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

Maximiliano Brizzio
Stevens Institute of Technology
CS 595: Information Security and The Law
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 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

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

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

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

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 that Yahoo waiting 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 because 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. Sure, an argument can be made that because the data was not 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 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

To bring things together in a 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 targets 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. **[TO BE CONTINUED]**