

The Impact of Data Disclosure

By Jason Andress – ISSA member, Colorado Springs, USA Chapter

This article discusses the mechanisms through which data is disclosed and exploited, both on an organizational and on an individual level, and covers steps that might be taken to mitigate such events.

Abstract

After years of research by security professionals, the challenge of unauthorized or unintended data disclosure continues to grow. Further complicating this issue is the criminal element, ready to exploit the disclosed data for fraudulent monetary gain. This article discusses the mechanisms through which data is disclosed and the exploitation of this data, both on an organizational and on an individual level. Also covered are steps that might be taken to mitigate such events.

When the term “identity theft” first appeared in the public eye in 1991,¹ it connoted that some individual’s data had been exposed in some fashion and that someone had exploited that exposure to nefarious ends, such as fraudulently acquiring financial accounts. In recent years, this phrase has been used to refer to an ever-expanding category of events, including exposures that do not lead to exploitation, a very common occurrence at present.

A variety of mechanisms can disclose data either intentionally or otherwise. For example, the loss by a credit card company of a shipment of backup tapes containing customer data is considered an unintentional data disclosure through loss (whether or not it is proven that someone was able to read those backup tapes). Subsequent malicious actions by an exploiter, such as the opening of fraudulent accounts in the name of the target, are considered data exploitation.

Data disclosure mechanisms

Disclosure means may be intentional or unintentional. Table 1 shows a list of mechanisms for disclosure.

Disclosure Mechanism	Disclosure Mechanism Examples
Data placed in public records	Names, addresses, marriages, deaths, births, real estate transactions, arrests, tax receipts, business licenses
Sending data via email	Personal data, file attachments
Sending data via instant messenger	Personal data, file transfers
Sending data via text message (SMS)	Personal data
Data intentionally shared over Peer-to-Peer file sharing (or other) networks	Music, video, software
Data gained via data gathering attacks (such as phishing)	Personal data, financial data
Data gained physically or logically by theft	Credit card numbers, social security numbers, financial data, correspondence, identity data, postal mail, breaches, deliberate employee action
Authorized users and/or trusted employees performing unauthorized disclosures	Financial data, passport records, restaurant employees surreptitiously copying credit card data, credit agencies, law enforcement agencies, insurance agencies
Data uploaded to video sharing sites	Police actions, crimes
Data gained through inference	Shopping patterns, utility use, access to various Internet sites
Data gained through surveillance	Wiretaps, network sniffing, video, shoulder surfing, eavesdropping post-it notes
Data physically or logically lost or discarded	Backup tapes, hard copy records, electronic documents, laptops, discarded hard drives, software or equipment misconfiguration

Table 1: Data Disclosure Mechanisms

Each category of data disclosure mechanism has a varying probability of causing a disclosure in a given situation. This probability can be ranked least to highest, from 0% to 100%.

¹ Elizabeth Neuffer, “One Name’s Double Life David Lombardi Lost His Id Cards, Acquired Coworker’s Debts and Crimes,” 1991 – http://www.linsky.org/html/prosecutor_det.php?id=133.

The cost of cleaning up the long-term results of such exploitation traditionally falls solely to the victim, not the discloser of the data.

Rankings are a best estimate given a particular environment and data.

Data disclosure is by no means an uncommon event. Consider these examples from the span of just one single week in August of 2009:²

- August 11, 2009 – loss – Bank of America Corp. – Account numbers for an unreported number of customers
- August 11, 2009 – theft – Citigroup Inc. – Account numbers for an unreported number of customers
- August 11, 2009 – theft – University of California-Berkeley – Names, social security numbers and dates of birth for 493 applicants
- August 13, 2009 – theft – National Guard Bureau – Names, social security numbers, payment amounts and dates for 131,000 Army National Guard personnel
- August 14, 2009 – theft – American Express – Undisclosed member account data for an unreported number of customers
- August 14, 2009 – unauthorized disclosure – Calhoun Area Career Center – Names, social security numbers, mailing addresses, telephone numbers, birth dates, school information for 455 students
- August 15, 2009 – theft – Northern Kentucky University – Names, social security numbers, other personal information on 200 students

This week was unexceptional as data disclosures go.

Limiting data disclosure mechanisms

Limiting data disclosure mechanisms can be a tricky prospect in the modern age of data ubiquity. This largely stems from the plethora of places that an individual item of data might be stored. For instance, the social security number for a person will likely be found in:

- One of more forms of identification personally carried
- Tax forms
- Educational paperwork
- Medical records
- Purchase records for certain items, such as automobiles or homes
- Records of legal proceedings

And this is by no means an exhaustive list. Keeping such an item of data from being disclosed requires that all parties holding the data do so, an exercise that becomes more unlikely, if not impossible, with each passing day.

Exploitation of data disclosures

Once exploited, unintended data disclosure can have many immediate costs such as legal or mitigation costs, but a too-often-ignored cost is the burden placed on the target of the disclosure.

Exploitation of disclosed data can range from credit-card purchases that must then be mitigated by the proper card holder (perhaps in conjunction with an insurance company), to fraudulently-acquired home mortgages and other debt, to improperly attributed arrest records. The cost of cleaning up the long-term results of such exploitation traditionally falls solely to the victim, not the discloser of the data.

To exacerbate the situation, a thriving underground economy based around stolen data exists, complete with suppliers, consumers, and marketplaces for exchanging (generally illicitly obtained) data. A given item of data may be exploited repeatedly by different parties. Table 2 shows what the prices for some of these data items are as summarized from a recent report by Symantec.³

Exploitable Data Item	Price (USD)
United States-based credit card with card verification number	\$1-\$6
United Kingdom-based credit card with card verification number	\$2-\$12
An identity (including U.S. bank account, credit card, date of birth, and government issued identification number)	\$14-\$18
List of 29,000 emails	\$300
Online banking account with a \$9,900 balance	\$3
Yahoo Mail cookie exploit – advertised to facilitate full account access when successful	\$3
Valid Yahoo and Hotmail email cookies	\$6-\$20
Compromised computer access	\$3-5
Phishing website hosting – per site	\$50-\$500
Verified PayPal account with balance (balance varies)	\$10-\$50
Unverified PayPal account with balance (balance varies)	\$12
Skype account	\$1

Table 2 – Advertised prices of identity data traded on underground economy servers

² Privacy Rights Clearinghouse, “A Chronology of Data Breaches,” 2009 – <http://www.privacyrights.org/ar/ChronDataBreaches.htm#CP>.

³ Symantec, “Symantec Internet Security Threat Report,” 2007 – http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_internet_security_threat_report_xi_03_2007.en-us.pdf.

The vast majority of the stored personal data likely enjoys no protection whatsoever and is routinely, if inadvertently, exposed.

Some of these data are protected by technologies such as Digital Rights Management (DRM), although DRM is now widely believed to be failing to have achieved its goals. Some data are protected by data classification systems and Mandatory Access Control (MAC) such as the system in use by the U.S. military (i.e., secret, top-secret, etc.). Some data are also protected by legal or regulatory means, though the majority of data that enjoy such protections in the U.S. are of a business nature. The vast majority of the stored personal data likely enjoys no protection whatsoever and is routinely, if inadvertently, exposed. In the last 36 months, the author's personal data was unintentionally exposed by financial, educational, and government institutions (among others) at least six times (according to postal mailed reports).

Impacts of severe data disclosures

In the case of a disclosure that has turned into a full-blown case of identity theft, the time and resources required to recover from the situation can be staggering. One study performed by the ID Theft Center shows that almost fifty percent of exploited identity theft cases are still unresolved after two years, and that four percent of cases linger without resolution for ten years or longer.⁴

The costs of data disclosure can be partitioned into two main categories: primary costs and secondary costs.

Primary impacts

Primary data exposure impacts are considered to be those borne in the year following the exposure, including:

- Legal costs
- Investigation
- Prosecution
- Fines/penalties

These costs can vary greatly according to the data exposed, from \$100 per record⁵ to \$100,000 or more per record.⁶ In general, the initial costs of a data exposure are relatively limited.

Secondary impacts

The secondary impacts of data exposure are those borne following one year after the exposure:

- Fraud
- Individual legal costs
- Court case/suit settlements
- Damage to reputation
- Marketing (reputation)
- Increased support calls
- New/changed security procedures
- Time lost
- Emotional trauma
- Higher prices/premiums/rates

While some secondary costs are easily measurable (e.g., lawsuit settlements or increased rates), the majority of these costs are more difficult to assess. Measuring the cost of damage to reputation or emotional trauma can be difficult, at best. Some of these costs are situational and may include less tangible values such as "reputation" and "emotional security." Damage due to secondary costs may take upwards of ten years to resolve.⁷

Reducing the impact of data disclosure exploitation

The burden of handling data disclosure exploitation generally falls squarely on the head of the individual victim. If such exploitation is not noticed and dealt with, the difficulty of resolving the situation will continue to increase over time. Most of the techniques center around the protection of the individual's information as held by the credit bureaus and the resulting credit score calculated from this information. The two major items of importance to detect fraudulent activity are monitoring credit information and monitoring financial transactions.

Monitoring financial transactions is the front line of defense when an individual's data has been exposed. In many cases, if such activity is reported to the financial institution in question immediately, the transaction can be reversed and the issue can be resolved in a matter of days. If the transaction is not noticed for weeks or months, it may be too late to take action easily, if at all.

In the area of credit monitoring, several products exist, some from the credit bureaus themselves, that provide such services. Most of these provide roughly the same set of features:

- Ability to view information reported to one of more of the major credit bureaus
- Notifications when certain activities take place, such as new accounts being opened or balances increasing by a certain percentage
- Insurance against identity theft, generally around \$20,000

Such services can be invaluable, as they enable the victim to take action swiftly when unauthorized activity is detected.

⁴ Identity Theft Resource Center, "Identity Theft: The Aftermath," 2003.

⁵ Ibid.

⁶ Zero Paid, "RIAA sues AllofMP3 for \$1.65 trillion," 2007 – [http://www.zeropaid.com/news/8175/RIAA+sues+AllofMP3+for+\\$1.65+trillion](http://www.zeropaid.com/news/8175/RIAA+sues+AllofMP3+for+$1.65+trillion).

⁷ Identity Theft Resource Center, "Identity Theft: The Aftermath," 2008 http://www.idtheftcenter.org/artman2/uploads/1/Aftermath_2008_20090520.pdf.

Fiscal aspects of data disclosure

The fiscal elements of data disclosure partition into two main categories: value and cost. Assessing the fiscal elements of data is not a formulaic process and, while guidelines assist in quantifying the value of data, ultimately requires judgment and experience.

Value

The overall value of the data is an aggregate of the ways that the fiscal value of a particular item of data is measured:

- **Market price** – The market value of the data: \$0.99 for an MP3; \$25 for a DVD movie
- **Business value** – The value of the data over time, presumably to its owner: customer lists, demographic data
- **Regulatory value** – Value gained from penalties for unauthorized use, for example intellectual property theft: \$100,000 per occurrence for copyright violation

Cost

The overall data-related costs are an aggregate of the ways that the fiscal cost of a particular item of data can be measured:

- **Acquisition** – Gathering or purchasing the data: public records, \$1.00/name magazine subscriber lists
- **Aggregation** – Assembling the data from a variety of sources: purchased records, customer lists, public records
- **Processing** – Transforming the data into a useful set: formatting, sorting
- **Validation** – Ensuring that the data is consistent and accurate enough for use: verifying names, addresses, phone numbers
- **Maintenance** – Keeping the data current and usable: re-verifying data, adding new data, removing old data
- **Destruction** – Properly disposing of data that is no longer useful or desirable: patient records, tax preparation data, electronic communication records
- **Regulatory cost** – Costs to stay in conformance with government and other regulations; penalties for violating data storage laws and rules: HIPAA violations, PCI compliance costs

While the value or cost of an individual item of data might be relatively easy to quantify based on the above factors, these factors are generally not considered in isolation from each other. For example, the value of a name, without the presence of other data will generally be small. When a mailing address supplements the name, the data gains considerable value, and may be of use to a variety of business and institutions. When a social security number is added to the name and address, the collection of data becomes more valuable, to the point of being considered sensitive data. For these items of critical data, the combined value or cost of the data may be greater than the sum of the values of the components.

The burden of handling data disclosure exploitation generally falls squarely on the head of the individual victim.

With so many institutions – medical, educational, retail, and a host of others – collecting data, many with little or no regulation, the situation of data aggregating to the point of becoming sensitive is nigh inescapable.

Reducing the costs of data disclosures

On a corporate or organizational level, much of the work involved in reducing the costs of data disclosures revolves around the storage of the data. In many of the cases reported in the media when a major loss of data has occurred, the data has been in unencrypted form on backup tapes, laptops, or any of a variety of portable storage devices. Most, if not all of these cases could have been prevented by the simple expedient of encrypting the data in question.

Another question that can be asked in many of these incidents is why data was allowed to be removed from a secure facility. Again, in many cases, if the data were to be stored in a secure central location and not allowed to be copied to portable devices, such loss of data would be considerably less likely.

Conclusion

While the occurrence of data disclosure and the possibility of ensuing identity theft do not escape the attention of the media and the public in general, the specific impacts of such disclosures often do. This appears to be due to the general lack of serious consequences to the disclosers of said data, often allowing them to quietly skulk away and continue poor business and security practices relatively unmolested.

Although laws are in place for certain types of data, i.e., medical, educational, etc., the U.S. lacks the sort of broad, overarching, body of law regarding the privacy and caretaking of data that some other countries enjoy. While it is true that there are bright spots in U.S. data disclosure and privacy law, such as California's SB-1386,⁸ it appears that the road to mitigating this issue will be a long one.

About the Author

Jason Andress is a tinkerer, rascalion, and all around geek. He works for a major software company, teaches graduate and undergraduate security courses, and enjoys a good game of Scrabble. He can be reached at jason.andress@gmail.com.



⁸ Senator Peace, "California Senate Bill 1836," 2002 – http://info.sen.ca.gov/pub/01-02/bill/sen/sb_1351-1400/sb_1386_bill_20020926_chaptered.html.