kriativ.tech

Edition Nº 8, Vol. I – 04-10-2020
ISSN Print: 1646-9976 | ISSN Online: 2184-223X |
http://www.kriativ-tech.com
DOI: 10.31112/kriativ-tech-2020-08-47

# Data: The most valuable commodity

Pedro Ramos Brandao
Coordinator Professor at ISTEC - Researcher Évora University (CIDHEUS) – pb@pbrandao.net

Manuel Rezende
MSc Student at ISTEC – manuelrezende@my.istec.pt

**Abstract:** *Throughout the 21st Century, corporate data breaches have become an increasingly common occurrence, progressively changing corporations and IT security specialist's focus from 'If' to 'When' the next significant event will occur. This article presents an insight into cybersecurity and why it has become the primary concern for corporations and institutions.*

***Keywords:*** *Data breach, data loss, data leakage, cybersecurity, financial loss, identity theft*

## I. Introduction

Although in a completely different setting and context, Gordon Gecko's character, portrayed by Michael Douglas in the 1987 motion picture 'Wall Street', stated that "The most valuable commodity I know of is information". Fast forward to today's digital economy heavily relying on the collection and storage of large amounts of business and personal sensitive data which needs to be protected and you have the new most valuable commodity: data. There are several aspects to be considered regarding data protection, ranging from user security awareness education to corporate data security culture and commitment.

Data breaches extend far beyond those widely recognized and used as keynote examples or general public information such as Yahoo (2013 and 2014), Equifax (2017), eBay (2014) or Sony PSN (2011) to name just a few. These occur with a frightening frequency and with staggering amounts of data stolen.

Two fundamental aspects of data breaches are data loss and data leakage whose handling is addressed very differently; Data loss may result from the intentional or unintentional, with or without malicious or wrongful intent handling and transfer of data by an inside user but is a vector that can be addressed in a fairly straightforward way through the use of Data Loss Prevention solutions and policies; Data leakage is usually intentional and results from the direct actions of external hackers. These are far more difficult to predict, detect and detain [1].

Data leaks often take a considerable amount of time to be detected and stopped, potentially providing the perpetrators with weeks, months or even years of free access within corporate systems. Conventional data breach response plans often take place after the fact conducting forensics and evidence gathering for further action but are no means of limiting or stopping a data breach [1].

A considerable percentage of data breaches target customer personal data as this has the potential of immediate own use or resell, resulting in heavy financial impact, not only for the customer whose data is stolen and used to whatever purpose but to the corporation as well.

Several compliance and regulation laws have been introduced regulating the use and storage of personal data in an attempt to mitigate risk and impact of data breaches through prevention and enforcing accountability for those who do not comply; one such example, regarded as the strongest and most comprehensive attempt to regulate the collection and use of personal data is EU's GDPR (General Data Protection Regulation).

Other types of data breaches target the victim's operational and business continuity capabilities, often through ransomware, for immediate financial gain but frequently result in data loss, whether the payment is met or not.

Information is a vital business asset and must be protected against internal and external threats at all cost. We have been witnessing the staggering amounts of money spent with cybersecurity in the last few years and the even larger amounts predicted for the next years but continue to find out that a significant number of data breach incidents are accomplished by relatively simple attack vectors or outright basic ones such as the exploit of long known security flaws that still are not being taken seriously or addressed in timely fashion by a large part of the industry, alongside common basic security policies and actions.

## II. Data breach threats

Data breach threats, as previously mentioned, occur with different players, in different circumstances and with a different set of intentions:

a) Unintentional insider threats result from inadvertent handling of data such as accidental publishing or transmission of sensitive data without proper encryption and/or usage of secure communications, lost hardware with data content or poorly stored credentials, disclosure of protected or sensitive data in public places.

b) Intentional insider threats that result from actions of sabotage or espionage perpetrated by disgruntled or unlawful employees

c) External intentional threats with the purpose of stealing or destruction of data and resources

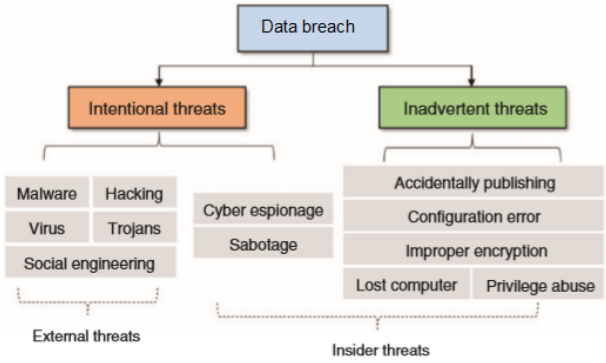


Figure 1: Classification of enterprise data breach threats [2]

The unintentional vector of data breaches is, perhaps, the one with the most straightforward set of rules, actions and policies to put in place and is the foundation of corporate cybersecurity; first and foremost, users need to be trained and made aware of information security culture and encouraged to embrace it. Many times, actions or behaviors seen as innocent or discounted as irrelevant are the foundation of unsecure or dangerous professional conduct regarding information security (who hasn't written a complex password on a post-it and stick it to the monitor, keyboard or wallet?).

As an example of an unintentional data breach, in October 2016 staff from the Australian Red Cross Blood Service accidentally exposed documents that contained 550,000 blood donors personal information on a public-facing website [2].

Another fundamental set of rules and policies lie with Data Loss Prevention systems; DLP functions inherently support compliance with laws and regulations through a detailed inventory of where personal data is, how it is used and how to best manage its access and movement. Its implementation identifies a tiered set of actions to meet its goals.

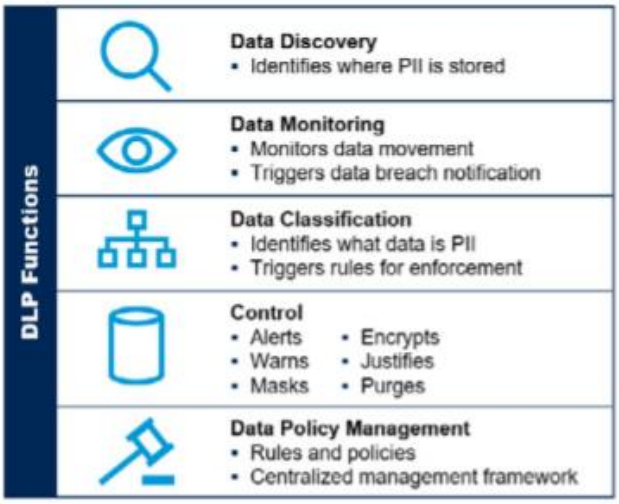


Figure 2: Data Loss Prevention Functions [3]

Besides clearly identifying what is and where is Personal Identifying Information data stored, puts in place a series of policies related to data access and movement monitoring with a set of corresponding alerts and triggers to suspicious or non-compliant actions. These policies greatly mitigate the risk of both intentional and unintentional actions and assume critical importance because, by its nature, internal breaches mostly involve users with legitimate access to the data and systems.

The clearly most complex scenario lies with the external intentional actions. These constitute the bulk of recently reported data breaches and make use of a diversity of attack vectors. Social engineering vectors have become increasingly sophisticated over the last few years, ranging from the common Phishing (email) to less known or usual techniques such as Vishing (phone call) or SMiShing (text message). In some cases, Impersonation also works; for instance, in the United States of America, a USPS employee (Federal employee) is often automatically trusted and granted access with little or no restrictions for deliveries inside corporations. An attacker, suitably dressed as one of these employees, could potentially gain access to unprotected or unsecured terminals which would allow, for example, the introduction of a worm or virus. Impersonation is known to have consistently worked with Penetration Testers.

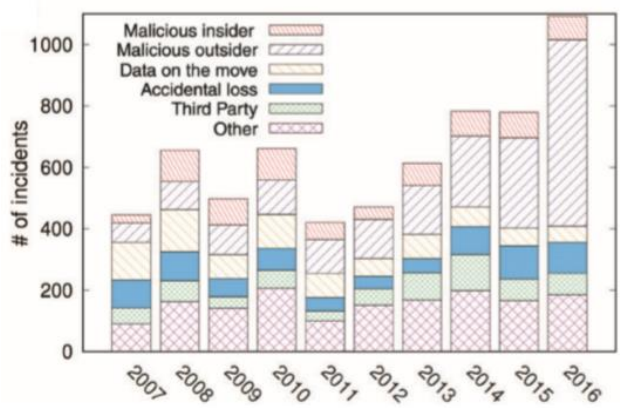


Figure 3: Data breach incidents in recent years by type of occurrence [2]

Besides compromised credentials obtained through social engineering, weak or stolen credentials also play a major role in this scenario. The initial attack vector of the J.P. Morgan Chase data breach succeeded through the use of credentials obtained from a leaked database that contained over 1 million usernames and passwords, some of which belonged to J.P. Morgan Chase Corporate Challenge employees. This, associated with a basic security flaw in one of the servers, provided access to the vast bank's network [4].

Other intrusion techniques make use of known vulnerabilities such as DNS spoofing and SQL injection or explore a window of vulnerability though the use of zero-day attacks.

Some less sophisticated attacks can be perpetrated with the goal of service disruption, such as a DDOS attack, but can also serve as a decoy for a more sophisticated action; this was the case of the Carphone Warehouse data breach in 2015 where a DDOS attack was launched to distract it's IT team while a coordinated attack was launched on its systems that resulted in the theft of 2.4 million customer records [1].

Often, hackers use a combination of methods to achieve their goals. In a case study attack perpetrated on the Target Corporation (2013) where 70 million customer's records were stolen. The attackers gained access to the network of a third-party vendor (Fazio Mechanical Systems) that was in charge of electrical consumption and temperature monitoring in Target stores and had remote

access rights to its network. Access credentials for Fazio's network were obtained through a Phishing attack and were used to access Target's network looking for vulnerable machines. A vulnerability was found on the POS (point of sale) systems and a data stealing malware (BlackPOS) was deployed to read the information that was later encrypted and moved to internal compromised systems and finally moved out to several dropsites [2][6].

The forensic investigation identified several key failures in Target systems implementation:

a) Failure to apply appropriate access control mechanisms on third-party partners.
b) Failure to segregate the payment systems from the rest of the network.
c) Failure to harden the POS systems allowing for the installation and configuration of unauthorized software.
d) Failure to investigate the security warnings issued by its own Firewall and IPS systems [2].

Several of Target's failures are identified and addressed by the Data Loss Prevention Functions.

## III. Information availability

Today's digital world invites for the disclosure of personal information on a large scale, not only for the convenience of E-commerce but also for social media. People increasingly voluntarily expose personal data on social media to a level hard to comprehend and with details that they would not share to a stranger face-to-face. Yet, there for all to see and for the social media site owners to collect, sufficient data is exposed to allow very accurate profiling and targeted advertising or, worse, to potentially be sold or stolen.

It wouldn't be on the average user's mind that the data he shared on Facebook could be used to steer a Presidential election campaign. That is exactly what happened with the Cambridge Analytica services hired by the Republican Party in the United States of America 2016 Presidential election. The Cambridge Analytica research was not particularly illegal or improper in itself, but it exploited a loophole in the terms and conditions of Facebook (which was later amended) allowing personal user data to be shared with an app publisher facilitating the data harvest. For all intents and purposes, several million records containing personal user information were used without explicit consent from the users.

This wasn't the first time that data analytics and mining were used in a political campaign. For instance, according to the MIT Technology Review, the 2012 Obama campaign used these methods through the download and install of a specific app (that used informed consent or opt-in) for data collection [5]

Besides the sharing of personal information on social media networks, the convenience of E-Commerce dictates that large amounts of personal data need to be shared with a vast networks of sites and vendors. The security of this data is of capital importance because, more often than not, is contains financial information in the form of credit card or other banking information.

Financial data is of particular interest to hackers as it provides direct and immediate access to banking information of thousands or millions of users in a single breach incident. Other highly targeted types of data are medical information and other items of personally identifiable information that can lead to identity theft. This is clearly illustrated by the percentage of businesses and medical/healthcare facilities targeted in recent years versus other industries.

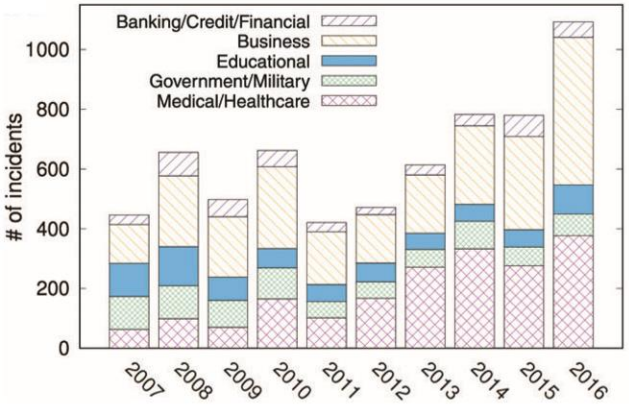


Figure 4: Data breach incidents in recent years by industry sector [2]

There is a new attack vector that has been emerging and potentially exploited in the last few years; the Internet of Things. Automated data collection and transmission from millions of deployed devices on smart-home or industrial applications have the potential to be exploited if not conveniently secured in what and how they transmit but also ensure that its endpoints are isolated from sensitive networks. Several of these devices use very basic forms of implementation and have extensive security flaws than can be exploited with relative ease. There are millions of devices around the world using inexpensive ESP32/ESP8266 Wi-Fi transmission devices working with minimal security and with well-known exploits that can be taken advantage of to infiltrate a home or corporate network if appropriate defensive measures are not applied.

In February 2020, a shocking and stunning CNN Business interview with Clearview AI's CEO Hoan Ton-That was aired illustrating their purpose and business model. According to a New York Times report from January, this start-up had been collecting in excess of 3B images from social media sites to add to its database [10]; Clearview AI developed a technology for scraping the internet for publicly available photos of persons (including Facebook, Twitter and other social media sites) and, without user consent, harvesting them and making them available in a commercial application selling a face recognition service, matching a capture with stored photos and tracking where, when and how they were published. The service turned out to be astonishingly accurate, with an interviewer's selfie taken at the moment finding several matches (and links to the respective websites) of the interviewer, even from several years back where its looks were considerably different or in instances wearing glasses or with shorter or longer beard . In an astonishing gingerly tone, Thon-That stated that at that time they had already about 600 customers, some of them law enforcement agencies to which they extended the courtesy of a 30-day free trial. In a 'famous last words' worthy statement, he then proceeded to ensure the interviewer 'we are not planning on selling this service outside law enforcement and the banking industry". Common sense would then ask 'how long will it take for them get hacked?'. Well, as it turned out, not long; Clearview issued a statement, still during February, that it had suffered a data breach but 'only' its customer database had been accessed and not the photo database or the technology behind it. Whether true or not, this database and service have the potential to severely impact privacy, either through its use by current or future customers or through a data breach (assuming it has not happened already in this instance).

## IV. Consequences

Most of the high-profile data breach incidents resulted in heavy financial losses to those organizations. After the Yahoo data breaches (2014 and 2016) in which more than one billion user accounts were compromised, the sale price to Verizon was reduced about $350 million from the planned sale price.

In the Target breach incident, an $18.5 million multistate settlement was reached to cover for a $10 million class-action suite in addition to free credit monitoring for all the affected customers and payment up to $10,000 to customers with evidence of losses directly linked to the data breach [6].

The average total cost of a data breach in the United States of America for the companies studied has grown from $3.54M in 2006 to $8.19M in 2019, a 130% increase over 14 years. The average total cost of a data breach in the healthcare industry was $6.45M, or 65% higher than the average total cost of a data breach [8]. Financial losses aren't by any means the full extent of the consequences but also carry a collateral damage of lost trust and reputation by consumers.

According to a PricewaterhouseCoopers enquiry and report, just 25% of the respondents believe most companies handle their sensitive personal data responsibly. Even fewer, only 15%, think companies will use that data to improve their lives. 87% of consumers say they will take their business elsewhere if they don't trust a company is handling their data responsibly [11]. The report also illustrated the degree of trust or perceived risk by industry, with some surprising outcomes.

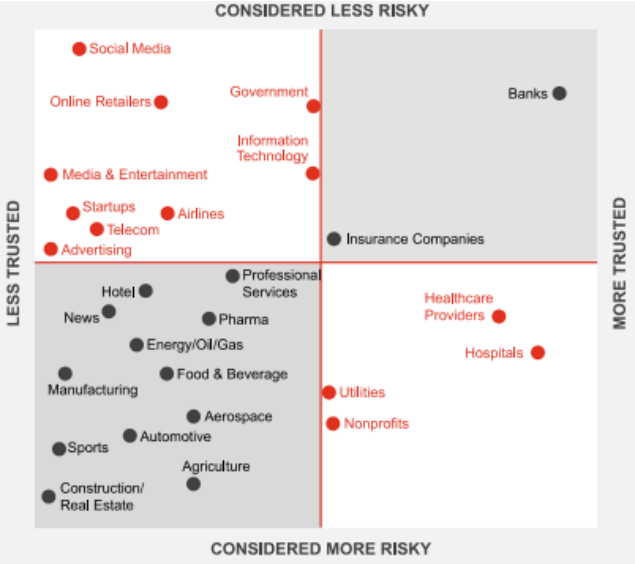


Figure 5: Which of the following industries or company types are the most trustworthy? [11]

Unfortunately, it seemed just a question of time for the consequences to cross the material loss boundary and enter the realm of direct cause for loss of life; The University Clinic in Düsseldorf, Germany, was hit with a ransomware attack on September 10th, 2020 perpetrated through a security flaw in a Citrix VPN system and had several systems compromised and inoperable. One of the affected systems was used for ambulance patient admission and caused the hospital to refuse admission for an urgent patient on the night of September 11th. The female patient, suffering from a life-threatening illness, had to be diverted to the nearest suitable hospital (Wuppertal, about 30Km away) but succumbed enroute. Germany's cyber-security agency, the Federal Office for Information Security (Bundesamt für Sicherheit in der Informationstechnik – BSI) was called in to shore up the hospital's systems. Its chief, Arne Schoenbohm, stated that the Citrix flaw had been known since Dec. 2019 and called on healthcare facilities not to delay IT security upgrades.

## V. Conclusion

The rising quantity of data breaches over the last few years have been driving global cybersecurity spending to unprecedented levels. According to a Forbes report, the global cybersecurity market is currently worth $173B, growing to $270B by 2026.

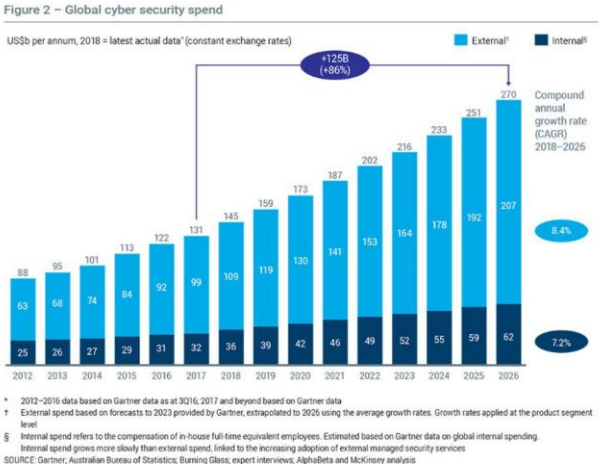


Figure 6: Global security spending [8]

New variables have now been introduced to this equation since the beginning of the global pandemic. Cyberattackers are quick to attack new unprotected threat surfaces created when tens of millions of employees started working from home. In a post-COVID-19 world, cybersecurity is as critical as Internet access itself [8].

A new approach to identity and trust verification has been fast growing and is regarded as the solution moving forward; Zero Trust Security. Remote workers identities and devices are the new security perimeter. This is ZTS was designed for, and the post-pandemic world will represent its trial by fire. The principle behind ZTS assumes that there are potential attackers inside and outside and no user or machine should be automatically trusted. Another principle of ZTS is the least-privilege access; Users are given access strictly to what they need to access. It also uses microsegmentation breaking up security perimeters into smaller zones to maintain separate access for different parts of the network. Multi-factor authentication is also a key aspect of augmented security.

In addition to user access control, ZTS also enforces strict controls on device access; it monitors how many different devices are trying to access their network and ensure that every device is authorized. This further minimizes the attack surface of the network. But ZTS alone, or any single other technology for that matter, doesn't solve the problem by itself. According to a study by Accenture, another technology has shown high efficiency on some parameters of cybersecurity; Artificial Intelligence. It has shown that enterprises who lead their industries in cyber resilience rely on A.I. to

reduce the number of successful attacks and deliver a more consistent quality of response [9].

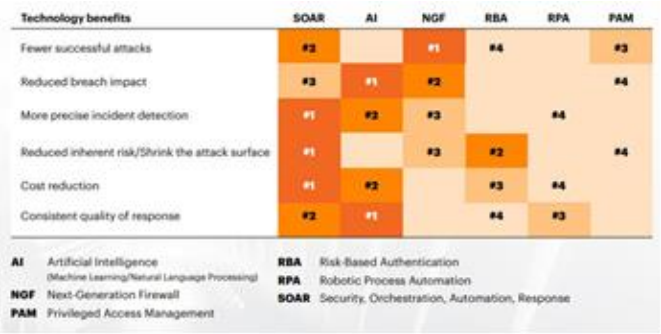


Figure 7: Technology performance [9]

It is clear that no single technology or approach to cybersecurity will achieve a zero-risk outcome. Common-sense (and evidence from past incidents) tells us that there is no such thing and it is of no use just to 'throw money at the problem'. High levels of cybersecurity can be achieved through a combination of factors, both human and technological but this needs to be done in a coordinate manner; People say 'hindsight is 20/20' but, as an example, in the Target breach case there was technology already in place issuing warnings that, apparently, were not taken seriously enough to warrant a deeper investigation. All of the other listed factors point to completely obvious and basic rules of IT security that many assume as such but, yet, so many fail to implement. There are surely hundreds or thousands of cybersecurity incidents that could have been prevented by applying the most basic and common-sense measures. In the recent case of Düsseldorf's University Clinic, failing to address a security flaw and its patching for almost a year ultimately resulted as being the linchpin for the first recorded loss of human life following a cyberattack.

Cybersecurity needs to be addressed as one of the most critical factors on today's connected world and must be seen as a multi-factor, multi-player challenge where only by combining a secure IT culture, user security awareness training and a combination of correctly implemented and maintained technologies will provide it a fighting chance, because, as we are all very aware by now, is it not a question of 'If' but 'When'.

# VI. References


[1] Ibrahim A., Thiruvady D., Schneider J. and Abdelrazek M. (2020). "The challenges of leveraging threat intelligence to stop data breaches". In Frontiers in Computer Science, August 2020. Volume 2, Article 36. doi:10.3389/fcomp.2020.00036

[2] Cheng L., Liu F. and Yao D. (2017). "Enterprise data breach: causes, challenges, prevention, and future directions". WIREs Data Mining and Knowledge Discovery. Published by John Wiley & Sons, Ltd.. doi: 10.1002/widm.1211

[3] Digital Guardian (2019). "The definitive guide to data loss prevention". Whitepaper 2019 edition

[4] Jeng A. (2015), "Minimizing Damage From J.P. Morgan's Data Breach" The SANS Institute

[5] Adams B., Clark A. and Craven J. (2018). "It is Free and Always Will Be: Trading personal information and privacy for the convenience of online services". Researchgate publication/324717676

[6] McCoy K. (2017) "Target to pay $18.5M for 2013 data breach that affected 41 million consumers". USAToday, May 23rd 2017 edition, Money section

[7] Reuters Editorial Staff (2020). "Prosecutors open homicide case after hacker attack on German hospital". Reuters Worls News, September 18th, 2020.

[8] Columbus L. (2020). "2020 Roundup Of Cybersecurity Forecasts And Market Estimates". Forbes Editor's pick, Apr 5th 2020

[9] Bissell K., Lasalle R. and Dal Cin P (2020) "Innovate for cyber resilience". Accenture Security, third annual state of cyber resilience.

[10] Hill K. (2020). "The Secretive Company That Might End Privacy as We Know It". The New York Times, January 18th 2020 edition.

[11] PricewaterhouseCoopers (2017). "How consumers see cybersecurity and privacy risks and what to do about it". PwC Consumer Intelligence Series: Protect.me