Corporate Stagnation: A History of Data Breaches and Their (Lackluster) Responses

Maximiliano Brizzio

Stevens Institute of Technology

CS 595: Information Security and The Law

Ms. JoAnna Luna

May 11, 2024

Abstract

This research paper aims to shed light on the impact of data breaches and their corresponding incident response strategies. Through brief historical analysis and three case studies–namely the 2013-2016 Yahoo! data breaches, the Facebook and Cambridge Analytica debacle, and the International Committee of the Red Cross' 2021 breach–along with U.S. government reports and news media articles, we seek an understanding of not only the relationship between breached companies and affected individuals, but also the hallmarks of an effective incident response initiative. Through analysis of each case study, we endeavor to comprehend the underlying motivations behind unfavorable actions taken by the incident response teams and their executive committees while proposing best practices to mitigate overall damage. It's imperative to note that the goal of this paper is not to answer, '*Why do data breaches happen?'* but instead it will attempt to respond to *'What happens after data breaches?'* and *'How should we respond?'*. By fulfilling this objective, the findings here may explain why many companies historically chose to abandon full disclosure, why perhaps it is better that they adopt the mentality, and who is responsible for securing the future today.

Corporate Stagnation: A History of Data Breaches and Their (Lackluster) Responses

Today we find ourselves in the height of the information age. New techniques, procedures, devices, and services for processing information are created and implemented every day. In fact, approximately 90% of the 120 zettabytes (120 trillion gigabytes) of the world's data has been generated in the last two years alone[1]. As the rate of information creation increases over time, it's clear that a rise in security concerns is correlated. However, as concerns over information security have drastically been exposed in the public debate of computing technologies, the practices of large corporations many times completely fail to meet the expectations of consumers. Through breaches that in some cases affect millions of individuals, specifically those of Yahoo.com, Facebook, and the Red Cross, we will analyze the salient features of the cyber-attacks, their responses, and why oftentimes those responses are completely lackluster in nature, sometimes bordering the line of malicious behavior.

Before we can begin our case studies, it's important to establish context through a brief history of data breaches. Technically speaking, the first breach in history took place in 1834 where two French thieves wiretapped France's telegraph system to steal important financial market information. In fact, at multiple points in time, 'pre-modern hackers' committed denial of service attacks where communication lines including phone lines were disrupted. It wasn't until 2005 that a data breach affected over a million consumers through the DSW Shoe Warehouse breach that exposed credit card information. That same year, George Mason University also

[1] Figure sourced from https://www.statista.com/statistics/871513/worldwide-data-created/

suffered an attack where tens of thousands of extremely sensitive PII (Personally Identifiable Information) of students were stolen, including SSNs. Since 2005, the number of data breaches steadily increased semi-linearly up to 2016, whereafter, followed by a dip in events of data compromises, breach counts exploded. From 2020 to 2021, there was an increase of ~800 breaches. From 2022 to 2023, the increase in breaches reached almost 2000 more instances[2].

In a time where the quantity of cyber-attacks has not only risen for almost two decades, but exponentiated in recent years, it's imperative that security policies implemented are comprehensive and malleable. Because threat-vectors change in the scene, falling out of relevancy and replaced by newer, more sophisticated attacks, it's impossible to see the future of information security. The best one can do to combat breaches relies on standardized practice, constant audits, and compliance tests.

Case Study 1: Yahoo!

If there were a list of the most impactful breaches in US history, Yahoo would certainly make one of the top spots of the list. The web services provider suffered two different breaches varying in scope over 2013-2014. The first, in August 2013, was then the biggest known breach of a single company's networks. In December 2016, over three years after the attack, Yahoo disclosed the 2013 breach along with a statement that over one billion user accounts had been compromised. The next year in October 2017, they corrected the statement and announced that every single user account had been compromised. All three billion. In 2017, the former CEO of Yahoo testified before congress that the perpetrator of the breach was still unknown. Such an impact is monumental in a vacuum, but what might come as more of a surprise is why Yahoo

[2] Figure sourced from https://www.statista.com/statistics/273550/data-breaches-recorded

chose to wait years before disclosing the breach. However, much more damage than the three billion compromised accounts took place through another breach.

In late 2014, Yahoo suffered a separate cyber-attack that resulted in a backup database containing sensitive information of over 500 million accounts being leaked. Among account names, email addresses, phone numbers, dates of birth, hashed passwords, and encrypted or unencrypted security question responses, the scope of the damage was as large, if not larger than the breach that took place a year earlier. Following the breach, the hackers fraudulently accessed email accounts to steal digital gift cards, boost search ranking of business interests, and pivoted towards attempting the seizure of Gmail accounts using the stolen credentials. In the end, the FBI "officially charged the 2014 breach to four men, including two that work for Russia's Federal Security Service (FSB)… The four men accused include Alexsey Belan, a hacker on the FBI Ten Most Wanted Fugitives list, FSB agents Dmitry Dokuchaev and Igor Sushchin who the FBI accused of paying Belan and other hackers to conduct the hack, and Canadian hacker Karim Baratov who the FBI claimed was paid by Dokuchaev and Sushchin to use data obtained by the Yahoo breaches to breach into about eighty non-Yahoo accounts of specific targets" (Yahoo! Data Breaches, n.d.). Only one of the four men accused was arrested, later being extradited to the US, pleading guilty to nine counts of hacking. During the investigation and trial, he admitted to accessing at least 80 compromised email accounts on behalf of Russian contacts.

Setting geopolitical factors aside, this breach was monumental in the public's view of information security, specifically in terms of corporate data handling. However, what's important to scrutinize in these breaches is not the methods in which the data was stolen; it's the method in which Yahoo chose (not) to respond.

On September 22, 2016, almost 2 years after the breach, Yahoo finally disclosed the 2014 breach. Through Yahoo's quarterly report filing to the SEC on November 9, 2016, they revealed that "As described above, the Company had identified that a state-sponsored actor had access to the Company's network in late 2014. An Independent Committee of the Board, advised by independent counsel and a forensic expert, is investigating, among other things, the scope of knowledge within the Company in 2014 and thereafter regarding this access, the Security Incident, the extent to which certain users' account information had been accessed, the Company's security measures, and related incidents and issues" (Yahoo! Inc., November 9, 2016). This directly contrasts Yahoo's own words in an SEC filing they made just two months earlier. In the *Preliminary Special Proxy Pertaining to a Sale* filing which was submitted September 9, 2016, they write "To the Knowledge of Seller, there have not been any incidents of, or third party claims alleging, (i) Security Breaches, unauthorized access or unauthorized use of any of Seller's or the Business Subsidiaries' information technology systems or (ii) loss, theft, unauthorized access or acquisition, modification, disclosure, corruption, or other misuse of any Personal Data in Seller's or the Business Subsidiaries' possession, or other confidential data owned by Seller or the Business Subsidiaries" (Yahoo! Inc., September 9, 2016). The timeline does not match up here, so the question left is why did Yahoo choose to wait so long to disclose the breach?

It is important to play devil's advocate in situations like these as any bias can disregard useful pieces of information. A major reason that Yahoo waited was perhaps that they wanted to take a moment to assess the situation. Investigation is required to check what data was stolen, how it was stolen, if the attacker is still in the system among other things. The issue with applying this idea to this case study though is that two years is an awfully long time to be

investigated. Even more so, Yahoo is shown to have concealed the breach not only by a lack of disclosure to the public, but through explicit contradictions made in company documents. An interesting note here is that Verizon was set to purchase Yahoo's business operations in the summer of 2016, and once the buyer found out about the breach, they lowered the original offer by $350 million, an expensive mistake. Without theorizing too much, a motive for why they would want to conceal the breach certainly exists. Whether the intent to deceive the public is there exists as a separate question entirely.

Case Study 2: Facebook

Nowadays, the announcement of a data breach on Facebook's network is almost yearly. Ironically, many of these data breaches take the form of Facebook breaching its own user's data rights. In the 2010s alone, Facebook suffered or mismanaged consumer data over 14 times. The last straw for the FTC (Federal Trade Commission) was July 2019 when they imposed a $5 billion fine for "undermining consumer's choices" (FTC, 2019), over 20 times larger than the previous largest data security penalty imposed. Now you might be asking yourself why I brought up data-mismanagement in a paper about data breaches. The reason is that when the mismanagement of data is done maliciously as seen throughout the history of Facebook, the effects on consumers make it no different than the effects stemming from breaches. The golden example of data-mismanagement closely tied to the $5 billion settlement comes from the Facebook-Cambridge Analytica fiasco. In 2018, Christopher Wylie, the whistleblower employee who himself admitted in an interview with The Guardian, he "ended up creating "Steve Bannon's psychological warfare tool" (Wylie, 2018). The ex-employee alleges that "the idea they bought into was to bring big data and social media to an established military methodology –

"information operations" – then turn it on the US electorate" (Wylie, 2018), 'they' referring to Steve Bannon, executive chairman of alt-right news platform Breitbart and later Trump's chief strategist, and Robert Mercer, a US hedge fund billionaire. Essentially, the purpose of this new technology was to use personal data of millions of people for political campaigning. The source of this data was Facebook. In the FTC press release relating to the lawsuit filed against Cambridge Analytica, they claim that the company "deceived consumers by falsely claiming they did not collect any personally identifiable information from Facebook users who were asked to answer survey questions and share some of their Facebook profile data" (FTC, 2019). The impact of the scandal is not limited to the settlements imposed by the FTC. Millions of people had their data maliciously stolen and used for purposes they directly chose against. Sure, an argument can be made that the data wasn't used for identity fraud as it commonly is with traditional breaches, but in principle, it follows the same form as any security incident and should be treated as such. Again, the impact isn't limited to the dollar amount fined to the companies, the trust of consumers breaks down in events like these. Now that the wake of the Cambridge-Analytica debacle is five years in the past, bringing up Facebook (or Meta) reminds people of bad data-handling practices and causes mistrust towards other companies, not just Facebook.

The golden example of data-mismanagement closely tied to the $5 billion settlement comes from the Facebook-Cambridge Analytica fiasco. In 2018, Christopher Wylie, the whistleblower employee who himself admitted in an interview with The Guardian, he "ended up creating "Steve Bannon's psychological warfare tool" (Wylie, 2018). The ex-employee alleges that "the idea they bought into was to bring big data and social media to an established military methodology – "information operations" – then turn it on the US electorate" (Wylie, 2018),

‘they’ referring to Steve Bannon, executive chairman of alt-right news platform Breitbart and later Trump’s chief strategist, and Robert Mercer, a US hedge fund billionaire. The purpose of this new technology was to use personal data of millions of people for political campaigning.

The source of this data was Facebook. In the FTC press release relating to the lawsuit filed against Cambridge Analytica, they claim that the company “deceived consumers by falsely claiming they did not collect any personally identifiable information from Facebook users who were asked to answer survey questions and share some of their Facebook profile data” (FTC, 2019). The impact of the scandal is not limited to the settlements imposed by the FTC. Millions of people had their data maliciously stolen and used for purposes they directly chose against. An argument can be made that even though the data was not used for identity fraud as it commonly is with traditional breaches, in principle, it follows the same form as any security incident and should be treated as such. Again, the impact is not limited to the dollar amount fined to the companies, the trust of consumers breaks down in events like these. Now that the wake of the Cambridge-Analytica debacle is five years in the past, bringing up Facebook (or Meta) reminds people of bad data-handling practices and causes mistrust towards other companies, not just Facebook.

Case Study 3: International Committee of the Red Cross

In a countering positive light, the International Committee of the Red Cross serves as the gold standard for incident response. On January 16, 2022, they announced to the public (not through an SEC Filing like Yahoo) that a “sophisticated cyber-attack [targeted] Red Cross Red Crescent data on 500,000 people” (International Committee of the Red Cross, 2022). The attack was discovered in mid-January and believed to have taken place sometime in December 2021

which compromised personal, confidential information of more than 515,000 ‘highly vulnerable’ people. It was later discovered that the attack most likely took place on November 9, 2021.

Thus, the timeline follows like this: on November 9, 2021 an unidentified threat-actor used a vulnerability in a tool used by the ICRC which allowed a “REST API authentication bypass with resultant remote code execution” (NIST). At some point towards the beginning of January 2023, the ICRC discovered that some sort of sophisticated breach took place, which was officially announced outside of the scope of the company on January 16th. Three days later they published an update on the breach, revealing what was stolen, who was affected, and the next steps in the incident response plan. Finally, a month from the original breach notification, on February 16th, they revealed much of the findings discovered after the analysis of the cyber-attack.

There seem to be two major differences between the Red Cross’s response and other companies, namely Yahoo. The first, which is most likely the most important, is that notification was almost immediate. Within the same month as the discovery of the attack, affected parties were made aware of exactly what kinds of data were exfiltrated. The second difference is the very nature of the announcement.

Comparisons: Juxtapositions between Responses

To begin, Yahoo’s breach response plan is quite difficult to pull information from as it could have taken place anywhere between the date of the December 2014 breach to the subtle announcement in 2016. Alongside this, the investigation of the breach took place behind closed doors meaning that public information on the breach is extremely limited, specifically with respect to the actions Yahoo took to mitigate damages. One key detail that can be extracted from

the fact that things happened so silently and away from the public eye, is that there were damages Yahoo was aware about and actively trying to mitigate. The damages I speak of here are the issues that arise from public perception of the company. A massive part of cyber-attacks that affect millions of users all from different backgrounds is that the perception of any company that undergoes such an attack means risking future investments and profits by releasing the information to the public. The very detail that best describes the incident response plan of Yahoo is that all was done in private, as opposed to the ICRC who almost from the very beginning, announced the attack and its scope. There are positives that arise from Yahoo's plan of concealing the cyber-attack but most of them protect the company solely. The negatives, of course, follow from the fact that affected consumers stay unaware of the potential damage coming their way by means of the exfiltrated data.

The example with Facebook is certainly trickier to break down in the same light. This is because the breach was perpetrated by the company itself, affecting it's users directly from the top-down. While the timeline of the breach did not start with a cyber-attack, future companies can learn from this example. The main lesson derived from Facebook is their treatment of consumer data, specifically the thought process is that consumer data is property of Facebook so it is their choice on how to command it. This corporate ideology is clearly dangerous to consumers which is reflected by the GDPR's lengthy restrictions on the use of user data, in all states it undergoes.

Path to Recourse

It's clear that there are certainly steps modern companies today can take to not only protect themselves from future cyber-attacks but also their customers. Breach notification laws exist for a reason: to defend the consumer. While a buffer period between discovery of an attack

and notification is necessary to exist as to prevent fear, uncertainty, and misinformation–companies need time to examine the events that took place so they can guide the public on steps to take to mitigate future damages, while patching their own current damages–that buffer period, in the case of Yahoo, cannot be needlessly long and indefinite. The sooner users are aware of a breach means that they can freeze credit scores, bank accounts, and employ other measures to stop the damage in its tracks.

The other way in which the cyber-world becomes a safer place is through corporate perception of data. Companies in many cases and industries require some form of data processing for internal modeling, and for the betterment of the consumer's lives. However, that should not mean that consumer data is manipulable in any way a company sees fit (i.e. for their own profit). This idea is directly reflected in the GDPR in all modes of data transfer. The regulation controls everything from how companies should secure consumer data to how they should transfer it.

While solutions can and in many cases already have been found for the problems that allowed these breaches to happen as they had, trusting companies to implement them in a vacuum will likely not result in much change. That is the entire reason that the GDPR was needed in the EU, and an equivalent is still needed in the US. It's clear that the party responsible for overseeing these changes is the government itself through regulation and policy. In a perfect world, Yahoo would have discovered the breach, announced it to its consumers, and taken further steps to prevent a cyber-attack of that scope from taking place again but the world we live in is not perfect. The company took deliberate inaction to protect its own interests first, directly affecting millions of users. Therefore, through government action alone can companies be trusted to take the proper, ethical approach to information security.

Specifically speaking with respect to the US government, they have been extremely slow historically to meet the demands of emerging technologies. From the time the Ford Model T was released, it took over 50 years for pedestrian protection acts to come about on a federal level, namely through the National Traffic and Motor Vehicle Safety Act in the mid 60s. From 1837, when the telegraph was invented, all the way through the creation of the first telephone and radios, up to 1934, the US government hadn't implemented regulation for communications, even as they quickly became the most tele-dense nation at the turn of the first world war. The FCC was created which proved as an effective measure through the Communications Act of 1934 to regulate the industry while protecting consumers from unlawful and unethical business practices. History seems doomed to repeat itself and in the case of information security, the everyday consumer can pray that the repetition is over and it's time for nation-wide policy to step in and take the reins of information security out of the hands of companies, and into the hands of the people.

References

Chin, Kyle. “Biggest Data Breaches in US History.” *UpGuard*, 5 Aug. 2022, www.upguard.com/blog/biggest-data-breaches-us.

Cox, Joseph. “Yahoo “Aware” Hacker Is Advertising 200 Million Supposed Accounts on Dark Web.” *Www.vice.com*, 2016, www.vice.com/en/article/aeknw5/yahoo-supposed-data-breach-200-million-credentials-dark-web.

“Form 10-K.” *Sec.gov*, 2016, www.sec.gov/Archives/edgar/data/1011006/000119312517065791/d293630d10k.htm.

FTC. “FTC Imposes $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook.” *Federal Trade Commission*, FTC, 24 July 2019, www.ftc.gov/news-events/news/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions-facebook.

FTC. “FTC Sues Cambridge Analytica, Settles with Former CEO and App Developer.” *Federal Trade Commission*, 23 July 2019, www.ftc.gov/news-events/news/press-releases/2019/07/ftc-sues-cambridge-analytica-settles-former-ceo-app-developer.

FTC. “Sophisticated Cyber-Attack Targets Red Cross Red Crescent Data on 500,000 People.” *International Committee of the Red Cross*, 19 Jan. 2022, www.icrc.org/en/document/sophisticated-cyber-attack-targets-red-cross-red-crescent-data-500000-people.

Harris, Kamala D. *Data Breach Report 2012*. 2013.

Heiligenstein, Michael. “Facebook Data Breaches and Privacy Violations: Full Timeline through 2021.” *Firewall Times*, 5 Oct. 2023, firewalltimes.com/facebook-data-breach-timeline/.

International Committee of the Red Cross. “ICRC Cyber-Attack: Sharing Our Analysis.” *International Committee of the Red Cross*, 15 Feb. 2022, www.icrc.org/en/document/icrc-cyber-attack-analysis.

International Committee of the Red Cross. “Cyber-Attack on ICRC: What We Know.” *Icrc.org*, 21 Jan. 2022, www.icrc.org/en/document/cyber-attack-icrc-what-we-know.

NIST. “NVD - Cve-2021-40539.” *Nvd.nist.gov*, 2021, nvd.nist.gov/vuln/detail/cve-2021-40539.

Office of Public Affairs. “U.S. Charges Russian FSB Officers and Their Criminal Conspirators for Hacking Yahoo and Millions of Email Accounts.” *Justice.gov*, 15 Mar. 2017, www.justice.gov/opa/pr/us-charges-russian-fsb-officers-and-their-criminal-conspirators-hacking-yahoo-and-millions.

Perlroth, Nicole. “All 3 Billion Yahoo Accounts Were Affected by 2013 Attack.” *The New York Times*, 3 Oct. 2017, www.nytimes.com/2017/10/03/technology/yahoo-hack-3-billion-users.html.

Peterson, Andrea. “Yahoo Discovered Hack Leading to Major Data Breach Two Years before It Was Disclosed.” *Washington Post*, 2016,

www.washingtonpost.com/news/the-switch/wp/2016/11/10/yahoo-discovered-hack-leading-to-major-data-breach-two-years-before-it-was-disclosed/.

U.S. Securities and Exchange Commission. “SEC.gov | Altaba, Formerly Known as Yahoo!, Charged with Failing to Disclose Massive Cybersecurity Breach; Agrees to Pay $35 Million.” *Sec.gov*, 24 Apr. 2018, www.sec.gov/news/press-release/2018-71.

“Wayback Machine.” *Web.archive.org*, 10 Jan. 2017, web.archive.org/web/20170110014942/investor.yahoo.net/secfiling.cfm?filingID=1193125-16-764376&CIK=1011006. Accessed 11 May 2024.

Wikipedia Contributors. “Yahoo!” *Wikipedia*, Wikimedia Foundation, 12 Jan. 2020, en.wikipedia.org/wiki/Yahoo.