A Virtue Ethics Analysis of the 2013 Yahoo Data Breach

STS Research Paper
Presented to the Faculty of the
School of Engineering and Applied Science
University of Virginia

By

Youssef Errami

April 11, 2020

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: _____Youssef Errami_____

Approved: _____ Date _____
Benjamin J. Laugelli, Assistant Professor, Department of Engineering and Society

Introduction

During December of 2016, the American web services provider Yahoo came out for the first time and said that more than one billion Yahoo accounts were compromised back in August of 2013. Further examination into this breach led Yahoo to believe that this hack was "state-sponsored", meaning that this breach was executed by a group of individuals who were sponsored by a forgien government (Thielman, 2016). About a year later, Verizon Communications, Yahoo's parent company, said that the 2013 hack affected all three billion Yahoo accounts, stealing billions of names, phone numbers, birthdates and passwords. This single data breach is considered the biggest data breach of all time (Perlroth, 2017).

Upon further investigation, it has been determined that Yahoo is not free from any blame for the occurrence of this data breach. Yahoo was still using an outdated data encryption scheme known as MD5 (Goodin, 2016). MD5's weaknesses have been exposed for a decade by the time that this breach occurred. With this, it is clear that Yahoo failed to protect user's data to the best of its ability and it can be inferred that they did not make security a high priority. This data breach is often brought up in software engineering courses during the lectures on security but what is often left out of this conversation are the morals that Yahoo embodied during this time of ignorance.

I believe that analyzing the morals of Yahoo from a virtue ethics perspective will be the best way to judge its morality. I will demonstrate this by showing that the actions, or lack thereof, undergone by Yahoo are morally unacceptable due to lacking two of Prichard's eleven 'Virtues for Morally Responsible Engineers': openness to correction and commitment to quality

(Prichard, 2001). These lack of characteristics embodied by Yahoo are evident through the manner in which they allowed for this breach to happen and in how they handled the breach.

Background

The main reason the Yahoo 2013 data breach happened due to Yahoo's using the MD5 algorithm for data encryption. MD5, which stands for Message-Digest algorithm 5, is a widely known cryptographic hash function. Hash functions are functions that map data to a specific value, where these values now represent the "hashed" version of the original data it was mapped from. These functions are used because they have a property known as "one-way". This property makes it "virtually impossible to recreate the input data from the hash alone", which is useful for sensitive data, such as passwords because even passwords that are very similar (such as *password123* and *password122*) will return very different hash values if they are passed in to the MD5 algorithm (Kiennert et al., 2015). However in 2004, it was shown that the MD5 algorithm failed one of the most important tests of a cryptographic hash function: collision resistance. Collision resistance is the idea that no two input messages should hash to the same value, but this was sadly the case with the MD5 (Hawkes et al., 2004). This vulnerability in the hash function is what allowed the hackers to breach Yahoo's private user data.

Literature Review

There are a great deal of scholarly articles that break down the aftermath of the Yahoo data breach. These articles focus on why Yahoo was attacked and what happened to Yahoo post-breach. In order to avoid similar problems in the future, it is often critical to answer both why and how this event occurred so that others can learn from this situation and make sure it

doesn't happen to them. With that said, these works avoid making any moral judgements about Yahoo, who played an integral role in this situation happening in the first place.

In *Digging Deeper into Data Breaches: An Exploratory Data Analysis of Hacking Breaches Over Time*, Hammouchi, Cherqi, Mezzour, Ghogho and El Koutbi dive deeper into the history of data breaches. They mention how Yahoo might have been targeted by this attack not due to their brand, but because the attackers were aware that they would be the least prepared for such an attack (Hammouchi, Cherqi, Mezzour, Ghogho & El Koutbi, 2019). They continue to mention that the information gained by the hackers in the Yahoo attacks include information such as names, telephone numbers, addresses, emails and passwords, which can lead wide scale repercussions such as this data being sold and used for mass spamming and advertising campaigns (Hammouchi et al., 2019). While Hammouchi et. al do recognize the massive repercussions of this Yahoo data breach, they don't judge the morality of Yahoo.

In *The Impact of Cyber Attacks On Brand Image*, Whitler and Farris discuss the aftermath of Yahoo's data breach from a brand image perspective. They mention how Yahoo's brand and business were affected significantly due to Yahoo's mishandling of the situation (Whitler & Farris, 2017). Yahoo's "slow disclosure of notification to users" and its "inconsistency in the proxy filings" lead to US lawmakers demanding clarity while also being bombarded with a multitude of lawsuits. At the time, Yahoo was also worried that this would derail its merger plans with Verizon because on the day they notified the public that this breach occurred, its stock price lost $1.5 billion in market value (Whitler & Farris, 2017). It was during these times in which Yahoo's public image was damaged significantly. Whitler and Farris do an

excellent job of explaining how Yahoo's brand was affected but they don't engage in any moral judgement of Yahoo.

While there is much to learn from how and why this horrific data breach took place, it is also important to hold the companies and engineers that create the software we use on a daily basis accountable for their actions. Future engineers can learn a lot from engaging in moral judgement of the actions committed by Yahoo. This paper will both engage in descriptive judgement of Yahoo's engineering practices and use a virtue ethics framework in order to morally judge the actions of Yahoo.

Conceptual Framework

I will analyze the morality of Yahoo using a virtue ethics framework. Virtue ethics is "an ethical theory that focuses on the nature of the acting person" where the theory "indicates which good or desirable characteristics people should have or develop to be moral" (van de Poel & Royakkers, 2011). It is important to recognize that this ethical framework is backed by the idea that humans are rational by nature and that they should use reason to determine how to live in a both moral and virtuous manner. Along with this note, it is understood that virtue is the mean between two extremes. An example of this would be how courageous action is the mean between a cowardice action and a reckless action. Virtuous activities involve having moral virtues. People aren't born with moral virtues, they are developed by the deeds and actions they commit (van de Poel & Royakkers, 2011). Virtue ethics is often used to analyze the morality of individuals but in this paper, the goal is to analyze the morality of Yahoo, which is a company. However, it is still possible to use virtue ethics as a framework if the notion of collective responsibility and the problem of many hands are considered. Collective responsibility is the "responsibility of a

collective of people" and was created to deal with "the intuition that there is more to responsibility in complex cases than just the sum of the responsibilities of the individuals considered in isolation" (van de Poel & Royakkers, 2011). This idea of collective responsibility can be used to help define the problem of many hands. The problem of many hands is "the occurrence of the situation in which the collective can reasonably be held morally responsible for an outcome, while none of the individuals can be reasonably held responsible for that outcome" (van de Poel & Royakkers, 2011). This can be applied to this case study by stating that no individual software engineer at Yahoo can be held morally responsible for this data breach but Yahoo as a collective can be held morally responsible. Therefore it is now possible to analyze the morality of Yahoo using a virtue ethics framework

Virtue ethics defines actions as morally acceptable if and only if they are the same actions that a virtuous agent would do if they were in that situation. In this paper, a virtuous agent will be one that abides by Prichard's eleven 'Virtues for Morally Responsible Engineers' (Prichard, 2001). The eleven virtues are shown in Figure 1.

1. Competence
2. Ability to communicate clearly and informatively
3. Cooperativeness (being a good "team player")
4. Willingness to compromise
5. Perseverance
6. Habit of documenting work thoroughly and clearly
7. Commitment to objectivity
8. Openness to correction (admitting mistakes, acknowledging oversight)
9. Commitment to quality

10. Being imaginative
11. Seeing the “big picture” as well as the details of smaller domains

Figure 1: Prichard's eleven 'Virtues for Morally Responsible Engineers'

It is important to mention that Prichard states how embodying all eleven of these virtues “is not sufficient” enough to claim that the engineer, or in the case of this paper, a company and its engineers, are morally responsible. However if they are lacking any one of these virtues, it is enough to detract from responsible engineering practice (Prichard, 2001).

For this paper, I will be connecting the idea of collective responsibility and the problem of many hands with Prichard's eleven 'Virtues for Morally Responsible Engineers' in order to provide moral judgement of the actions of Yahoo. In the analysis section of this paper, I will analyze the actions made by Yahoo with respect to two of the virtues described by Prichard: openness to correction and commitment to quality (Prichard, 2001).

Analysis

Yahoo has acted in a morally irresponsible way due to its lack of two of the eleven virtues described by Prichard: openness to correction and commitment to quality. As mentioned before, missing any one of these values makes it impossible to engage in responsible engineering practice according to Prichard (Prichard, 2001). From a virtue ethics perspective, Yahoo were not virtuous agents due to the decisions they made that led to this data breach being possible and its actions should be considered morally irresponsible. The subsequent subsections of this

analysis section will involve taking each one of these missing virtues and detailing the actions and decisions that Yahoo made that indicate the absence of said virtue.

Openness to correction

The first virtue missing from Yahoo's engineering practice is an openness to correction. Prichard broadens the meaning of openness to correction by stating that this also means "admitting mistakes" or "acknowledging oversight" (Prichard, 2001). The portion this paper is concerned with is Yahoo's inability to properly admit its mistakes.

Yahoo's massive data breach occurred in August of 2013, but it wasn't publicly announced to the world that this was the case until December of 2016. Almost 3.5 years passed by until there was a public statement from Yahoo indicating that its system was breached and that nearly three billion accounts were affected by the breach (Perlroth, 2017). With all of that said, it came out that Yahoo found out about this data breach back in 2014 and did not say anything about it to the public (Trautman & Ormerod, 2017). Yahoo has refused to publicly respond to why it took them three years to notify the public about this data breach or why they did not let the public know sooner (Rohlf, 2017). This refusal to publicly comment on this breach shows that Yahoo does not value users safety enough to let them know as soon as they found out about the breach. If Yahoo embodied the openness to correction virtue, they would've notified the public as soon as they found out instead of waiting three years to notify the public.

This period of time between when Yahoo found out about the data breach and its public statement revealing the data breach shows a significant lack of openness to correction. This three year period that they spent not letting the public know that its data has been breached could have affected billions of users other private accounts. Studies have shown that "59% of people use the

same password everywhere" (Truta, 2018). With a leak of this magnitude, where the hackers gained access to the three billion user's passwords and emails, hackers could have theoretically gained access to many of the users other accounts for other websites. With 59% of people using the same password everywhere and having access to their email and password through the data breach, the hackers could now try and gain access to users accounts on other websites using the email and password they gained from the breach and it would theoretically work more than half the time.

As I have argued, the three years that it took Yahoo to disclose this data breach to the public is unacceptable and a direct violation of an openness to correction. Some might think that the three year wait period was necessary in order to truly investigate the system internally before letting the public know that this event has occurred in order to try and get ahead of the problem at hand. But this view fails to consider the importance of notifying the public in a timely fashion. There is no excuse to wait three years to announce the biggest data breach in history. The amount of people who had their private information leaked to these hackers are immense and many of them are still dealing with the backlash of these hacks to this day. The long-term effects of this breach are massive due to the data that was breached has the possibility of being used for "identity theft or other impersonation scams" and the worst part is that "once such personally identifiable information gets lost, furthermore, it can never be retracted, posing long-term risks especially in relation to static data, such as birth dates" (Kirk, 2017). Having this kind information available to hackers can lead to widespread identity theft. Identity theft can ruin peoples lives to the extent that it can never be repaired. People's lives may be ruined due to Yahoo's inability of letting its users know that their data has been compromised. With this data

breach affecting billions of accounts, the scale at which how many people were affected is massive, and waiting three years to publicly announce this tragedy is unacceptable.

If Yahoo told the public as soon as they realized that there was a data breach, users could have started to take the precautionary steps towards securing their information and changing their passwords on the other web applications they may use. The reality of the situation is with Yahoo waiting three years to make a public announcement, most of the damage has probably been done. All users can do is secure their information as best as they can and hope that this data breach does not come back to haunt them in the form of identity theft or some other scam. Yahoo's lack of admitting its mistake in a timely fashion is representative of its lack of openness to correction, which is not representative of what a virtuous agent would do.

Commitment to quality

It is evident that Yahoo did not commit to quality. Security and privacy go hand in hand with the software quality. Security and privacy measures implemented in a software system are there strictly to maintain the safety of said system and its users private data. Poel and Royakkers describe four different strategies that engineers can follow in order to ensure safe products: Inherently safe design, Safety factor, Negative feedback mechanism and Multiple independent safety barriers. Inherently safe designs involve "an approach to safe design that avoids hazards instead of coping with them". A safety factor involves "a factor or ratio by which an installation is made safer than is needed to withstand either the expected or the maximum (expected) load." Negative feedback mechanisms are mechanisms that "if a device fails or an operator loses control assures that the (dangerous) device shuts down." Lastly, multiple independent safety barriers involve a chain of "safety barriers that operate independently of each other so that if one fails the others do not

necessarily also fail" (van de Poel & Royakkers, 2011). The three safety strategies that are relevant here are inherently safe design, safety factor and multiple independent safety barriers where a negative feedback mechanism is less relevant to this case study and software security in general. Yahoo failed to implement any of the three relevant safety strategies.

Yahoo's failure to engage in an inherently safe design when it came to its security measures is highly morally irresponsible. Inherently safe designs involve avoiding hazards instead of coping with them and Yahoo decided to forgo this strategy and cope with them instead. Yahoo simply failed to employ adequate security measures due to its lack of up-to-date security measures. Back in 2013, Yahoo was still using the MD5 data encryption scheme (reference Background section for more details), which had a major vulnerability, known as collision resistance, that was made public nearly a decade before the incident occured. It is impossible that Yahoo nor any of the highly educated engineers that they have were unaware of this groundbreaking discovery of this vulnerability that warranted security researchers from the Carnegie Mellon University's Software Engineering Institute to end up suggesting that MD5 "should be considered cryptographically broken and unsuitable for further use" ("MD5 vulnerable to collision attacks", 2008). Given this information, Yahoo should have changed its encryption scheme right away. Knowing that this encryption scheme is broken and choosing not to do anything not only is horrible engineering practice, but also violates IEEE Code of Ethics for software engineers as they are no longer acting in the best interests of their clients and users (Code of Ethics: IEEE). This is a prime example of how instead of trying to avoid hazards, Yahoo attempted to cope with them instead and it ended up coming back to haunt them.

Yahoo also failed to implement a safety factor strategy when it came to the safety and security of its system. A safety factor strategy involves usually creating systems that are stronger, or in our case safer, than they need to be in order to avoid safety issues in the future. Yahoo failed to execute this strategy which is evident through its unwillingness to update its security measures to a more cryptographically secure encryption scheme as opposed to sticking with a cryptographically broken MD5. One of the basic principles of data security is "the need to stay abreast of technological developments and maintain satisfactory security controls" (Rohlf, 2017). Yahoo failed to do this by relying on an outdated data encryption scheme instead of maintaining satisfactory security controls. Instead, they put billions of accounts at risk by refusing to make its system safe and secure for users.

Yahoo also did not have any sort of multiple independent safety barriers implemented anywhere into its system. This severe lapse in judgement could have saved billions of accounts from being breached. Having multiple levels of security barriers could have prevented the hackers from being able to breach into Yahoo's system or it also could have given engineers enough time to realize that hackers are trying to breach its system. The benefits of having independent safety barriers could have been massive for Yahoo. Even with its usage of MD5, another independent safety barrier could have saved the company from a massive data breach but instead, they relied on one cryptographically broken safety barrier.

Lastly, another way that Yahoo failed to show commitment to quality was through its lack of commitment to making security a company priority. There were reports at the time that said that when "Yahoo's security team requested new tools and features to strengthen Yahoo's security, they were turned down because Yahoo was concerned such requests were too costly or

complicated" (Rohlf, 2017). As Yahoo grew larger and larger, security seemed to still be on the backburner of matters they were concerned with as the company didn't prioritize security. Yahoo was and still is valued at over a billion dollars in net worth and they were still worried about security being "too costly or complicated". The fact that Yahoo forwent security and the safety of its users in order to save costs shows an incredible lack of commitment to quality and is an obvious contradiction to what a virtuous agent would do.

Security, safety and quality are all intertwined facets when it comes to building large pieces of software, therefore a commitment to quality also means a commitment to safety. As shown through the examples listed in this section, it is abundantly clear that Yahoo failed to commit to providing its users with a more secure system to make its users data safe from data breaches. Therefore, due to Yahoo's lack of commitment to safety, it is safe to say they lacked the virtue of commitment to quality.

Conclusion

Although the blame for the 2013 Yahoo data breach is often placed and pointed directly at the hackers who executed these attacks, it is possible to place blame onto the plates of Yahoo themselves as well. The actions and decisions Yahoo made throughout its engineering teams reveal notable shortcomings with respect to the two virtues, openness to correction and commitment to quality, that are necessary to be a morally responsible engineer, or in this case, a morally responsible engineering company. With a virtue ethics framework, it can be seen that the actions undergone by Yahoo are morally irresponsible and not an example of what it means to be a virtuous agent.

This case study shows society that companies should be held partially responsible for attacks on its own software if they had the ability to prevent this from happening. Companies have the moral responsibility to protect its users private data and should be adequately judged when they fail to do so. After reading this paper, companies should reevaluate their engineering process and make sure that their engineering teams embody the virtues for morally responsible so that they can be the virtuous agents we need in our society.

Word Count = 3670

References

Code of Ethics: IEEE Computer Society. (n.d.). Retrieved from https://www.computer.org/education/code-of-ethics

DiChristopher, T. (2016, July 25). Verizon to acquire Yahoo in $4.8 billion deal. Retrieved from https://www.cnbc.com/2016/07/25/verizon-to-acquire-yahoo.html

Goodin, D. (2016, September 22). Yahoo says half a billion accounts breached by nation-sponsored hackers. Retrieved from https://arstechnica.com/information-technology/2016/09/yahoo-says-half-a-billion-accounts-breached-by-nation-sponsored-hackers/

Hammouchi, H., Cherqi, O., Mezzour, G., Ghogho, M., & El Koutbi, M. (2019). Digging deeper into data breaches: An exploratory data analysis of hacking breaches over time. *Procedia Computer Science*, *151*, 1004–1009. doi: 10.1016/j.procs.2019.04.141.

Hawkes, P., Paddon, M., & Rose, G. G. (2004). Musings on the Wang et al. MD5 collision. *IACR Cryptology ePrint Archive*, *2004*, 264.

Kiennert, C., Bouzefrane, S., & Thoniel, P. (2015, April 10). Authentication systems. Retrieved from https://www.sciencedirect.com/science/article/pii/B9781785480041500031?via=ihub

Kirk, J. (2017, February 22). Yahoo takes $350 million hit in Verizon deal. Retrieved from https://www.bankinfosecurity.com/yahoo-takes-350-million-hit-in-verizon-deal-a-9736

MD5 vulnerable to collision attacks. (2008, December 31). *Carnegie Mellon University's Software Engineering Institute,* Retrieved from https://www.kb.cert.org/vuls/id/836068/

Perlroth, N. (2017, October 3). All 3 billion Yahoo accounts were affected by 2013 attack. Retrieved from https://www.nytimes.com/2017/10/03/technology/yahoo-hack-3-billion-users.html

Pritchard, M. (2001). Responsible engineering: The importance of character and imagination. *Science and Engineering Ethics*, *7*(3), 391–402.

Rohlf, M. (2017, March 14). Yahoo data breaches: A lesson in what not to do. Retrieved from https://www.bytebacklaw.com/2017/03/yahoo-data-breaches-a-lesson-in-what-not-to-do/

Thielman, S. (2016, December 15). Yahoo hack: 1bn accounts compromised by biggest data breach in history. Retrieved from https://www.theguardian.com/technology/2016/dec/14/yahoo-hack-security-of-one-billion-accounts-breached

Trautman, L. J.; Ormerod, P. C. (2017). Corporate directors' and officers' cybersecurity standard of care: The yahoo data breach. *American University Law Review, 66(5),* 1231-1292.

Truta, F. (2018, May 3). 59% of people use the same password everywhere, poll finds. Retrieved from https://securityboulevard.com/2018/05/59-of-people-use-the-same-password-everywhere-poll-finds/

van de Poel, I., & Royakkers, L. (2011). Ethics, technology, and engineering: An introduction. Hoboken, NJ:Blackwell Publishing Ltd.

Whitler, K. A., & Farris, P. W. (2017). The impact of cyber attacks on brand image. *Journal of Advertising Research*, *10*(2501). doi: 10.2501/JAR-2017-005