

Trinity College Trinity College Digital Repository

Senior Theses and Projects

Student Works

Spring 2018

The Deception of Data: How Millennials' Political Views on Data Guide Their Use and Understanding of Facebook

Griffin Drigotas

Trinity College, Hartford Connecticut, griffin.drigotas@trincoll.edu

Follow this and additional works at: <https://digitalrepository.trincoll.edu/theses>



Part of the [Digital Humanities Commons](#)

Recommended Citation

Drigotas, Griffin, "The Deception of Data: How Millennials' Political Views on Data Guide Their Use and Understanding of Facebook". Senior Theses, Trinity College, Hartford, CT 2018.

Trinity College Digital Repository, <https://digitalrepository.trincoll.edu/theses/692>

TRINITY COLLEGE

SENIOR THESIS

**THE DECEPTION OF DATA: HOW MILLENNIALS' POLITICAL VIEWS ON DATA
GUIDE THEIR USE AND UNDERSTANDING OF FACEBOOK**

Submitted by

Griffin J. Drigotas

27 April 2018

In partial fulfillment of the requirements
for Honors in American Studies 2018

Director: Professor Scott Gac
Advisor: Professor Jack Giesecking
Second Reader: Professor Thomas Wickman

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	2
INTRODUCTION: The Facebook Fallacy	3
CHAPTER ONE: Identifiers	32
CHAPTER TWO: Facial Recognition	55
CHAPTER THREE: Geolocation	76
CONCLUSION: For a Different Future	94
APPENDIX	100
BIBLIOGRAPHY:	101

ACKNOWLEDGEMENTS

Thank you . . .

To Professor Giesecking . . . for always making me question the things we take for granted and for first introducing me to the study of data. Without you and your guidance, this project would be nothing. I appreciate every moment of your time that has dedicated to my success over the past several years.

To the American Studies Department . . . for providing me an academic home for the past four years. To Professor Scott Gac, Professor Davarian Baldwin, Professor Christina Heatherton, Professor Diana Paulin, Dr. Erin Valentino, and Professor Tom Wickman, I thank you each for providing me a unique glimpse into the inner workings of the complicated country we live in. Your time, dedication, and knowledge has never ceased to amaze me; I appreciate each and every one of the faculty members who has helped me along this journey.

To my family . . . for your words of encouragement. You have each provided a sense of inspiration, positivity, and motivation over the last year which I appreciate immensely.

To my participants . . . for your time, patience, energy, and thoughtfulness. Without you, this project would have lacked critical perspective and intelligence.

To Crowe's Nest . . . for keeping me sane over the past year. Thank you for keeping the music down when I needed focus and blasting it when I needed a break.

INTRODUCTION: THE FACEBOOK FALLACY

Topic of Research

As I open my computer, the soft glow of the screen lights up my face. The hum of the cooling fan fills the silence of my room as I trace my finger across a few inches of trackpad, I select the Google Chrome application. Instinctively, my finger looms over the letter “F,” and I know Facebook will be suggested immediately by my browser as my online destination. Click. Facebook loads, and my eyes are drawn towards my “friends” in the form of their most recent photos, achievements, and differing political opinions, all displayed on my screen. I see my “friends: online—people I see every day or someone I met once over a beer studying abroad in Vienna—just a click away from connecting. Each time I logon, the Facebook interface conveys a safe and friendly website, which gives us the ability to link up quickly and effectively, perhaps even deeply.

Yet behind the blue, white, and grey site the public knows so well, algorithms churn. Their overwhelming power is constantly tracking our every move as we navigate through the cyber world. The data is stored away and analyzed. Don’t you want these new shoes you clicked on last week? Would you like to allow your friend to tag you in that picture from last weekend? Or isn’t it time you read (or skimmed) the last politically-driven article your grandmother posted? Whatever you decide to do, from the move of your mouse to the click on another hyperlink, — it waits. It collects. It is always watching. Of the over two billion daily active Facebook users—nearly one-third of the world’s population—few realize the vast amount of data they give away. Each click, like, search, and tag is recorded and allows the Facebook algorithm to make calculated decision on a plethora of data points, all which

seem too personal for comfort.¹ This is all made possible by the user accepting (probably unknowingly) the Terms of Service. The way young people use and make sense of this new and debated form of privacy and its online public requires attention.

In our ever-increasing digital world, interconnectedness is easily found, or so websites like Facebook lead us to believe. Every time each of us logs on to a computer, we are being tracked and analyzed by an unseen and looming presence of algorithms that produce findings that sort us, sell to us, persuade us, and report on us.² We are even encouraged to and do share about ourselves and others. In so doing, daily technological advancements reshape our ideas and abilities for interconnectedness by providing a place for the ability to record, distribute, and analyze information.

Policy, governance, and law can barely keep up with these changes in both perception and understanding. As a result, there is a growing fear of being monitored, along with substantial evidence of the questionable ethics and legalities of this surveillance.³ Discussions of diminishing privacy rights dominate the media, particularly within the context of the United States and its democratic claim to privacy. With each day, new information and opinions are shared with the public, and the way the largest market of applications like Facebook, young people, requires increased attention.

I am one of those “young people.” Since the age of 14 and now 22-years old, I use Facebook mainly as a way of entertainment. I have never posted a status, but I am still on it each and every day. In my many years of having an account, and never sharing any

¹ Josh Constine, “Facebook now has 2 billion monthly users ... and responsibility,” TechCrunch, June 27, 2017, <https://techcrunch.com/2017/06/27/facebook-2-billion-users/>.

² Pariser, *The Filter Bubble*.

³ Daniel Trotter, *Social Media as Surveillance: Rethinking Visibility in a Converging World* (London: Routledge, 2012): 23.

information in posts, I thought that I was immune to the taking of privacy. However, I learned that I was very wrong. The ways data is taken is secretive, concealed, and vast. It is this realization—and the fear and wonder it set off in me—that put me on track to study the intricacies of Facebook privacy. At a time when heightened political divides rage in the United States, I believe studying Facebook’s privacy, or lack thereof, through a political lens will help me more clearly articulate what privacy means today and therefore what it means to be a politically-engaged American today. It is important for Facebook users to understand what is actually happening with these advancements, realize its implications, and come up with ways to use it for good.

Law scholar Julie E. Cohen breaks down the discussion of personal data into distinct categories with an eye towards the context of US democracy: owning, choosing, knowing, speaking, and becoming. She argues that “the debate about data privacy protection should be grounded in an appreciation of the conditions necessary for individuals to develop and exercise autonomy,” and further, “that meaningful autonomy requires a degree of freedom from monitoring, scrutiny, and categorization by others.”⁴ In other words, personal data is more than just a series of numbers, words, and images, but rather a social entity whose power is being governed. Cohen concludes by “calling for the design of both legal and technological tools for strong data privacy protection.”⁵ These tools will be an integral part of the future of online digital democracy. With these in place, Facebook may become a far safer, personal, and less intrusive space.

⁴ Julie Cohen, “Examined Lives: Informational Privacy and the Subject as Object,” *Stanford Law Review* 52 no 5 (2000): 1373.

⁵ *Ibid.*

Another key aspect for my work, Cohen's discussion of cyber law however, is not so much put within a law context, but rather one that involved the ideas of the place/space and the ongoing production of networked space. She makes clear, which I find important, that space does not refer to the abstract idea of it. Space in this sense is experienced by embodied human cognition.

My key argument is that political party affiliation does not cause any *change* when it comes to the use of Facebook. Rather, political leanings only shape *opinions* on what can and should be done with data. While the differences in these opinions are vast, the goal of this research is to make clear this political bifurcation – perhaps realization of differences may be the first step towards the healing of a torn democracy.

I argue that my findings contribute to the American Studies discipline because our society is becoming increasingly data-driven. There is little work completed which has examined Facebook specifically, and none that use the popular social network as a means of understanding political thought, or vice-versa. Most recently, the intersection of Facebook and politics became evident. On March 17, 2018, just weeks after my research had concluded, the Cambridge Analytica controversy was made public when expose's in both the *New York Times* and *The Guardian* revealed how the data from millions of different Facebook users ended up being given and exploited by Cambridge Analytica. Although important, this thesis needs to be read with the understanding that the event had not yet been made public, and thus, my subjects and I were not aware, and my research methods were not influenced. This topic of Cambridge Analytica and its political significance will be discussed in my epilogue. I am hopeful that these findings can lead people to take action for healing the vast and immoral collection of data.

Main Arguments & Contributions

My study seeks to identify the evidence behind these perceptions found in conducted interviews with fellow Trinity College students. Trinity College is a small liberal arts institution located in Hartford, Connecticut. It provides an interesting background due to its status as a private campus, within a bustling and diverse urban setting. I conducted interviews to help me to understand the detailed nature of the relationship between Facebook data, cultural artifacts (newspapers, blogs, etc.), and the actual (and often vague) rights to privacy found in U.S. government legislation. In comparing academic and journalistic findings with my interview transcripts, I was especially interested in tracing how Millennials perceive and react to data privacy. Millennials interest me because they are the only individuals whose entire adult lives have been partially online, and they are the group of people who are most likely to spur change in this country. I found that these understandings are shaped by political leanings. I argue that there would be distinct differences between how the left and the right not only understand data privacy, but how they react to its effect on the political climate of the United States.

In this thesis, I explore the ways in which the relationship between three actors: political perceptions, the data they share, and Facebook have on the American identity in its current contentious climate. I expected that there will be a vast difference between how the left and the right view the topic and understand the solution of Facebook privacy. Further, I thought that there would be a great deal of difference in the data I collect when it comes to ads and information that the subjects see when they interact with the app. These differences stem from the drastically varying views each party hold true.

I found that there were indeed differences. Each political party were aware of current situations facing the United States but, their thought process differed when it came to how the data should be handled. This opened my eyes to not only the data issue, but the actual differences between political parties. Individual statements were eloquently said and well thought out, they just held key beliefs which demonstrate how contentious this topic truly is.

Methods

My primary research site is Facebook. This website is a social media platform which allows users to connect and share virtually with their friends, families, and others they may know. Facebook has been under an extreme amount of scrutiny in the last few years in regard to the amount of privacy it actually gives its users versus what it appears to offer. Through emails, recommendations, and in-class announcements, I recruited eight Facebook users to take part in my study. I sought to have this group of individuals be diverse political leanings (right and left), gender, and at least two racial identities. This will allow my subjects and thus, my result, to be the most fortified and accurately represent the actual perceptions of personal Facebook privacy.

A series of two interviews followed by a survey will be my primary method of collecting data and understanding the perspectives on the interaction of users with Facebook. Six users took part in both interviews. This longitudinal data collection allows me to track how perceptions of privacy online change. I observed whether participants' opinions on the subject matter changes when they gain knowledge about the inner workings of Facebook, and why or why not they do change. By asking the participants the same questions

twice, over a period of time, and after they have interacted with specific types of media pertaining to the subject matter, then I can truly see how perceptions are changing or staying the same.

My first round of interviews introduced participants to the topic in two sections. My participants first answered general questions which helped me understand their political thoughts, their identifying information, and their use of Facebook. It was important for the questions not to favor a party or system of belief; there is certainly tension within and surrounding political party's stances and people may answer particular questions based on the current social standing around the answer, rather than basing answers on actual beliefs. Thus, I will make sure that the subjects know their responses will be held in complete anonymity, so people will answer the questions more honestly. The second part of this first interview will be the asking of general questions to get a baseline of understanding. These questions include:

- What is your primary use for Facebook?
- Does the news you see on Facebook reflect your political beliefs – why or why not do you think this is the case?
- What personal data do you want protected?
- Do you think your data is secure?
- Do you own your own personal data?
- In today's increasingly connected and digitized world, does privacy actually exist?
- What does privacy mean to you?
- What precautions do you take online?
- What types of ads do you see?
- Can you tell me what you know about targeted ads?
- Can I look at your ads you see?

All of these questions provide a baseline of understanding and will kick off discussion.

Another important approach to understanding interactions with Facebook is by looking at the actual feeds of my subjects. The similarities and differences I observe will be

eye opening. There will be profound and important data in what my subjects think they see, as their own Facebook world is never compared to someone else's.

Between the two interviews, I asked participants to read or listen to the article "Didn't Read Facebook's Fine Print? Here's Exactly What It Says," by Amanda Scherker. This *Huffington Post* article breaks down the Facebook Privacy Policy and describes the clear ways in which we are giving our data to Facebook every single time we use the site. Its simple language allows the participants to get a solid grasp on the difficult material. This will then lead to the final interview. These questions are an attempt to finally understand how their minds have changed regarding the subject. I think that at this point I will have the data to make educated conclusions about how the perceptions of my subjects have changed, based on the subjects' political leanings.

Finally, I sent out a survey to my participants to fill in for any questions I may still like to find out the answers to. This will begin with the same general questions as the first, as I wish to see if any opinions have changed, or which have been solidified. Then, questions with more specificity, specifically focused on location data and visuals. This survey is aimed at delving deeper into the issues and being a conversation starter in looking specifically at the issue of Facebook. I hope that the participants will reflect upon their own experiences to foster intelligent and meaningful discussion on the issue.

I chose to research three types of data. While most studies talk about "data" in a vague way, I noticed in my conversations with friends and family, and in my research, that people had different thoughts about privacy in regard to personal information (like an address or SSN) versus sharing their location or tagging themselves in an image. I wanted to find out why there was a difference between how people thought about these types of data. Through

my interview process, it became clear that these differences came down to expectations within each, and the types of real world implications that can occur. Why the difference? I wanted to dig into this further, so I pressed my subjects on specific topics to gauge their level of understanding and reactions.

In my data analysis, I read across the interviews to look for patterns of thought on each of these types of data and then reflect on them separately by data type to see if my participants may feel they need to protect one sort of data over the other. This mattered because the historically intrinsic differences in opinions between political thoughts, when put in the framework of a modern technological platform will help in understanding the future of social media interaction. This entire process allowed me to understand a baseline for understanding, concrete ways people processed the thought of sharing data, and a guide for educating people about data privacy which will play a key role in my conclusion.

Site & Period

Facebook is a social network website which allows its users to connect, share, and interact digitally with friends and family. The website was originally created by Mark Zuckerberg for college students in 2004 but, has since developed into a network of over 1.4 billion daily active users.⁶ Facebook's place within has evolved throughout its existence. It is now the most popular social site used and is the new norm for online interaction. My study is timely and is important as the website is becoming more and more engrained in the

⁶ "Facebook Reports Fourth Quarter and Full Year 2017 Results," January 31, 2018, <https://investor.fb.com/investor-news/press-release-details/2018/Facebook-Reports-Fourth-Quarter-and-Full-Year-2017-Results/default.aspx>.

everyday lives of Americans, and thus has unknown impacts on the information we see and receive, based on our own personal beliefs.

The Facebook Privacy Policy is a document which outlines the overall protection of its users, and for the purpose of this thesis, specifically the collection and distribution of personal information. It details the ways in which data is actually surrendered and processed by Facebook. This document is important because it is a primary source directly from Facebook. They put into words what the algorithm is constantly processing but won't reveal what is actually in the algorithm.

Today, Facebook and its place within the political climate in the United States is one of contention. Within the last several years, as political tensions have ramped up, as have the amount of news, both real and fake, that is displayed on users' feeds. The two entities are now so tied together that people now receive much of their news from Facebook, a source which lacks the ability to automatically determine posts to be fact or fiction. This has been going on for several years now, they cannot be separated from one another. The algorithm Facebook has in place provides people with the content that they *want* to see, and hides differing opinions.⁷ I sought the evidence behind this idea: what people see, why they see it, and how they react to the reasoning behind it all.

Literature Types

My literature will primarily be drawn from three different types of sources, each with their own temporality in terms of research. The first type of literature is peer-reviewed

⁷ Eli Pariser, *The Filter Bubble: How the Web Is Changing What We Read & How We Think*, (New York: Penguin, 2012): 9

journal articles and scholarly monographs. These assisted in helping me understand how experts think about the issues, through an academic lens, and forced me to critically examine the data that I collected in my interviews and helped me think about them through new contexts. Secondly, since this topic is a very new, modern, and a continuously changing issue, for each scholarly article I explored many news and blog articles. Although not peer reviewed, they provided me with new and up to date opinions on current matters.

Due to the large amount of information that is released each day, I narrowed the news outlets to four major news sources that cover the technology industry from varying left and right political perspectives: *Gizmodo*, *Huffington Post*, *The New York Times*, and *FOX*. This helped me visualize how the topic is changing; especially because many of the articles are being written by the same authors. Their perspectives changed in real-time as new information and data was shared regarding this relatively new topic. News articles provided indication of the both the nature of society and specific aspects of the culture of the United States at the current moment. In continuing this thought, news articles are a product of the societies in which they are produced, and therefore may offer a limited perspective. Another limitation, is that news sources all have political agendas. However, this may have helped in understanding perspectives on the topic, in reflection, as long I am sure to be cognizant of that matter.

My final type of literature were actual primary documents. The most major of these being the actual Privacy Policy of Facebook. Due to its extreme length, I was not able to read the entire Policy. To put its length in perspective, it would take 76 days to read to read all of

the Terms and Conditions we sign, across all of the internet, in an average year.⁸ Even so, I ensure, through other readings, that I am familiar with the most important, shocking, and pertinent information that I need from it.

Literature Review

Because of the cutting-edge nature of my research topic, I primarily drew on scholarship in the fields of communications, American studies, and sociology, as well as primary research documents from news sources and historical documents. Across these disciplines, I identified three themes through which to shape my reading and analysis of the interviews regarding the data collected and analyzed by Facebook: the identifiers (name, phone number, email, and so on), visuals (images and films), and location data tracking. These three themes, when researched through the three different types of literature returned a massive number of quality sources which detailed many different aspects within the topic of Facebook data privacy.

To begin, this topic is extremely democratic, which is why my first source is directly from the United States Constitution:

Amendment I: Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.⁹

Amendment IV: The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by

⁸ Amanda Scherker, "Didn't Read Facebook's Fine Print? Here's Exactly What It Says," *Huffington Post*, July 23, 2014, http://www.huffingtonpost.com/2014/07/21/facebook-terms-condition_n_5551965.html.

⁹ U.S. Constitution. Amendment I.

Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.¹⁰

It is important to begin here to lay the foundational framework for the questions guiding this thesis. The democratic foundation laid forth by the Constitution, and specifically these two amendments play a significant role in personal privacy.

Although the notion of digital democracy, networked selves, or even the internet were not present at this time, the significant thought towards personal privacy were still instilled deep within the foundation of the United States government. My thesis is primarily focused on political perceptions of privacy and therefore it is imperative to have a solid foundation of knowledge on the ideals that the U.S. Constitution lay forth. Although democracy is not a theme of my essay, it has a focus throughout. Many of the authors I have read allude to these amendments and I presume people's perceptions will as well.

All of this conglomerated Facebook data is known as "big data." While this term is often thought of as just a great amount of data, which it is, it is so much more. As technology scholars danah boyd and Kate Crawford discuss in their article, Big Data is far less about the actual size of the collected data but its capacity to search aggregate, and cross reference these large data sets.¹¹ However, in the case of Facebook's data, they have both the mass and the ability to search this data. It is important to keep this type of data in mind, as it is primarily the type of data being discussed throughout this project.

Throughout all of my research, I realized that while there is so much literature on the topic of Facebook's data collection, the concept of data is never truly defined and explored.

¹⁰ U.S. Constitution. Amendment IV.

¹¹ danah boyd and Kate Crawford, "Critical Questions for Big Data," *Information, Communication & Society* 15, no. 5 (2012): 663.

The term “data” is often generalized but I saw a distinction and wanted to explore the intricacies. With that, I decided to break down the data Facebook collects into three distinct categories: identifiers, visual, and geolocation.

I rely on John Cheney-Lippold’s article, “A New Algorithmic Identity,” to guide my research. He draws on the work of historian Michel Foucault who argues that states deploy regulatory power through population control in order to supplement disciplinary power.¹²

Cheney-Lippold argues for:

The process of identification, at least in the online world, becomes mediated by what I term soft biopolitics, as user identities become tethered to a set of movable, statistically-defined categorizations that then can have influence in biopolitical decisions by states and corporations.¹³

Given Facebooks close relationship with the U.S. Government, among others, I see the process of soft-biopolitics playing out throughout this thesis.

Identifiers

The first theme of this thesis revolves around the Facebook algorithm and the identifying data that it collects. As many of the authors share, the notion of data collection all begins with the algorithm that Facebook has running. As journalist Robinson Meyer explains in the left-leaning *The Atlantic* magazine: Facebook has lost knowledge regarding what its algorithm is capable.¹⁴ Although this is explained through a marketing mindset, Meyer describes that there are extreme social implications as well:

¹² Michel Foucault, *Security, Territory, Population: Lectures at the Collège de France, 1977-78*. (New York: Springer, 2007): 70.

¹³ John Cheney-Lippold, "A new algorithmic identity: Soft biopolitics and the modulation of control," *Theory, Culture & Society* 28, no. 6 (2011): 164.

¹⁴ Robinson Meyer, “Could Facebook Have Caught Its ‘Jew Hater’ Ad Targeting?” *The Atlantic*. September 15, 2017. <https://www.theatlantic.com/technology/archive/2017/09/4>

After all, the average American spends 50 minutes of their time there every day. Facebook's algorithms do more than make the platform possible. They also serve as the country's daily school, town crier, and newspaper editor. With great scale comes great responsibility.¹⁵

Facebook does not understand what their algorithm has the capacity to generate and makes me question what good can come from it. This is important because Facebook really should know the power of something that can have so much impact on the everyday lives of people.

In another recent article on the tech site *Gizmodo*, journalist Franklin Foer furthers this discussion of its algorithm and its massive, overbearing size in his clearly named article "How Facebook Tricks You into Trusting Algorithms." Facebook has the power to do so much, while at the same time might have gotten out of hand. The algorithm, according to Facebook, has many benefits, such as increasing voter turnout in the 2016 Presidential Election. Foer records Facebook founder Mark Zuckerberg's fantasy: "that this data might be analyzed to uncover the mother of all revelations, 'a fundamental mathematical law underlying human social relationships that governs the balance of who and what we all care about.'"¹⁶ Overall, algorithms currently have an impressive power and soon will possess a possible terrifying power if not legislated and overseen properly.

But what happens when you tie together data protection and what is actually happening within Facebook's algorithm when it results not only in targeted ads but with real world social implications of this practice? With political implications discussed, it provides pertinent information to my interviewees, and the notion of digital democracy. Alyza Sebenius, another *Atlantic* author taken Foer's article one step further, when she writes: "The

¹⁵ Ibid.

¹⁶ Franklin Foer, "How Facebook Tricks You into Trusting Algorithms," *Gizmodo*, Accessed September 22, 2017, <https://gizmodo.com/how-facebook-tricks-you-into-trusting-algorithms-1810792161>.

decisions we make regarding the way that tech companies are a market for advertising affect, in a very real sense, what kind of digital democracy we are going to build.”¹⁷ A recent example would be companies posting job ads to younger audiences only. Julia Angwin’s *ProPublica* article details this extremely undemocratic practice.

The actual technical process of this practice will be discussed below, but companies such as Amazon, Verizon, UPS and Facebook itself have been targeting certain audiences for potential job opportunities.¹⁸ Understandably, many companies want young, upcoming, and motivated employees, but this practice leaves out a massive hole in the potential workforce. For example, Facebook recently ran a career ad which only reached people between the ages of 25 to 60.¹⁹ Is this practice against the law? According to Angwin’s article, “Several experts questioned whether the practice is in keeping with the federal Age Discrimination in Employment Act of 1967, which prohibits bias against people 40 or older in hiring or employment. Many jurisdictions make it a crime to “aid” or “abet” age discrimination, a provision that could apply to companies like Facebook that distribute job ads.”²⁰ Based on this example, the data collection and its use is a discussion which needs to be examined closer, as major democratic implications have taken place.

From the discussion of the Facebook algorithm, a line must be drawn to their Privacy Policy. These two pieces are integral to the understanding how data is actually surrendered

¹⁷ Alyza Sebenius, “Should Facebook Ads Be Regulated Like TV Commercials?” *The Atlantic*, September 14, 2017, https://www.theatlantic.com/technology/archive/2017/09/facebook-ads-free-speech/539736/?utm_source=nl-atlantic-daily-092017&silverid=MzEwMTkwMTM1NzAxS0.

¹⁸ Aleksandra Korolova, “Privacy Violations Using Microtargeted Ads: A Case Study,” *Data Mining Workshops* (2010): 475.

¹⁹ Julia Angwin, “Dozens of Companies Are Using Facebook to Exclude Older Workers from Job Ads,” *ProPublica* Accessed January 10, 2018, <https://www.propublica.org/article/facebook-ads-age-discrimination-targeting>.

²⁰ *Ibid.*

and processed by Facebook. The most important piece of literature to this section is from the left-leaning *Huffington Post* article, which is also included in the interview process: “Didn’t Read Facebook’s Fine Print? Here’s Exactly What It Says.” As previously mentioned, this breaks down the Privacy Policy and allows for true understanding of the lengthy and technical document. Scherker, a *Huffington Post* journalist, makes clear that “Nothing you do on Facebook is private. Repeat: Nothing you do on Facebook is private.” And that you agree to this from the first moment you sign up: with your name and gender selection.²¹ From there, the data points collected and determines are both lengthy, and invasive. Several of the most and obtrusive are expectant-parent types of credit cards, presence of children in the household, and several different health categories. This article really is a great way to understand the absurdity that is the Facebook Privacy Policy. As it reveals that it does not protect your privacy, quite at odds with what its name suggests. It is beneficial as it makes clear that Facebook knows a shocking amount about you and have tendrils which reach far outside the application to surveil you as well.

In contrast, there are exactly zero right-leaning FOX news articles written on the topic, informing their readers on the intricacies of Facebook’s data collection. This is important to note because it is the first glimpse into the differing views that political parties have on this confusing and loaded topic.

Communications studies scholar Seeta P. Gangadharan wrote several articles about digital inclusion and surveillance. Gangadharan brings to light important details towards the notion that very few studies have addressed the manner and extent by which digital tracking and targeting have impacted, and most likely forwarded inequality within society. A key term

²¹ Scherker, *Huffington Post*.

in this article is the “digital divide”: the gap between demographics and regions with regard to the use of modern technology.²² Gangadharan goes into great detail to break down the digital divide and its complicated social, racial, and economic implications. She furthers this discussion by examining the digital inclusion downside. In brief, she finds “an interaction between social status of marginalized individuals and a particular type of Internet tailored for and targeted at the marginal user”²³ (Gangadharan, 14). This is important because the digital divide is a political issue and will certainly be a loaded topic when conducting interviews. Different political influences will challenge this idea in drastically different ways.

Media policy researcher Jisuk Woo discusses the self-privacy on “the network.” She suggests that the only way to ensure privacy in the interactive digital environment is by, “allowing affirmative acts of secrecy and deception regarding identity and identification.”²⁴ Through the identification of key words such as anonymity, interactivity, and power, Woo makes clear the developmental changes of telecommunications technology and the ways in which information is collected, managed, and shared. At the same time, she brings to light the lack of practical options which individuals may have to conceal their identity online. She offers a unique perspective on this issue by drawing a connection between the right to conceal your information and identity, and how that right is an important tool in influencing distribution of both social and political power.

²² Daniel Greene, “Discovering the Divide: Technology and Poverty in the New Economy,” *International Journal of Communication*, 10 (2016): 1215.

²³ Seeta P. Gangadharan, “The Downside of Digital Inclusion: Expectations and Experiences of Privacy and Surveillance among Marginal Internet Users,” *New Media & Society* (November 9, 2016): 14.

²⁴ Jisuk Woo, “The Right not to be identified: privacy and anonymity in the interactive media environment,” *New Media & Society* 8 no. 6 (December 1, 2006): 949.

Catherine Crump, a privacy activist and professor at Berkley Law School, looks at it through the lens of the First and Fourth Amendment of the United States Constitution. Both of these critical pieces to the U.S. Constitution have a great deal to do with privacy, and the basic rights of the people. Within the context of Facebook, Crump makes clear that the distinction between 'data retention' and 'data preservation' is important when looking at the constitution, especially since it was far from an actual thing when the constitution was written. Crump writes that data retention is a much broader concept which "aims to change the context of Internet activity."²⁵ From this, it is made clear that data retention would diminish both privacy and anonymity, but when put within the laws of the Constitution they are protected. It will be interesting to look at the politics of Facebook and apply them through this concept.

Visual

The second theme of my thesis is the visual side of Facebook. The majority of this literature review section revolves around the facial recognition algorithm that Facebook has accumulated through the years. Further, it will describe the types of data that this algorithm can collect, and how invasive that may be to the personal privacy of individuals. I was especially interested in visual data because I find it to be the closest link between virtual and real when it comes to the idea of being. Finally, an interesting piece to look at is whether or not this practice is actually morally wrong, or if it is deemed okay because the users agree to this when they sign up for the website.

²⁵ Catherine Crump, "Data Retention: Privacy Anonymity, and Accountability Online," *Stanford Law Review* 56 no 1 (October 2003): 194.

Author and journalist Bryson Masse opens the article, “What’s the Worst That Could Happen with Huge Databases of Facial Biometric Data?,” with a background of geometric data history and then plunges into technical definitions given by experts in the fields of law, technology and facial recognition in an effort to answer two questions: “What can happen when we combine the large amount of facial biometrics data with a potentially imperfect system?” and “What sort of societal implications would there be if you were recognized by someone, anywhere and everywhere you went?”²⁶ These questions are imperative to this study, as the answers provide real-life impacts that facial recognition algorithms can have on the public. These implications are often not realized by people who use Facebook, whether that be due to simply not knowing, because they do not care, or another option, that still remains to be understood.

One of the “panelists” in Masse’s article, Christopher Dore, a partner at Edelson PC law firm in Chicago, responds to Masse’s questions while tying it directly to Facebook. Dore says, “Facebook has the largest database of recognition data in the world, period,” and “when you have a situation where a company is holding a database of that type, there are a lot of concerns that come up,” including, “what are they going to do with it?”²⁷ Dore mentions that Facebook is currently only really using this feature to tag people in images, but they have the power to do so much more. This data stored could be sold off to stores for marketing purposes, surveillance, and identification.²⁸

²⁶ Bryson Masse, “What’s the Worst That Could Happen with Huge Databases of Facial Biometric Data?” *Gizmodo*, September 11, 2017. <http://gizmodo.com/what-s-the-worst-that-could-happen-with-huge-databases-1802696698>.

²⁷ *Ibid.*

²⁸ *Ibid.*

The Economist published an article very similar to Bryson Matte's article as it brings up many of the same implications that facial recognition may hold. They begin by making the comparison between fingerprint recognition and facial recognition. Much like my own thinking about data privacy and the power of algorithms, when first reading this comparison it seemed odd, but the more I thought about it, makes perfect sense why they begin with this. Fingerprint recognition has become the standard for security and is accepted by many, not as an invasion of their privacy. Even though this is a way to be tracked and known, people seem to be completely fine with it. It is a standard of security that has become enmeshed and accepted within society. However, as the author of this article writes: "One big difference between faces and other biometric data, such as fingerprints, is that they work at a distance. Anyone with a phone can take a picture for facial-recognition programs to use."²⁹ Basically, the application for facial recognition is far easier to be used without the identified knowing it. Facial recognition can identify people without consent, and thus, can be the basic for controversy.

Another article from *The Economist* speaks directly at a type of identifier that can be taken from the facial recognition software: sexuality. The article titled, "Advances in AI are used to spot signs of sexuality," is a shocking read which takes a hotly contested social topic and places data and technology behind it. The author writes:

AI's power to pick out patterns is now turning to more intimate matters. Research at Stanford University by Michal Kosinski and Yilun Wang has shown that machine vision can infer sexual orientation by analyzing people's faces. The researchers suggest the software does this by picking up on subtle differences in facial structure. With the right data sets, Dr. Kosinski says, similar AI systems might be trained to spot other intimate traits, such as IQ or

²⁹ "What machines can tell from your face," *The Economist*, September 9, 2017, <https://www.economist.com/news/leaders/21728617-life-age-facial-recognition-what-machines-can-tell-your-face>

political views. Just because humans are unable to see the signs in faces does not mean that machines cannot do so.³⁰

These are several data points which are possible to spot simply through face biometrics. This means, that through Facebook's algorithm they are able to tell an incredible amount of information about you, simply based off of your face. In fact, Facebook excitedly proclaims that they need not see your full face to recognize you.³¹ I would like to take this one step further: Facebook is also collecting data based on your posts, activities, soon-to-be mentioned location, and everything else you do. Once these data points are all combined and run through the Facebook algorithm, Facebook most likely has the ability to know way more about you than you would care to have them. Thus, when comparing to Byson Masse's article, this data has the possibility of being sold to any company or institution around the world.

Facial recognition software is one of the most important processes within the Facebook sphere. Most of its success is due to the implementation of the of *Face.com*. This service, when put in perspective on Facebook, tagged over 400million images in one month out of Facebook's 10 billion photo archive.³² Further, real-world identifications have become a part of the program. The images online are inextricably tied to reality. This is due to the fact that "Facebook members are expected (as delineated in the Terms of Service) to use their "real identity" for their Facebook profile, and thus considerable evidence that a large

³⁰ "Advances in AI are used to spot signs of sexuality," *The Economist*, September 9, 2017. <https://www.economist.com/news/science-and-technology/21728614-machines-read-faces-are-coming-advances-ai-are-used-spot-signs>.

³¹ Nicholas Pinto, et al., "Scaling up Biologically-Inspired Computer Vision: A Case Study in Unconstrained Face Recognition on Facebook," *Computer Vision and Pattern Recognition Workshops* (2011): 36.

³² Ariane Ellerbrok, "Playful Biometrics: Controversial Technology through the Lens of Play," *The Sociological Quarterly* 52 no. 4 (2011): 535.

percentage of individuals do so.”³³ Therefore, the biometric data collected can be assumed as verifiable identities, which has a significant impact in this data post-collection.

As this research reveals, moral versus ethical questions are present when looking at this type of data. Should this data be treated as immoral because it is about certain people and who they are? Or, Is it a completely independent type of data? The users willingly gave raw information to Facebook, and the algorithm they designed was able to make assumptions – terrifying assumptions – but assumptions nonetheless about who they people are, where they are from, and the specific details about their lives.

Another moral and ethical question lies within the topic of race. The majority of facial recognition software is written by white and Asian male engineers and therefore reflect features that are most known to them. This means that often times, the faces of black individuals. Now, given the intrusiveness of facial recognition, why is this a negative thing? Even if facial recognition classifications are deemed accurate, their use can undoubtedly perpetuate discrimination and exclusion. Race or ethnic classification can be used by advertisers to exclude showing certain product to a protected class, such as African-Americans.³⁴ It is a clear discriminatory aspect to this technology, which has the potential to cause many social issues within the United States.

Location

Location tracking on Facebook is considered perhaps one of the most personal forms of data that Facebook collects. It can reveal your current whereabouts, most visited places,

³³ Ellerbrok, *The Sociological Quarterly*.

³⁴ Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," *Conference on Fairness, Accountability and Transparency* (2018): 24.

and can predict future movements. The majority of the research done on this topic has been found in the reviewed journal *GeoJournal*, which is devoted to the spatially integrated social sciences and humanities.

In an article by Anthony Stefanidis, Andrew Crooks, and Jacek Radzikowski, professors of geography and geoinformation, they take a look at the basic building blocks of social media and how it “supports a greater mapping and understanding of the evolving human landscape.”³⁵ The authors main argument is that the “emergence of ambient geospatial analysis represents a second step in the evolution of geospatial data availability, following on the heels of volunteered geographical information.”³⁶ This extremely technical document has a strong focus in making the connection between ambient geospatial data and turning it into knowledge. Along these lines, there is significant dialogue which draws the connection between the data and actual human activities: “this emergence of ambient geospatial analysis represents a second step in the evolution of geospatial data availability, following on the heels of volunteered geographical information.”³⁷ Further, people have started to make progress in highlighting the issue of privacy relinquishing when they share locational information.”³⁸

This therefore begs the question: where does Facebook come into play in this conversation? In other words, what does Facebook want to do with this data? It is said by many authors I have read, that Facebook uses this geolocation data to increase the number

³⁵ Anthony Stefanidis, Andrew Crooks, and Jacek Radzikowski, “Harvesting Ambient Geospatial Information from Social Media Feeds.” *GeoJournal* 78, no. 2 (2013): 319.

³⁶ Ibid.

³⁷ Ibid, 320.

³⁸ Ibid.

of users.³⁹ Of course, as noted in Facebook's recently released article, "Open population datasets and open challenges," they wish to increase the internet availability in developing countries,⁴⁰ but the fact remains that their true motive is to increase their user population. Geographic Information System (GIS) researcher, Troy Lambert sees immense positive impacts of Facebook geolocation data. He presents ways in which this information can map "population density and crisis mapping" which "can be used to direct aid and aid workers. Epidemiologists use them to map outbreaks of disease, predict its spread, and develop strategies to contain them."⁴¹ This leads me back to the question: should people really be worried about their location being tracked? Although objectively "creepy," in the words of one of my participants, does it pose a threat to their personal privacy?

Journalist Kashmir Hill tells the story of how the location tracking can go wrong:

Last week, I met a man who suspected Facebook had tracked his location to figure out who he was meeting with. He was a dad who had recently attended a gathering for suicidal teens. The next morning, he told me, he opened Facebook to find that one of the anonymous parents at the gathering popped up as a "person you may know."

The two parents hadn't exchanged contact information (one way Facebook suggests friends is to look at your phone contacts). The only connection the two appeared to have was being in the same place at the same time, and thus their smartphones being in the same room. The man immediately checked the privacy settings on his phone and saw that

³⁹ Cho Eunjoon, et al., "Friendship and Mobility: User Movement in Location-Based Social Networks," *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2011): 1083.

⁴⁰ Tobias Tiecke, "Open population datasets and open challenges," *Facebook*, November 15, 2016, <https://code.facebook.com/posts/596471193873876/>.

⁴¹ Troy Lambert, "GIS and Artificial Intelligence Used to Build Facebook's World Population Map," *GIS Lounge*, March 1, 2016, <https://www.gislounge.com/gis-artificial-intelligence-used-build-facebooks-world-population-map/>.

Facebook "always" had access to his location. He immediately changed it to "never."⁴²

This story brings to light very serious privacy breaches that the location services may unknowingly provide its users. There are many more stories, but this particular one highlights the issues present. This topic is extremely important to investigate because Facebook may unknowingly and unintentionally be revealing people's identities in the wrong settings. Perhaps Facebook simply does not care, or they are. This data they collect seems to be directed at to self-important roles: increasing the connections of current users and increasing their overall active user population.

Chapter Summaries

This thesis will be comprised of three chapters based on three potential sets of data Facebook collects: identifying, visual, and location. Within each of these chapter, general subsections will be present in an effort to extrapolate more information from both my research and from the interviews. In general, these subsections will be the implications of that data on the perceptions of: private vs. public, policy, and the political impact.

After this general introduction to my study, supporting literature, and arguments, each chapter takes the same structural approach. I first examine the literature on each type of data collected by Facebook, then I read the Privacy Policy in regard to this data, and, finally, I bring in my participants' thoughts on this data sharing. Each chapter examines a different type of data shared on Facebook which, in order, are: identifiers, visual, and geolocation.

⁴² Kashmir Hill, "Facebook is using your phone's location to suggest new friends—which could be a privacy disaster," *The Splinter*, June 28, 2016, <https://splinternews.com/facebook-is-using-your-phones-location-to-suggest-new-f-1793857843>.

The first chapter dives deep into the raw personal data Facebook collects including your name, gender, phone number, email, and so on. It defines the process that goes into giving this information to Facebook and what happens to this data once put online. Then, a careful examination and explanation of the interviewees thoughts on the public vs. private side on this topic will be detailed, in comparison to scholarly research completed. Although closely tied to the second subsection which will break down the ever-changing Privacy Policy, this portion will be a more elementary and emotional look at how people begin to react.

The second chapter focuses on the visual data that Facebook collects. This includes a primary focus on the biometric data that is collected and possibly dispersed by Facebook. The first subsection will look at the ethics behind collecting this type of data, and what the severe implications of sharing it are. I will then take a look at the notes within the Privacy Policy and see where they are being clear and transparent and the places in which they are hiding their true motive. Further, this section will take a look at other primary sources, including U.S. Legislation that has an impact on this type of data protection. The third section of this chapter will be an analysis of my interviews and the answers that pertain to the topic of visual data. Different political parties handle privacy, especially privacy such as this in very different ways. However, I think that Millennials might transcend the historical party beliefs and may offer interesting and viable perspectives on this highly controversial subject.

My third chapter examines geolocation data. I pay close attention to the data specifically from Facebook, because often times the studies I have read present data that is collected from different studies and sources. I look specifically at data that is collected, primarily from mobile devices. Few people realize that their phones have setting that must

be turned off in order to not have Facebook track their location constantly. This is the type of data that I think people realize is being collected the least; the settings are hidden away. This may be the most obtrusive type of data. The third subsection will take into account the fact that location tracking on Facebook is considered perhaps one of the most personal forms of data that Facebook collects. It can reveal your current whereabouts, most visited places, and can predict future movements. I am sure that people will be very wary of this type of data collection and will be worried that they can always be found and traced. The interviews conducted will reveal this information and will answer questions surrounding the topics of space, place, and the implications the digital world.

Data privacy on Facebook is only on this rise and will soon have severe implications of the social and economic lives of its users. This topic must be explored, in an effort to uncover users' opinions and perhaps shed more light on this issue. Most people know that data is being collected, but few realize the expansive nature of the collection and severe impact it will soon have. My hope is that this thesis will make Millennial Facebook users realize the true nature of their online presence, and hopefully help them in being safer within this digital environment.

CHAPTER ONE: IDENTIFIERS

"The power to destroy a thing is the absolute control over it."
- Frank Herbert, *Dune*

Data, its processing, and eventual utilization is what drives Facebook forward as a corporation, trend-setter, and culture-definer. The data contributes to not only their economic success, but their significant place within society: Facebook is the most popular form of social network worldwide. But, what data is being collected about individuals, how is it being collected, and what happens to that data once it is stored? In other words, how is data about each person's identity, or what I call "identifiers," used by Facebook, and how does it shape these people's lives, particularly their privacy or lack thereof? And finally, how does the action of sharing one's identifiers turns into a politically driven issue, economics, through the creation of "free labor" in this process, both socially and economically? These questions are extremely important in understanding the larger ramifications this has on personal data and the user experience on this social network.

Identifiers are pieces of information that are given willingly to Facebook when users sign up. Many of the points are required, and Facebook will not allow their service to be used by people who do not supply them with this information. Facebook is a site where users basically recreate themselves online, so that it seems natural that this information would be surrendered.

Mark Zuckerberg, the founder, CEO, and Chairman of Facebook, acknowledged that "People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people, and that social norm is just something that's evolved over time."⁴³ In fact, through his mass amount of users, he was able to change a social norm, he noted in a speech at *Tech Crunch's* "Crunchie Awards," that changing a Privacy Policy for 350 million users isn't the type of action many companies would do. But Facebook

⁴³ *Terms and Conditions May Apply*. Documentary. Cullen Hoback. (2013. Hyrax Films).

decided that the new terms would be the social norms now, and they, according to Zuckerberg, “just went for it.”⁴⁴ It is clear that Zuckerberg (at this time) didn’t have many worries about the social implications of data collection, only that he felt as though he had the ability to rewrite entire social norms for a population of social media users.

I am particularly interested in studying identifiers on their own because they are the most personal form of data that Facebook collects. Some of the main identifiers collected can be listed as follows:

Age	Phone Number	Country	Income
First Name	Email Address	Hometown	Relationship Status
Last Name	City	School	Political Leanings
Race	State	Interests	Employer

These particular points mark some of the most important to Facebook users’ human and real-life identity. When thinking about who you are as a person, these are the data points that you usually share, as they are your identity’s most basic building blocks, at least according to machine learning. These points are shared when users first sign up for Facebook, as well as each and every time they interact with other users’ posts or add to their profile. The majority of the data collected on Facebook is raw and personal, and alone it seems minuscule and non-important. As I previously mentioned, Facebook collects points including your name, gender, phone number, email, etc. I remain fascinated and in awe, just like the FBI and other government agencies, that this type of data is willingly and easily given

⁴⁴ Mark Zuckerberg, *The Crunchies*, *Tech Crunch*, January 10, 2010.

to Facebook by its users, often without any forethought.⁴⁵ Facebook knew, as the director of the documentary, *Terms and Conditions May Apply*, that “Anonymity wasn’t profitable.”⁴⁶ Facebook is based on its users, and therefore they knew that it was their users that would make the company money.

Much of my own understanding about how Facebook works came from *DataEthics*, a think-tank based in Europe with a goal of promoting data ethical products and services and an effort to “provide knowledge, collaboration and consultation.”⁴⁷ In a recent study, the authors write, “Every one of over 1 billion Facebook users, digital workers, work averagely 20+ minutes per day on liking, commenting, and scrolling through status updates. That is more than 300.000.000 working hours of free digital labour per day.”⁴⁸ The term “digital workers” is especially poignant when thinking about this process. Every user is an unknowing Facebook employee, contributing to their success. This “free” labor in itself is incredible in terms of how people agreed so easily; however, when put in context of the lack of knowledge or realization on the topic, it is truly astonishing. Many realize that they are supplying data to Facebook, but few realize the importance this data, and the impact each user has. The process of data collection from user interaction to the final stage of utilization is extremely complex. When asked about what actually happens to their data, my participants responded with the basic knowledge that the data is collected and then targeted back towards them. Neither party truly knew, which makes me think that the real issue with data privacy is a lack of understanding.

⁴⁵ *Terms and Conditions May Apply*, Cullen Hoback.

⁴⁶ *Ibid.*

⁴⁷ “About,” *DataEthics*, <https://dataethics.eu/en/about/>.

⁴⁸ “The Massive Data Collection by Facebook – Visualized,” *DataEthics*, June 26, 2017, <https://dataethics.eu/en/facebooks-data-collection-sharelab/>.

To begin, I must break down the differences between private versus public, with reference to Facebook use and data collection. Especially in the online sphere, there tends to be a feeling of invincibility. My participants overlook the online dangers and seem to disregard the knowledge that they are putting a great deal of valuable information out in the open with little to no protection. However, they still provide information, even though they do not feel comfortable. It is an interesting dichotomy which will be discussed shortly. Throughout the interview process, I made the realization that nearly everyone was worried about their data getting in the hands of the “they.”

As I have previously mentioned, the amount of time and energy that would be needed to fully grasp the Privacy Policy as a whole is extreme. It would take 76 days to read all of the privacy policies you encounter in a year.⁴⁹ Law Professors, Alicia McDonald and Lorrie Cranor, discuss the economic implications of it. They “present a range of values and found the national opportunity cost for just the time to read policies is on the order of \$781 billion.”⁵⁰

Median Length of a privacy policy: **2,518 Words**
 Average time to read that policy: **10 minutes**
 Number of privacy policies you encounter in a year: about **1,462**
 Number of work days it would cost to read those: **76**
 Number of hours it would take US users to read them: **53.8 billion**
 Hypothetical National opportunity cost of reading
 privacy policies: **\$781 billion**

Aside from these shocking numbers, this information will help me in understanding the economic implications and the thoughts my interviewees have on this topic.

⁴⁹ Alexis Madrigal, “Reading the Privacy Policies You Encounter in a Year Would Take 76 Work Days,” March 1, 2012 *The Atlantic*, <https://www.theatlantic.com/technology/archive/2012/03/reading-the-privacy-policies-you-encounter-in-a-year-would-take-76-work-days/253851/>.

⁵⁰ Aleecia M. McDonald, and Lorrie F. Cranor, “The Cost of Reading Privacy Policies,” *A Journal of Law and Policy for the Information Society* 4, no 3 (2008): 2.

The political discussion of economics is a clear issue and is a place where I believe there will be clear divides in my results. Each political party maintains their own beliefs, especially when it comes to economics. The main difference lies within markets. The left finds their beliefs held in monitored markets while the right is far keener on free markets, and this is also clear in my research. This drastically differing view on this topic, when put within the context of understanding the data held within Facebook identifiers. In looking at this topic, I wanted to answer several key questions and understand how much subjects would respond: Is the cost of giving these identifiers away worth it? Is the price tag on privacy policies (\$781 billion) worth it? Do they provide us with any protection at all? How do beliefs on socioeconomic issues impact beliefs on this topic?

As much as they do not agree, they also have a great deal in common: their labor. To begin the understanding of economics, I turn to Karl Marx's theory on "free labor," which has been used widely in internet studies, social media studies, and digital studies to analyze the production of knowledge.⁵¹ In order to apply this theory to this research, there are three terms, which make up the process of labor that need to be broken down and defined. The process, on an elementary level is as follows: The personal activity of man (work itself), the subject of that work, and its instruments.⁵² Each of these three play a significant role in not only Marx's theory, but the understanding of how Facebook handles identifying data and the labor involved.

⁵¹ Karin Fast, Henrick Örnebring, and Michael Karlsson, "Metaphors of Free Labor: A Typology of Unpaid Work in the Media Sector," *Media Culture & Society* 36 no. 7 (2016): 964.

⁵² Karl Marx, *Capital a Critique of Political Economy. Vol. I, Book One, The Process of Production of Capital* (London: Electric Book Co., 2001): 258.

Debbie Kasper, a sociologist, thinks that the main source of worry, when it comes to social media, is the that there has recently been a drastic increase in the amount of literature.⁵³ Given that this article was written in 2007, even more worry has grown in recent years, months, and even days, mainly due to actual real-life data breaches. I found tens of thousands of blog posts, news items, and academic articles, chapter, and books in my own search, and I expect that number will continue growing.

At first, I didn't realize how my subjects were discussing the topic of data collection because I, too, had no complete understanding of what was actually happening to the data once it was collected. Being able to just say "they" have my data made sense to me too. The "they" that my participants referred to mostly made up the companies and governments Facebook is selling and sharing this identifying data to. There was little worry of Facebook having the data, but where the data was going afterwards. Also, many were worried about breaches in security where hackers could potential enter the system without validation.

In the first round of interviews, the only data my participants assumed that Facebook gathered and used were identifiers, more specifically the "simple" ones (name, age, phone number). Given this seeming simplicity of "sharing," there wasn't much worry for they saw these types of data is being unimportant. However, given the ways such data and its presentation can shape the political climate and everyday disposition—and even the elections—of the United States and its international beyond, personal data becomes more and more valuable as the days pass us by.⁵⁴

⁵³ Debbie V.S. Kasper, "Privacy as a Social Good." *Social Thought and Research* 28 (2007): 171.

⁵⁴ Kent C. Berridge and Terry E. Robinson, "What is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Salience?" *Brain Research Reviews* 28 (1998): 313.

But, before any sort of analysis can take place, political or otherwise, we must first be able to understand the process of the data collection as its intricacies help guide this conversation. However, the “they” was still a vague conglomerate for my participants. I found that after introducing them to some readings on the topic in between their two interviews with me, my interviewees were better able to understand the topic and have concrete opinions on what is actually happening with this type of data. Overall, my participants were able to more articulately present their thoughts on privacy as a whole and with more specificity to the process and implications the collection of data is having on the United States as a whole.

With the vague “they” looming in most of the world’s imaginations as possessing their data—if Facebook users even imagine anyone has their data—what emerges is a tenuous and new relationship between the private and the public. I argue that the public and the private do not exist within the Facebook world. Of course, posts may be hidden from family members, certain users can be locked, but the data collected is free to take by Facebook, for whatever use they may seem fit. This conversation regarding private and public must be brought to the forefront of conversations.

The discussion around online privacy of identifiers could be extremely simple, but, as I will show, corporations who collect this data make it difficult to ascertain their actual legal rights and possible acts of privacy. At the same time, corporations (and governments too) encourage the use of social media and make it seem like both a required social and networking act of the 21st century. Again, most people do not realize the extent to which data is taken and used against them; if they did, it remains to be seen if they would use these services. They provide the data and thus, they provide the free labor for Facebook.

This chapter is broken into three distinct parts. In the first, I describe Facebook's data collection and include key technical definitions about this information and the processes that are in place. In the second, I explore these practices in conjunction with the Facebook Privacy Policy. Finally, I discuss how these identifiers have an economic impact with reference to Marxist theory of free labor to think further about how it further produces political leanings and opinions.

Facebook Data Collection:

Facebook is constantly collecting your data. Every use, every click, and every movement on the site is recorded. Your likes for the number of years you have been of Facebook (and now dislikes), views, comments, and even interactions with friends, are recorded. Even if you use a different website that has a Facebook button on it to "share" or "like" specific pages and you are signed in to your Facebook in another tab or window or on your phone, your data is being recorded.⁵⁵ As this data piles up, algorithms are applied to discern a great deal of information about who you are in terms of various identifiers, and this information is connected to what you like, what you look like, where you go, and so on. Data points ranging from home size to your buying activity can be understood based on actions on Facebook.⁵⁶ I searched through many academic journals, and even Facebook's publications, but nowhere could I find how many data points they collect. However, I had an idea. I turned to *Facebook Business* to a section titled, "Choose Your Audience." On this single page lone I was able to record over 100 different data points by which advertisers can determine their target

⁵⁵ "Privacy Policy," August 9, 2005, <https://web.archive.org/web/20050809235134/http://www.facebook.com:80/policy.php>.

⁵⁶ Scherker, *Huffington Post*.

audience.⁵⁷ This made me realize not only the massive extent of their identifying data storage, but also how far they go to hide from plain sight the in-depth nature of their collection. This is important because it gives the first glimpse into the public and private sides to their motives.

The majority of the research completed by *DataEthics* was far too technical for a non-computer science minded individual to know. An article posted on the *DataEthics* website breaks down data collection on Facebook in an extremely technical and thorough, as well as accessible and visual manner: “Facebook Algorithmic Factory.”

This “Factory” runs through four distinct phases: data collection, storage, algorithmic processing, and finally, targeting. The researchers were able to map out the data collection on Facebook by looking at the Facebook data policy, input fields on Facebook, cookie and pixel technology on 3rd party websites, Facebook-owned companies’ policy, Facebook Vendors, service providers, and other partners and Facebook Ireland Ltd Report of Audit from 2011. From these outlets, researchers make the differentiation between data collection within and outside the Facebook platform” (see Figure 1 for *DataEthics*’ visual model).⁵⁸

Having the ability to properly define these distinctions between what happens inside of the Facebook realm and what happens outside is imperative to this mapping. Without that distinction, it is hard to accurately trace data, and thus, lose validity on the process as a whole. The researchers at *DataEthics* stated that the process of data collection is a result of “the crossing between technology and society in an effort to better understand the new,

⁵⁷ “Choose Your Audience,” *Facebook Business*, <https://www.facebook.com/business/products/ads/ad-targeting>

⁵⁸ “The Massive Data Collection by Facebook – Visualized,” *DataEthics*.

emerging forms of privacy-related risks, network neutrality and security threats.”⁵⁹ The different stages presented are complicated to explain in detail, and equally as complicating to produce visually, thus, it is important to use the two modes in conjunction. Examining this process affords a greater understanding of the technology and will move my study in the direction of understanding the politically-driven personal perceptions on the topic. After an in-depth analysis of this diagram and its corresponding descriptions over some weeks, I supply a summary analysis here. The first stage is aptly titled: Data Collection and is where the entire process begins. The collection of data stems from several different outlets: Actions and Behaviors, Account and Profile Information, Data Collection: Online Trackers, Device Information (type of phone, tablet, or computer device used, internet service provider, and browser software), and Outside Facebook Domain. These different ports of collection are based on outlets from within Facebook and from a digital footprint outside of its walls. Your “digital footprint” is the information about you that exists from all of your online interactions and use, which can include points such as your IP address (the exact address from where you are when you log on in the world) or operating system.

Looking specifically at Account and Profile information (where the majority of identifiers are held), the churning algorithm stores away identifiers. This research project points towards many specific types of data. After completing my interviews, it was clear that the following identifiers were the most important to Millennials at Trinity College:

Contact and Basic Info:	Mobile Phone, Religious Views, Political Views, Linked Accounts, Address, etc.
Family and Relationships:	Family members and Relationship Status.

⁵⁹ Ibid.

Life Events: Travel Experiences, Habits, Weight Loss, Illness, Changed Beliefs, Tattoos, Important Dates, etc.⁶⁰

These are all examples of what I term identifiers. This type of data, as I discussed earlier, is the type that users do not realize the significance, but it must be surrendered in order to even have a Facebook in the first place. As John Kleinig, et al., explains in his book, *Security and Privacy*: “Most information posted on [Facebook] is associated with the user’s real identity. In social networks such as Facebook, pseudonymity is discouraged through site usage norms and the need to provide a valid email address to register.”⁶¹ This information is inextricably tied with the real identity and thus the being of the user. The line between public and private life is so closely tied together, it is often difficult to tell them apart.

Throughout my interviews, I realized that the life events portion of identifiers were the types of data that my subjects were the most interested in Facebook having knowledge of. Of course, they knew that your name and email address were stored but, they were shocked to find out that information regarding health and family members was being taken, tracked, and made assumptions on. One particular left-leaning interviewee, William, stated his surprise in the amount of data that was actually being collected and understood:

When you actually detail what they are collecting, it’s crazy. Gambling is on the list, allergy sufferer, here are my meds, they know I have them ‘cause I buy them on Amazon. Ignorantly, I thought that Facebook kept to themselves and weren’t as spread throughout the internet.

Another left-leaning interviewee, Benjamin, stated: “It’s surprising that they can figure things out about your living arrangement. Presence of roommates, parents, that was

⁶⁰ Ibid.

⁶¹ John Kleinig and Peter Mameli, Seumas Miller, Douglas Salane and Adina Schwartz, “Surveillance Technologies and Economies,” In *Security and Privacy*, 134.

surprising. Political views aren't hard to figure out, based on what they are posting." It seems from Millennials I spoke to that there is an awareness of data collection, but when it comes down to the specifics, they are shocked and irritated. People are aware that their basic identifiers are collected, but more "personal" or private elements, such as health, are a big surprise. This means that their political views on the left side of the spectrum worry far more about social implications than those of economics. They care about their identity and who they are as people. Having health be a big concern showed me that they had a fear of people perhaps being treated differently based on conditions they may have. It is a social issue.

With the topic of discussion focused on Identifiers, my participants were mostly worried about the personal implications regarding the basics of their identity. There were no references made on the topic of race, gender, sexuality, or class profiling.

Users may originally put this information out to share with our friends, because we want "friends" closest to us to know this information. However, we often do not realize that data is all being stored and analyzed to piece together who you really are as a person. It is important to note, as *DataEthics* does, the significant difference between account information and profile information: account information is basically static information that is rarely updated, depending on the direct input you give. Profile information is more of a variable because it can contain misleading or faulty information.⁶² For example, when first having a Facebook listed many of my friends as family members, I made up particular relationship statuses. Much of my information was false, which could lead Facebook to misrepresent who I am as a person. Daily activities and behaviors on Facebook are "dynamic

⁶² "Immaterial Labor and Data Harvesting," *Share Foundation*, August 21, 2016, <https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/>.

and represents what you like, share, create and interact with in real time.”⁶³ In other words, Facebook could use algorithms to determine what is untrue, but, then again, such algorithms can make assumptions about us and be shared with or sold other corporations, governments, or individuals.

The final aspects of data collection come from your overall digital footprint. To further my previous definition, the “digital footprint” is collected from online trackers, your device information, and Facebook services found on other sites, such as Instagram or WhatsApp. When analyzing your digital footprint and your cookies in conjunction, Facebook’s algorithms look for and reveal behavioral patterns online.⁶⁴ Cookies are small pieces of code used to store information on web browsers. They can be used to store and receive identifiers and other information on computers, phones, and other devices.⁶⁵ Facebook has no opposition to this practice; they clearly state in their current privacy policy clearly that they take information from third party partners such as your experiences and interactions, especially with advertisers and use it for any purpose they see fit.⁶⁶ Overall, this portion of the process sets the stage for the next steps, and ultimately, the most worrisome portions. This section focuses strictly on collection, which in essence is harmless, but paves the path for identities to be exploited.

The second and third stage of the Facebook algorithmic Factory are Storage and Algorithmic Processing. It is best to discuss these in conjunction because the processes are closely tied together. All of the data previously discussed is categorized in “stores,” as well

⁶³ Ibid.

⁶⁴ Ibid.

⁶⁵ “Cookies & Other Storage Technologies,” *Facebook*, <https://www.facebook.com/policy/cookies/printable>.

⁶⁶ “Data Policy,” *Facebook*.

as specific information regarding each data point such as time and location. These stores (action, profile, edge and content) are the massive warehouses which keep all of your raw data. In this state it is both useless, as no true meaning can be discerned.

When the raw data is filtered into this next phase, it is where the algorithmic process comes into play. Algorithms, in simple terms, are simply chronological lists of instructions for computers to carry out operations.⁶⁷ Although that sounds simple enough, the best way to introduce algorithms is to realize their complexity. This is best summed up in the title of Adrienne LaFrance's *The Atlantic* article: "Not Even the People Who Write Algorithms Really Know How They Work." This opaque quality to algorithms is especially true of one that has the immense size of Facebook's. It is an interesting dichotomy: algorithms rely on its capabilities, without knowing what exactly it is capable of.⁶⁸ The data that had been held in "stores" is transferred along to the algorithmic processing. When the data is stored and analyzed through algorithms, Facebook can discern information about who you are as a person and will then choose what ads, statuses, and political information to share with you.

The algorithmic process is the most difficult, complex, and important piece of this system, and it is broken down into six distinct sections. To begin this, the most important place to begin is to actually define an algorithm, its purpose, and how it works. As *DataEthics* writes:

What is an algorithm? Although for the purpose of storytelling it would be much more appealing to attribute algorithms with some superpowers, in most cases, we speak about some really amazing piece of code that applies some advanced statistical or analytical methods. The definition of an algorithm is: A procedure for solving a mathematical problem in a finite number of steps that frequently involves repetition of an operation; broadly: a step-by-step

⁶⁷ Taina Bucher, "Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook." *new media & society* 14, no. 7 (2012): 1172.

⁶⁸ Meyer, *The Atlantic*.

procedure for solving a problem or accomplishing some end especially by a computer.”⁶⁹

I include this long definition for two reasons. The first, it lays forth the basic understanding that algorithms are confusing and, in many ways, without an in-depth understanding, they can come off as being a supernatural device. They have the capability to complete unfathomable problems, while not showing themselves at all. Algorithms are hidden behind the veil of website design, but drive companies like Facebook forward. The second important piece to the definition is that it lays forth a basic definition that is easily understood. In looking at many of the definitions online, they were far too technical and unnecessarily confusing. *DataEthics* makes clear that algorithms are lines of code which has the ability to solve problems and decipher information mathematically.

This portion on the algorithm is best understood as the middle point between the storage phase and the targeting phase. The algorithm makes determinations on your data which in turn is forwarded to the located most important to the data discovered.

For my research, I want to highlight one, if not the most invasive part of this algorithm: Inferring Users Household Income. This process begins with the algorithm receiving information from user profiles, posted content, and external sources about users. By filtering through different data points such as place of work (current and past), educational institution, life events, family relations, and marriage status, the algorithm has the ability to map the user into a particular income bracket.⁷⁰ It may seem as though this determination could be arbitrary but, Facebook has a solution for that. They further take into account the false information people may post online and analyze behavior, visited websites,

⁶⁹ “Immaterial Labor and Data Harvesting,” *Share Foundation*.

⁷⁰ *Ibid*.

and purchases. None of my participants had any clue about this, and were all greatly shocked, especially my right-leaning participants. This is most likely due to a previous point I made, which is that they are far more worried about the economics of data, which include their personal monetary information.

All of these data points, when combined, allow Facebook to make an extremely educated guess on the actual income a user may have. Further, the algorithm has a machine learning function which has the ability to detect, based on all of the other factors, when faulty information is given, or when people have forgotten to update their location, place of work or relationship status.⁷¹ Overall, this algorithmic process allows Facebook to understand who we are as people and our social and financial status. Even though people may block others from seeing certain pieces of their information, Facebook still can take and decipher all of your data as they wish.

The Privacy Policy

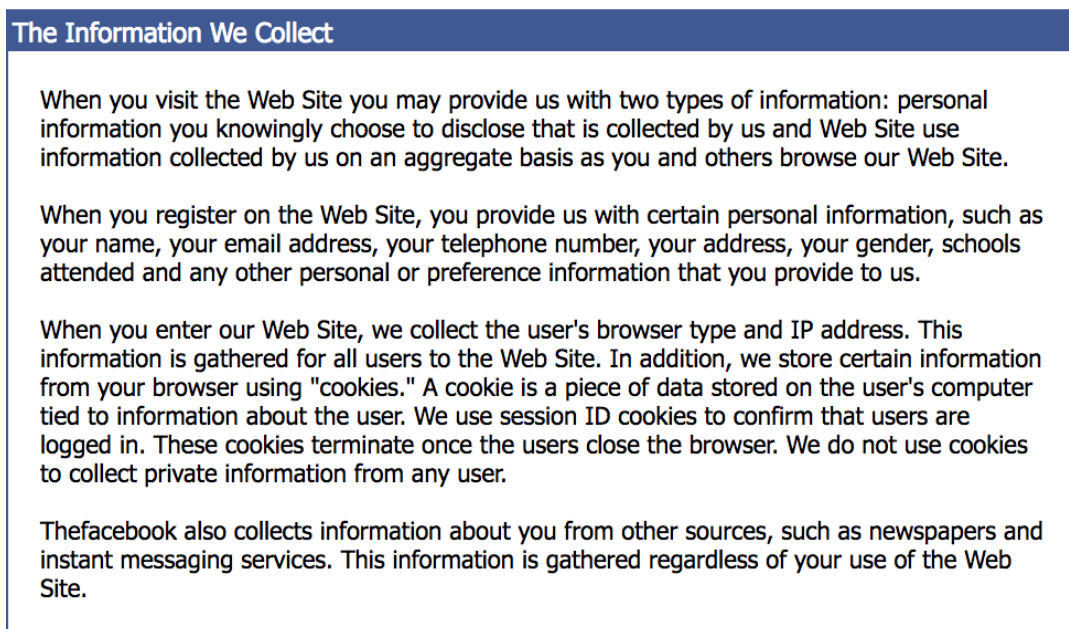
This conversation on the algorithmic process us brings to a discussion of Facebook's Privacy Policy. This document found on their website has grown and evolved drastically over the years and provides a glimpse into the legality of their practices of data collection and utilization.

In order to fully understand the Privacy Policy, we must go back to their beginning. In an effort to find the first Privacy Policy published by Facebook I turned to the *WayBack Machine*. This open source online service allows its users to trace archived websites, and to actually use them as well. The first privacy policy released by Facebook, archived by the

⁷¹ Ibid.

WayBack Machine, was published on August 6th, 2005. Then called “The Facebook” or “Thefacebook,” and geared exclusively towards elite college and university students. The first line reads: “Thefacebook Privacy Policy is designed to assist you in understanding how we collect and use the personal information that you provide to us and to assist you in making informed decisions when using Facebook web site located at www.thefacebook.com (the “Web Site”).”⁷² It is clear that at this point in time, the small company was trying to be as clear and concise as possible when it came to privacy. Its small number of users were solely college students who were using the site to connect with students from their universities and others as well.

With that, in this first policy, Thefacebook company make mention of several different types of data they collected, and how they use it:



*Figure 1: "The Information We Collect"*⁷³

⁷² "Privacy Policy," August 9, 2005, <https://web.archive.org/web/20050809235134/http://www.facebook.com:80/policy.php>.

⁷³ "Privacy Policy," August 9, 2005, <https://web.archive.org/web/20050809235134/http://www.facebook.com:80/policy.php>.

The key term in this document is “cookies.” These little bits of information follow users around the internet, multiplying as we interact with different websites. Facebooks’ Policy gives a brief introduction to this topic, but they state that their only use is to confirm the user is logged in to the browser. However, it seems as though they have a different mechanism for collecting data when off of Facebook. They collect data about you “regardless of your use of the Web Site.” Therefore, TheFacebook has alternate technological systems in place to track users online, and to learn more about their actions.

In addition to the brief explanation of the information they collect, they also make mention of the changes they can, and will, make (Figure 2).

Changes in Our Privacy Policy

We reserve the right to change our privacy policy at any time. If we do this, we will post the changes to this policy on this page and will indicate at the top of this page the policy's effective date. We therefore encourage you to refer to this policy on an ongoing basis so that you understand our current privacy policy.

Figure 2: "Changes in Our Privacy Policy"⁷⁴

From the beginning of the site’s use, Facebook has had the ability to change their privacy policy and just keep it on this page, and not publicize it. In other words, they will post the changes, but it is not often that users are looking at the privacy page to learn more about how their data is being protected. Before starting this project, I had never looked at the Privacy Policy on Facebook before. The fact that it never occurred me to look—or millions of others

⁷⁴ Ibid.

my age, as well as those younger and older—is definitely an eye-opening example of how Facebook users come to understand their own data privacy.

Over the years, the Facebook Privacy Policy has changed. It has grown in both length and depth. I definitely attribute this as a response to the growing concerns of their users. The new privacy policy, specifically its portion on data, has included how they use the information, how it is collected, how it is shared, how will changes be notified, and interestingly, how information can be managed or deleted. The descriptions for each of these are extensive and make up an extremely small portion of the privacy policy as a whole.

In order to really understand how digital labor plays out on Facebook, it must be placed within the proper context. According to the *ShareLab* research, we can look at Facebook through two different types of society: information and algorithmic. The article makes clear a difference between the two, but I find them inextricably tied together when looking at it through the lens of Facebook specifically.

Type of Society	Who is performing the labor	Objects of Labor	Instruments of Labor	Product
Information Society	Human Workers	Information, Knowledge	Offices, Computers	Business, educational, intellectual products and services
Algorithmic Society	Algorithms	Digital content, digital footprint, metadata	Social networks, digital platforms, devices.	Profiles, patterns, anomalies, predictions

*Table 1: Immaterial Labor*⁷⁵

⁷⁵ “Immaterial Labor and Data Harvesting,” *Share Foundation*.

When referencing back to *DataEthics'* "Algorithmic Factory," we can understand these two types of society and apply them directly to the identifiers: an information society and an algorithmic society. These types of society drastically differ; however, when put into the of Facebook, they must be discussed in conjunction. Information societies force citizens to be life-long, self-directed learners with filtering skills and tools perhaps even more powerful than finding skills and tools.⁷⁶ Written in 1999, I find the argument outdated. I argue that society has taken an information society and entities like Facebook have used the knowledge to transform us into an algorithmic society. Humans are the ones who give Facebook identifying data, but this is where things get complicated – there is one key difference. The raw materials in the process that include data, content, and metadata are considered the objects of labor as they are created by humans, but the actual labor is performed by algorithms.⁷⁷ There is a very fine line between giving the data, and working on the data, the information society can be seen as a catalyst in this process. In this, privacy is not static. Its breach takes form when the data is attempted to be understood. When it is run through algorithms and presented to companies for their use.

When asking my participants about whether or not they should be able to opt out of data collection, the discussion of free labor was raised. Participants offered very different responses based on political affiliations. When asking a left-leaning research participant, William, "Do you think you should be able to opt out of data collection?" He replied, "Yes, most definitely. I think that your data is your property and there is definitely a considerable amount of, it's not a black and white thing. They can sell your photos and videos and info to

⁷⁶ Gary Marchionini, "Educating Responsible Citizens in the Information Society," *Educational Technology* 39, no 2 (1999): 17.

⁷⁷ "Immaterial Labor and Data Harvesting," *Share Foundation*.

companies – you won't be compensated at all." Whereas, a right-leaning Samantha made the statement that they believe they should be able to do anything they wish with your data. You sign up for the website, you agree to their terms, and they should be able to do anything they wish to do with it. Facebook is its own independent company, and there should be no government or social pressure for what they do. This shows that the right-leaning participant has an extremely economic lens on this topic, and sides firmly with business privacy.

From this, it is clear that the left sees the algorithmic society, as a capitalistic process to make these companies richer. Michael Samway, a policy advisor at the United States Department of the Treasury, discusses how the only way to combat this is to seek change within the business structure. Taking data is a way to make money but, the authors argue that financial success can be attained by learning from their experiences and continuing to build a better company. However, a "better company" requires responsible decision-making around human rights continued to be a core part of the business. They are that this will help global obligations to users, as well as increasing financial success.⁷⁸ Further, Samway uses smart diplomatic business practices as the way to conduct business, rather than deception – human rights are important, and the government intervention impacts this.⁷⁹ In this, clear solutions to these capitalistic tendencies are discussed. There seem to be simple ways in which social media companies, such as Facebook, can still make the money they need to make, while increasing the quality in how they treat their users.

⁷⁸ Michael A. Samway and Warren Ryan, "The Internet, Human Rights, and the Private Sector," *Georgetown Journal of International Affairs* 15, no 1 (Winter/Spring 2014): 25.

⁷⁹ *Ibid.*

Identifiers are the most basic form of data that Facebook takes and utilizes. It is clear that there is a definite process this data takes. The process makes clear the extreme scale of the algorithmic factory. The Facebook Privacy Policy details the fact that things can change without informing their users. People must take it upon themselves, to take the time and do research to find out what they need to think about in terms of privacy. Further, the Privacy Policy is directed at what they know is most important: data. Finally, when looking at the labor of Facebook through a labor perspective, it is clear that the idea of “free” labor causes stirs politically. Both sides of the political spectrum understand that free labor is an occurrence, but they respond to it in very different ways, a telling sign of political stances in the United States. This discussion matters because the historical and contemporary intrinsic differences in opinions between political thoughts, when put in the framework of a modern technological platform will help in understanding the future of social media interaction.

The left and the right of the political spectrum respond to this data collection very differently. This is important because it demonstrates the major issue that is facing this specific issue, and the climate of the United States’ political scene as a whole. The two-party system is severely impeding progress and holding society back. I believe that this research will help open the eyes of close-minded individuals and make them realize that there is a solid knowledge basis around everyone’s beliefs; realization is the first step towards progress, and this research is helping bridge that gap.

CHAPTER 2: FACIAL RECOGNITION

“Facial recognition software can pick a person out in a crowd, but a vending machine can't recognize a bill with a bent corner.”

- Jeff Dwoskin

On December 19th, Facebook announced new and optional tools to “help people better manage their identity on Facebook using face recognition.”⁸⁰ To announce these tools, Facebook released a mass notification to the entire Facebook community:

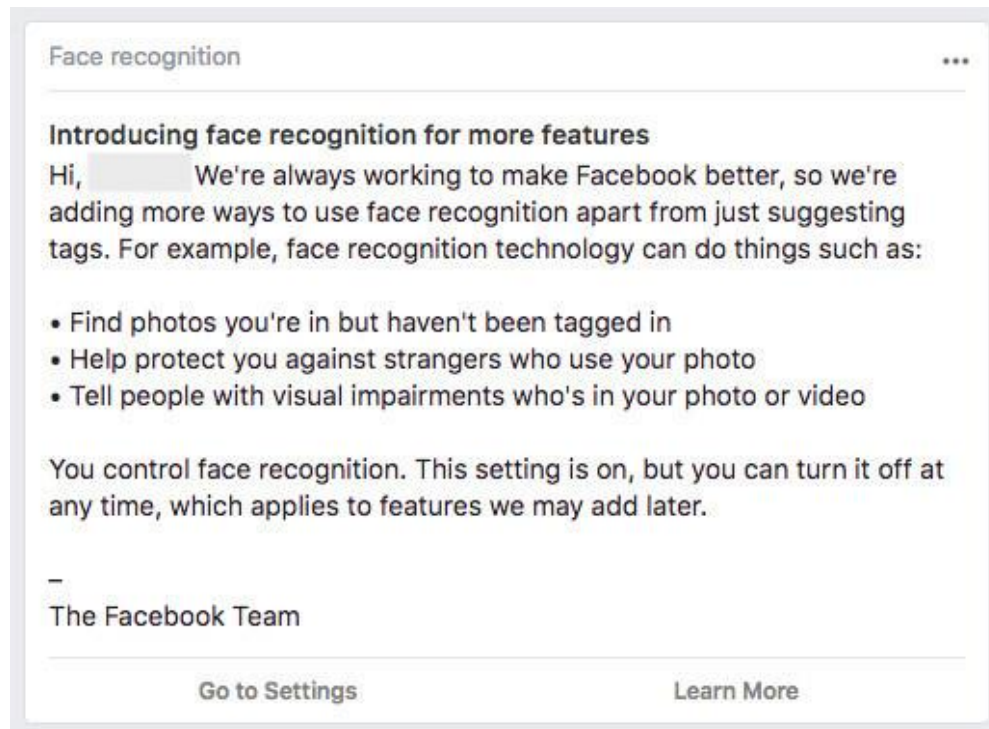


Figure 3: Facebook's New Facial Recognition

This notification briefly made mention of the new features that were being added to Facebook and which are geared towards the use of facial recognition technology.

I argue that this seemingly unimportant notification was actually very important for three reasons. First, it provides a rare example of Facebook actually informing its users of new technologies that would be automatically turned on. Many of the settings Facebook has,

⁸⁰ Joaquin Quiñonero Candela, “Managing Your Identity on Facebook with Face Recognition Technology,” *Facebook Newsroom*, December 19, 2017, <https://newsroom.fb.com/news/2017/12/managing-your-identity-on-facebook-with-face-recognition-technology/>

per their Privacy policy, are hidden and need to be searched for in order to be realized. Second, it shows that Facebook frames what I and many scholars, politicians, journalists, and citizens understand to be invasions of privacy in a very positive light. Facebook does not want to highlight their invasive tendencies. But, if we play along with Facebook's plan, the third and final reason is that this feature was only been rolled out to users located in the United States. The technology blog *Verge* reporter Russell Brandom explained, "The feature [wouldn't] roll out to Facebook users in Canada or the European Union because of stricter regulations on the collection of facial data in those regions."⁸¹ This geographic discrepancy brings to light a key conversation that needs to happen when discussing privacy: legislation.⁸² The United States protects their citizens only to an extent, and this political line must be analyzed.

Users of Facebook find facial recognition data personal, even more so than identifiers like name, address, email, and so on. With the many varying definitions of what exactly facial recognition data is, I clarify a definition for this chapter. For a simple definition, it is a computer application automatically identifying or verifying a person from an image or a video. It is a combination of systems: biometric, image processing, computer vision, and machine learning.⁸³ Yet, facial recognition is not simple. After pouring over a plethora of journal articles, newspaper sources, and technical process drawings, I determined that the most important explanation of facial recognition, drawing from a *Gizmodo* article:

⁸¹ Russell Brandom, "Facebook is starting to tell more users about facial recognition," *The Verge*, February 27, 2018, <https://www.theverge.com/2018/2/27/17058268/facebook-facial-recognition-notification-opt-out>

⁸² Lucía Tello, "Intimacy and «Extimacy» in Social Networks. Ethical Boundaries of Facebook," *Comunicar* 21, no. 41 (2013): 210.

⁸³ Ming Yang, "Face Recognition Systems," *Slideshare Lecture*, July 17, 2014, <https://www.slideshare.net/milkers/lecture-10-ming-yang-face-recognition-systems>.

“Not all biometrics are the same. If you wanted to take my fingerprint from me, you have to get me to touch an object that needs something. If you want to do an iris scan you probably have to get pretty close to me to do that. Facial recognition operates at a distance. One of the things that comes with the facial biometrics is that it’s remotely capturable and is capturable without your consent.”⁸⁴

The important takeaway here is that biometric data (biologically driven data points), happens at a distance. No one necessarily be present for facial recognition data to be collected. It can be collected silently, without anyone’s knowledge. And further, with Facebook, it is collected through every user’s labor. Users are giving the data to them with each picture posted, tagged, and commented upon.

Unlike identifying data, Facebook gives us the option to “opt out” of facial data collection, but with considerable difficulty. Using a computer, the process takes four steps, and five on mobile devices. More importantly, no one knows that you have this ability, i.e. Facebook does not advertise this “opt out” option. None of my interview subjects knew that this was a possibility and assumed their facial recognition data was given up permanently. Also, even though I now know that this is the case, why do I not go into Facebook right now and shut off this setting? Admittedly, there is a part of me that likes being able to be tagged in pictures automatically. The instantaneous effect of learning a picture has been posted with me in it is an addicting one, which I don’t want to lose.⁸⁵

Facebook Newsroom is the site that updates users about functionality and business related to Facebook and is written by Facebook. In a post on the site published on December 19, 2017, titled “Hard Questions: Should I Be Afraid of Face Recognition Technology?” the

⁸⁴ Masse, *Gizmodo*.

⁸⁵ Indeok Song, et al., "Internet Gratifications and Internet Addiction: On the Uses and Abuses of New Media." *Cyberpsychology & Behavior* 7, no. 4 (2004): 384.

Deputy Chief Privacy Officer of Facebook, Rob Sherman, addresses the concerns over facial recognition software. This paper reveals that Facebook knows that its many users and society as a whole are uneasy with facial recognition. Sherman writes that “this tension isn’t new. Society often welcomes the benefit of a new innovation while struggling to harness its potential.”⁸⁶ Although I can tell that this article is staging Facebook’s technology is a positive light, I sympathize with one point, which made me pause. He cites the Kodak camera in the late 19th century as a similar situation. This inexpensive, consumer-grade camera equipment came to market and allowed photography to be available to the masses. This “new terror for the picnic” singlehandedly changed the way history was documented. While this technology could have been restricted by society, it was instead regulated by “prohibiting stalking or letting people sue for invasion of privacy,” rather than requiring licenses to use ‘camera technology’ or written consent forms before a person could appear in a photo.”⁸⁷ Thus, what resulted was that people familiarized themselves with these cameras, social norms evolved, and “the world decided that the benefits of personal photography far outweighed the risks.”⁸⁸

After further consideration, Sherman’s comparison is mind-boggling. Being a photographer myself, I had never stepped back to really think about how people may have reacted when cameras were placed in the hands of the consumer. People were now able to take picture, and capture moments, people, and objects indefinitely. I question facial recognition but, I have no desire to turn this setting off because the hidden facial recognition

⁸⁶ Rob Sherman, “Hard Questions: Should I Be Afraid of Face Recognition Technology?” *Facebook Newsroom*, December 19, 2017, <https://newsroom.fb.com/news/2017/12/hard-questions-should-i-be-afraid-of-face-recognition-technology/>.

⁸⁷ *Ibid.*

⁸⁸ *Ibid.*

is something that is so engrained in my mind. How has Facebook engrained this ideas into not only my mind, but the vast majority of its users?

Facebook has the facial recognition setting automatically turned on. As one of my right-leaning interviewees, Charlie, discussed, there are many windows that pop up, no matter the social network that I click “Accept” without reading. Such a response is just the same as accepting terms and conditions. Charlie instinctively accepts these suggested structures of his life and his data without an actual understanding of what terms he is accepting. Although Facebook supposedly has no control over this action by their users, they structure a sense of forced acceptance.⁸⁹

There is also a serious lack of understanding with the implications of the data being taken, stored, and utilized. As discussed extensively in the previous chapter, the Facebook sphere is not contained, it has tendrils which are far reaching into not only the web, but society itself by collecting data wherever you wander when the site or app is open.⁹⁰ Thus, in order to enlighten people on this topic, a discussion of the technical process is in order (with simplified diagrams to follow), as this may open the eyes of users in an effort to keep their wits about them when in contact with new policies online.

The Basic Process

In order to understand the process of how Facebook’s facial recognition system works, we must turn to several sources. Like the process of collecting identifiers,

⁸⁹ Robert E. Wilson, Samuel D. Gosling, and Lindsay T. Graham, "A Review of Facebook Research in the Social Sciences," *Perspectives on psychological science* 7, no. 3 (2012): 212.

⁹⁰ Arnold Roosendaal, "Facebook Tracks and Traces Everyone: Like This!" *Tilburg Law School Legal Studies Research Paper Series* no. 3 (2011): 8.

understanding how Facebook collects and analyzes visual data is complicated, and, in fact, I found it more difficult to understand than the identifying data collection process. A research document published by Facebook's AI research team breaks down the technical construction and overall implementation of the facial recognition software. I turn to this publication for a basic understanding, and then to more technical documents to understand the actual process.

Facebook employs a system known as DeepFace, which was created by a Facebook research group. Simply put, it is a facial recognition system which has "closed the majority of the remaining gap in the most popular benchmark in unconstrained face recognition and is now at the brink of human level accuracy."⁹¹ When the software was tested, it was trained on a dataset which included over four million images, of over 4,000 of their users. Per Facebook, the "method reaches an accuracy of 97.35% . . . reducing the error of the current state of the art by more than 27%, closely approaching human-level performance."⁹² Notably, Facebook makes no mention of the range of races, ages, genders, ethnicities, and so on in this data sample of photos which is used to represent so many.

This process' basic structure can be understood through four distinct phases: (1.) Detect (2.) Align (3.) Represent and (4.) Classify (As seen in figure 2). Although the technical aspects of this facial recognition software are unimportant to this thesis, the basic premise of its design is necessary in understanding its extreme implications. Thus, this simple diagram below will suffice.

⁹¹ Yaniv Taigman, Ming Yang, Marc' Aurelio Ranzato, and Lior Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," *Facebook AI Research*, (June 24, 2014): 1.

⁹² *Ibid.*

The process begins with “detection.” In this phase, the faces of the individuals present in the pictures are found and made aware to the system. Next, the faces are extracted and perfectly aligned. In this, the defining and of facial landmarks occurs and builds the number of landmark detectors the algorithm is able to detect. Also, this step allows the images all to be standardized to allow the process to work faster, and with better results. Representation then occurs. At this stage, the biometric data harvested is broken down into numbers and categories. This feeds into the classification process where the data is cross referenced with other data points to detect the individual pictured.

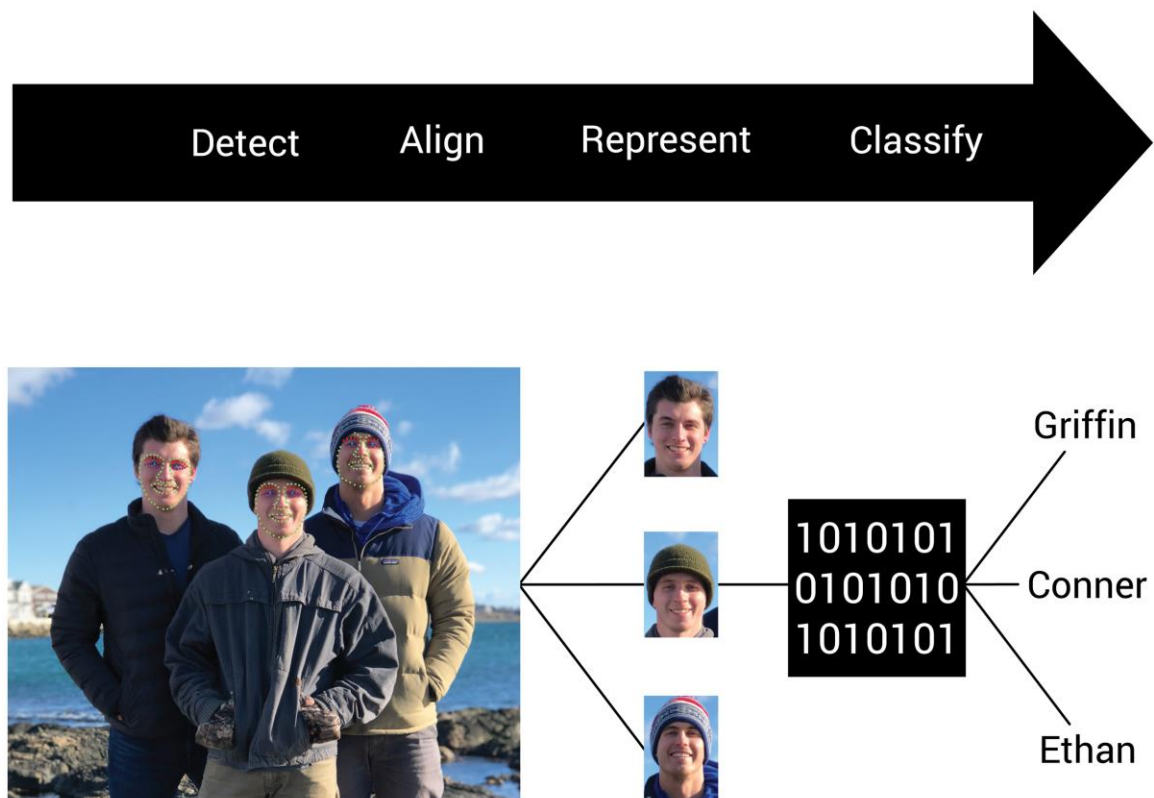


Figure 4: Facebook's Facial Recognition Pipeline

Yaniv Taigman, et al., the designers of the software and algorithms of DeepFace wrote about their work on Facebook's own AI research blog. They realize the previous limitations

of past facial recognition software, and believe that the software Facebook currently employs has and will overcome them:

“An ideal face classifier would recognize faces in accuracy that is only matched by humans. The underlying face descriptor would need to be invariant to pose, illumination, expression, and image quality. It should also be general, in the sense that it could be applied to various populations with little modifications, if any at all.⁹³

While the technical process is interesting, it is far more important and relevant to examine the social issues this process intersects with. The authors go on:

Our work demonstrates that coupling a 3D model-based alignment with large capacity feedforward models can effectively learn from many examples to overcome the drawbacks and limitations of previous methods. The ability to present a marked improvement in face recognition, attests to the potential of such coupling to become significant in other vision domains as well.”⁹⁴

The idea of three-dimensional technology and its implications will be discussed far more in-depth shortly but, this quote provides a realization that Facebook actually has the ability to create three-dimensional representation of its users. They can literally recreate our bodies using this technology and understand human anatomy on a level that has previously been unattainable.

In this, what are the key social issues that present themselves when discussing facial recognition software? We have briefly mentioned several, but in order to understand where Millennial politics differ, we must explore this topic with more depth. In the article, “What’s the Worst That Could Happen with Huge Databases of Facial Biometric Data?” author and journalist Bryson Masse opens by asking: “What can happen when we combine the large amount of facial biometrics data with a potentially imperfect system? ... What sort of societal

⁹³ Ibid, 8.

⁹⁴ Ibid.

implications would there be if you were recognized by someone, anywhere and everywhere you went?"⁹⁵ The discussion and answer to these questions is imperative to this study. The answers provide real-life impacts that facial recognition algorithms can have on the public. Whether it's due to simply not having the knowledge or because they do not care, the implications are often not realized by the users of Facebook.

Relatedly, my interviewees often referenced one of the most major applications of facial detection: the threat to global, national, and local security. I will soon be discussing security when it comes to both international and domestic terrorism, but in brief, national security and the visual significance tied is important to mention now. This discussion has its focus on the different types of security and how the public views its ever-increasingly normalized implementation. An *Economist* article published in September of 2017 the many global uses for facial recognition that may seem positive or menial:

America facial recognition is used by churches to track worshippers' attendance; in Britain, by retailers to spot past shoplifters. This year Welsh police used it to arrest a suspect outside a football game. In China, it verifies the identities of ride-hailing drivers, permits tourists to enter attractions and lets people pay for things with a smile.⁹⁶

But, so what, right? Our fingerprints are used as an identifying piece of biometric information—why is that not controversial? The *Economist* authors of this article begin by making the quality comparison between fingerprint and facial recognition, which parallels the earlier comparisons 19th century Kodak and image collection today. Much like my own thinking about data privacy and the power of algorithms, both comparisons seemed odd, but it also makes perfect sense why journalists begin here. Fingerprint recognition has become

⁹⁵ Bryson Masse, *Gizmodo*.

⁹⁶ "What Machines can tell from your face," *The Economist*.

the standard for security and is accepted by many, not as an invasion of their privacy. It is used in airports, it is used if you are arrested, and it is even used to unlock smart devices.⁹⁷

Even though this type of biometric data is considered a way to be tracked and known, people seem to be completely fine with it. It is a standard of security that has become enmeshed and accepted within society. However, as *The Economist* article contends, “One big difference between faces and other biometric data, such as fingerprints, is that they work at a distance. Anyone with a phone can take a picture for facial-recognition programs to use.”⁹⁸ The application for facial recognition is far easier to be used without the identified knowing it. It works behind the scenes and can identify people without consent. These are contributing factors to the reasons why facial recognition is a challenging subject, and thus, can be a basis for controversy.

I presented my interviewees with the question, “How do you think we as a culture have consented to his behavior of giving our data to Facebook?” Some of my subjects took to the now historical idea of the creation of phonebooks, and how there was no real issue with this when people began to have the name, address, and phone numbers of strangers in their houses. Of course, I was not alive at this time, but I have never heard about any uproar over this event as well as the normalization around this form of data collection and sharing.

What was different? One of my subjects, left-leaning William, pointed towards the propensity for repression that Americans experience. Citizens of the United States are “going to repress the ugly parts of it and just take the more beneficial parts because it’s easier to do that. Throughout your day, you’re not worried about your intellectual property on the

⁹⁷ Salil Prabhakar, et al., “Biometric Recognition: Security and Privacy Concerns.” *IEEE security & privacy* 99, no. 2 (2003): 40.

⁹⁸ “What Machines can tell from your face,” *The Economist*.

internet, or your data rights. It's much easier to get away from the world on Facebook." In other words, Americans have become complicit because they're getting something in return. Many of my subjects had thought about deleting their Facebook's but ended up not because there are definite advantages to having an account which cannot be replicated anywhere else in life. In this, we surrender the data that we hold most important, in return for being able to connect with people in ways which we previously were unable. Both political parties agreed on this topic, meaning that this topic transcended their fundamental beliefs.

Further, when looking specifically at facial recognition, there are many data points which are possible to spot simply through the collection of face biometrics. Facebook's algorithm reveals any and all information they maintain about you, simply based off of your face. In other words, when combined with the identifiers, based on your posts, activities, soon-to-be mentioned location, and everything else you do, and then run through the Facebook algorithm, they most likely have the ability to know far more about you than you would care to have them know.

One of the most well-known of these seemingly private data points is, shockingly, sexuality. Although this topic can be categorized as an identifier, Facebook is purportedly using its facial recognition software to discern their users' sexuality, an intimate detail. As I will describe below, the findings are dubious, but the media attention and response given by academics means that such algorithms are assumed to be correct, whether they are or not. Research conducted at Stanford University by Michal Kosinski and Yilun Wang has shown that "faces contain much more information about sexual orientation than can be perceived and interpreted by the human brain," the authors further argue that their "findings advance

our understanding of the origins of sexual orientation and the limits of human perception.”⁹⁹ On top of this, Kosinski, et al. mention that “machine vision can infer sexual orientation by analyzing people’s faces.”¹⁰⁰ This argument suggests, whether correctly or not, that this facial recognition software can pick up extremely subtle differences in facial structures. Additionally, given that companies and governments are increasingly using computer vision algorithms to detect people’s intimate traits, their findings expose a threat to the privacy and safety of gay men and women.¹⁰¹

Much larger than the original DeepFace study dataset, this study relied on the following data: 130,741 images of 36,630 men, 170,360 images of 38,593 women which were downloaded from an American dating website. After their software reduced the images to ones that held single-face with sufficient clarity, they were left with, “35,326 pictures of 14,776 people, with gay and straight, male and female, all represented evenly.”¹⁰² These images were fed into a piece of software called, “VGG-Face” which represented each person as a number. Kosinski and Wang call these “faceprints.” Using a “simple predictive” model, the software then found correlations between the images and their owners’ sexuality. After this occurred, the model was run and outperformed humans at distinguishing between gay and straight faces.¹⁰³

When shown one photo each of a gay and straight man, both chosen at random, the model distinguished between them correctly 81% of the time. When shown five photos of each man, it attributed sexuality correctly 91% of the time. The model performed worse with women, telling gay and straight apart with 71% accuracy after looking at one photo, and 83% accuracy after five. In

⁹⁹ “Advances in AI are used to spot signs of sexuality.” *The Economist*.

¹⁰⁰ *Ibid*.

¹⁰¹ Michal Kosinski and Yilun Wang, “Deep neural networks are more accurate than humans at detecting sexual orientation from facial images,” *Journal of Personality and Social Psychology*, 1.

¹⁰² “Advances in AI are used to spot signs of sexuality.” *The Economist*.

¹⁰³ Kosinski, et al., 13.

both cases the level of performance far outstrips human ability to make this distinction. Using the same images, people could tell gay from straight 61% of the time for men, and 54% of the time for women. This aligns with research which suggests humans can determine sexuality from faces at only just better than chance.¹⁰⁴

In this, the researchers find there to be some possible explanations for the performance of their model. The most primary of this is biological. They cite that fetuses, while developing in the womb are exposed to various levels of testosterone which play a significant role in the development of facial structures and may have a role in determining sexuality. Facial structures are, obviously, the key attention of the research. Specifically, the nose, eyes, eyebrows, cheeks, hairline and chin for determining male sexuality; the nose, mouth corners, hair and neckline were more important for women.¹⁰⁵

Unsurprisingly, this study has severe limitations and profound controversies. First, with the images coming directly from an online dating service, there is a far greater likelihood that they would be “revealing of sexual orientation.”¹⁰⁶ Also, there is a 91% accuracy rate which only applies when the two images being referenced has one individual who is shown to be gay – they reference that the accuracy of this testing outside of the lab would be far lower.¹⁰⁷

The researchers of the study have made clear that their study was not geared towards outing those who may be gay, but rather to make clear the potential power that machine learning possesses.¹⁰⁸ This erosion of privacy is inevitable and the dangers of it must be

¹⁰⁴ “Advances in AI are used to spot signs of sexuality.” *The Economist*.

¹⁰⁵ Kosinski, et al., 18.

¹⁰⁶ Ibid, 26.

¹⁰⁷ Ibid, 31.

¹⁰⁸ Ibid, 7.

understood.¹⁰⁹ There are parts of the world “where being gay is socially unacceptable, or illegal, such software could pose a serious threat to safety. Dr. Kosinski is at pains to make clear that he has invented no new technology, merely bolted together software and data that are readily available to anyone with an internet connection.”¹¹⁰ I found it interesting to also point out that the technology used was rather rudimentary. I believe that if Facebook had the desire to do this, there may be severe implications. The Economist article details that if a company obtains the right data sets (which Facebook undoubtedly possesses), similar AI systems might be trained to spot other intimate traits, such as IQ or political views. Just because humans are unable to see the signs in faces does not mean that machines cannot do so.¹¹¹

Many disagree with the accuracy of these findings as reductionist and sexual profiling akin to other sorts of profiling. However, what’s important to think about here is how much corporations like Facebook believe these findings. This this research definitely has been met with some controversy, which is telling of the social implications facial recognition software holds. However, the researchers have made clear that their study was not geared towards outing those who may be gay, but rather to make clear the potential power that machine learning possesses.¹¹² This erosion of privacy is inevitable and the dangers of it must be understood.¹¹³ There are parts of the world “where being gay is socially unacceptable, or illegal, such software could pose a serious threat to safety. Dr. Kosinski is at pains to make clear that he has invented no new technology, merely bolted together software and data that

¹⁰⁹ Ibid, 8.

¹¹⁰ Kosinski, et al., *Journal of Personality and Social Psychology*, 18.

¹¹¹ “Advances in AI are used to spot signs of sexuality.” *The Economist*.

¹¹² Ibid, 7.

¹¹³ Kosinski, et al., *Journal of Personality and Social Psychology*, 8.

are readily available to anyone with an internet connection.”¹¹⁴ I found it interesting to also point out that the technology used was rather rudimentary. I believe that if Facebook had the desire to do this, there may be severe implications. The Economist article details that if a company obtains the right data sets (which Facebook undoubtedly possesses), similar AI systems might be trained to spot other intimate traits, such as IQ or political views. Just because humans are unable to see the signs in faces does not mean that machines cannot do so.¹¹⁵

In furthering this, surveillance studies scholar, Ariane Ellerbrok, details the important process of how facial recognition software came to be within the Facebook sphere. As previously mentioned, most of its success is due to the implementation of the of *Face.com*. This service, when put in perspective on Facebook, tagged over 400 million images in one month out of Facebook’s 10 billion photo archive.¹¹⁶ Further, real-world identifications have become a part of the program, and part of the problem. The images online are inextricably tied to reality. This is due to the fact that “Facebook members are expected (as delineated in the Terms of Service) to use their “real identity” for their Facebook profile, and thus considerable evidence that a large percentage of individuals do so.”¹¹⁷ Therefore, the biometric data collected can be assumed as verifiable identities, which has a significant impact in this data post-collection. When comparing this to the discussion of sexuality, if Facebook were to use this software for malicious purposes, they have the ability to do so,

¹¹⁴ Ibid, 18.

¹¹⁵ “Advances in AI are used to spot signs of sexuality.” *The Economist*,

¹¹⁶ Ellerbrok, *The Sociological Quarterly*.

¹¹⁷ Ibid.

and the results would not stay contained within the confines of Facebook. No, this data would most likely be sold, shared, and distributed for purely capital gains.

Also, facial recognition software is sometimes, if not often, racist. For example, in 2015 a photo app created by Google identified black people in some photos as gorillas.¹¹⁸ Perhaps facial recognition software should not be made for public use until it is guaranteed to not make racist gestures such as this. It divides the United States visually and causes more problems than it solves. Anthropologist Shaka McGlotten argues for a form of “black data,” as a response to big data that stereotypes, categorizes, and enacts violence against people of color. McGlotten’s suggested response is to create a blackout in data, to go dark, to opt out.¹¹⁹

Notably, Facebook has always included “gender” options, in order to collect and categorize more data. But, in February of 2014, they professed their radical advancements in gender politics with the addition of 54 new and progressive gender categories. Communications scholars Rena Bivens and Oliver Haimson argue that these seem radical but are actually intermediaries that entrust “social media platforms with a considerable degree of control over the generation of broader categorization systems, which can be wielded to shape the perceived needs and desires of both users and advertising clients.”¹²⁰ As forward thinking as they their gender politics seem, in 2014, Facebook deleted the pages of drag queens, native American, first nation, and indigenous people.¹²¹ The former group was not allowed to have two profiles, “their birth identity and their drag identity,” and the latter’s

¹¹⁸ Frederick A. Miller, Judith H. Katz, and Roger Gans, “The OD Imperative to Add Inclusion to the Algorithms of Artificial Intelligence,” *OD Practitioner* 50, no. 1 (2018): 7.

¹¹⁹ Shaka McGlotten, “Black Data,” In *No Tea, No Shade* ed. Partick Johnson (Durham: Duke University Press, 2016): 263.

¹²⁰ Rena Bivens and Oliver L. Haimson, “Baking gender into social media design: how platforms shape categories for users and advertisers.” *Social Media+ Society* 2, no. 4 (2016): 1.

¹²¹

names were deemed to unbelievable to exist. As much as Facebook seems forward thinking, they instead reinforce the norms of gender, sexuality, and race with each click.

One final social and economic issue that I will cover here is income. Once again, this is an identifier I previously discussed, but it is important to note that there are visual algorithmic tools employed by Face book within this category. Facebook's algorithm uses recognition techniques to map income brackets. This takes the form of image analysis to recognize brands the users wear on photos they upload, and how often certain brand names are used in posts and searches on their website.¹²² This of course, has many social implications. The brand names that people wear (or do not wear) are often socially tied with income level. To infer this information about people is dangerous. Income is a dividing topic, and the targeting of certain ads may contribute to the separation of society.

There have been many companies and groups who have implemented facial recognition software; however, it is Facebook's who is truly the best. I suggest that this is mainly due to the vast amount of free labor they have at their fingertips.¹²³ Training a system is extremely complicated, complex, and expensive, and domain specific data is an alternative to the costly human data collection and labeling.¹²⁴ When you think about it, Facebook users have done all of the work for them. When Facebook first began tagging pictures, only the detection process was in existence. We, the users, did all of the work for Facebook. This "free" labor which has previously been discussed extensively is once again present. We supplied

¹²² Immaterial Labor and Data Harvesting," *Share Foundation*.

¹²³ Mary L. Gray, et al., "The Humans Working Behind the AI Curtain," *Harvard Business Review* (January 9, 2017): <https://hbr.org/2017/01/the-humans-working-behind-the-ai-curtain>.

¹²⁴ Masi I., Trần A.T., Hassner T., Leksut J.T., Medioni G., "Do We Really Need to Collect Millions of Faces for Effective Face Recognition?" *Computer Vision – ECCV 2016*, September 16, 2016, 593.

Facebook with the data they required and has created a facial recognition software that far exceeds the FBI's.¹²⁵ I would argue that the FBI's recognition software compared to Facebook's can be understood with the following graphics:

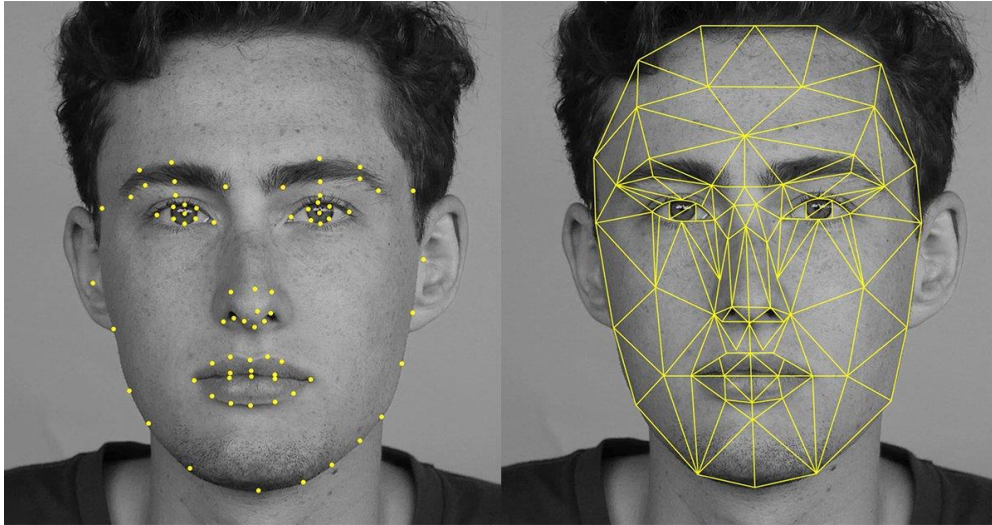


Figure 5: FBI 2D

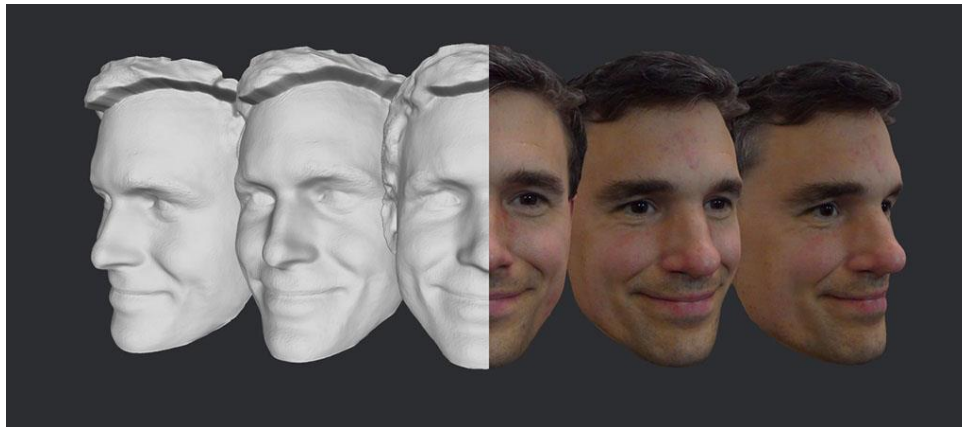


Figure 6: Facebook 3D

¹²⁵ Jennifer Lynch, "FBI Plans to Have 52 Million Photos in its NGI Face Recognition Database by Next Year," *Electronic Frontier Foundation*, April 14, 2014, <https://www.eff.org/deeplinks/2014/04/fbi-plans-have-52-million-photos-its-ngi-face-recognition-database-next-year>.

The majority of the FBI's images in their database are two dimensional. They rely on images from passports, licenses, mugshots, and other forms of identification where the images are lined up forward facing. Whereas Facebook has the ability to train their software with virtually every single angle an individual post's online. With its users tagging every image, and thus confirming an identity, I would argue that Facebook has the undeniable ability, if they needed, to create 3D composites of their users for their own purposes. This allows for far more applications in the real world, especially pertaining to the topic of security. People may be able to not only be identified via their facial features, but an entire body rendering based off of Facebook images.

Prominent digital law partner Christopher Dore says, "Facebook has the largest database of recognition data in the world, period," and "when you have a situation where a company is holding a database of that type, there are a lot of concerns that come up," including "what are they going to do with it?"¹²⁶ Dore mentions that Facebook is currently only really using this feature to tag people in images, but they have the power to do so much more. This data stored could be sold off to stores for marketing purposes, surveillance, and identification.¹²⁷ Especially when it comes to this 3D potential, there should be many worries if the government were to force Facebook to hand over not only its software, but their entire database of images.

Most simply put, the facial recognition software Facebook has at its disposal is fairly terrifying. The extreme implications it could have on the world are extreme and there is little standing in their way. Many of my interview subjects were aware and worried about their

¹²⁶ Masse, *Gizmodo*.

¹²⁷ Masse, *Gizmodo*.

images and their privacy setting, but not in reference to Facebook's data collection at all. They were mostly using privacy setting to keep photos hidden from certain members of their friend group and family. Perhaps if people were made aware of the data that is being collected, there would be more worry directed towards the Facebook side of things. As I have mentioned time and time again, we, as users of Facebook need to realize the importance of our data. We need to not underestimate its power, and the significance of each click, like, and tag we perform. The data is being collected, and although it is not being utilized to its full (and alarming) capacity, it has the potential to be.

CHAPTER THREE: GEOLOCATION

Ron Swanson: "So it learns information about me? Seems like an invasion of privacy."

April Ludgate: "Dude, if you think that's bad, go to Google Earth and type in your address."

*Ron Swanson: *Throws computer in dumpster**

- Parks & Rec S4 E9

No one wants their location tracked – your exact location is the ultimate form of privacy; geospatial data (data with a location) provides the intimate connection between the virtual and physical world. If a device knows where you are, then anyone can find your location. This has become clearer in recent times, as Facebook launched their new service within its *Messenger* app which allows users not only to text but also to share their location with their “friends.” However, seemingly harmless, if you forget to turn off your location services then your location data is continuously collected and utilized by Facebook. This constant gathering of millions if not billions of users’ locations repeats Facebook’s pattern of providing a “free” service to its users. Facebook presents such services in ways they know their users will be excited about, while Facebook benefits from the data that the software allows them to collect. The motivation is hidden behind an illusion of technological progress and connection, and until examined or read about, remains hidden.

There are two specific types of location data that is collected: *position-aware services* rely on a device’s knowledge of its own location while *location-tracking services* find their basis on other parties tracking a specific user’s location.¹²⁸ Both of these services are important to examine, especially as to which people are more worried about. Tech design scholar Louise Barkhuus and human-computer interaction scholar Anind Dey look at these two types of location data collection methods in an attempt to understand which type people are more worried about. This discussion of these two types of collection, of course, brings to light the notion of private vs. public, which is also a look into real life vs. reality. Both of these

¹²⁸ Louise Barkhuus and Anind Dey, "Location-Based Services for Mobile Telephony: A Study of Users' Privacy Concerns," *Interact*, vol. 3 (2003): 702.

dichotomies play role in the political perceptions of Facebook's data collection and allows for a greater understanding of what privacy means to American Millennials.

Location services are primarily derived from mobile devices. Based on my review of the literature and interviews with participants, many think this data was gathered intentionally over time, as companies realized that mobile devices are constantly turned on in the literally mobile pockets of individuals. Location data is important, and it is valuable to companies as well as governments. Although regulated and protected from this in the European Union through Law 32002L0058 Section 14, the United States has no laws protecting the privacy of cell phone users. In fact, the US has no laws protecting the privacy of its data besides a law on online child pornography.¹²⁹

The collected and analyzed data can be aggregated and linked to real people.¹³⁰ This of course, is a worry because the publishing or sharing of anonymized location data has the ability to lead to privacy risks.¹³¹ Even with laws present, tech scholar Hui Zang and Jean Bolot, VP of Research & Innovation at Technicolor Artificial Intelligence Lab (AI Lab) in Silicon Valley, detail that the anonymization of location data does not even work because enough locations reveal where people are. In order to reduce privacy risks, at the bare minimum, the collected data needs to be coarse in either time domain or space domain.¹³²

¹²⁹ Iris A. Junglas and Richard T. Watson. "Location-Based Services." *Communications of the ACM* 51, no. 3 (2008): 67.

¹³⁰ Carlo Ratti, Dennis Frenchman, Riccardo Maria Pulselli, and Sarah Williams, "Mobile landscapes: using location data from cell phones for urban analysis," *Environment and Planning B: Planning and Design* 33, no. 5 (2006): 740.

¹³¹ Hui Zang and Jean Bolot. "Anonymization of location data does not work: A large-scale measurement study." In *Proceedings of the 17th annual international conference on Mobile computing and networking*, ACM, 2011. 155.

¹³² Ibid.

Further, software-sorted geographies (S-SG) play a role in both this discussion and the shaping politics in advanced societies. The processes are multifaceted, complex and ambivalent; however, they have recently been found to have separated groups of people along the lines of privilege and marginalization.¹³³ This is a clear political and social issue which must be addressed when examining the use of geospatial data.

When reflecting on the vast amounts of identity and visual data I have discussed and adding in geospatial data within the context of software-sorted geographies makes it seem as though Facebook has the ability to know exactly who we are as people. Our identities, our photos, and our locations have long been a private concern, but with the advent and mass growth in popularity of both mobile devices and social networks, our data is slowly becoming a pseudo-public entity as it becomes owned and analyzed by corporations, shared with governments, and pilfered by hackers. My interview subjects reacted to this “inevitable” data collection in very different ways, again along political leanings. Participants brought up a range of social issues as they related to locational data, including the differences between public and private understandings to data, the Flint water crisis, and, a tragically recurring event, school shootings. The political implications of these will be discussed shortly.

Section One: Private versus Public & Reality versus Virtual

Just three years ago, the *Harvard Business Review* printed research findings from Timothy Morey, the VP of innovation strategy at frog (a global product strategy and design firm), Theodore “Theo” Forbath who is the global VP of digital transformation at Cognizant

¹³³ Stephen D. N. Graham, “Software-Sorted Geographies,” In *The People, Place and Space Reader*, eds. Giesecking, Mangold, Katz, Low, Saegert (New York: Routledge, 2005): 133.

(an American IT consulting service), and Allison Schoop, an associate Strategy Director at frog. While looking at users' perceptions of their data and its use, Morey, et al. note that:

Consumers worry about how their personal data is gathered and used, they're surprisingly ignorant of what data they reveal when they're online, and most companies opt not to enlighten them. This dynamic erodes trust in firms and customers' willingness to share information.¹³⁴

In other words, the transparency between consumer and corporation is nonexistent. The users have no true comprehension of what their data reveals, which is a great concern.

In the Barkhuus and Dey article, they predicted that that location-based services would eventually be the most common form of context-aware computing, which are applications used to track users' locations¹³⁵ By understanding the level of worry within my users and the technical aspects of this computing, I can then lay forth new uses for the data, in a positive manner as well as tips for Facebook users to protect their data as much as possible, based off of their specific worries.

In both interview sections, I asked my subjects the following question: "Where do you define the line between privacy in real life and privacy online?" This question proved to be difficult, as few had thought about this distinction. The line between public life and online life is so often unclear, that people struggle to even think about the distinction.

Barkhuus and Dey also predicted that as "mobile telephony" becomes increasingly common as a handheld computing platform, "location-tracking of mobile phones enables location-based services to spread outside closed environments."¹³⁶ The differentiation

¹³⁴ Timothy Morey, Theodore Forbath, and Allison Schoop, "Customer Data: Designing for Transparency and Trust," *Harvard Business Review*, (May 2015): 9.

¹³⁵ Peter Ljungstrand, "Context awareness and mobile phones," *Personal & Ubiquitous Computing* 5, no. 1 (2001): 58.

¹³⁶ Barkhuus, 702.

between public and private comes down to the geospatial advancements in technology. Especially within the Millennial population, life before this technology is widely unknown and unremembered. Millennials, myself included, don't remember a time before Google maps, and therefore don't remember when we could send our physical location virtually. This process and constant interchangeability between data existing in real life (your physical location) to then being shared virtually (Facebook Messenger) and then back to the physical (finding someone's location) is common-place, and people rarely reflect on that matter.

They also found that people are positive towards the *location-based services* or *position aware services*, as long as they perceive them to be useful. We also found that *location-tracking services* generate more concern than position-based ones.¹³⁷ As a reminder, location-based services are those that rely upon a device's knowledge of its own location while *location-tracking services* find their basis on other parties tracking a specific user's location. The key difference between these being the physical device, and the third-party applications, such as Facebook. This discovery lends well to my research, as it proves that people are more worried about a website such as Facebook claim and use the data, rather than simply a device. While the device may be storing this information, it is Facebook who has the ability to actively analyze and learn for the collection process.

I realized my interviews prompted people of my generation to challenge their previous thoughts and relay the first-time reflection on this topic; it was eye-opening. In one particular interview with a white, left-leaning male individual, privacy was understood to be tangible and abstract, depending on its context. In real life, its tacit: "You have your house, your room, your belongings, there is a tactile nature to it. Whereas in the internet there is a

¹³⁷ Barkhuus, 712.

comfortable illusion of privacy which I think allows people to keep coming back to use it.” This showed me that there was a definite understanding that privacy online is different. This subject furthered this by saying that there always a “key” to unlock the door on the internet. In fact, “I hold no belief that my email addresses or even my back account is safe from someone who wants to get into it.” These personal identifiers were a worry, but there was no discussion referencing location data. It wasn’t until prompted that he made any comment. “I imagine they're collecting it even when I try to make it so they can't.” On top of this, he expressed that their agency in real and virtual life were polar opposites: “I can close my doors and lock them up, but you only have the illusion of that online. People can still get on it, and if they couldn’t there wouldn’t be an industry to protect your privacy online.” This subject understood that there really is only an illusion online.

Another subject, left-leaning Vickie, discussed the idea of “space” as a form of privacy. Space, both physical and virtual were important, but when talking about “space” there was much more concern over privacy in real life, with reference to her house and room. To her, online was a private space, hidden from her parents, peers, and potential employers, and you “need to be much more careful about what you’re putting out there.” However, she made no mention to the data being collected by Facebook, with potential to be exploited. Rather, she was far more worried about her peers intruding on this space.

Privacy, as I will discuss more in-depth soon, to right-leaning Samantha, was in reference to international terror and financials. The worry of hackers was definitely real to them as well. Right away, they focused on issues of location, as Samantha stated, “location privacy is important to me. I know that if I tag myself as being in Prague then they would know I was there, but I don’t want them to currently know and store my location at all times,

especially with hackers these days.” It is clear that the right-wing participants are worried about international issues. She immediately focused on a different country as the origins of her worry.

Participants in both left and right political parties were aware of location playing a role, but there was no in-depth understanding like most of the US population. They all also accepted and/or assumed that there was nothing they could really do about it. In fact, my findings lined up directly with a study conducted by human-computer interaction scholars Mark Ackerman, et al. Their study found that while there is a concern for privacy among Millennials, it greatly depends on two factors: what type of information users give up, and the usefulness of the application to the user.¹³⁸ In other words, people’s responses to this data collection has a direct correlation to the usefulness for them. If the user is benefiting from the data collection, then they have no worries.

The idea of “usefulness” of data is a topic that my participants discussed time and time again. Facebook’s motives are apt to be hidden for users while they focus on fulfilling their own desire in using the Facebook app or site. For example: one of my interview subject, referenced “data” as the places he had been, events he had attended, and the locations he had checked into. For him, this was deemed “okay”, because these interactions with the site were useful, and not perceived as dangerous. This means that he does worry about his data as long as he could forget about it when he *gains* something from Facebook. It seems to be a major psychological advantage that Facebook maintains over their users.

¹³⁸ Ackerman, Mark S., Lorrie Faith Cranor and Joseph Reagle. “Privacy in e-commerce: Examining User Scenarios and Privacy Preferences.” *EC* (1999): 1-8.

I want to break this down further by looking specifically through his description of attendance of an event listed on Facebook. Events on Facebook are tagged with many different pieces of data: what it is, where the event is located, what time it is at, who will be there, who are the hosts, etc. For example, several weeks ago I marked myself as attending *Trintoberfest*, an event here at Trinity College. From me simply pressing the button “Going,” Facebook can understand what type of people are attending based off of content posted, mutual friends, and proximity to the event location. From here, targeted marketing can occur (as detailed in Figure 7.).

As discussed in Chapter One, Facebook data is held in different types of *stores*. In this case, the data for the events are derived from the following: User Profile, Edge Store, and Content Store. From here, the data is gathered. The previously discussed data points are analyzed in conjunction. Temporal Proximity Analysis (metric that measures the distance, in temporal units, between a user interested in an event and the time of the event), Event History Analysis (examines previous attendance history of the users in association with the retrieved user profiles), and Event Inference Module (determines the users that may be inferred to attend described in a targeting event criteria).¹³⁹ All of these pieces of analyzed data are then fed into an ad targeting module. A visual representation is below:

¹³⁹ “The Massive Data Collection by Facebook – Visualized,” *DataEthics*.

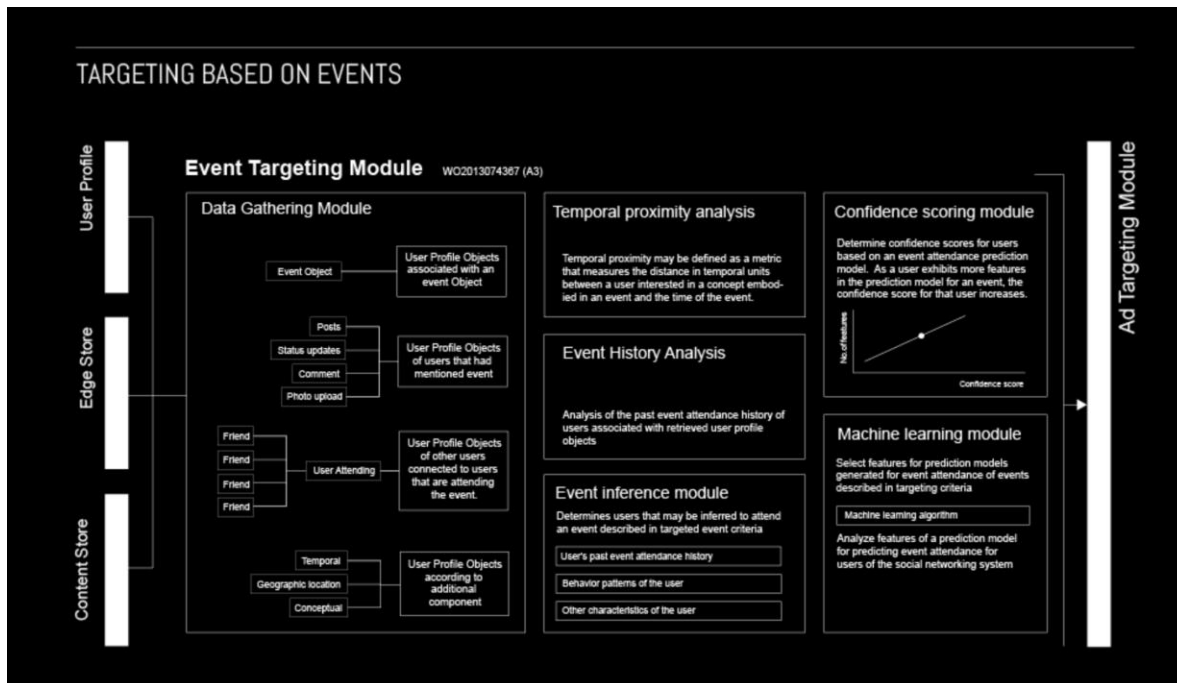


Figure 7: Targeting Based on Events¹⁴⁰

First, it is clear that this process is primarily based off of location data. Although there are identifiers present, the main focus is on location. It is also clear, that there are motives beyond simply making the user happy. In fact, through all of the research I have pored over, there has really been to the user's satisfaction as a driving motive. Only the notion that the satisfaction of users' keeps ulterior motives hidden.

Many of my participants like many other Facebook users ask: so what? Many people reveal their location, but what's the big deal? I often heard in my interviews, from several different individuals, that they know they are only a data point, and they don't put anything online that is self-incriminating.

One of the biggest issues related to the geolocation conversation is issues of control. Computer Science scholars Alexandra-Mihaela Olteanu, et al, sought to quantify

¹⁴⁰ Ibid.

interdependent privacy risks with location data. They write that “a user’s location privacy is no longer entirely in her control,” and that, “the individual location information disclosed by other users significantly affect her own location privacy.”¹⁴¹ In other words, anonymity is actually impossible.¹⁴² Even though you are simply a data point, there are many ways for your identity to be found through the use of your location data collected by Facebook. The majority of documents dealing with privacy only make reference to identifiers.¹⁴³

Location data is important to protect because it makes a quality comparison between identifying data and location data. As Barkuus and Dey write, “Identity has several aspects to it and we consider a person’s position to be a specific attribute of identity, like full name and social security number. The major difference between location and most other attributes is that location changes continually and is mostly relevant to mobile computing.”¹⁴⁴ Location privacy is undoubtedly important, and its constant protection means even more. In order to explore this more in depth, I want to turn attention towards the technical aspects of location data collection, as this will lead the conversation directly into several social impacts and possible solutions to the problem.

It’s Time to Check-In: Promoting Geolocation Data Collection on Facebook

¹⁴¹ Alexandra-Mihaela Olteanu, Kévin Huguenin, Reza Shokri, Mathias Humbert, and Jean-Pierre Hubaux, “Quantifying Interdependent Privacy Risks with Location Data,” *IEEE Transactions on Mobile Computing* 16, no. 3 (March 1, 2017): 841.

¹⁴² Hui Zang et al. 155.

¹⁴³ Marc Langheinrich, “A Privacy Awareness System for Ubiquitous Computing Environments,” *Proceedings of UbiComp* (2002): 237.

¹⁴⁴ Barkhuus, 713.

On a very basic level, Facebook's location collection is based off of its users' images, created data files, and device locations, including your specific geographic locations using GPS, Bluetooth, or Wi-Fi signals.¹⁴⁵ Although I briefly mentioned a process earlier examining the data surrounding an event, I want to provide you with an overview of the entire location data collection process.

On November 18, 2009, now general partner at Google Ventures, M.G. Siegler wrote an article in the popular tech blog *TechCrunch*, titled, "Location Is the Missing Link Between Social Networks and The Real World." It is thus unsurprising that less than a year later, Facebook introduced their "Check-In" function which allowed users to use the GPS already installed on their phone to let their "friends" of their exact location. Soon enough, new articles were published with titles such as "How Facebook Will Own All Your Location Data" and "The Five Stages of Facebook Grief." These articles were at the forefront of this topic and provide an early glimpse into the public's perception of location services. In hindsight, it is remarkable that the collection of data, without proper protection, has continually grown without much pushback from its users. People, time and time again, are allowing the social media superpower to control the data, without any real complaint.

From the inception of Facebook collecting location data from its users, the process has looked something like this diagram I created.

¹⁴⁵ "Data Policy," *Facebook*.

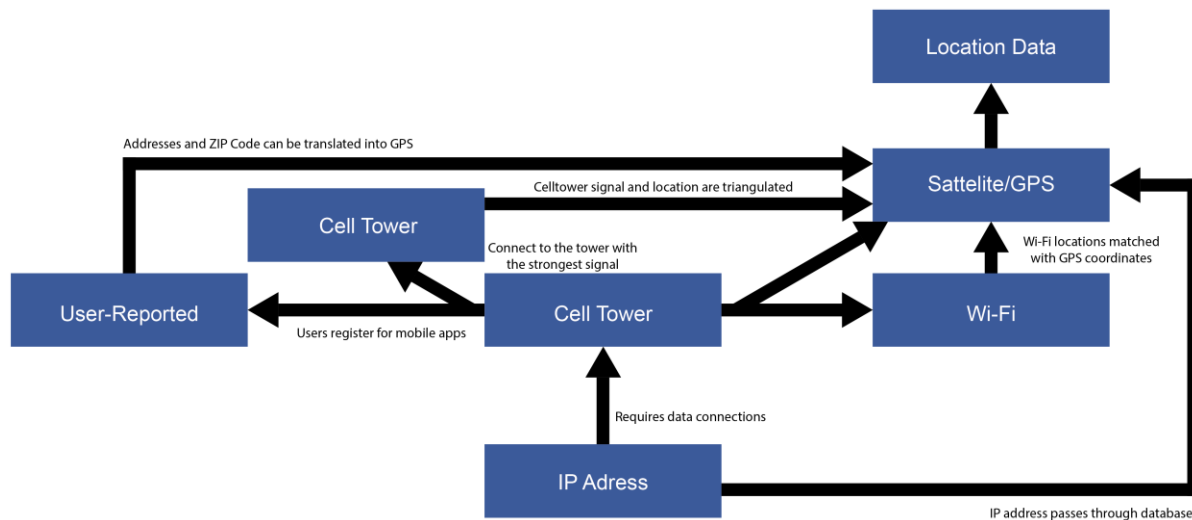


Figure 8: How Location Data is Collected

As previously mentioned, the majority of Facebook's collection is done through cell phone data. What I like most about this diagram is it not only accounts for not only the data transmitted by cell phones, but the now intrinsic connection between mobile devices and mobile apps. This means that the user reported data, in these apps, are being collected and factored into the location data. It is no longer merely location data, rather now a conglomerate of geospatial and geo-identifying data. Each one of the entities' pictures plays a key role in data collection, and also shows that the data can and is being collected constantly. Facebook gives the option to opt out of this data, but other companies have fallen into public eye as not truly following the setting prescribed by their users, and it did so via one of the pictured mechanisms: cell towers.¹⁴⁶

¹⁴⁶ Shannon Liao, "Google admits it tracked user location data even when the setting was turned off," *The Verge*, November 21, 2017, <https://www.theverge.com/2017/11/21/16684818>.

Given the technology available to bypass the settings and continue to collect data, it's difficult to think that Google has been the only company to do this. The impact of this technology and the resultant information is profound, both to personal safety and security, and more general rights to privacy. It also has a great deal to do with social justice.

Section Three: Flint Water Crisis & Other Issues

In connecting the impressive technology with my previous discussion of private versus public, I wish to bring up several recent issues brought up by my interviewees. The first, being the Flint Water Crisis, which has plagued the patrons of Flint, Michigan since 2012, and the second being school shootings. Both of these are issues which were discussed extensively by my interview subjects, with reference to the ways in which Facebook data should be used. Both political factions brought these topics up, but the way they were talked about and discussed prove extremely different.

Throughout all of my interviews, the discussion of location data has evoked precise attention to recent events in the United States that must be addressed. This has shown me that students at Trinity College are in-tune with current events, at least domestically, and have the ability to make connections between these topics through the lens of data collection. The key difference in student perceptions however, is where their political beliefs come into play. The same topics and themes were brought up, but individuals, depending on their political leanings, differed in their understanding and reaction of the topics.

Geographers David Swanlund and Nadine Schuurman recently wrote an article titled: "Mechanism Matters: Data Production for Geosurveillance." This article breaks down geosurveillance mechanisms into three distinct classifications: geolocation, unique

identification, and the surveillance medium.¹⁴⁷ Through this distinction, the authors detail that “we, as subjects, did not choose yet are increasingly forced to negotiate,” and that the, “mechanisms are both numerous and highly complex and are only one component within large ecosystems of geosurveillance, making privacy ever more evasive.”¹⁴⁸ They mention, as I have stated throughout this project, that the key to find any prospects for intervention is to understand the basis of the mechanisms at hand.

Given the basis of understanding that has now been attained, it is important to fulfill the mission of this project: to provide a solution, through the examples of current events, that can help not only the users of Facebook to be smarter with their actions online and help them realize what is actually happening, but also make a stand against this company. To force Facebook to increase their level of transparency with their users, and to promote new, updated, and bipartisan legislation or technological solutions will protect the rights of American citizens.¹⁴⁹

Swanlund and Schuurman cite a Canadian Broadcasting Corporation (CBC News) article published on January 30, 2014. This article, they mention, provides examples which epitomize impingements on geo-privacy and how these intrusions prove the combination of both large corporations with governments as means of representing the prevalent corporate and government control over data.¹⁵⁰ But, this raises an important question: if this data

¹⁴⁷ David Swanlund and Nadine Schuurman, “Mechanism Matters: Data Production for Geosurveillance,” *Annals of the American Association of Geographers* 106 no. 5 (2016): 1064.

¹⁴⁸ Ibid.

¹⁴⁹ Finn Brunton and Helen Nissenbaum, “Political and Ethical Perspectives on Data Obfuscation,” in *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology* (New York: Routledge, 2013): 168.

¹⁵⁰ Swanlund, et al., 1066.

collection is unstoppable, and we are already accustomed to its nature, can it be used for good? Can there be significant positive benefits as well?

Both left and right-leaning subjects brought up the Flint Water Crisis, an event which began in 2012 and is continuing today. An event which has impacted the town of Flint, Michigan by exposing 6,000 to 12,000 children to extreme amounts of lead.¹⁵¹ An event which led to a public health state of emergency. William explained that he believes most of the data is dealt with via capitalistic tendencies. For example, “CDC tracks where the flu is breaking out based on Google I imagine they buy data from Facebook as well. I imagine they’re using the data to map the entire world in a 4D way. And they’re not keeping that to themselves, they’re selling that.” When pressing him further for more of his thoughts, he noted his thinking about large corporations. He acknowledged that they have more power than the government because they don’t play by the same rules – they don’t have to. “I don’t know what insidious things Facebook is getting up to but, I’m pretty sure it’s not altruistic. I don’t think Zuckerberg is collecting the data, so he can *fix* the Flint Water Crisis or track where the next may be. If he is, he’s going to hedge his bets against people wherever it will happen. It won’t be used to right wrongs in the world. If it will be, someone will benefit beforehand. It will be used in a capitalistic way.”

In opposition to this viewpoint, the right-leaning individuals don’t find Facebook responsible for doing anything with the data. Of course, they sympathize with the individuals affected, but they stand by private companies and their rights. As Charlie stated, “I’m fine with data collection – allows company to be more effective and efficient. Facebook is a

¹⁵¹ Maggie Fox, “Flint Water Crisis: Feds Expand Programs to Help Kids Affected by Lead,” *NBC News*, March 2, 2016, <https://www.nbcnews.com/storyline/flint-water-crisis/flint-water-crisis-feds-expand-programs-help-kids-affected-lead-n530556>.

private company, and no one is forcing you to take part in their data collection. You sign up, you're agreeing to the Terms and Conditions, and they can use it as they wish." While another, Stephen, stated, "I think it's fine. No one is forcing you to use the site, it is a voluntary commitment. It's a private company and the goal of any private company is to maximize profit and they can do that by using private information, in a legal manner." When it comes to national security, the right-leaning individuals maintain their focus on international terror groups. Facebook has users around the world, and data should be used if its analysis detects a national security issue. Democrats however, made no mention to international terror. Rather, they raised concern with domestic issues, such as school shootings.

In this, there are clear divides in political perceptions on proper use of this location data. The right sees differences in these two events. They do not view the Flint Water Crisis as a major societal issue, and only see the economic implications held within. Thus, they do not believe that privacy should be breached. The left-wing however sees them both as ways in which data should be shared. If there is an issue, social, health, or violent, the data should therefore be used to correct the situation.

Further, in terms of mass shootings, the issue of gun control was raised. Facebook undoubtedly has the ability to make, with an educated guess, who and where an individual with the potential for violence is. The left wing sees fit that the location data should be released, so the person in question can have their guns seized and submitted for mental health assistance. Whereas, the right-wing still failed to agree to Facebook releasing this data. Instead, they understood the basic rights of individuals and their ability to bear arms.

Location data is undoubtedly a controversial topic. Nobody wants their location stored constantly, but we are in too deep now to object. The key differences in political thought regarding this topic, is the following:

The left doesn't want their data taken or stored. But, they realize that it is. If it is being stored, there should be alternate uses for it other than targeting ads and selling it. The data is a social issue and has the potential to save the lives of individuals and if used properly, can actually better society in the long-run.

The right doesn't want their data taken either. But, just as the left-wing, they realize that it is. In this, they respect the private company and have no problem with their utilization. Its users sign up for the site, and therefore are allowing for the data collection to occur. Data is an economic issue and has the potential to greatly improve the United States economy and increase market value. The data, according to the law, belongs to Facebook, and they respect the ownership as it is.

This is important because it is a new and unique contribution to the study of politics and the impact of a relatively new part of social life. This research may open the doors for new legislation to come forth which will hopefully bind the country in agreement on the issue of privacy. I believe that a key issue when it comes to politics is the two-party system. There are two drastically different belief systems in place, but the research I have conducted may allow people to read how the other side reacts, and to understand their thinking. It is important to note that nobody is wrong with their assessment of Facebook's data collection procedures, but the political system must come together, share their ideas, and come to a reasonable and collaborative solution.

CONCLUSION: FOR A DIFFERENT FUTURE

It has been made clear that identifying, biometric, and location data are all being collected and, currently, there is no quick fix to stop that from happening. Further, I do not believe that the data collection will ever cease to occur. However, I propose that there are solutions, which may help alleviate the problem., and spark social change. These ideas stem both from my readings and interviews. I am not an expert, but these issues are dire and require intervention on behalf of democracy.

I end by offering several solutions to the issue of data protection. They each fall within vastly different categories: educational and technological, but each possibly contribute unique and modern solutions to this study.

Privacy Policy

As republican senator John Kennedy bluntly told Mark Zuckerberg in the Senate hearing on April 10, 2018: “Your user agreement sucks . . . the purpose of your user agreement is to cover Facebook’s rear end. It’s not to inform your users about their rights.”¹⁵² This is an obvious first step that Facebook can take to increase transparency between the corporation and its users. Kennedy informed Zuckerberg that the Policy should be rewritten in plain English so average American people can understand it. Further, Kennedy questioned Zuckerberg on his willingness to expand right to delete data, expand knowledge of where data goes, and expand right to control who gets your data. To all of these, Zuckerberg was open and seemingly interested in doing what was right for the users of his social network.

¹⁵² *Committee on Energy and Commerce, Full Committee, hearing entitled “Facebook: Transparency and Use of Consumer Data.”*

Making the Privacy Policy more accessible for Facebook's users would be a clear first step in the right direction and may help in putting off the regulation of Facebook.

Education System

Secondly, education is key. This project has demonstrated that there is an extreme lack of understanding data, specifically when it comes to Facebook, but such a lack of understanding about data is widespread. If people knew the extent of the issues, then they may act in a more responsible manner online. Given this, I believe one solution may be to provide an education system based around these ideas.

First and foremost, I do not believe that people should not be allowed to use Facebook without prior education, but I believe it would be very beneficial to all users. People should maintain their individual rights to use the websites they want to use. It's common to hear students in college complaining about how they do not know how to file their taxes, pay bills, or other common life-skills. Similar to these, as social media becomes more and more popular, there should be educational systems put in place to teach people about the online world, and the dangers present. Teaching individuals as early as possible is better, and this curriculum should be offered in middle schools and high schools. This will allow students to maintain a basis of knowledge before they start to use social media, or right when they are beginning. People will be aware of data collection and clicking through Terms and Conditions will no longer be a worry.

Technology System

The second possibility is in response to the current technology which governs our data. The current situation is a churning algorithm geared to collect the most amount of data and process it into quantifiable information, and in the end, make money. In this technological solution, I connect most with an article published by computer scientist Marc Langheinrich, "A Privacy Awareness System for Ubiquitous Computing Environments," in which he presents a technological solution to the issue of extreme and exploitative data collection by providing a more concrete rational system. Langheinrich provides a common-sense approach which relies mostly on user accountability.¹⁵³ He discusses, realistically, the concept of anonymization technology, which is used to prevent observers from discovering the source of online communications.¹⁵⁴

Langheinrich goes on to stress that while this type of technology, as well as encryption, can make tracing identities nearly impossible, we need to worry about the societal implications this may lead to. He writes, "Unless we want to abandon our current social interactions completely and deal only behind digital pseudonyms in virtual reality with each other, we must realize that our real-world presence cannot be completely hidden, nor perfectly anonymized."¹⁵⁵ This system just isn't realistic; as our society turns towards social media more and more and the line between virtual and reality draws closer, anonymization technologies are dangerous. It would lead to a loss of personal identity in the real world.

¹⁵³ Langheinrich, 238.

¹⁵⁴ Neal Leavitt, "Anonymization Technology Takes a High Profile," *Computer* 42, no. 11 (2009): 15.

¹⁵⁵ Langheinrich, 237.

In this, Langheinrich presents a privacy awareness system (pawS) which strikes a reasonable balancer between anonymization and encryption technologies.¹⁵⁶ The system, pawS, provides collection and processing tools which allow data collectors (Facebook) to effectively communicate their collection and procession details to us, and help them keep their promises. Held on accountability, this system blends the technological side of computing with society. It is a balance which few systems currently in place allow for. If implemented correctly, this system may be able to change the face of data collection in a very positive manner. However, on a side-note, I believe that Facebook should only change their process to this type if they want to, the corporation should not be forced.

The main point of this thesis was to prove the political views on the topic of Facebook's data collection which are widely varied, depending on party affiliation. These different solutions offered are examples of bipartisan first steps towards a main goal that perhaps everyone can agree to: data protection. If implemented correctly, they may not only have the ability to help people learn of the dangers, but also support the technological side of the process. Perhaps Facebook can implement these to create a safer user network and avoid government regulation.

People need to learn what is actually happening. This issue is not politically driven, but its alteration and future are. This project has made me realize that there is such little knowledge on this issue. Specifically, at Trinity College, people are more informed than the majority of individuals in the Millennial population. This is a further worry, and an education system must be implemented across the country. Thus, Millennial social networks users will

¹⁵⁶ Ibid, 238.

become the first of many generations who have been trained in the context of a digital world. These more informed citizens will soon be making impactful decisions in this country. Thus, the educational system proposed will have a trickle-down or, even better, exponential nature; over the next several years it will virtually change the face of political opinion on Facebook data collection.

Appendix**Participants:**

Name	Age	Political Stance	Interview #1 Date	Interview #2 Date
Alex	22	Democrat	February 22, 2018	March 8, 2018
Benjamin	22	Democrat	February 21, 2018	March 20, 2018
Brandon	22	Democrat	February 21, 2018	-----
Charlie	22	Republican	February 23, 2018	March 8, 2018
Samantha	21	Republican	February 22, 2018	March 8, 2018
Stephen	24	Republican	February 25, 2018	-----
Vickie	21	Democrat	February 25, 2018	March 3, 2018
William	20	Democrat	February 21, 2018	March 5, 2018

Bibliography:

"About," *DataEthics*, <https://dataethics.eu/en/about/>.

Ackerman, Mark S., Lorrie Faith Cranor and Joseph Reagle. "Privacy in e-commerce: Examining User Scenarios and Privacy Preferences." *EC* (1999): 1-8.

"Advances in AI are used to spot signs of sexuality." *The Economist*. September 9, 2017. <https://www.economist.com/news/science-and-technology/21728614-machines-read-faces-are-coming-advances-ai-are-used-spot-signs>.

Angwin, Julia. "Dozens of Companies Are Using Facebook to Exclude Older Workers From Job Ads." *ProPublica* Accessed January 10, 2018, <https://www.propublica.org/article/facebook-ads-age-discrimination-targeting>.

Barkhuus, Louise, and Anind K. Dey. "Location-Based Services for Mobile Telephony: A Study of Users' Privacy Concerns." In *Interact*, vol. 3, pp. 702-712. 2003.

Berridge, Kent C. and Terry E. Robinson. "What is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Saliency?" *Brain Research Reviews* 28 (1998): 309-369.

boyd, danah and Kate Crawford. "Critical Questions for Big Data." *Information, Communication & Society* 15, no. 5 (2012): 662-679.

Brandom, Russell. "Facebook is starting to tell more users about facial recognition." *The Verge*, February 27, 2018. <https://www.theverge.com/2018/2/27/17058268/facebook-facial-recognition-notification-opt-out>.

Brunton, Finn and Helen Nissenbaum. "Political and Ethical Perspectives on Data Obfuscation." In *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology* (New York: Routledge, 2013): 164-188.

Bucher, Taina. "Want to Be on The Top? Algorithmic Power and the Threat of Invisibility on Facebook." *new media & society* 14, no. 7 (2012): 1164-1180.

Buolamwini, Joy, and Timnit Gebru. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." *Conference on Fairness, Accountability and Transparency* (2018): 1-116.

Candela, Joaquin Quiñero. "Managing Your Identity on Facebook with Face Recognition Technology." *Facebook Newsroom*. December 19, 2017. <https://newsroom.fb.com/news/2017/12/managing-your-identity-on-facebook-with-face-recognition-technology/>.

Cheney-Lippold, John. "A new algorithmic identity: Soft biopolitics and the modulation of control." *Theory, Culture & Society* 28, no. 6 (2011): 164-181.

Cho, Eunjoon, Seth A. Myers, and Jure Leskovec. "Friendship and Mobility: User Movement in Location-Based Social Networks." *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2011): 1082-1090.

Constine, Josh. "Facebook now has 2 billion monthly users ... and responsibility." TechCrunch. June 27, 2017. <https://techcrunch.com/2017/06/27/facebook-2-billion-users/>.

Cohen, Julie. "Examined Lives: Informational Privacy and the Subject as Object." *Stanford Law Review* 52 no 5 (2000).

Crump, Catherine. "Data Retention: Privacy Anonymity, and Accountability Online." *Stanford Law Review* 56 no 1 (October 2003): 191-229.

"Cookies & Other Storage Technologies." Facebook.
<https://www.facebook.com/policy/cookies/printable>.

"Data Policy," Facebook, September 29, 2016, https://www.facebook.com/full_data_use_policy.

Ellerbrok, Ariane. "Playful Biometrics: Controversial Technology through the Lens of Play." *The Sociological Quarterly* 52 no 4 (2011): 528-547.

"Facebook Reports Fourth Quarter and Full Year 2017 Results." January 31, 2018.
<https://investor.fb.com/investor-news/press-release-details/2018/Facebook-Reports-Fourth-Quarter-and-Full-Year-2017-Results/default.aspx>.

Fast, Karin, Henrick Örnebring, and Michael Karlsson. "Metaphors of Free Labor: A Typology of Unpaid Work in the Media Sector." *Media Culture & Society* 36 no. 7 (2016): 964-978.

Foer, Franklin. "How Facebook Tricks You into Trusting Algorithms." *Gizmodo*. Accessed September 22, 2017, <https://gizmodo.com/how-facebook-tricks-you-into-trusting-algorithms-1810792161>.

Foucault, Michel. *Security, Territory, Population: Lectures at the Collège de France, 1977-78*. (New York: Springer, 2007).

- Fox, Maggie. "Flint Water Crisis: Feds Expand Programs to Help Kids Affected by Lead." *NBC News*. March 2, 2016. <https://www.nbcnews.com/storyline/flint-water-crisis/flint-water-crisis-feds-expand-programs-help-kids-affected-lead-n530556>.
- Gangadharan, Seeta P. "The Downside of Digital Inclusion: Expectations and Experiences of Privacy and Surveillance among Marginal Internet Users." *New Media & Society* (November 9, 2016): 597-615.
- Graham, Stephen D. N. "Software-Sorted Geographies." In *The People, Place and Space Reader*, eds. Gieseeking, Mangold, Katz, Low, Saegert (New York: Routledge, 2005): 133-138.
- Gray, Mary L. and Siddharth Suri. "The Humans Working Behind the AI Curtain." *Harvard Business Review* (January 9, 2017): <https://hbr.org/2017/01/the-humans-working-behind-the-ai-curtain>.
- Greene, Daniel. "Discovering the Divide: Technology and Poverty in the New Economy." *International Journal of Communication*, 10 (2016): 1212-1231.
- Hill, Kashmir. "Facebook is using your phone's location to suggest new friends—which could be a privacy disaster." *The Splinter*, June 28, 2016, <https://splinternews.com/facebook-is-using-your-phones-location-to-suggest-new-f-1793857843>.
- "Immaterial Labor and Data Harvesting." *Share Foundation*. August 21, 2016. <https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/>.
- Junglas, Iris A., and Richard T. Watson. "Location-Based Services." *Communications of the ACM* 51, no. 3 (2008): 65-69.
- Kasper, Debbie V.S. "Privacy as a Social Good." *Social Thought and Research* 28 (2007): 165-189.
- Korolova, Aleksandra. "Privacy Violations Using Microtargeted Ads: A Case Study." *Data Mining Workshops* (2010): 474-482.
- Kosinski, Michal and Yilun Wang. "Deep neural networks are more accurate than humans at detecting sexual orientation from facial images." *Journal of Personality and Social Psychology*, 246-257.
- Lambert, Troy. "GIS and Artificial Intelligence Used to Build Facebook's World Population Map." *GIS Lounge*. March 1, 2016. <https://www.gislounge.com/gis-artificial-intelligence-used-build-facebooks-world-population-map/>.
- Langheinrich, M. "A Privacy Awareness System for Ubiquitous Computing Environments." *Proceedings of UbiComp* (2002): 237-245.

Leavitt, Neal. "Anonymization Technology Takes a High Profile." *Computer* 42, no. 11 (2009): 15-18.

Liao, Shannon. "Google admits it tracked user location data even when the setting was turned off." *The Verge*. November 21, 2017. <https://www.theverge.com/2017/11/21/16684818/google-location-tracking-cell-tower-data-android-os-firebase-privacy>.

Ljungstrand, P. "Context awareness and mobile phones." *Personal & Ubiquitous Computing* 5, no. 1 (2001): 58-61.

Lynch, Jennifer. "FBI Plans to Have 52 Million Photos in its NGI Face Recognition Database by Next Year." *Electronic Frontier Foundation*. April 14, 2014. <https://www.eff.org/deeplinks/2014/04/fbi-plans-have-52-million-photos-its-ngi-face-recognition-database-next-year>.

Kleinig, John and Peter Mameli, Seumas Miller, Douglas Salane and Adina Schwartz. "Surveillance Technologies and Economies." In *Security and Privacy* (Canberra: ANU Press 2011).

Madrigal, Alexis. "Reading the Privacy Policies You Encounter in a Year Would Take 76 Work Days." March 1, 2012 *The Atlantic*. <https://www.theatlantic.com/technology/archive/2012/03/reading-the-privacy-policies-you-encounter-in-a-year-would-take-76-work-days/253851/>.

Marchionini, Gary. "Educating Responsible Citizens in the Information Society." *Educational Technology* 39, no 2 (1999): 17-26.

Marx, Karl. *Capital a Critique of Political Economy. Vol. I, Book One, The Process of Production of Capital* (London: Electric Book Co., 2001).

Masi I., Trần A.T., Hassner T., Leksut J.T., Medioni G. "Do We Really Need to Collect Millions of Faces for Effective Face Recognition?" *Computer Vision – ECCV 2016*. September 16, 2016, 1-18.

Masse, Bryson. "What's the Worst That Could Happen with Huge Databases of Facial Biometric Data?" *Gizmodo*. September 11, 2017. <http://gizmodo.com/what-s-the-worst-that-could-happen-with-huge-databases-1802696698>.

McDonald, Aleecia M., and Lorrie F. Cranor. "The Cost of Reading Privacy Policies." *A Journal of Law and Policy for the Information Society* 4, no 3 (2008): 1-22.

McGlotten, Shaka. "Black Data." In *No Tea, No Shade* ed. Partick Johnson (Durham: Duke University Press, 2016): 262-286.

- Miller, Frederick A., Judith H. Katz, and Roger Gans. "The OD Imperative to Add Inclusion to the Algorithms of Artificial Intelligence." *OD Practitioner* 50, no. 1 (2018): 6-12.
- Meyer, Robinson. "Could Facebook Have Caught Its 'Jew Hater' Ad Targeting?" *The Atlantic*. September 15, 2017. https://www.theatlantic.com/technology/archive/2017/09/on-facebookadvertisers-can-show-their-ads-only-to-jew-haters/539964/?utm_source=nl-atlantic-daily-092017&silverid=MzEwMTkwMTM1NzAxS0.
- Morey, Timothy, Theodore Forbath, and Allison Schoop. "Customer Data: Designing for Transparency and Trust," *Harvard Business Review*, (May 2015): 1-11.
- Olteanu, M., K. Huguenin, R. Shokri, M. Humbert and J. P. Hubaux. "Quantifying Interdependent Privacy Risks with Location Data." *IEEE Transactions on Mobile Computing* 16, no. 3 (March 1, 2017): 829-842.
- Pariser, Eli. *The Filter Bubble: How the Web Is Changing What We Read & How We Think*, (New York: Penguin, 2012).
- Pinto, Nicolas, Zak Stone, Todd Zickler, and David Cox. "Scaling up Biologically-Inspired Computer Vision: A Case Study in Unconstrained Face Recognition on Facebook." *Computer Vision and Pattern Recognition* (2011): 35-42.
- Prabhakar, Salil, Sharath Pankanti, and Anil K. Jain. "Biometric Recognition: Security and Privacy Concerns." *IEEE security & privacy* 99, no. 2 (2003): 33-42.
- "Privacy Policy." August 9, 2005. <https://web.archive.org/web/20050809235134/http://www.facebook.com:80/policy.php>.
- Ratti, Carlo, Dennis Frenchman, Riccardo Maria Pulselli, and Sarah Williams. "Mobile landscapes: using location data from cell phones for urban analysis." *Environment and Planning B: Planning and Design* 33, no. 5 (2006): 727-748.
- Roosendaal, Arnold. "Facebook Tracks and Traces Everyone: Like This!" *Tilburg Law School Legal Studies Research Paper Series* no. 3 (2011): 1-10.
- Samway Michael A. and Warren Ryan. "The Internet, Human Rights, and the Private Sector." *Georgetown Journal of International Affairs* 15, no 1 (Winter/Spring 2014): 25-32.
- Scherker, Amanda. "Didn't Read Facebook's Fine Print? Here's Exactly What It Says." *Huffington Post*, July 23, 2014, http://www.huffingtonpost.com/2014/07/21/facebook-terms-condition_n_5551965.html.

- Sherman, Rob. "Hard Questions: Should I Be Afraid of Face Recognition Technology?" *Facebook Newsroom*. December 19, 2017. <https://newsroom.fb.com/news/2017/12/hard-questions-should-i-be-afraid-of-face-recognition-technology/>.
- Song, Indeok, Robert Larose, Matthew S. Eastin, and Carolyn A. Lin. "Internet Gratifications and Internet Addiction: On the Uses and Abuses of New Media." *Cyberpsychology & Behavior* 7, no. 4 (2004): 384-394.
- Stefanidis, Anthony, Andrew Crooks, and Jacek Radzikowski. "Harvesting Ambient Geospatial Information from Social Media Feeds." *GeoJournal* 78, no. 2 (2013): 319-338.
- Swanlund, David and Nadine Schuurman. "Mechanism Matters: Data Production for Geosurveillance." *Annals of the American Association of Geographers* 106 no. 5 (2016): 1063-1078.
- Taigman, Yaniv, Ming Yang, Marc' Aurelio Ranzato, and Lior Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification." *Facebook AI Research*. (June 24, 2014): 1-8.
- Tello, Lucía. "Intimacy and «Extimacy» in Social Networks. Ethical Boundaries of Facebook." *Comunicar* 21, no. 41 (2013): 206-213.
- Terms and Conditions May Apply*. Documentary. Cullen Hoback. 2013. Hyrax Films.
- Tiecke, Tobias. "Open population datasets and open challenges." *Facebook*, November 15, 2016. <https://code.facebook.com/posts/596471193873876/open-population-datasets-and-open-challenges/>.
- "The Massive Data Collection by Facebook – Visualized." *DataEthics*. June 26, 2017. <https://dataethics.eu/en/facebooks-data-collection-sharelab/>.
- Trottier, Daniel. *Social Media as Surveillance: Rethinking Visibility in a Converging World* (London: Routledge, 2012).
- U.S. Constitution. Amendment I.
- U.S. Constitution. Amendment IV.
- Wilson, Robert E., Samuel D. Gosling, and Lindsay T. Graham. "A Review of Facebook Research in the Social Sciences." *Perspectives on psychological science* 7, no. 3 (2012): 203-220.

- “What machines can tell from your face.” *The Economist*. September 9, 2017.
https://www.economist.com/news/leaders/21728617-life-age-facial-recognition-what-machines-can-tell-your-face?utm_source=newsletter&utm_medium=email&utm_campaign=sendto_newsletter&stream=top-stories.
- Woo, Jisuk. “The Right not to be identified: privacy and anonymity in the interactive media Environment.” *New Media & Society* 8, no 6 (December 1, 2006): 949-967.
- Yang, Ming. “Face Recognition Systems.” *Slideshare Lecture*. July 17, 2014.
<https://www.slideshare.net/milkers/lecture-10-ming-yang-face-recognition-systems>.
- Zang, Hui, and Jean Bolot. "Anonymization of location data does not work: A large-scale measurement study." In *Proceedings of the 17th annual international conference on Mobile computing and networking*, (2011): 145-156.
- Zuckerberg, Mark. “The Crunchies,” *Tech Crunch*, January 10, 2010.