: Why Americans Want to Rewrite the Rules.* Washington, DC: Pew Internet and American Life Project.

Hall, S., Critcher, C., Jefferson, T., Clarke, J., & Robert, B. (1978). *Policing the Crisis: Mugging, the State, and Law and Order*. New York: Palgrave Macmillan.

Howard, P. N. (2002). Network ethnography and the hypermedia organization: New media, new organizations, new methods. *New Media & Society*, **4**(4), 550–574.

Howard, P. N. (2006). *New Media Campaigns and the Managed Citizen*. New York: Cambridge University Press.

Howard, P. N., Carr, J., & Milstein, T. J. (2005). Digital technology and the market for political surveillance. *Surveillance and Society*, **3**(1).

Jordan, T., & Taylor, P. (1998). A sociology of hackers. *The Sociological Review*, **46**(4), 757–781.

Jordan, T., & Taylor, P. (2004). *Hacktivism and Cyberwars: Rebels With a Cause?* New York: Routledge.

Milne, G. R., & Culnan, M. J. (2004). Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of Interactive Marketing*, **18**(3), 15–29.

Mnookin, R., & Kornhauser, L. (1979). Bargaining in the shadow of the law. *Yale Law Journal*, **88**, 950–968.

Nissenbaum, H. (2004). Hackers and the contested ontology of cyberspace. *New Media & Society*, **6**(2), 195–217.

Peck, J., & Tickell, A. (2002). Neoliberalizing space. *Anitpode*, **34**(3), 380–404.

Rose, N. (1999). *Powers of Freedom: Reframing Political Thought.* Cambridge: Cambridge University Press.

Samuelson, P. (2003). Digital rights management (and, or, vs.) the law. *Communications of the ACM*, **46**(4), 41–45.

Skibell, R. (2002). The myth of the computer hacker. *Information, Communication and Society*, **5**(3), 336–356.

Skibell, R. (2003). Cybercrimes & misdemeanors: A reevaluation of the computer fraud and abuse act. *Berkeley Technology Law Journal*, **18**(3), 909–944.

Thomas, D. (2002). *Hacker Culture*. Minneapolis: University of Minnesota Press.

United States v. Kevin Mitnick. (1998). 145 F.3d 1342 (9th Circuit).

United States v. Thomas. (1996). 74 F. 3d 701 (6th Circuit).

Universal City Studios, Inc. v. Reimerdes. (2000). 111 F.Supp.2d 294 320 (Federal Superior Court).

Wilson, M. (2003). Chips, bits and the law: An economic geography of Internet gambling. *Environment and Planning A*, **35**(7), 1245–1260.

Zook, M. (2003). Underground globalization: Mapping the space of flows of the Internet adult industry. *Environment and Planning A*, **35**(7), 1261–1286.

About the Authors

Kris Erickson is a doctoral candidate in Geography at the University of Washington, where he studies the role of computer subcultures in the regulation of cyberspace. His dissertation research on the hacker community explores the rise of the computer security profession and its implications for global Internet governance.
Address: Department of Geography, 408 Smith Hall, Box 353550, University of Washington, Seattle, Washington, 98195 USA

Philip N. Howard is an assistant professor in the Communication Department at the University of Washington. He researches the role of new media technologies in political communication in advanced democracies, and the role of new media technologies in solving social problems in developing countries. Currently, he directs the World Information Access Project (www.wiareport.org).
Address: Department of Communication, 141 Communications Building, Box 353740, University of Washington, Seattle, Washington, 98195 USA