

The Paradox of Digital Trust

A Research Paper Submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia - Charlottesville, Virginia


In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering


By

James Foster

Spring, 2022

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:  _____ Date 20 April 2022
James Foster

Signed:  _____ Date 10 May 2022
Richard D. Jacques, Ph.D., Department of Engineering & Society

Introduction

Consumer trust is vital for the success of any business. If people distrust a company or their practices then they will avoid both purchasing their goods and using the services provided by them. This has been the dynamic for thousands of years. However, in the age of industrialization and with the rise of mega-corporations, this practice is becoming more difficult to uphold. Companies today, especially online, are diversifying and carving out large portions of the market for themselves. Acquisitions and mergers are becoming more commonplace and large companies are able to expand immensely. Instances such as Facebook's purchase of WhatsApp for \$19 billion (Covert) allow companies to hold more market share and influence over consumers than ever before. This makes it much more difficult for users to avoid using the services provided by these companies. This results in the first paradox of digital trust identified in this paper: users purchasing goods and utilizing the services provided by entities they deem untrustworthy. Users are also impacted by the big data that mega-corporations populate their servers with. This data being the perfect target for hackers who want to steal and sell it for profit, emulating the revenue model of the corporations themselves. Company protocols regarding users' personal data are unclear and there seem to be no common guidelines that companies must follow regarding the security of this data. There is no clear cut method to determine where personal information is being stored and whether it will be deleted if a user closes their account with that company. There is also little knowledge about the protections in place that might mitigate the possibility of a hack on a corporate server. All of this culminates in the second paradox of digital trust of this paper: the fact that users' personal information is being stored on corporate servers. A commonplace practice with unfounded impetus. This paper aims to describe these two paradoxes in further detail and provide guidance as to the next steps society should take to prevent them from enduring into the next generation of the big technology industry.

Paradox I: Utilizing the services of entities deemed untrustworthy

People purchase goods and utilize services provided by entities they recognize as untrustworthy. This has been a common practice in society since the beginnings of job specialization. The difference is that now it seems almost impossible to carry on day to day without using services provided by technological giants like Google, Amazon, and Facebook. The utility of the services along with the multitude of active users influence people to conform and become active users of those services as well. This, by itself, is nothing out of the ordinary, the real problem is when a company's growth continues even in the wake of data breaches and scandal.

Facebook is a platform for connecting with others and sharing personal information. It created the novel concepts of "likes" and user feeds that helped it grow exponentially. It has seen its monthly users continually increase every year since its launch (Richter) even while the company has dealt with dozens of data breaches and user privacy concerns. These stories are headlines, yet they seem not to sway the layperson's opinion on Facebook because there is simply no other social media service that has interconnected the world like Facebook has, it is a unique and unavoidable service. Another tech giant, Amazon, is a one stop online superstore that has some of the fastest parcel shipping in the country. It has been mired with allegations of anti-competitive practices and poor treatment of workers. One such anti-competitive practice is Amazon studying the data of successful businesses that use its platform and deducing which products they could produce in-house for a lower cost (Tonner). This is the backbone of the Amazon Basics line of low-priced items that resulted in small businesses being muscled out of the Amazon ecosystem. Even with this news public, Amazon has seen revenue steadily increase each year over the past decade (Coppola). This is because, like Facebook, Amazon provides a unique service that no one else has been able to replace. A final example is that of Google; a company that provides a search engine to help users navigate through the internet. However, behind the scenes, user searches are recorded and individuals are profiled so that advertisements, provided by Google,

encountered online are personalized to that user. Google joins Amazon and Facebook by also providing a service that has been unmatched for a decade and has had constant growth since it was introduced to the internet. These technology behemoths are akin to monopolies at this point, each in their own industry. All of these businesses are exemplary of the first paradox of digital trust: they have grown so large and provide services so beneficial to the average person, that these people cannot resist the temptation to employ them, even when the businesses in question are rife with scandal. Certainly, both the businesses and people using their services can recognize this paradox and make individual changes to reduce its frequency.

Corporate giants such as Facebook, Amazon, and Google have an ethical obligation to be trustworthy and respectful of their users. Here are some possible avenues for rebuilding genuine trust with their users. First, these businesses can hire penetration testers to regularly let them know where their servers are weakest and how to harden them to stave off data breaches and malicious hacks. They can also engage in more trustworthy business practices and encourage more fair competition in the marketplaces they moderate. All user accounts and preferences should begin with the most privacy focused settings turned on. This makes it so that laypersons can have the personal data protections they deserve by default, without having to conduct any research to disable them. Businesses should also create concise, simplified versions of their privacy policies so that users can quickly read and understand them. Including practices like these into the business flow of these large companies would immensely raise public trust in them and likely not affect their revenue. Users are the second half of this paradox and have obligations to themselves to protect their privacy and personal data. Some advice for them is that they should put in the effort to research alternative businesses that conform to better specifications of user privacy and data security. For instance, users can read corporate privacy policies or at least know the core tenets of privacy that the business follows to help determine whether to support that business. By doing this, users are effectively boycotting the more dangerous businesses and coercing them to implement better business practices. This is a difficult choice to make, however,

since the novelty and utility of big technology company services are hard to do without. This makes it even harder for users to influence business practices, as most will not abstain from using the services provided by businesses they recognize as untrustworthy. Thus, data breaches will continue to occur as businesses who use the client-server model choose to not reevaluate their privacy policies. This might be avoided if these businesses did not have to store users' personal information on their corporate servers.

Paradox II: Personal information stored on corporate servers

Ever since the creation of the internet, the predominating online service architecture has been the client-server model (Yadava). Users, clients, connect to servers via the internet and query for resources and data. Take for instance an email service. An email user logs in to an online portal that verifies their account details and shares all the emails associated with that account with them. Logins can be verified since the account data for every user of the service is stored on the email server. The emails do not persist on the user's computer so they must be fetched from the email server each time the user logs in or an email is sent to the user. The email server is an example of a centralized system that can serve many clients. It can usually serve millions of client requests immediately and its hardware must scale to meet these demands. This architecture is the most widely used because of its relatively simple setup and low maintenance. Facebook, Amazon, and Google all use this architecture for certain services they provide. The paradox here is that personal data is private, yet it is copied and exists for an indefinite amount of time on corporate servers, innately unsafe locations for storing data. A drawback of this architecture is the multiple attack vectors that exist for malicious actors since users must always retrieve their sensitive personal data from the remote server. Man in the middle attacks and data breaches are the most common. A man in the middle is where an attacker sniffs packets that are sent from the client to the server and can read their data if unencrypted. Data breaches are direct attacks on a server that hosts a large amount of data. This is usually personal data related to users who use the service that the business that owns the server provides.

With such simple yet devastating attacks available, it is bewildering to imagine how this architecture has managed to still be widely adopted even today. This paper identifies two main reasons. The first is that companies themselves feel little impact from data breaches other than a social stigma that develops against them due to their poor data security. Monetarily, they do not care if personal info is stolen since they have backups of the data and can still sell it and the stats derived from it for a profit. For instance, when the FTC fined Facebook \$5 billion over the ramifications of their data breaches, the Facebook stock price went up (Stewart). Investors understand that the larger a business is, the higher the chance that their servers are hacked, however the legal penalties for disregarding crucial data security seems to only elicit a small fine for companies earning trillions in revenue. Further, these fines placed on companies as a result of data breaches usually trickle down to affect shareholders and workers (Solove), which lessens the intended severity of the fines. Many businesses have recognized and believe that data breaches will occur regardless of any security protocols; so, it is preferable to pay a one-time fine rather than reinvent their data security from the ground up. The second reason that the client-server model persists is that consumers are numb to the effects of data breaches and feel as if they have no control over their data's privacy. Along with this comes the group think where consumers believe they are obligated to share their data if they want to use online services (Imperva). One common example of this is web browser cookies. Nearly every modern website uses cookies to persist certain settings to make the user experience more consistent when visiting the same page again or being logged in with an account. Most users accept the cookies without even determining what the cookies will be used for and what data they expose from the client's web browser to the remote server. This is antithetical to the privacy concerns that users have with their data. There is a trust deficit where "86% of consumers are worried about data theft and its consequences" (Imperva), but they choose to ignore this axiom when it is convenient for their web browsing. Web browser cookies are just one in a myriad of ways for users to lose their data by their own compliance. Due to the prevalence of technologies like cookies and the simplicity with which they can be activated and put user data at risk, 50% of

consumers believe that they cannot possibly check each provider's privacy and security track record to keep track of their sensitive data and how it's protected (Imperva). It is important to raise the awareness of these concerns among users, but systemic changes in business data security would be far more effective.

There are several preferable alternatives and modifications to the client-server model from the consumer's perspective that would increase its security and viability as a reliable architecture. The first and most costly solution would be to switch services over to a peer-to-peer architecture. This alternative architecture allows users of a service, peers, to connect directly to each other with no intermediary server. Each peer has all the data it needs to send to other peers stored locally on it so there is no need to query a corporate server for resources (Bauwens). The BitTorrent protocol and other file sharing services utilize the peer-to-peer architecture to obfuscate and decentralize data transfers. If there is no centralized server, it is a lot harder to perform man in the middle attacks or data breaches since peer to peer connections spin up unpredictably and are hardened covert channels as opposed to servers which are hosted at a particular IP address for an indefinite amount of time and are not innately secured. Another alternative could be a novel digital identity system, where users own their identity and the bits representing it on their devices and businesses simply reference the data when it is needed. In this scenario, identity data does not have to be copied and stored on a remote server, but can be queried for as if the server was a user requesting a resource on a user's device that acts as a server. Along with this, it would have to be legally punishable to copy a person's digital identity data, especially if a business was the thief. A significant benefit from this setup is that users would be able to toggle access to their digital identity on a case-by-case basis, according to which entities the user is comfortable with referencing their data. Both this solution and peer to peer architecture solution would certainly end this paradox as personal data would now be stored in a logical location inside its owner's devices. Attackers would have a much more difficult time trying to extract data from billions of individual devices rather than one corporate server. A last, radical suggestion for businesses using the

client-server model is to not collect data. If there is no data being collected by a business then there is no data that can be breached by malicious actors (Stewart). This, of course, is an implausible solution as big data companies make a majority of their revenue from collecting and selling data about their users. The magnitude of this solution would require a complete reevaluation of corporate data collection and provide an alternative source of income for big data firms which is out of the scope of this paper.

Conclusion

A new age of information is on the horizon as the concept of centralization and technologies under its flag, such as the client-server model, come under scrutiny. Digital trust is becoming increasingly more important as time goes on and people become aware of just how valuable their data is. In a world with logical and ethical data privacy policies, the two paradoxes mentioned in this paper should simply not exist. As users take ownership over their data and businesses reevaluate the ethics of selling big personal data these paradoxes will dissipate. It is paramount for users to recognize the power they hold and put their personal data behind a lock and key, since selling user data is the real money maker for many businesses. Peer to peer technologies emulate the flow of real-life interactions where data is passed from one entity to another without a centralized intermediary and giving both peers equal footing with regards to controlling the flow of data, providing a logical alternative to the client-server model. Referenced digital identities are one way for users to easily recognize their power and hands control over personal data to the rightful owners. Now is also the time for businesses to invest in new data privacy technology and write new protocols for handling user data. The cost of revealing security strategies to raise user trust should be evaluated as much less than that of the fines and social stigma associated with data breaches and attacks that result from careless privacy policies. This will require the government to create stricter corporate guidelines to protect consumer privacy and harsh legal repercussions for not heeding them. The time for the establishment to show that they value individual consumer rights over that of corporations is now. Consumers have a higher technological

literacy rate now than ever before and this has given them the ammunition to put pressure on those of whom they loan their personal data to. Ultimately, the dynamic between users and businesses will have to drastically change to provide new ways to handle personal data that respect the owner of the data and do not exploit their privacy for the sake of profit.

References

1. Bauwens, M., Kostakis, V., & Pazaitis, A. (2019). *Peer to Peer: The Commons Manifesto*. London: University of Westminster Press. DOI: <https://doi.org/10.16997/book33>
2. Coppola, D. (2022, May 3). *Net revenue of Amazon from 1st quarter 2007 to 1st quarter 2022 (in billion U.S. dollars) [Graph]*. Statista. Retrieved from <https://www-statista-com.proxy01.its.virginia.edu/statistics/273963/quarterly-revenue-of-amazoncom/>
3. Covert, A. (2014, February 19). *Facebook buys WhatsApp for \$19 billion*. CNNMoney. Retrieved from <https://money.cnn.com/2014/02/19/technology/social/facebook-whatsapp/index.html>
4. Heiligenstein, M. X., Jodi, & Robert. (2022, March 21). *Facebook data breaches: Full timeline through 2022*. Firewall Times. Retrieved from <https://firewalltimes.com/facebook-data-breach-timeline/>
5. Imperva. (2022). (rep.). *No Silver Linings Report*. Imperva.
6. Richter, F. (2021, February 4). *Infographic: Facebook keeps on growing*. Statista Infographics. Retrieved from <https://www.statista.com/chart/10047/facebooks-monthly-active-users/>
7. Solove, D. J., & Hartzog, W. (2022). *Breached!: Why data security law fails and how to improve it*. Oxford University Press.
8. Stewart, E. (2022, April 21). *Companies lose your data and then nothing happens*. Vox. Retrieved from <https://www.vox.com/the-goods/23031858/data-breach-data-loss-personal-consequences>
9. Tonner, A. (2016, April 27). *Amazon is quietly building its own private label Empire*. The Motley Fool. Retrieved from <https://www.fool.com/investing/general/2016/04/27/amazon-is-quietly-building-its-own-private-label-e.aspx>
10. Yādava Subhāsha Candra, & Singh, S. K. (2009). *An introduction to client/server computing*. New Age International (P) Ltd., Publishers.