The Struggle over Digital Privacy in the United States

An STS Research Paper
presented to the faculty of the
School of Engineering and Applied Science
University of Virginia

by

Alex Kwong

May 9, 2023

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.

Alex Kwong

STS Advisor: Peter Norton

The Struggle over Digital Privacy in the United States

Recent developments in machine learning (ML) offer new possibilities in automation that may transform numerous economic sectors and boost organizational efficiency. Yet ML also threatens to accelerate problematic trends, including invasion of personal privacy, unimpeded user data collection and monetization, and propagation of targeted misinformation (Kerry, 2020). Big data is in demand because recommendation systems are profitable. Services and products such as Amazon's e-commerce platform, Netflix's streaming platform, and Meta's advertisement system generate personalized recommendations (Leung et al., 2019).

In the European Union, the General Data Protection Regulation governs how personal information is collected and processed, but in the United States no laws protect data privacy (Levenberg, 2022). Data privacy advocates demand regulations to restrict the collection and use of personal data, but they face powerful opposition from data collectors. Data collectors' success at thwarting regulation, however, is not due solely to their financial resources and their political influence. They also skillfully deploy ideas and values such as free enterprise, user responsibility, personal convenience, and user experience optimization to legitimize their material interests.

Review of Research

Jain et al. (2016) propose means of securing big data to protect privacy. Mehmood et all. (2016) discusses technical techniques that protect big data (Mehmood). These techniques include access restriction, data encryption, different cloud storage options, and data processing precautions. He discusses the state-of-the-art privacy preservation mechanisms and explores possible future directions related to privacy preservation.

Narayanan (2016) discusses the risks associated with big data analytics. Narayanan discusses the idea that released data will get increasingly vulnerable to re-identification as time passes. Unlike the two previous pieces of research, this paper discusses big data policy and how policymakers should work to create more comprehensive guidelines to regulate data privacy protection.

The papers above and the majority of other research on big data privacy tend to focus on the technical aspects of data privacy protection techniques. There is a lack of information and review on the struggle over data privacy in the United States.

The Consumer Is the Product

Many online products and services are offered free of charge in exchange for user data (Tabora, 2019). In such a business model, we can say that "the consumer is the product" (Tabora, 2019). The term is descended in part from a 2010 observation: "If you are not paying for it, you're not the customer; you're the product being sold" (Lewis, 2010). For example, Meta, which owns social media platforms such as Facebook and Instagram, provides free products but assembles consumer data into ad-targetable groups and sells them to advertisers (Wagner, 2018). With user data, companies can forego slow, labor-intensive, and unreliable surveys. Instead, they

can automatically track consumer habits and interests and use the data to develop new products and improve existing ones. Big data can be sold for advertising, for market research, and for product development.

Data Privacy

Under federal law, data collection is unregulated. This means companies and organizations are not required to notify users when their data is shared with other entities by federal law. Agencies such as the Federal Trade Commission (FTC) provide guidelines for consumer privacy. These guidelines cover basic protections such as requiring informing consumers when their data is collected and giving the choice of how it is used (Levenberg, 2022). Some industries have additional FTC guidelines that adhere to laws such as the healthcare industry with the Health Insurance Portability and Accountability Act of 1996 (HIPAA).

According to the Pew Research Center, 63 percent of Americans say they understand very little or nothing at all about the laws and regulations that protect their data privacy (Atske, 2019). Of U.S. adults surveyed, 70 percent said they believed their personal data was less secure than it was 5 years earlier. Only 6 percent believed their data was more secure. This lack of confidence in the ability of organizations to protect personal data must be addressed if consumers are to feel secure in their online interactions.

Reasons for Big Data Distrust

Facebook Data Breach 2019

In 2019, phone numbers, full names, locations, and email addresses were made available in a public database from a Facebook data breach (Bowman 2021). Facebook chose not to notify over 530 million users that their data had been compromised. After the data was leaked to the public, Facebook stated that they would still not notify users individually. Although financial information, health information, or passwords were not included in the leak, Adam Levin, a cybersecurity expert and consumer protection advocate, says that "Scammers can do an enormous amount with little information from us" (Bowman 2021). This leak resulted in Facebook having to pay a $5 billion settlement to the FTC.

Marriot Data Breach 2014

In September 2018, an internal security tool flagged suspicious activity within the Marriot Starwood brand guest database (Fruhlinger 2020). It was found that user data had been compromised in a 2014 attack. Up to 500 million guest records were compromised which contained sensitive information like credit card and passport numbers. Marriot stated that they would cover passport replacement costs and fraudulent credit card expenses but has not provided any other type of compensation.

LinkedIn Data Scraping 2021

In 2021, the information of 700 million LinkedIn users was advertised for sale on a darknet forum (IdentityTheft.org 2023). By using LinkedIn's API, "TomLiner" was able to scrape email addresses, full names, phone numbers, physical addresses, geolocation records, and more from publicly accessible data. Unlike the previous examples which involved security

systems being compromised, this incident highlighted the lack of regulation regarding consumer privacy. While no sensitive information was exposed, the information is useful in phishing scams and social engineering attacks.

Push for Big Data Acceptance

Entities that are pro big data aim to push their agenda by highlighting the benefits of big data, addressing concerns, educating the public, collaborating with policy makers, and investing in big data research and development.

Council for Big Data, Ethics, and Society

One such entity is the Council for Big Data, Ethics, and Society (CBDES). The CBDES is a group of scholars, industry leaders, and policy makers that aims to address the legal and social issues with the collection and use of big data. It was founded in 2014 and is funded by the National Science Foundation (NSF).

Their agenda is to inform the public about the use of big data. The CBDES provides recommendations and guidelines for responsible usage of big data which include user privacy, data security, addressing bias, transparency, and accountability. They also conduct their own research and publish reports on those same topics. These include papers addressing the importance of transparency, informed consent, and data protection.

Data & Society

Data & Society is a research institute that studies the social and ethical implications of emerging technologies. The organization was founded in 2013 to investigate the intersection of technology and society and to provide recommendations for policymakers, industry leaders, and

the general public. Data & Society researches the collection and use of data by government agencies and private companies and how it impacts individual privacy and civil liberties.

Electronic Frontier Foundation

The Electronic Frontier Foundation (EFF) is a leading advocate for digital privacy rights. The EFF's agenda includes government surveillance, data breaches, corporate data collection, and online tracking. The advocacy has made efforts to protect privacy rights via taking legal action against government surveillance programs and providing legal aid to individuals whose private data is compromised.

Digital Advertising Alliance

The Digital Advertising Alliance (DAA) is a non-profit organization formed in 2009 which focuses on key issues related to online advertising. Its agenda is to promote responsible data collection for online advertising purposes with a focus on respecting consumer privacy and providing transparency around data usage.

American Civil Liberties Union

The American Civil Liberties Union (ACLU) is an organization dedicated to defending civil liberties and protecting individual rights. Originally founded by activists and lawyers concerned about government overreach in 1920, the ACLU now focuses on areas such as free speech, racial justice, LGBTQ rights, criminal justice reform, and privacy rights. As one of the key areas of focus, the ACLU is committed to protecting individual privacy rights. Its agenda includes protecting consumer data privacy, defending digital rights, and promoting transparency of how data is collected and used.

Data Privacy Advocacy Strategies

These advocacies and many others alike utilize similar strategies to push their agendas forward.

Public Education

One method that these advocacies use to push their agendas forward is public education. The goal is to inform consumers about their privacy rights and what options they have regarding them. Public campaigns help raise awareness about data privacy and the potential risks associated with the misuse of big data. By educating the public, advocacies can help individuals make more informed decisions regarding their own data.

The Digital Advertising Alliance (DAA) runs a public education campaign named "Your AdChoices" (Digital Advertising Alliance, 2012). Its goal is to provide information about how online advertising works and how consumers can control their online advertisement preferences. Its website, YourAdChoices.com, contains a tool that allows users to opt out of targeted advertising from participating companies.

The Electronic Frontier Foundation (EFF) runs several public campaigns to educate the public on digital privacy. One example is their "Surveillance Self-Defense" campaign which aims to educate individuals on their digital privacy rights. It also hosts workshops and events that teach topics such as digital security for activists and how to protect privacy on social media.

Although public education may not directly impact the politics of big data and data privacy policy, advocacies are able to push their agendas by increasing the coverage of data privacy knowledge among the public.

Litigation

Litigation is another method advocacies use to push their agendas. By taking direct legal action, advocacies are able to defend data privacy rights and establish more comprehensive legislation.

In 1997, the EFF and ACLU challenged the Communications Decency Act of 1996 which resulted in the victory in *Reno v. American Civil Liberties Union* (1997) (Hudson Jr., 2017). The ACLU argued that the Communications Decency Act of 1996 violated the First Amendment right of free speech. The ruling established the precedent that regulation of online content must be consistent with the First Amendment. The two advocacies also challenged the Child Online Protection Act of 1998 which resulted in the victory in *Ashcroft v. American Civil Liberties Union* (2004) (Hudson Jr., 2017). This case follows the same principle of protecting the First Amendment in the context of online privacy.

In 2016, the EFF and the Electronic Privacy Information Center (EPIC) supported Apple in a legal battle against the FBI's demand that Apple create a custom operating system to unlock an iPhone that was seized during the investigation of the San Bernardino terrorist attack (Electronic Privacy Information Center, 2016). The EFF argued that allowing so would set a dangerous precedent for government access to personal data.

These actions taken by advocacies have had direct impact on data privacy policy and legislation which have set new standards for big data regulation and consumer privacy protection.

Industry Leading Companies

Companies such as Google, Meta, and Amazon stand at the center of the struggle over digital privacy in the United States. These entities have the most stake in the discussion as the majority of their business models revolve around the collection and use of big data (Vigderman, 2023).

Meta

Meta has been criticized for its data practices, specifically for its ad-targeting system. Vice president of advertisement Rob Goldman stated, "You are entitled to your opinion, but we don't sell people's data. Period. That's not a dodge or semantics, it's a fact." (Monroe, 2019). Instead, Meta states that they categorize users and employ an ad-targeting system which sells the ability to target and advertise to specific categories. Meta claims that it is important to make the distinction between selling private data versus ad-targetable categories.

Meta has also faced criticism for poor data security. In 2016, the personal data of 50 million Facebook users were compromised due to a security vulnerability (Wagner, 2016). The breach was the third security issue Facebook had faced between June and October 2016. A major incident is the Cambridge Analytica scandal. In 2018, it was revealed that political consulting firm Cambridge Analytica had obtained the data of nearly 30 million users in 2014 (Confessore, 2018). The data was used to create targeted political ads during the United States 2016 presidential election. The scandal resulted in increased criticism of Meta’s data privacy practices and has been a major talking point in the conversation of the need for better data privacy regulations.

Mark Zuckerberg in Washington

Following the Cambridge Analytica scandal, Mark Zuckerberg was called to testify before the senate to answer questions about Facebook's data privacy practices (Wichter, 2018). During the hearing, Zuckerberg acknowledges the mistakes and discusses steps that will be taken to better protect user data.

In his opening statement, Zuckerberg stated that he believed that the company's responsibility for the first 10 or 12 years was to build tools to empower people to do good things. However, he stated that he has realized that a more proactive role needs to be taken to ensure those tools are used for good. Senator Bill Nelson asked if "you consider my personally identifiable data the company's data, not my data?" Zuckerberg replies by saying "the first line of our terms of service say that you control and own the information and content that you put on Facebook." He then describes the ad-targeting system and highlights the fact that users are able to turn off personalized advertisements which follows what is stated in the terms of service.

Senator John Kennedy pointed out that the Facebook's user agreement's purpose was not to inform users of their data and privacy rights but rather to legally safeguard themselves. He says, "the purpose of that user agreement is to cover Facebook's rear end. It's not to inform your users about their rights." While the user agreement may state the rights that individuals have concerning their data, it is naïve to believe that the majority of individuals will properly read it.

Zuckerberg used this hearing as an opportunity to address the data privacy concerns of Facebook as well as to acknowledge the mistakes made. He claimed that Facebook would be taking greater measures to ensure the security of consumer data but acknowledged that "people will measure us by our results." Following this hearing, Meta discovered a security breach that

exposed the personal data of 50 million users in 2018, was exposed for storing hundreds of millions of user passwords as plaintext in 2019, and had the personal data of over 500 million users leaked in 2021 (TeamPassword, 2022).

The Geopolitics of TikTok

Until recently, all major social media platforms popular in the United States such as Facebook, Twitter, and Instagram were owned by American companies. However, TikTok is a subsidiary of ByteDance, a Chinese company, which has sparked fears over TikTok's relationship with the Chinese Communist Party. This has caused concerns over the security of American's data privacy and beliefs such as Tiktok is "under the control of the Chinese government (Maruf, 2022).

While TikTok executives claim that the app is not a security threat and that no information has been shared with the Chinese government, leaked audio has shown repeated cases of ByteDance employees in China accessing private US-based TikTok data (Baker-White, 2022). Because of these data privacy concerns, there have been many attempts to restrict the use of TikTok in the United States such as the attempt by the Trump administration in 2020 (Trump, 2020). In February 2023, the White House declared that all United States federal agencies had 30 days to delete TikTok from all government-issued devices (Hadero, 2023).

Shou Chew in Washington

On March 23, 2023, TikTok CEO Shou Chew appeared before Congress to discuss United States data privacy and TikTok's ties to China (Thorbecke, 2023). During the hearing, Chew is asked about TikTok's relationship with ByteDance and China, the measures TikTok can take to improve data privacy protection, and the transparency of TikTok's data policies. Chew

makes the following commitments: “We will keep safety, particularly for teenagers, as a top priority for us. We will firewall protect the U.S. data from unwanted foreign access. TikTok will remain a place for free expression and will not be manipulated by any government. We will be transparent and we will give access to third-party independent monitors to remain accountable for our commitments.”

In 2017, China implemented a law which requires companies to hand over any personal data that the Chinese government deems relevant to China’s national security (Guardian, 2023). Congresswoman Anna Eshoo referenced this and asked Chew how the United States and its citizens can be sure that their data will not be handed over to the Chinese government if requested. Chew states that Project Texas, an ongoing effort “to move American data to be stored on American soil, by an American company overseen by American personnel”, addresses this issue.

Much of the concern stems from the belief that the CCP owns everything in China. Thus, if ByteDance is in China, and ByteDance owns TikTok, then the CCP and China own TikTok and its data. Chew worked to address this misconception by stating that ByteDance is “not owned or controlled by the Chinese government.” He stated that it is a private company with “60% of the company being owned by global institutional investors, 20% is owned by the founder, and 20% is owned by employees around the world. ByteDance has 5 board members and 3 of them are American.”

Conclusion

The struggle over digital privacy in the United States is an ongoing effort to find a balance between big data collection and usage and protecting the privacy of the American public. Companies try to realize their material interests even in the face of harsh regulation. The

government and policymakers must establish this legislation to regulate the emerging technologies to protect the public's best interests. This proves to be a difficult task when those in power to develop the legislation do not understand the technology and circumstances that surround the conversation around consumer data privacy well. However, like in any other private sector or industry, the struggle between the material interests of businesses and its legislation will continue indefinitely.

References

Atske, S. (2019, November 15). Americans and privacy: Concerned, confused and feeling lack of control over their personal information. *Pew Research Center: Internet, Science & Tech*.

Baker-White, E. (2022, June 17). Leaked Audio From 80 Internal TikTok Meetings Shows That US User Data Has Been Repeatedly Accessed From China. *BuzzFeed News*.

Bowman, E. (2021, April 10). After data breach exposes 530 million, Facebook says it will not notify users. *NPR*.

Confessore, N. (2018, April 4). Cambridge Analytica and Facebook: The scandal and the fallout so far. *The New York Times*.

Digital Advertising Alliance. (2012). Digital Advertising Alliance (DAA) announces 'your adchoices' consumer education campaign. *DigitalAdvertisingAlliance.org.*

Electronic Privacy Information Center. (2016). Apple v. FBI. *EPIC*.

Fruhlinger, J. (2020, February 12). Marriott Data Breach FAQ: How did it happen and what was the impact? *CSO Online*.

Guardian. (2023, March 17). The tiktok wars – why the US and China are feuding over the App. *The Guardian*.

Hadero, H. (2023, March 1). Why TikTok is being banned on Gov't phones in US and beyond. *AP NEWS*.

Hudson, D.L., Jr. (2017, Dec.). Electronic Frontier Foundation. *The First Amendment Encyclopedia*.

Jain, P., Gyanchandani, M., & Khare, N. (2016, November 26). Big Data Privacy: A Technological Perspective and Review. Journal of Big Data, 3(1). https://doi.org/10.1186/s40537-016-0059-y

Kerry, C. (2020, February 10). Protecting privacy in an AI-driven world. *Brookings*.

Lewis, A. (2010, August 26). reader comment on Jordan M. Rhaomi (2010, August 26). User-Driven Discontent. *MetaFilter*.

Leung, C. K., Kajal, A., Won, Y., & Choi, J. M. (2019). Big Data Analytics for personalized recommendation systems. *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*. https://doi.org/10.1109/dasc/picom/cbdcom/cyberscitech.2019.00190

Levenberg, K., Pittman, F. P., & Shamir, S. (2022, August 7). *Data Protection Laws and Regulations Report 2022-2023 USA*. International Comparative Legal Guides International Business Reports.

Maruf, R. (2022, July 3). An FCC regulator wants TikTok removed from app stores. Here's how a company executive responded. *CNN*.

Mehmood, A., Natgunanathan, I., Xiang, Y., Hua, G., & Guo, S. (2016). Protection of Big Data Privacy. *IEEE Access*, *4*, 1821–1834. https://doi.org/10.1109/access.2016.2558446

Monroe, A. (2019, February 15). Fact check: Does facebook sell your personal data? *The Arizona Republic.*

Narayanan, A., Huey, J., & Felten, E. W. (2016). A precautionary approach to big data privacy. *Data Protection on the Move*, 357–385. https://doi.org/10.1007/978-94-017-7376-8_13

Tabora, V. (2019, November 1). In Big Data, the consumer is the product. *Medium*.

TeamPassword. (2022, December 1). Facebook hacks: A history of security breaches at Facebook. *TeamPassword*.

The linkedin data breaches: What to do and who was affected. *IdentityTheft.org*. (2023, March 20).

Thorbecke, C. (2023, March 23). Tiktok CEO in the hot seat: 5 takeaways from his first appearance before Congress. *CNN*.

Trump, D. J. (2020, Aug. 6). Executive Order on Addressing the Threat Posed by TikTok. *Trump White House Archives*.

Vigderman, A. (2023, February 6). The Data Big Tech companies have on you. *Security.org*.

Wagner, K. (2018, September 28). Why should anybody trust Facebook with their personal data? *Vox*.

Wichter, Z. (2018, April 12). 2 days, 10 hours, 600 questions: What happened when Mark Zuckerberg went to Washington. *The New York Times*.