Data Breaches in Europe:
Reported Breaches of Compromised Personal Records in Europe, 2005-2014

Philip N. Howard and Orsolya Gulyas

CMDS Working Paper 2014.1

Center for Media, Data and Society
School of Public Policy
Central European University

October 2014

Table of Contents

Tables

Figures

I. Executive Summary

A growing number of massive data breaches are degrading the personal privacy of people around the world. Data security and privacy policy are ongoing concerns in Europe. However, it can be difficult to assess privacy breaches in Europe, since many of the biggest incidents involve people and organizations from around the world. This working paper offers early descriptive statistics and analysis of the first cross-national, systematized event log of data breaches in Europe. The data is available for download at http://cmds.ceu.hu/.

Methodology. The sample frame includes major media news reports on compromised personal records and is unique for:

- sampling 28 European Union member countries, plus Norway and Switzerland;
- sampling from 2005 through the third quarter of 2014;
- sampling credible news sources in national languages;
- high-level social science standards for event database construction, with multiple sourcing, inter-coder reliability tests, recoding, and specific exclusion criteria.

Findings. A data breach is defined as any incident involving the loss or exposure of digital personal records. Personal records are defined as a) data containing privileged information about an individual that cannot be readily obtained through other public means, and b) information only known by an individual or by an organization under the terms of a confidentiality agreement. Preliminary analysis reveals that over the last decade:

- Some 229 data breach incidents involved the personal records of people in Europe. Globally, all these incidents resulted in the loss of some 645 million records, though not all of these breaches involved people in Europe. We confirmed 200 cases involving people in Europe, and 227 million records lost in Europe-specific breaches.

- The total population of the countries covered in this study is 524 million, and the total population of internet users in these countries is 409 million. Expressed in ratios, this means that for every 100 people in the study countries, 43 personal records have been compromised. For every 100 internet users in the study countries, 56 records have been compromised.

- 51 percent of all the breaches involved corporations and 89 percent of all the breached records were from compromised corporations. 41 percent of the incidents involved clear acts of theft by hackers, but 57 percent of the incidents involved organizational errors, insider abuse, or other internal mismanagement. 2 percent of the data breaches were unspecified.

- The sophistication and detail of journalistic coverage of privacy and personal data has increased, but is largely driven by “mandatory reporting” rules in particular countries. In other words, we know most about data leaks in countries where organizations are required to report that personal records have been compromised.

II. Reports of Data Breaches in Europe

1. Introduction

The internet, mobile phones, and a host of other new information technologies have allowed more and more people to conduct the business of their personal lives over digital media. And even people who are not heavy technology users are tracked, surveilled, and surveyed electronically, revealing facts about their attitudes, behaviors, and other life details. . Many of those activities, such as banking, shopping, e-government, social networking, and emailing, require disclosure of personal data. The data citizens or companies store online ranges from email or postal addresses, login information or passwords to sensitive personal information, including bank and credit card account information. As more activities take place online, more data is stored in servers. This situation poses challenges for maintaining privacy and data safety.

Almost everything we know about privacy violations in Europe comes from news reports from across Europe (expressed in various languages) of specific breaches. So, what can we learn from the collective coverage of the highest quality media reports?

Privacy policy and data protection are of concern to policymakers in Europe. Countries such as Germany, the United Kingdom, and Ireland are implementing strict rules for information management and data protection. Due to software vulnerability, mismanagement or human error data is frequently stolen or lost. Attempts have been made to measure the cost of data loss. However, few of these attempts have been systematic efforts to measure the scale of data breaches across the continent. The following paper provides new knowledge on the scale and quantity of data breaches, in line with Europe's unique values on privacy and surveillance.

Non-governmental organizations and data protection authorities acknowledge that there has been a steady increase in data breaches. Although Europe is moving towards a unified policy of data protection and requirements for reporting breaches, there is a dearth of information about exact cases and incidents. Not only are there few news accounts of big-picture trends in data breaches, public policy researchers have little comparative data to work with. Several years ago, another comparative event database revealed that 1.9 billion records were compromised between 1985 and 2006 in the United States. In 2006, this meant that for every hundred U.S. adults 875 personal records had been breached.[1]

For us, the lack of organized event records is both an empirical obstacle and an opportunity to generate new knowledge about data and privacy protection. This study investigates a decade of records to help assess both the changing volume and character of data breaches, and the way in which those breaches are reported to the public in Europe.

2. Methodology

To understand the trends in data breaches, we built an original event database of incidents as reported by credible, multilingual news outlets in Europe. The database we built and on which our findings are based includes all the cases reported on the internet in which personal data of European citizens—those of the 28 European Union Member States plus Norway and Switzerland—were compromised during the period 2005–2014. These incidents include data

[1] Erickson, Kris, and Philip N. Howard. (2007). "A case of mistaken identity? News accounts of hacker, consumer, and organizational responsibility for compromised digital records." *Journal of Computer-Mediated Communication* 12(4): 1229-1247.

breaches, leaks, and identity thefts, and cover only the last decade since few cases were reported prior to 2005.[2]

Since our study focuses on cases whose victims were European, instances in which European citizens compromised data of overseas citizens do not appear in our database, even though we found dozens of such examples. Although they are privacy violations, we also did not include cases of surveillance, as we are concerned with how existing data is handled rather than how data is collected. The database, moreover, does not include instances of non-personal data breaches. Neither were we concerned with hacker attacks that attempted to steal money (e.g., transferring money from one bank account to another) but otherwise did not compromise personal data. Incidents involving phishing or malware were excluded because their overall impact was impossible to estimate. We were not concerned with privacy breaches of paper-based records.

We were concerned with the compromised personal records of only the countries we were studying. There were four cases of breaches on embassies that were excluded from this analysis because they were the embassies of governments outside Europe and the data lost involved nationals from outside Europe. If a news report included details about the kinds of documents compromised or the number of gigabytes of data put at risk, the case was coded but excluded from analysis. Only cases with personally identifiable information of people living in the European Union, Norway and Switzerland were analyzed. Given the diversity of reporting styles and variations in information completeness, we implemented the standard rubric for codifying and quantifying news reports. If the news reported "thousands" then the value 3,000 was recorded. If the news reported "tens of thousands" then the value 30,000 was recorded. Broad, nonspecific reports of phishing and malware attacks were excluded. Reports about the compromised personal records of European Union citizens by the national security agencies of other countries (such as the United States) were also excluded. Second hand reports and unsourced reports were excluded.

We looked for reports of compromised personal data on specialized websites, the LexisNexis Academic database, and the Google News archive. Appendix A identifies the coding variables, and Appendix B identifies the list of specialized websites consulted. Appendix C presents a country-by-country breakdown of incidents. We also utilized the language skills of the project team[3] by having researches conduct Google searches using relevant terms, a strategy that yielded information on specialized national and news websites. For the rest of the countries in our sample we enlisted the help of Google Translate to conduct domain name searches. . As much as possible, we relied on credible news reports. We did not use sources that contained aggregated data because of the possibility of overlaps among them and different conceptualizations of what constitutes a data breach. We also aimed to find more than one source for each of the cases. Overall, the research team of eleven people spent 450 hours identifying and evaluating data-breach reports. As is common with event datasets, a standard set of incident descriptors was devised. They include:

[2] The countries included in this case list include: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, and the United Kingdom.

[3] The language-specific searches include Bulgarian, Croatian, Czech, Dutch, English, Estonian, Finnish, French, German, Hungarian, Italian, Latvian, Lithuanian, Polish, Slovakian, Slovenian, and Romanian.

- when a data breach was reported,
- when it occurred,
- which country and organization was affected by it,
- to which sector the affected organization belongs,
- how strong was its impact (measured as the total number of people, records, gigabytes, or emails compromised);
- what kind of data was compromised;
- whether the data was stolen, mismanaged, or lost; and exactly what kind of breach happened.

In the cases of hackings, we identify the country of the attackers when known. Moreover, we classified instances of whistleblowing.

Every case had its own complications. An example is the case of German tax authorities obtaining records about German nationals from an employee of a Swiss bank. From the German perspective, such an event may be seen as a case of whistleblowing; from the Swiss perspective, it involves illegally compromised data. Assessing the impact of cases involving multiple countries was another challenge. Through regular coder training sessions, inter-coder reliability scores, and consulting multiple sources, we were able to fact check incidents in a variety of ways, and uncovered dozens of cases, in which malware is known by security experts to have compromised personal records. While we have been able to count the number of these cases, we have no way to evaluate their impact on individual privacy.

In addition, many kinds of cases were excluded from this analysis for methodological reasons. If too many details were unknown or the sourcing was questionable, the case was not included. Many dramatic cases simply did not qualify. Breaches of Bitcoin wallets, print files, browsing histories, unspecified forms of data, laptop computers, USB sticks, contact lists, call histories, and photo archives were not included unless personally identifiable information had been compromised. In one case, the Dutch government's annual budget was stolen before its release. In another, an attack of the dating website www.beautifulpeople.com allowed "ugly people" to be admitted. In one dramatic breach, a large value of carbon credits were stolen, but no personal information was lost. While many of these cases are interesting examples, only those breaches meeting our strict criteria were included.

3. Findings: Descriptive Statistics

Table 1 identifies some basic descriptive statistics that reveal both the nature of this event dataset and some important trends. All in all, there were 229 reported incidents in which the personal records of at least a few people in Europe were breached. Over all, around 641 million email addresses, names, passwords and other kinds of personally identifiable information were compromised, though most reports do not specifically identify the proportion of victims residing in Europe. Many reports, however, list countries in which there were known victims. European states were specifically identified 267 times in the event dataset.

The regulatory environment defines transparency. The stricter the reporting requirements in a given country, the greater the number of cases identified and breaches described. Also, many cases are known because organizations are obliged to report data breaches in some European countries. .

Table 1: Quick Fact Table	Values
Total Number of Breaches Involving European Targets	229
Total Volume of Breached Records Across All Incidents	641,979,541
Number of Times a Specific Country in Europe Was Identified as Target	267
Number of Global Breaches Involving European Targets	29
Volume of Records From Global Breaches that Impact People in Europe	415,012,618
Volume of Records From Europe-Specific Breaches	226,966,923
Total Number of Breaches Involving European Targets	229
Total Volume of Breached Records Across All Incidents	641,979,541
Number of Times a Specific Country in Europe Was Identified as Target	267
Number of Global Breaches Involving European Targets	29
Volume of Records From Global Breaches that Impact People in Europe	415,012,618
Number of People Living in Study Countries	523,730,791
Number of Internet Users in Study Countries	408,583,658
Volume of Records from Europe-Specific Breaches per 100 People	43
Volume of Records from Europe-Specific Breaches per 100 Internet Users	56
Number of All Breaches in Which Attacker Was Unspecified	48
Number of All Breaches in Which There was No Attacker	116
Number of All Breaches in Which The Attacker Was Known and Specified	65
Number of All Breaches in Attacks Originating in the EU	45
Number of Cases in Attacks Originating in the UK	10
Number of All Breaches	274
Number of All Breaches in Which There Was an Attacker, The Attacker Was Specified, and the Attacker Was Not in Europe	20
Percent of All Breaches in Which Attacker Was Known As Being In Europe	69
Percent of Breaches Revealed by a Whistleblower	2
Percent of Breaches in Which Attacks Originated in the UK	15
Percent of Europe-Specific Breaches Originating in the UK	24
Percent of All Breaches Involving Corporations	51
Percent of All Breaches Involving Hackers	41
Percent of All Breaches Involving Insider Abuse, Missing Hardware, Accidental Exposure Online, or Administrative Error	57

Some breaches have an impact on people around the world. There were 29 of these global breaches, many of which involved major credit card companies or data mining firms that are incorporated in the United States but store the data of people living in Europe. These global breaches accounted for 415 million personal records, though again news reports rarely identified the volume of breaches impacting people in Europe. This does mean, however, that breaches affected 226 million records in Europe.

Since there are 524 million people living in the study countries and 227 million compromised records, the ratio of compromised records to people is 43 to 100. Since there are 409 million internet users living in the study countries, the ratio of compromised records to internet users is 56 to 100. On the basis of news reports, however, little is known about the distribution of privacy violations according to race, gender, class, or other social categories. More is known about the distribution of security breaches by country.

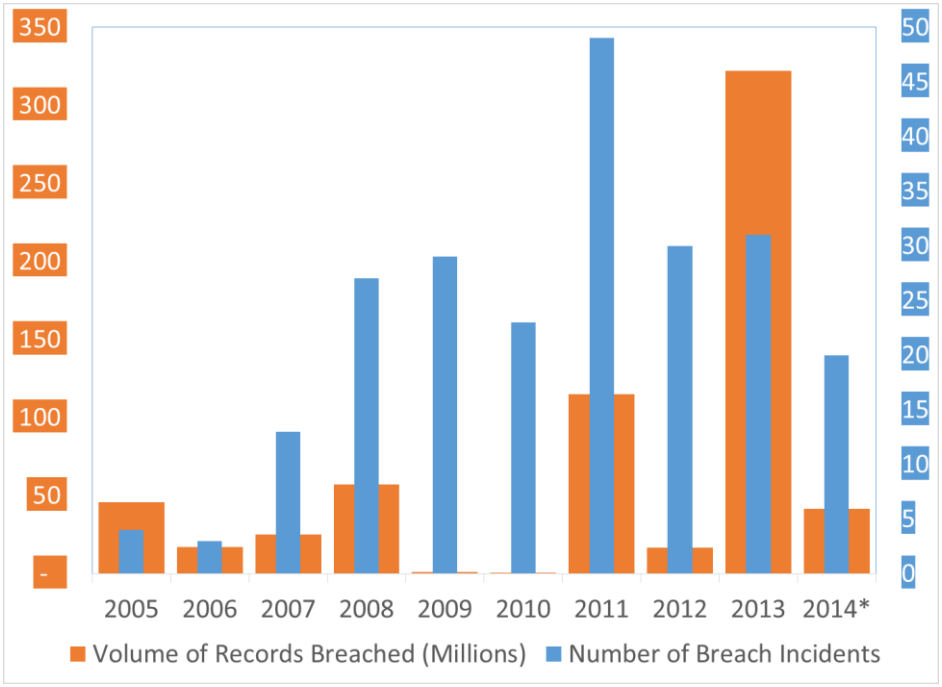
Some news reports provided additional information allowing us to code for trends across all reports. All the incidents in this event dataset involve targets residing in Europe. While some news reports provided s comparable data and context, the scope and quality of information in these reports reported. Many of the incidents resulted from organizational error. Other

involved an external attack, but the identity and location of the attacker may not have been known. Among all the cases—both global and Europe-specific—in which the breach of personal records involved an attack, 15 percent of the cases were launched from the United Kingdom. Among the cases in which the targets resided in Europe, 24 percent originated in the United Kingdom.

This study did not analyze incidents of privacy violations by government security agencies around the world, such as cases revealed by whistleblowers Chelsea Manning and Edward Snowden. Our research revealed that only 2 percent of the reported incidents involved a whistleblower.

The last decade has seen a significant increase in breaches. There are three reasons for this trend. While it is true that more and more people have been putting personal information online, journalists have become better at reporting about breach incidents, and more and more governments have required that the victims of breaches be informed.

Figure 1: Volume and Number of Breach Incidents, 2005-2014



*Inclusive to Third Quarter 2014

Which Countries in Europe Are Impacted the Most?

Table 2 identifies the countries in Europe with the most personal information breaches over the last decade. Germany, Greece, Netherlands, Norway and the United Kingdom have had unusually high numbers of incidents, and large volumes of records breached.

Table 2: Severity of Breach Patterns, Top 5 Country Targets	Compromised Records Per 100 People	Compromised Records Per 100 Internet Users
Germany	68	79
Greece	81	140
Netherlands	23	24
Norway	80	83

United Kingdom	220	245

According to the best publicly available data on incidents in Europe, organizations in these countries are doing a poor job managing personal information, and are targets for cybercrime. Again, much of what we know is shaped by reporting requirements. It is likely that the media's coverage of data breaches has improved over time, but does not include all incidents. Nonetheless, the per capita trends across countries suggest that as a national average, the ratio of compromised records to people in the UK is 2:1. The ratio of compromised records to internet users is more than 1:1 in Greece and 1:2 in the UK.

Several countries, including Croatia, Estonia, and Slovenia, had no reported incidents of citizens losing data. We expect that some citizens in these countries have been impacted by the large global breaches, but that their absence from the event dataset can be explained by our sampling frame. Either journalists did not specifically mention these countries as being the targets of attack or the incident reports from these countries were of dubious quality.

a) How and Why Personal Records Are Compromised

To understand how and why personal records were being compromised we came up with two coding schemes. The first was a simple, three-category typology of breaches. Each case was coded for whether the data was stolen, lost, or mismanaged (exposed online or mismanaged in an organizational accident). **Table 3** demonstrates that by this typology, some 57 percent of incidents involved theft, and 570 million records were stolen over the last 10 years.

Table 3: Type of Loss, Three Categories	By Number of Incidents	Percent	By Number of Records	Percent
Lost	20	9	34,980,276	5
Mismanagement	77	34	36,751,944	6
Stolen	131	57	570,079,321	89
TOTAL	228	100	641,811,541	100

Many of the reports of stolen data, upon further inspection, were cases in which a disgruntled employee or company insider stole the data. In these cases, a large part of the story involved security issues. So a more nuanced six-category typology was developed, often using the same keywords used by technology reporters and the organizations revealing a breach. The coding system in **Table 4** reveals that while 42 percent of the cases clearly involved external attack by criminal hackers, the majority of cases involved problems internal to the organization: insider abuse or theft, hardware that the organization either lost track of or lost to theft, and administrative errors. Some organizations mistakenly uploaded personal records online.

Table 4: Type of Loss, Six Categories	By Number of Incidents	Percent	By Number of Records	Percent
Administrative Error	22	10	33,171,867	5
Exposed Online	49	22	2,381,386	0
Insider Abuse or Theft	25	11	12,150,489	2
Missing or Stolen Hardware	29	13	37,273,276	6
Stolen - Hacker	94	42	556,106,552	87
Unspecified	4	2	656,413	0
TOTAL	223	100	41,739,983	100

b) The Organizations Breached Most Often

Breaches impact different kinds of organizations. Not all are commercial firms, but most are. **Table 5** reveals that businesses accounted for the vast majority (89 percent) of the organizations reporting a breach, losing some 538 million personal records. Government offices were the next largest target. Almost a quarter of the incidents involved public agencies, but these breaches tended to be much smaller in scale. All in all, only a few of the breaches involved non-profit groups, the military, medical facilities, or educational organizations

Table 5: Data Breach by Type of Organization Compromised	By Number of Incidents	Percent	By Number of Records	Percent
Commercial	117	51	538,349,868	89
Educational	11	5	80,221	0
Government	55	24	59,173,346	10
Medical	18	8	9,337,197	2
Military	8	3	917,001	0
Non-profit	12	5	1,818,765	0
Unknown	8	3	32,303,143	5
TOTAL	229	100	06,940,632	100

4. Findings: Europe-wide Comparisons

Across the dataset, we identified two transnational phenomena involving particular types of attackers and targets.

a) Attacks by Anonymous

Anonymous is an international "hacktivist" network. The group conducts disruptive online activism, in the form of cyber attacks on many kinds of individuals and organizations. Somewhere between an affiliation of clubs and a network of makeshift teams, *Anonymous* chapters in various European countries have breached numerous personal-record collections. The targets of *Anonymous-Spain* has included Sony Corps, Spanish banks, governments and other actors with devastating results such as the theft of 77 million account details from the Sony network. Three Spanish nationals were ultimately arrested for this crime. *Anonymous-Spain* also leaked 5 gigabytes of financial documents of the People's Party, Spain's largest right-wing political organization.

Anonymous-Sweden hacked into the official website of Sweden's National Board of Health and Welfare in retaliation for a police raid of the office of the Stockholm-based web-hosting company PRQ. *Anonymous-United Kingdom* revealed thousands of British email addresses and encrypted passwords, including those of defense, intelligence and police officials as well as politicians and NATO advisers.

Among the huge database of private information exposed by self-styled "hacktivists" are the details of 221 British military officials and 242 NATO staff.

Another incident involved *Anonymous-Italy*, a group that hacked the system of the National Anti-Crime Computer Centre for Critical Infrastructure Protection (CNAIPIC - *Centro Nazionale Anticrimine Informatico per la Protezione delle Infrastrutture Critiche*), stealing 8 gigabytes of confidential documents. *Anonymous-Italy* also exposed a public figure of crimes. Roman Catholic cleric Don Giacomo Ruggeri was suspended of his duties and arrested under charges of child abuse by the Italian police after his personal emails were made public.

b) Breaches of Crypto-Currencies

A recent and serious type of data leaks have been hacker attacks against crypto-currencies. These are open-source peer-to-peer digital currencies, which use public ledgers to keep track of transactions and wallet balances. While the protocols, protected by cryptography, are resilient against attacks, their services are often vulnerable.

The primary goal of crypto-currency hacks is monetary (digital coin) theft, but data leaks inevitably occur as a direct or indirect result of these intrusions. Given that transaction ledgers are public, crypto-currencies are pseudonymous rather than anonymous. If a hacker gets hold of coins associated with a particular wallet, he or she is then able to track the transactions back through the ledger. Many public addresses (or wallets) can be linked to an individual or entity. Therefore, ledgers may yield valuable intelligence about or even the identity of its owner. But the crypto-currency pipeline flows both ways. When money is stolen, it can often be traced back to the hacker, as long as the attacker does not use money-laundering services, (which admittedly are difficult to employ for large amounts of coins.)

Because of the pseudonymous nature of crypto-currency networks and because this economy operates underground in what is commonly called the “deep web”, it is often difficult to determine the geographic origin or extent of an attack. Listed below is an overview of several notable cases of bitcoin thefts over the past three years. (Note: several of these cases were not part of the dataset analyzed for this report.)

The first notable crypto-currency theft linked to Europe was the Bitomat.pl hack in June 2011. The Polish exchange, which was the third largest at that time, lost approximately 17,000 bitcoins worth over €150,000. In April 2013, an unknown attacker managed to reset the password of the French exchange, Bitcoin Central, through its hosting provider's web interface which effectively locked it out of its own site. As a result, the attacker, after initiating a server reboot, successfully stole hundred bitcoins worth tens of thousands of euros.

In August 2013, the UK-based bitcoin wallet Blockchain.info lost 50 bitcoins worth thousands of euros after an attacker exploited vulnerability in the JavaScript random number generator. In October 2013, a group calling themselves “The Hole Seekers” attacked the online forum BitcoinTalk. While the full extent of the attack is not known, the hackers may have gained access to a database containing user information, including passwords.

A November 2013 attack on the bitcoin payment processor BIPS resulted in the theft of €750,000. Thieves hacked into Poland-based Picostocks in late November 2013 and made off with 6,000 bitcoins worth approximately €4.5m. During the same period, approximately 4000 wallets within the Czech exchange Bitcash.cz lost 480 bitcoins, amounting to a theft of roughly €74,000. (Experts believe attackers hacked into Bitcash.cz through its web interface.) Another Poland-based company—Bidextreme.pl—was attacked in the same month with hackers stealing an undisclosed amount of bitcoins. In early February 2014, thieves attacked bitcoin hardware manufacturer Cointerra, gaining access to email records and leaving customers vulnerable to phishing attacks. Poland's leading bitcoin exchange, Bitcurex, was hacked in March 2014, but only a small portion of its operational or “hot-wallet” balance was stolen.

Lastly, the deep web service known as Sheep Marketplace with servers reportedly based in the Czech Republic lost 96,000 bitcoins worth €165m in an attack, which is now considered the largest theft in bitcoin history. It is not clear whether this attack was masterminded by hackers or insiders within the company.

5. Findings: Unusual Cases Listed by Country

There are unusual examples of data breach, where information was lost or published in an unorthodox way. One example occurred in Denmark where a PowerPoint Presentation that included the personal information of HIV patients was published online. Another incident happened in the UK when a staff member of an educational institution lost a camera containing sensitive information, namely photographs of the passports of job applicants. Another case took place before the 2011 Bulgarian elections when the Ministry of Foreign Affairs accidentally published online the names and addresses of the permanent residences of Bulgarian nationals living abroad, making these citizens vulnerable to theft and burglary. In another incident, the sports organization, FC Manchester City, initiated investigations against a rival club for allegedly hacking confidential records of its players.

While the cases of data breaches are many, several unusual country stand out as being particularly egregious, or indicative of trends. These cases include:

a) Belgium

Hackers in Belgium have reportedly been responsible for attacks against foreign embassies of both European and non-European countries, and for illegally obtaining private data from international organizations.

b) Bulgaria

Few incidents of data leaks have been identified in Bulgaria. However, this country has reported widespread identity theft and misuse of personal data. In 2011, the head of Bulgaria's Computer Crimes and Intellectual Property Department of the Ministry of Interior's Chief Directorate for Combating Organized Crime stated that these crimes doubled in Bulgaria between 2006 and 2010. The number of Bulgarians using the Internet also increased during this period.

c) Czech Republic

. In November 2011, an administrative error exposed records of 893 Roma recipients of a governmental stipend for studies and vocational training. The dataset included the names of the students as well as the amount each was awarded, an incident that stirred both publicity and controversy given the high levels of anti-Roma discrimination that reportedly exists in the Czech Republic.

d) France

In 2012, the American Chamber of Commerce in France was attacked by a hacker collective known as DeleteSec. As a result, hundreds of email addresses and passwords were compromised. Before the intrusion, the hackers claimed they warned the Chamber about an SQL injection error, informing them of a possible security threat before breaking into the system. According to DeleteSec, the Chamber not only ignored the warning it responded with a rude message. This case illustrates an important lesson to be learned: overconfidence can be costly in the digital era.

e) Netherlands

Condoms will protect a user during intimacy, but sadly not online. In 2009, a Dutch website that enabled young people to order condoms free of charge proved vulnerable. Customer data (name and address) of every young person who made an order could be easily found out on line. This data leak affected about 10,000 people.

f) Italy

Data breaches by hackers amounted to the majority of cases in Italy. One incident involved personal data theft from Sony Italy carried out by a Turkish hacker group called Turkish Ajan, which leaked confidential personal records of Sony customers to the public. The other four cases involving hacking were linked to the Italian wing of Anonymous, which announced its intention to demonstrate the weak digital security measures taken by government bodies. In 2011, they attacked the Italian police on two occasions stealing login information to the police computer system.

Anonymous launched a larger operation in 2012 and stole about 3500 records from the state police reports, mobile phone numbers, personal emails, information on salaries, and even soft-porn pictures were found in the compromised dataset.

g) Slovak Republic

In 2003, Orange, the Slovak telecommunication company, had problems securing the privacy of its customers. Almost 900,000 personal records of subscribers to the Orange telecom company were exposed online, including phone numbers (both listed and unlisted), names, addresses and birth dates. Interestingly, Orange France compromised over 2 million personal data records in 2014.

h) Ireland

The Irish Department of Social and Family Affairs lost up to 400,000 records between 1985 and 2014. Some of these incidents involved stolen laptops, while others were the result of insider abuse. The breaches included the loss of sensitive personal data such as social security numbers.

i) United Kingdom

Due to strict legislation against data leaks and careful monitoring, the UK has uncovered and reported an enormous number of data breach cases, both paper-based and digital. A large number of breaches were attributed to carelessness, by either the owners or handlers of records. But most cases involve administrative errors and mismanagement, such as not erasing the hard drives of old computers offered for re-sale.

Despite the extra care with which medical records are treated, dramatic cases of data breaches involving confidential medical information were reported. In London, a private clinic contracted a tech company to computerize its patient records, which included confidential details of the patients' conditions, identities, addresses and dates of birth. However, after scanning the documents, this UK tech company sub-contracted other facets of the project to a company in India where local employees offered the records for sale, primarily to insurance companies and marketing agencies specializing in health products. Hundreds of thousands of personal medical records of UK citizens have been leaked to Indian companies in this way, even though under the UK Data Protection Act, it is illegal to send such documents outside the EU unless appropriate security is guaranteed. This case is a good example of the difficulty of classifying information of cross-sector breaches for the purposes of research.

Another example of such breaches involved the Surrey and Borders Partnership NHS Foundation Trust, an organization that provides care and services in the areas of mental health, drug and alcohol abuse, learning disabilities, and so on. Consequently, it is difficult to define the organization's sector, as it is both educational and medical. For this research, the

organization's sector was determined on a case-by-case basis with many cases involving multiple levels of responsibility and many different organizations.

6. Conclusion: Moving Forward with Mandatory Reporting

Public policy oversight of personal records is evolving quickly. These preliminary findings demonstrate the size and complexity of the problem, and the positive value of mandatory reporting, for both public awareness and policy making. In March 2014, the European Parliament voted to support a new General Data Protection Regulation that created a complicated and strict legal framework for processing personal data. The decision to back the renewed data protection plan was triggered by a number of high-profile incidents of personal data loss across Europe, which made the question of secure handling of personal information a priority.

In 2009, the EU made a breach notification law as part of its Directive on Privacy and Electronic Communications, or the E-Privacy Directive. The directive was to be implemented as law in the member states by May 2011. A new regulation on mandatory personal data breach disclosures came into force in August 2013, building on the provisions laid out in the E-Privacy Directive. According to the regulation, telecom operators and Internet service providers are obliged to notify national authorities of "any theft, loss or unauthorized access to personal customer data, including emails, calling data and IP addresses. Details concerning any incident, including the timing and circumstances of the breach, nature and content of the data involved, and likely consequences of the breach, must be reported."

In addition, the report must be made within 24 hours of the detection of the incident or within three days of being alerted to the breach. Telecommunications firms and Internet service providers must also report on the measures they took to address the breach. Each company must also directly inform all subscribers whose data may have been compromised. One aim of this legislation is to incentivize companies to improve the encryption and security of personal data. If companies comply with specific security measures and recommendations made by the European Commission, they may be exempt from having to report all data breaches.

These regulations have been criticized for the burden it would place on authorities due to the expected high number of breach notifications. Companies have suggested that regulations should also involve categorizing data breaches according to the level of security risk they pose, in order to avoid 'notification fatigue' both for clients and operators. Furthermore, operators wish to control the communication of data breaches to their clients as to avoid negative impacts on their brand as much as possible.

In January 2014, the EU's Justice Commissioner called for larger fines against companies breaching European data privacy laws. New proposals currently under debate in the European Parliament involve the establishment of a single EU regulator with the authority to issue fines.

The new legal framework for European data protection is likely to go through more changes as negotiations continue between the European Commission, European Parliament and the Council of Ministers. The regulations, which include progressive privacy measures, such as limits on 'profiling' requirements, using clear and plain language in privacy policies, obtaining the explicit consent of data subjects on processing any form of their personal data, etc. are not likely to come into effect before 2016.

Currently, the UK has a mandatory reporting requirement in place for organizations such as telecoms and Internet providers "who provide a service allowing members of the public to send electronic messages." They are required to notify the Information Commissioner's Office (ICO), an independent authority dealing with information rights, within 24 hours of uncovering a data breach. Other types of companies are not required by law to report any incidents. However, the ICO has established reporting data breaches as a 'best practice,' and has made it clear that it will take unreported data breach cases extremely seriously.

The Netherlands is planning to pass its own mandatory reporting bill. However, the latest amendment to the proposal made in April 2014 would only require data breaches to be reported when the breach has led to seriously adverse consequences. Such wording is not only vague, it also erodes the purpose of mandatory reporting, which is to compel companies to prevent data breaches through the implementation of rigorous security measures.

According to the EU's new data breach regulation that came into effect in August 2013, providers of publicly available electronic communication services in all member states—including telecommunications firms and Internet service providers—must notify national authorities of data breach cases within 24 hours of an incident. However, how seriously each member state takes this otherwise binding regulation may vary. Most European countries have consumer data protection offices with websites providing information on how individuals and organizations can protect themselves. But not every country has established an archive of data-breach incidents, or a clear system for reporting these incidents.

The cases listed in this report reveal how nuanced, complex, and diverse data-breach scenarios can be, and illustrate the difficulty in making good policies that reflect the complicated nature of the issue without posing limitations—legal or otherwise—on our use of Internet Communication Technology and curbing the many advantages they offer in everyday life.

III. About the Project

1. Correspondence

Please direct correspondence to Philip N. Howard, Director, Center for Media, Data and Society, Central European University, Nador 9, Budapest, 1051, Hungary, howardp@ceu.hu, @pnhoward.

2. About the Authors

Philip N. Howard is director of the Center for Media, Data and Society and a professor in the School of Public Policy at Central European University. He is also a professor at the University of Washington and a fellow at the Tow Center for Digital Journalism at Columbia University. He is the author, most recently, of *Democracy's Fourth Wave? Digital Media and the Arab Spring*. Currently, he is writing *Pax Technica: How the Internet of Things May Set Us Free or Lock Us Up*, a book about the future of global information politics for Yale University Press. He blogs at http://philhoward.org and tweets from @pnhoward.

3. Research Team

This research was conducted by the Spring Media Practicum at the School of Public Policy of Central European University: Gulnara Alimbayeva, Roxana Damian, Tamilla Dauletbayeva, Orsolya Gulyas, Zintis Hermansons, Tautvydas Juskauskas, Attila Mester, Róbert Papp, Radka Pudilova, Marija Stojanovska Rupcic.

IV. Institutions and Funding

1. The Center for Media, Data and Society

The Center for Media, Data and Society is the leading center of research on media, communication, and information policy in Central and Eastern Europe. Based in the School of Public Policy at Central European University, CMDS produces scholarly and practice-oriented research addressing academic, policy and civil society needs. CMDS research and activities address media and communication policy, social media and free expression, civil society and participation, fundamental communication and informational rights, and the complexities of media and communication in transition.

2. The School of Public Policy

The School of Public Policy (SPP) at Central European University, in the words of CEU's founder, George Soros, is a "new kind of global institution dealing with global problems" through multi-disciplinary study of public policy, innovative teaching and research, as well as meaningful engagement with policy practice.

3. Central European University

Central European University (CEU) is a graduate-level, English-language university accredited in the U.S. and Hungary and located in Budapest. The university offers degrees in the social sciences, humanities, law, public policy, business management, environmental science, and mathematics. CEU has more than 1,500 students from 100 countries and 300 faculty members from more than 30 countries.

V. Appendix A: Case and Variable Definitions

Variable Name	Definition
Administrative error	Accidentally disclosing private data, for example by misplacing hardware, or by selling hardware that had not been wiped of identifiable information.
Attacker country	Location from which the breach originated. Country-to-individual or individual-to-country cases have been taken into account; country-to-country and government-sponsored attacks on other governments are not included.
Compromised records	Collections of electronic personal records that have been breached by third parties through illegal or negligent acts. The cases where data is sold to third parties for marketing purposes without users' informed consent are not taken into consideration as compromised records.
Data exposed online	Personal records are made accessible either by publishing online, software error or accidental disclosure.
Electronic personal records	Data containing privileged information about an individual that cannot be readily obtained through other public means; this information is only known by an individual or by an organization under the terms of a confidentiality agreement. Examples include individual personal credit histories, credit card numbers, account numbers, medical records, social security numbers, grades earned in school.
Incident (list of incidents)	A case where one or more electronic personal records were compromised through negligence or theft.
Hacker	Intruder deemed responsible for compromising records.
Mismanagement	Exposing private records online, leaking data due to administrative error or using data for activities not related to the work of the organization.
Phishing	Cases where victims are deceived into voluntarily revealing their personal information.
Security breach in an organization	Accidental exposure of personal records online, inside abuse or theft, missing or stolen hardware, administrative error.
Target country	The country of residence for the people who had personal records compromised.
Unknown (reference to type of compromise of the records)	A case where an estimation of the compromise has not yet been made or it is impossible to be made.
Unspecified (reference to type of compromise of the records)	A case of unwillingness to disclose information about the type of compromise occurred.
Whistleblower	A person or network who discloses alleged wrongdoing or illegal activity occurring in an organization like law or rule violation, fraud, health and safety violations and corruption. The whistleblower attribute was reserved for cases where the source of the breach was described as serving the public interest.

VI. Appendix B: Sources

Specialized Databases
http://www.databreachtoday.eu/
http://datalossdb.org/
http://www.dataprotection.ie/
https://www.huntonprivacyblog.com/archives/
http://ico.org.uk/news/latest_news
http://www.infosecurity-magazine.com/
http://www.insideprivacy.com/data-security/data-breaches/
http://nakedsecurity.sophos.com/
http://www.pogowasright.org/
http://www.privacy-europe.com/blog/
http://www.scmagazineuk.com/
http://seclists.org/
http://thehackernews.com/

VII. Appendix C: Country-Specific Breaches

Country	Population	Internet users	Number of Breaches Involving Each Country	Volume of Breaches Exclusively Involving That Country	Records Per Person	Records Per Internet User	Breaches Originating In This Country
Austria	8,526,429	7,135,168	9	683,731	8.02	9.58	2
Belgium	11,144,420	9,441,116	4	9,700	0.09	0.10	1
Bulgaria	7,167,998	4,083,950	5	64,678	0.90	1.58	0
Croatia	4,272,044	2,780,534	0	-	0.00	0.00	0
Cyprus	1,153,058	726,663	1	-	0.00	0.00	0
Czech Republic	10,740,468	8,322,168	8	159,538	1.49	1.92	1
Denmark	5,640,184	5,419,113	6	32	0.00	0.00	1
Estonia	1,283,771	1,047,772	0	-	0.00	0.00	1
Finland	5,443,497	5,117,660	7	428,300	7.87	8.37	1
France	64,641,279	55,429,382	15	2,782,428	4.30	5.02	1
Germany	82,652,256	71,727,551	28	56,422,711	68.27	78.66	3
Greece	11,128,404	6,438,325	4	9,016,885	81.03	140.05	1
Hungary	9,933,173	7,388,776	2	55,146	0.56	0.75	1
Ireland	4,677,340	3,817,491	12	916,934	19.60	24.02	1
Italy	61,070,224	36,593,969	7	74,601	0.12	0.20	3
Latvia	2,041,111	1,560,452	2	3,500	0.17	0.22	0
Lithuania	3,008,287	2,113,393	3	107,475	3.57	5.09	0
Luxembourg	536,761	510,177	1	-	0.00	0.00	0
Malta	430,146	173,003	1	-	0.00	0.00	0
Netherlands	16,802,463	16,143,879	31	3,868,446	23.02	23.96	2
Norway	5,091,924	4,895,885	6	4,060,032	79.73	82.93	3
Poland	38,220,543	25,666,238	8	787,066	2.06	3.07	2
Portugal	10,610,304	7,015,519	3	657	0.01	0.01	0
Romania	21,640,168	11,178,477	3	5,000	0.02	0.04	2
Slovakia	5,454,154	4,507,849	6	2,401	0.04	0.05	0
Slovenia	2,075,592	1,501,039	0	-	0.00	0.00	0
Spain	47,066,402	35,010,273	9	15,444	0.03	0.04	3
Sweden	9,631,261	8,581,261	6	90,250	0.94	1.05	2
Switzerland	8,157,896	7,180,749	3	1,000	0.01	0.01	1
United Kingdom	63,489,234	57,075,826	77	139,666,768	219.98	244.70	10

Howard, P. (2014). Data Breaches in Europe: An Analysis of Reported Breaches of Compromised Personal Records in Europe. *Center for Media, Data and Society Central European University. Working Paper 2014.1*. 24 pp. Budapest, Hungary. Retrieved from cmds.ceu.hu.