The Dark Data Cycle: How the U.S. Government Has Gone Rogue in Trading Personal Data from an Unsuspecting Public

A thesis presented

by

By Melissa Carrie Oppenheim

to

The Department of History of Science
in partial fulfillment of the requirements
for an honors degree in History and Science

Harvard University
Cambridge, Massachusetts
March 2012

The Dark Data Cycle: How the U.S. Government Has Gone Rogue in Trading Personal Data from an Unsuspecting Public

By Melissa Oppenheim

ABSTRACT

While the historical and legal relationship in the U.S. between privacy expectations and privacy protections has remained a continuously evolving topic, this thesis demonstrates how this relationship has grown increasingly muddled since the mid-twentieth century. Contrary to existing law, this thesis argues that a new phenomenon, the "Data Cycle," has developed in the twenty-first century whereby governmental entities employ private companies as middlemen to buy and sell individuals' data. Politicians in particular are extremely interested in obtaining information about their constituents. This thesis contends that the Data Cycle has developed due to a shift in informational power between the government and private companies, permitting the government to indirectly acquire large amounts of personal information from individuals online. Although Americans remain concerned about their waning privacy protections, individuals' increasing addiction to the very technologies that spur the Cycle has pushed privacy to a crossroads in 2012.

KEYWORDS

Privacy, Technology, Internet, Databases, Data Economy, Personal Information, Data.Gov, Voter Registration, the Privacy Act of 1974, Science and Society

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Initially, I received a grant from Harvard's Institute of Politics to conduct a summer research project on privacy and the Internet. Now a History of Science Thesis, this work would not have been possible without the constant guidance of my wonderfully supportive concentration adviser Leandra Swanner and my dedicated, invaluable mentor Jim Waldo. I cannot overstate my sincerest gratitude to each of these individuals, both of whom transcended well beyond their respective roles of adviser and mentor.

I would additionally like to thank the following individuals who agreed to provide interviews for this thesis: Clay Johnson, Laura Quinn, Jim St. George, and Paul Zak. This work would not have been the same without your knowledgeable perspectives.

Further, I would like to recognize and thank the following individuals from the greater technology, academic, and legal community who spoke with me about my thesis on a more informal basis: Allan Friedman, Chris Hoofnagle, Clay Johnson, Vivek Kundra, Nicco Mele, Roy Oppenheim, Latanya Sweeney, and Jonathan Zittrain.

For their continued support, my utmost appreciation goes to my family. You have always been my number one fans. I would also like to thank my thesis buddies and close friends for their constant cheer and encouragement. I am especially grateful to Ellen Pilelsky, Charles Lano, Esther Hsiang, Carly Dickson, and Samantha Reiser for your sharp eyes and selfless time and energy spent reviewing my work.

Importantly, I would like to thank Amy Howell and Harvard's Institute of Politics for providing me with the Summer Thesis Research Grant. I am, lastly, thankful for the History of Science Departmental faculty, in particular, Alice Belser and Christopher Phillips, in addition to the Harvard depository system, the Eliot House Dining Hall staff (Domus!), and the Starbucks in the Garage for our early mornings and late nights together.

DEDICATION

To my parents –
Who have only ever dreamt the world for me,
You inspire me everyday.
Thank you for everything.

LIST OF FIGURES

INTRODUCTION

> Computers are here to stay. So are large organizations and the need for data. So is the American commitment to civil liberty… Our task is to see that appropriate safeguards for the individual's rights to privacy, confidentiality, and due process are embedded in every major record system in the nation, particularly the computerizing systems that promise to be the setting for most important organization uses of information affecting individuals in the coming decades.[1]
>
> – ALAN WESTIN AND MICHAEL BAKER

Similar to having to pay a toll on a highway, in order to participate in some of the most basic democratic activities, Americans[2] must pay a toll to the government[3] in the form of personal information. In order to vote or acquire a driver's license, individuals must provide the necessary documents and identification at the government's behest to receive permission to participate in specific activities or receive certain services or benefits.

These types of information tolls are not new. Americans have complied with paying such tolls for over two hundred and thirty-five years. Following the Constitutional Convention of 1787, the United States government developed record-keeping practices to facilitate popular elections and decennial censuses, as mandated by the U.S. Constitution.[4] The popular census of 1790 marked the first time in which the government attempted to conduct a "systematic and universal gathering of information about its citizens."[5] Since this humble beginning, the process of exchanging information to receive a service or benefit has become second nature.

---

[1] Alan Westin and Michael Baker, *Databanks in a Free Society: Computers, Record-Keeping and Privacy* (New York: Quadrangle Books, 1972), 25.

[2] Throughout this thesis I refer to "American individuals" to include all individuals living in the United States.

[3] Throughout this thesis I refer to "government" to include local, state, and/or federal government, unless otherwise specified. Due to space limitations, this paper will not necessarily make a distinction between such entities. Differentiating between these governmental actors is not germane to the relevance of this thesis. This is particularly true in the area of voting, the focus of my case study, where the federal government has created an overlay of how states must collect and maintain their data.

[4] James Waldo, Herbert S. Lin, Lynette I. Millet, eds., *Engaging Privacy and Information Technology in a Digital Age* (Washington, DC: The National Academies Press, 2007), 354, accessed May 20, 2011, http://books.nap.edu/openbook.php?record_id=11896&page=R1.

[5] The 1790 Census Survey only sought to record information from free males above the age of sixteen. For more detail, see Waldo, Lin, and Millet, *Engaging Privacy,* 354.

Currently, individuals may not realize that when they provide required information to the government, they pay a much bigger toll than has historically been the case. While the U.S. government has traditionally possessed the most information concerning its population, the government is quickly being eclipsed by corporations that profit from collecting and selling troves of data in a rapidly expanding "data economy."[6] In the data economy, individuals' personal information has accrued a "currency of value," provoking the government and private companies to routinely trade different pieces of individuals' information.

In this thesis, I present a new pattern of personal information sharing, the "Data Cycle" (or "Cycle"), that has emerged between the public and private sectors in the twenty-first century. The Data Cycle reveals an increasingly commonplace relationship between individuals, the federal government, and private companies, whereby the federal government and private companies regularly buy and sell robust amounts of Americans' data gathered from an unsuspecting public. While I argue that individuals fall victim to the Cycle's occurrence, I contend that Americans' twenty-first century online sharing habits and dependency on communication technologies is actually a crucial ingredient to the Cycle's formation and increasing strength. As individuals continue to share unprecedented amounts of personal information online, a data economy has developed where private parties glean more information about those individuals than ever before and subsequently, provide such information to the government. I argue that

---

[6] Brad Lockwood, *Domestic Spying and Wiretapping* (New York: The Rosen Publishing Group, 2007), 16, accessed April 24, 2011, http://books.google.com/books?id=WoiNocUQRNkC&pg=PA16&dq=domestic+spying+to+new+heights&hl=en&ei=eZavTcKmKYP30gGS9oCgCQ&sa=X&oi=book_result&ct=result&resnum=1&ved=0CD4Q6AEwAA#v=onepage&q=domestic%20spying%20to%20new%20heights&f=false; Erick Schonfeld, "The Privacy Dilemma," *TechCrunch*, January 28, 2009, accessed November 18, 2010, http://techcrunch.com/2009/01/28/the-privacy-dilemma/.

although Americans' informational privacy is clearly waning in light of the Data Cycle, individuals may surprisingly not defend their perceived rights to information control, as scientific studies have evidenced that humans are growing neurologically addicted to the very activities that fuel the Cycle's existence.

This thesis draws attention to the symbiotic relationship that has emerged, whereby individuals are dependent on the same technology and services that have spawned and support the Data Cycle. The U.S. government continues to seek to attain personalized data for a multitude of legitimate reasons including, but not limited to, national security, law enforcement, disaster response and relief, redistricting, voting, tax collection and filing returns. But, increasingly governmental entities are turning to private companies that specialize in data collection and aggregation to obtain robust amounts of individuals' personal information. Private companies possess resources, trade secret protections, and regulatory freedom that the government lacks.

Ultimately, the laws intended to protect Americans from the government's abuses of information technology have either grown obsolete through technical advancements or have been ignored due to strong political and economic interests. Specifically, the government is not permitted to arbitrarily collect more data than it needs and/or append its own databases with privately collected information due to the Privacy Act of 1974 ("the Act"), the only major U.S. federal law that protects an individual's direct right to informational privacy. The Act restricts the collection and use of an individual's data by federal agencies and prohibits the federal government from amassing personal information unless the agency has a proper purpose for doing so.[7] Importantly, all federal

[7] As U.C. Berkeley Law School Professor Chris Hoofnagle details in his 2004 report "Big Brother's Little Helpers," the government continues to hire private data brokers in order to access stockpiles of personal

agencies, with the exclusion of law enforcement agencies that collect Americans' personal information must only use the collected records in "routine use," or in other words "with respect to… the use of such record for a purpose which is compatible with the purpose for which it was collected."[8] Pursuant to the Act, individuals have protections regarding the Executive branch and federal agencies' ability to collect, maintain, and use their personal information. Unlike corporations, start-up companies, and other online and offline services with which we share our personal information on a regular basis, the U.S. government supposedly is bound by federal law to handle our information in restricted ways.[9] However, this thesis demonstrates that through the Data Cycle, the U.S. government circumvents these limitations by hiring private sector companies as middlemen to collect and share data that the government is legally restricted from obtaining itself.[10]

By examining how specific technological advances, such as increases in computing power, processing speeds, and connectivity, have functioned in a changing sociopolitical, legal, and economic landscape since the mid-twentieth century, we learn how and why the Cycle has developed and how we are all – tacitly or consciously – contributing to its reality. This thesis substantiates and documents that while individuals are victims of the

---

information that would otherwise be legally inaccessible. Chris Hoofnagle, "Big Brother's Little Helpers: How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement," *The Electronic Privacy Information Center* (Summer 2004), accessed August 20, 2011, http://epic.org/privacy/choicepoint/cp_article.pdf; Waldo, Lin, and Millet, *Engaging Privacy,* 137.

[8] "Definitions: Routine Use," in the Privacy Act 1974 (5 U.S.C. § 552a), Electronic Privacy Information Center, last modified January 3, 2005, acessed October 15, 2011, http://epic.org/privacy/laws/privacy_act.html.

[9] Pursuant to the "routine use" clause, although the government is not directly repurposing (or adapting or giving a new purpose or use to) the data itself, it is allowing for others to do so – and often, on an open-scale. While the government is ostensibly following its own rules, it is promoting exactly what the Act sought to prevent. *Merriam-Webster Dictionary and Thesaurus Online*, s.v. "repurpose," accessed December 8, 2011, http://www.merriam-webster.com/dictionary/repurpose.

[10] Hoofnagle's report is comprised of successfully acquired Freedom of Information documents from over nine federal agencies. Hoofnagle, "Big Brother's Little Helpers," 16.

Cycle, we are the ones empowering its operation. Both the information that we provide as citizens to the government as a requisite to engage in activities *and* the information that we pay as consumers to private online services have contributed to the Cycle's growth. The government routinely ignores the law in order to access more data on its constituents by employing private entities to acquire detailed personal information.

In order to clearly unfold this argument throughout the thesis, my strategy is to identify the Data Cycle, exemplify its occurrence, and then investigate how and why the Cycle is growing more powerful and commonplace in the twenty-first century.

To this end, Chapter One presents and explains the Data Cycle, an original mapping of how the federal government is sidestepping the Privacy Act by transacting with private sector entities. Chapter One focuses on how the government and private companies take advantage of an unsuspecting and/or apathetic public to produce rich datasets of personal information. In particular, private entities frequently enter into business relationships with the government, allowing for the combining of privately harvested personal information with the government's publicly collected personal data. The chapter illustrates how the Cycle works in practice through exploring a case study in the area of voter registration data.[11] Prior to delving into how the Cycle has affected this process, I analyze the government's voter registration data collection practices prior to the advent of computing technology in order to demonstrate how changes in technology, Americans' sharing habits, and the law have dramatically affected the government's participation in the Data Cycle in 2012. Through this case study, I expose how the Data Cycle allowed Members of Congress ("MCs") in 2010 to purchase

[11] I strategically chose to investigate how voter registration data was being collected, compiled, stored, and sold to various entities, due to the raw data's availability and wealth of public records.

enhanced information about their constituents using tax dollars. By analyzing government documents and reports and congressional members' Member Representational Allowances, this chapter demonstrates how private companies and governmental entities engage in trade practices in the Data Cycle that are almost entirely veiled and unregulated.

Chapter Two addresses why the government has entered the Data Cycle and how its active participation in the Cycle, which blatantly evades the Privacy Act of 1974, has rendered the government's actions rogue. This chapter argues that the Data Cycle's emergence illuminates a shift in informational power between the government and private sector in the twenty-first century. By discussing how private companies – and eventually the public – began to widely adopt government-funded technologies, such as the computer and the Internet, this chapter presents how private companies were able to eventually produce a digital information market based on freely exploiting Americans' new online sharing habits. As companies began to grow increasingly powerful and knowledgeable, a realized shift in informational power has encouraged the government to hire private companies in order to indirectly access typically forbidden information. Although the Privacy Act prevents the government from repurposing Americans' collected information, the government has begun hiring whomever and whenever in order to hold onto a notion of informational superiority.

Chapter Three contends that privacy in America in the twenty-first century is alive but currently at a crossroads. While individuals claim they are interested in maintaining personal privacy in the Information Age,[12] their actions, which may be

[12] The "Information Age" (or "Computer Age" or "Information Era") refers to a period of time starting in the last twenty-five years of the twentieth century as computers and computer networks began making

driven by developing neurological addictions and misunderstandings of how private companies sell their information, speak differently. In this chapter, I contrast the public's concern for privacy in the mid-twentieth century with those currently held in the face of the Data Cycle. Ultimately, I proffer that the Data Cycle would ordinarily just be another attempt by political powers to obtain individuals' personal information. However, the Data Cycle is different, as political interests in the twenty-first century are able to achieve their goal due to individuals' neurological addiction and social misunderstanding. Both Chapters Two and Three employ contemporary historical sources from the mid-twentieth century to early twenty-first century, government reports, scientific studies, and official legal documents to support these arguments.

The Conclusion of this thesis hypothesizes how the continuation of Data Cycle will pose serious, counterproductive implications for society. By examining the consequences of the unanticipated adoption of Social Security Numbers as unique identifiers in the 1960s as a prior example, I offer that the government may similarly adopt the Data Cycle without weighing its long-term effects. Additionally, the Conclusion reinterprets the age-old theory of technological determinism or the idea that technology is the sole driver of culture and history and discusses its relevance to the Data Cycle and humans' addiction to instant communication and access to information. This thesis concludes that although the Data Cycle extracts information from a largely unsuspecting public, individuals have become so dependent on the technology that supplies data to the Data Cycle, they may not even find such practices objectionable.

Throughout I contemplate how the contemporary government initiative Data.gov,

---

access to information easily accessible. David Holtzman, *Privacy Lost: How Technology Is Endangering Your Privacy* (San Francisco: Jossey-Bass, 2006), 4, 151.

a website launched in May 2009 to provide the public with raw government data,[13] has perhaps indirectly incentivized the largest repurposement of individuals' data in history, given that the site's publicly funded collection of datasets are available for anyone at any time for any purpose to freely download. Additionally, this thesis exposes how there are no clear rules governing how information may be shared between private and public entities. As Harvard Law School Professor John Palfrey contends, "We are racing ahead quickly with the development of new technologies while the institutions – legal and otherwise – designed to protect user privacy have lagged behind."[14,15] While the government is ostensibly following the technical language of the Act, it is promoting, encouraging, and participating in activities contrary to the Act's intent.

Interestingly, author and privacy expert Daniel Solove's posited in a 2002 law review article, "[T]he government routinely pour[s] [personal] information into the public domain…" through posting information and/or entire databases on the Internet and/or providing it to companies, organizations, and/or individuals who request it.[16] Solove tells his readers, "Imagine… that this information [that the government compels of its residents] would be traded among hundreds of private-sector companies that would

---

[13] Heretofore "raw data" will refer to data that has not yet been analyzed nor processed.
[14] John Palfrey, "The Public and the Private at the United States Border with Cyberspace," *Mississippi Law Journal* 78 (2008): 243, accessed October 24,2010, http://nrs.harvard.edu/urn-3:HUL.InstRepos:2554026.
[15] For instance, in April 2011, security researchers discovered that Apple's iPhone tracks and stores one's location(s) on a hidden file both on the phone itself and copied onto the owner's computer when the two are synchronized. Further, in March 2011, *The New York Times* reported how cell phone companies in general do not typically disclose how much information they collect, as they track a customer's whereabouts as part of their business to ostensibly route calls to the closest cell phone towers. In fact, in the U.S. the telecommunications industry does not have to report the information they collect since such data collection is supported by law enforcement and national security efforts. Charles Arthur, "iPhone Keeps Record of Everywhere You Go," *The Guardian*, April 20, 2011, accessed January 1, 2012, http://www.guardian.co.uk/technology/2011/apr/20/iphone-tracking-prompts-privacy-fears; Noam Cohen, "It's Tracking Your Every Move and You May Not Even Know," *The New York Times,* March 26, 2011, accessed December 14, 2011, http://www.nytimes.com/2011/03/26/business/media/26privacy.html.
[16] Daniel Solove, "Access and Aggregation: Public Records, Privacy, and the Constitution," *Minnesota Law Review* 86 (2002): 1138-1139.

combine it with a host of other information such as one's hobbies, purchases, magazines, organizations, credit history, and so on."[17] Solove conjectures that after public entities share Americans' personal information with private companies, his readers should imagine what would happen if governmental bodies could then purchase back the databases in their supplemented form. Abruptly, Solove commands his readers, "Stop imagining."[18] In 2002, Solove tells his readers to wake up, as the interactions he is describing are actually beginning to occur throughout the U.S.[19]

In 2012, this phenomenon is not only commonplace, but also gaining momentum exponentially, as an unaware public continues to embrace technological advancements that outpace legal constraints. This thesis presents that through the Data Cycle this phenomenon is occurring and regardless of our awareness, we are all unwittingly participating subjects in this experiment.

[17] Solove, "Access and Aggregation," 1138-1139.
[18] Ibid.
[19] Ibid.

CHAPTER ONE

The Emergence of a New Data Cycle

> It is the custom of Americans to believe that no 'lady-or-the-tiger' choice has to be made between science and liberty. For 200 years, in the tradition of Franklin and Jefferson, we have hammered out legal rules that allowed each successive wave of invention to realize its potential, but also required each to be brought under the rule of law. Sometimes it took a while for the principles of regulation to become clear, and we have come to realize that the awesome effects of contemporary technology give us less lead time for social learning and regulatory response than we had in earlier eras. But that is the challenge we face.[20]
>
> – ALAN WESTIN

In 2009, a team of Princeton University professors led by Edward Felten, Director of Princeton's Center for Information Technology Policy and a frequent Congressional witness, published the white paper "Government Data and the Invisible Hand." In this paper, the researchers advocate that the federal government should provide public, raw data "that is easy for others to re-use."[21] The Felten paper ultimately argues that in order for public data to meaningfully benefit the public, "the federal government must reimagine its role as an information provider" and widely share its publicly collected data with private parties so that it may be made useful.[22]

Later that same year, President Obama and then Chief Information Officer of the U.S. Vivek Kundra expressed similar sentiments to the ideas proposed in Felten's paper. Specifically, the Obama Administration designed a plan to create "an unprecedented level of openness in government" by moving government data to the World Wide Web

---

[20] Alan Westin, "A Policy Analysis of Citizen Rights: Issues in Health Data Systems," NBS Special Publication, U.S. Department of Commerce, and National Bureau of Standards (Washington, DC: 1977), 41.

[21] The paper calls for the U.S Government to provide raw data to the public and allow private sector companies that are "better suited" to present such information through effective interactive means to the public. David Robinson et al., "Government Data and the Invisible Hand," *Yale Journal of Law & Technology* 11 (Fall 2009): 170, accessed October 2, 2011, http://ssrn.com/abstract=1138083.

[22] Robinson et al., "Government Data and the Invisible Hand," 170.

("Web").[23] As part of the Administration's "Open Government Initiative,"[24] the Administration launched the website Data.gov, an ongoing project that seeks to encourage the release of as much government data to the Web as possible without compromising national security or individual privacy, in May 2009.[25] In less than one year of the website's launch, over 118,000 datasets had been made available online for free to anyone with Internet access.[26,27] Through this initiative, the government's data is available in "machine-readable format… ready to be processed, recombined, mashed up, and displayed visually."[28] Data.gov goes much further than prior related efforts by seeking to fundamentally alter the presumption that government data ought to remain secret. At its core, Data.gov seeks to make "federal sector data more accessible and usable."[29]

While government transparency advocates such as Clay Johnson, author of *The Information Diet,* view Data.gov as one small step in the right direction of open information provided by the federal government, Data.gov raises new questions, signaling an

---

[23] The World Wide Web is a system of interlinked hypertext documents accessed via the Internet (not to be confused with being the Internet). Using a Web browser, one may visit websites on the World Wide Web and navigate between them via hyperlinks. For more on the Web, see Chapter Two, page 55; Karim Lakhani, Robert Austin, and Yum Yi, "Data.gov," *Harvard Business School Case Study,* May 23, 2010, accessed September 8, 2011, http://www.data.gov/documents/hbs_datagov_case_study.pdf.

[24] Pursuant to a memorandum produced by the Obama administration, the Open Government Initiative intended to direct executive departments and agencies to take certain actions to implement the principles of transparency, participation, and collaboration.

[25] Data.gov offers over 200,000 internally vetted data sets to anyone for free online. "Data.gov," Data.gov, 2009, accessed December 8, 2011, http://www.data.gov/.

[26] In a phone conversation with government and information transparency advocate Clay Johnson, Clay noted that most of Data.gov's data actually comes from one huge dataset from the EPA's toxic release inventory; it is just spliced and diced by year and state to make the website seem more robust. Clay Johnson (Author, Technology and Transparency Advocate), phone interview by Melissa Oppenheim, November 4, 2011, 11am EST, transcript in *Appendix B,* "Oral History Interviews."

[27] Lakhani, "Data.Gov."

[28] Ibid.

[29] Ibid.

underlying change in modern privacy trends.[30] Exemplified by Data.gov, a recognizable pattern has emerged in the twenty-first century whereby the federal government has dramatically altered its data collection practices and treatment of governmental records, using and re-using publicly collected information in unprecedented ways.

This chapter identifies the Data Cycle, an increasingly common interaction that has emerged in the twenty-first century between individuals, the government, and private companies.[31] After explaining how the Data Cycle functions in the abstract, I focus on one line of investigation to concretely demonstrate the Cycle's existence. As an exemplar, I chose voter registration data, due to the abundance of publicly accessible records, as an area ripe for investigation. This chapter highlights how Americans must provide personal information to the government as a requisite to partake in certain activities. However, most individuals do not know that this information then enters the Data Cycle where it is shared between the government and private companies. While Americans must comply by providing such personal information to the government, this chapter and subsequent chapters additionally argue the Cycle is actually made most powerful through individuals' intensified, discretionary sharing of personal information online with private companies – many of which later enter business relationships with the government.

Moreover, this chapter provides a comparative analysis on how a single area of government data collection and maintenance existed before and after the digitization of voter registration records and the advent of the Internet. This comparison demonstrates that the government began sharing citizens' voter registration data with various third

[30] Clay Johnson also believes that although a delicate balance exists between privacy and transparency, as of November 2011 society protects privacy too much and is not nearly transparent enough. Johnson, interview.

[31] Latanya Sweeney (Professor of Computer Science and Privacy at Harvard University), in person conversation with Melissa Oppenheim, October 4, 2011, 4pm EST, Harvard University, Boston, MA.

parties even before the major technical advancements existed to aid in this practice. Thus, this chapter provides a foundational background by exploring how both technical computer advancements and non-technical socio-economic, political, and legal events have paved the way for how the government has grown to collect, share, and re-acquire citizens' data in the twenty-first century.

Lastly, throughout the chapter, I discuss the Data Cycle's relevance to the Privacy Act of 1974, the only major U.S. federal law that protects an individual's direct right to information privacy and information use from governmental agencies. The Data Cycle highlights the unregulated, antiquated nature of the twenty-first century's legal infrastructure.[32] By the end of Chapter One, the reader shall therefore understand the Data Cycle, how the American public remains generally unaware of its practice even though we are participants, and how technology within a historical context set the groundwork for this phenomenon to emerge.

The Data Cycle

Since the founding of the United States, the government has compelled its residents to provide various pieces of personal information in order to participate in certain activities or receive certain services or benefits. For example, in order to vote, or buy a home or car, Americans must provide the government with some required information. For over two hundred years, the government has been collecting individuals' information for such intentional purposes.

Public records, especially those in the form of punched cards and paper copies, were historically not collected with the intention of later being aggregated (*or* combined,

[32] Waldo, Lin, and Millet, *Engaging Privacy,* 137.

clustered, or gathered in a mass or whole)[33] and repurposed. Specifically, the Privacy Act, which limits the collection, maintenance, and use of an individual's data by federal agencies, also possesses a "routine use" clause, which explicitly restricts the federal government from repurposing information without the individual's consent.[34,35] The Act explains that collected information may only be used "for a purpose which is compatible with the purpose for which it was collected."[36] In fact, all federal agencies - with few exemptions such as law enforcement and intelligence activities – that collect Americans' personal information must only use the collected records with respect to the "routine use" clause.[37] However, due to an array of sociopolitical, economic, and technical advancements touched on in this chapter and unfolded in subsequent chapters, the U.S. government continues to enter new types of relationships with the private sector that encourage the indirect repurposing of information despite the intent of the Privacy Act.

In studying the changing relationships between the public and private sectors over the first decade of the twenty-first century, I offer the discovery of a new, recurring cycle: the Data Cycle. Specifically, the Data Cycle (*see Figure 1*) describes the way in which the government and third parties have grown to collect, handle, and augment government-

---

[33] *Merriam-Webster Dictionary and Thesaurus Online*, s.v. "aggregate," accessed December 8, 2011, http://www.merriam-webster.com/dictionary/aggregate.

[34] Waldo, Lin, and Millet, *Engaging Privacy*, 137.

[35] According to the Privacy Act of 1974, "Each agency that maintains a system of records shall— 1) maintain in its records only such information about an individual *as is relevant and necessary to accomplish a purpose* of the agency required to be accomplished by statute or by Executive order of the President; (2) collect information to the greatest extent practicable directly from the subject individual when the information may result in adverse determinations about an individual's rights, benefits, and privileges under Federal programs; (3) inform each individual whom it asks to supply information… (A) the authority… which authorizes the solicitation of the information and whether disclosure of such information is mandatory or voluntary; (B) the *principal purpose or purposes* for which the information is intended to be used; (C) *the routine uses* which may be made of the information, as published pursuant to paragraph (4)(D) of this subsection…" For full text, see Privacy Act of 1974, 5 U.S.C. § 552a (1974), The U.S. Justice Department, accessed November 2, 2011, http://www.justice.gov/opcl/privstat.htm.

[36] "Definitions: Routine Use," Privacy Act of 1974.

[37] Privacy Act of 1974, 5 U.S.C. § 552a (1974).

mandated data from Americans in the first decade of the twenty-first century. The cycle has the following three distinct nodes:

(1) The government mandates and collects certain pieces of information from Americans under the premise of using that information for a specific purpose.[38]

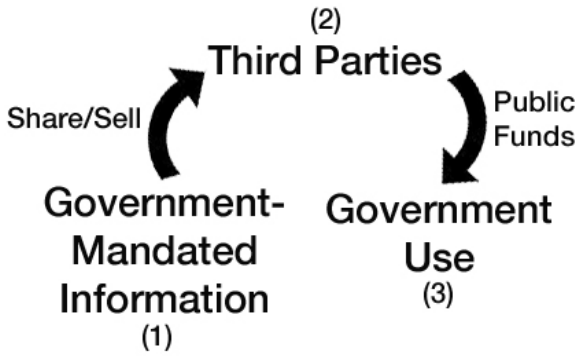

Figure 1: Melissa Oppenheim, "The Data Cycle," October 2011.

(2) This information is later given or sold to private third parties for other purposes, unlimited in scope. Many private companies enhance or augment the government's collected data with additional data points by associating the government's datasets with externally gathered personal information. As individuals increasingly share large amounts of personal information online, private companies capture this information and use it to enhance government datasets and turn a profit.[39]

(3) The information originally collected by the government from individuals, is later re-purchased or re-acquired in its enhanced version by the government using taxpayer dollars. The government agency that acquires this information may use the enhanced

[38] In some instances government-mandated data that is required of individuals in order to participate in a certain activity is collected by a third party rather than by the government itself. As detailed in *Appendix A,* such is the case in dealing with collecting Americans' pharmaceutical information.

[39] For example, as reported in the official initial public offering prospectus for Facebook to the U.S. SEC, more than 845 million people use Facebook (or more than one in every thirteen people on Earth). In the U.S. 155 million people use the site (~50% penetration). Mark Zuckerberg, the company's founder, has infamously created the "Zuckerberg Law" whereby he revealed that the company is observing that each year not just on Facebook.com but across the Web, individuals are sharing double than they did the year before. U.S. Securities and Exchange Commission, "Form S-1 Registration Statement Under the Securities Act of 1933: Facebook Inc." (Washington, DC, February 2012), accessed February 12, 2012, http://s3.documentcloud.org/documents/288894/facebook-prospectus.pdf; Hansell Saul, "Zuckerberg's Law of Information Sharing," *The New York Times,* November 6, 2008, accessed February 10, 2012. http://bits.blogs.nytimes.com/2008/11/06/zuckerbergs-law-of-information-sharing/.

version for purposes other than that for which the information was initially collected, and the cycle continues again.

In other words, for the past decade, publicly collected information, required of individuals to engage in democratic activities and receive specific services, has been computerized, and then sold or given to various entities within the private sector – thus, repurposed in potentially countless ways. Typically, private companies that receive government-collected information proceed to augment the dataset by associating the information with additional data points (such as consumer behaviors, credit ratings, magazine subscriptions, fishing licenses, online activity and profiles, etc.). As individuals continue to share more personal information online with private companies, the Data Cycle strengthens. Individuals' shared personal information is harnessed by private companies and used to make the government's original dataset more robust and therefore, more valuable. The enhanced version of the government's original dataset is then sold to various parties – and of particular interest, parties on behalf of the government. Often the government not only buys back the information it once sold or gave to the private sector, but also does so by using taxpayer funds to acquire the data for a new purpose.

Case Study: Voter Registration Data

Although the Data Cycle is occurring in a number of different industries (*see Appendix A)*, the remainder of this chapter focuses specifically on investigating how the Cycle is affecting the treatment of citizens' voter registration data. I have selected to closely examine only one prime case study in order to concretely and thoroughly illustrate how the Cycle functions in practice. The voter registration data trail is uniquely ripe for scrutiny because many Members of Congress have used their Member Representational Allowances ("MRAs") to obtain enhanced voter registration data by hiring private

companies.[40] Since MCs must disclose to the public how they spend their MRAs while in office, the details of their specific purchases is ascertainable.

Within this section, I conduct both a comparative historical analysis and a close examination of how individuals, private data companies, and the government participated in the Data Cycle in 2010. This case study reveals how the government has shared citizens' voter registration data with various third parties since as early as the 1920s – much before practical computers, databases, and the Internet even existed. However, after the advent and adoption of modern technologies, the extent to which the government shared voter registration data in 2010 has dramatically proliferated. This case study illuminates how a combination of technical advancements, a weak and outdated legal infrastructure, the rise of a new industry, and powerful political interests have contributed to the Cycle's development in the twenty-first century.

Historical Baseline

Prior to the advent of computers or networking capabilities, the federal government still collected and shared citizens' voter registration data with external entities. This section, which serves as a historical baseline to provide context to compare the current state of affairs, exemplifies how technological advancements made over the mid-twentieth century enabled and amplified the government's already existing desire to re-purpose and re-use citizens' voter registration information.

As a starting point of comparison for this case study, I chose to inspect the American government's Progressive era election reforms due to two important factors.[41]

---

[40] MRAs are the official budgets of MCs used to conduct "official and representational duties to the district from which elected." The Glossary of Terms of the Statement of Disbursements. U.S. House of Representatives, s.v. "member representational allowance," Disbursements.House.Gov, accessed October 1, 2011, http://disbursements.house.gov/glossary.shtml.

[41] The Progressive era in the U.S. occurred largely from the 1890s to the 1920s.

First, throughout the Progressive era the federal government administered voter registration processes without the use of modern technologies, such as computers, databases, and connectivity. Since such facilitative technologies were not in existence during the early 1900s, the Progressive era's election reforms provide insight into how longstanding political, legal, and economic factors have contributed to the Cycle's development. Additionally, during the Progressive era the government mandated the permanent voter registration of its citizens for the first time. Specifically, during the 1890s to 1920s, election reform spread throughout the states as a means to change the way elections were executed. An important part of these reforms included changed voter registration procedures.[42]

In 1929, Political Science Professor at the University of Washington Joseph P. Harris conducted a study reflecting on Progressive era voter registration reforms, published by the Brookings Institution as the "Election Administration in the United States." Harris describes that with the increase in immigration and the rise of "large cities" after the Civil War, election frauds became increasingly commonplace. Additionally, during this time, "drunkenness, disorder, violence, and bribery" characterized polling locations.[43] Although Massachusetts and many other New England states had enacted registration laws in the early 1800s to specifically prevent illegal voting and violence at the polls, the rest of the states did not follow suit until after the Civil War when it became apparent that registration was crucial to conduct an acceptable election.[44] Therefore, from 1860 to 1890, almost all of the remaining states adopted some type of a voter

[42] Committee on State Voter Registration Databases et al., *Improving State Voter Registration Databases Final Report* (Washington, DC: The National Academies Press, 2010), 59.
[43] Joseph P. Harris, "Election Administration in the United States," The Brookings Institution and Institute for Government Research Studies in Administration, no. 27 (1934): 18, accessed September 23, 2011, http://www.nist.gov/itl/vote/upload/chapter1.pdf.
[44] Ibid., 20.

registration protocol – even though in some states, the requirement just applied to citizens living in large cities.[45]

In the early 1900s, registration laws were tightened. States began to implement the "permanent registration of voters," resembling the process that is required of citizens in the twenty-first century. Voter registration and identification was a process that was initially instated in the 1920s and 1930s in order to ease the process of election administration and again, directly combat voter fraud. In 1930, about thirty states had made the procedural change to permanent voter registration in order to improve their registration systems.[46]

However, with this change to permanency, the government began to quietly use citizens' recorded information in ways other than solely for election administration and voter fraud prevention. In particular, since the 1920s the government has provided voter registration data to the quasi-public entities of political campaigns.[47] Political campaigns have utilized voter registration data for nearly a century in order to better target individuals for donations and/or support.[48] Additionally, the government has regularly

---

[45] Ibid., 18.
[46] Ibid., 22.
[47] "Voter Privacy in the Digital Age," The California Voter Foundation, June 9, 2004, accessed September 29, 2011, http://www.calvoter.org/issues/votprivacy/pub/voterprivacy/introduction.html.
[48] While the focus of the Data Cycle refers to how the federal government in its official capacity is circumventing the Privacy Act of 1974 by repurposing Americans' data through the use of private party intermediaries, the Cycle is also occurring in the political campaign arena. For example, the private company NGPVAN is a campaign software firm that has a national contract with the Democratic National Party ("DNC") and ran the Obama campaign's data platform in 2008 and is doing so again in 2012. The firm acquires voter lists that have been compiled and standardized by the DNC (originating from state election offices). Once NGPVAN obtains this data they then do "hygiene checks" and then enter it into its software system. After the voter list information is entered, NGPVAN appends this information with other types of information (e.g., hunting license registries) and purchases data from other firms such as Catalist. Campaign-related information produced from fieldwork (e.g., canvassing, donating, and phone-banking) is also added to the database. While this information is not sold directly to the federal government, any Democratic candidate including incumbents running for election or re-election may use it throughout the entire election cycle. Jim St. George (Co-owner of NGPVAN), interview by Melissa Oppenheim. November 11, 2011. 11:30am EST, NGPVAN Offices, Somerville, MA, transcript in *Appendix B*, "Oral History Interviews;" "Voter Privacy in the Digital Age," The California Voter Foundation.

employed voter registration lists to help establish jury pools.[49] Notably, only analog technologies existed during the Progressive era prohibiting the government from widely using and re-using such information. But even with just paper and pencil collection and recording paradigms, strong non-technological factors influenced the government's superfluous uses of such information.

The Move to Digitize Voter Registration Data

Although technology such as processing power, storage capacity and connectivity significantly advanced over the latter half of the twentieth century, the government did not officially start deploying such technologies to administer voter registration protocols until the early 2000s. This section details how modern technologies and policies have enabled the government to act on its desire to repurpose citizens' voter registration data.

Throughout the 1960s, 1970s, and even throughout the 1980s, the government ran on a largely paper-based collection and record maintenance system even though new useful technologies had already emerged.[50] As discussed further in Chapter Two, prior to the 1980s, the notion of "computing" generally referred to mainframe computers that were largely purchased only by big corporations or the U.S. government for military purposes.[51] From the beginnings of the "Electronic Numerical Integrator and Computer" ("ENIAC") used by the military at the University of Pennsylvania to calculate ballistic equations in the 1940s, and the "Electronic Recording Method of Accounting" ("ERMA") used by Bank of America to automate their bookkeeping systems in the 1960s, computing machines were only used to collate and organize a limited number of efforts.

---

[49] Committee on State Voter Registration Databases, *Improving State Voter Registration Databases,* 92.
[50] U.S. Congress, Office of Technology Assessment, "Federal Government Information Technology: Electronic Record Systems and Individual Privacy," (Washington, DC: Diane Publishing, 1986), 2, accessed October 24, 2011, http://books.google.com/books?id=jrP7XBOZUK4C.
[51] Stephanie Dick, "History of Computing" (lecture, Sever Hall 306, Harvard University, Cambridge, MA, February 9, 2011).

Concurrently, over the course of the 1960s, the U.S. Defense Department's Advanced Research Projects Agency ("DARPA") was developing packet-switched networks, the beginnings of the modern day Internet. Throughout the 1960s to the 1980s, new types of standards and protocol infrastructures were tested and implemented. In 1990, Tim Berners-Lee who was working as a consultant at CERN (or the European Organization for Nuclear Research) invented a linked hypertext organizational system for CERN's network, which later came to be known as the World Wide Web. The diffusion of the Internet and the Web occurred largely over the course of the late 1990s.[52]

The government's adoption of computing technologies and the Internet impacted the ways in which it wished to collect and handle citizens' voter registration data. This was reflected in policy changes where, in 1993, the National Voter Registration Act ("NVRA"), also known as the "motor voter law," sought to make voter registration more accessible and laid out rules for how states should maintain lists.[53] Specifically, the Act mandated that the states ensure that "accurate and current voter registration rolls are maintained."[54] After NVRA passed in the early 1990s, efforts increased to upgrade and maintain voter registration databases.[55] For example, in 2002 Congress passed the Help America Vote Act ("HAVA"), which mandated that all local and state election offices computerize their voter registration lists.

---

[52] David Mowery and Timothy Simcoe, "Is the Internet a US invention? An economic and Technological History of Computer Networking, Research Policy," *Science Direct* 31 (December 2002): 1369-1387, accessed October 24, 2011, http://www.sciencedirect.com/science/article/pii/S0048733302000690.
[53] "Statewide Voter Registration Database," PEW Center on the States, Electonline.org, and the Constitution Project (March 2002): 1-16, accessed September 24, 2011, http://www.pewcenteronthestates.org/uploadedFiles/Statewide%20Voter%20Registraion%20DB.pdf.
[54] HAVA also provides registrants the ability to remove their names from the roll should they desire to do so. See U.S. Election Assistance Commission, "SUBCHAPTER I–H—NATIONAL VOTER REGISTRATION," accessed September 20, 2011, http://www.eac.gov/assets/1/workflow_staging/Page/27.PDF.
[55] Committee on State Voter Registration Databases, *Improving State Voter Registration Databases*, 61.

Importantly, up until HAVA's "computerization mandate" in 2002, voter registration forms and voter lists were generally held on file at an election office in "practical obscurity" – meaning the documents were available in limited formats and locations that greatly limited access and sharing.[56] The notion of "practical obscurity" usually protected information from widespread distribution and repurposement by the mere fact that documents were only available on paper in specific, physical locales.[57] Even in 2002, only ten states had implemented a unified database system; the majority of states had mediocre compilation databases, and thirteen states had not even created statewide databases.[58]

Also in 2002, the California Voter Foundation conducted the seminal, state-by-state study, entitled "Voter Privacy in the Digital Age," which analyzed state governments' voter registration data gathering and privacy practices. The study sought to specifically shine light on impending voter privacy issues and offer recommendations as to how states should best handle citizens' data.[59] The study was conducted prior to the passage of HAVA and is regarded as a "benchmark of pre-HAVA data collection and dissemination practices."[60] The study's pre-HAVA statistics evidence how some states were already sharing voter registration data even before the federal computerization

---

[56] "Supreme Court Justice John Paul Stevens defended the notion of practical obscurity in a 1989 decision limiting journalists' access to criminal records compiled and computerized by government agencies. Stevens wrote, 'There is a vast difference between the public records that might be found after a diligent search of courthouse files, county archives and local police stations throughout the country and a computerized summary located in a single clearinghouse.'" See "Voter Privacy in the Digital Age," The California Voter Foundation.
[57] Ibid.
[58] "Statewide Voter Registration Database," PEW Center on the States, 2.
[59] Kim Alexander and Saskia Mills, "New Study Examines Voter Data Privacy in the US," The California Voter Foundation, May 28, 2004, accessed September 29, 2011, http://www.calvoter.org/news/releases/052804release.html.
[60] Ibid.

mandate. Thus, the study is extremely useful, as it illuminates how HAVA's computerization mandate affected the sharing practices of state election offices.

The California Voter Foundation found that prior to HAVA's mandate in 2002, twenty-two of the fifty states (or fifty-one including the District of Columbia, as the district possesses its own voter registration form) allowed unrestricted access to voter lists for commercial and non-commercial purposes.[61] While twenty-nine states in 2002 prohibited commercial use of voter lists, few had set a policy on how non-commercial entities (e.g., academics, journalists, and non-profits) could use and re-use the information.[62] Further, all states granted political groups and individuals' access to voter lists.[63]

The government was already broadly sharing citizens' data with other entities even before HAVA was enacted. As we learn in Chapters Two and Three, this discovery speaks to how non-technological factors such as longstanding sociopolitical, legal, and economic pressures have shaped the government's participation in the Data Cycle and disregard for the law. While voter lists were unreliable and difficult to re-use before they were computerized, the government's Progressive era practices expose how the government has been sharing voter information with other entities since the 1920s and 1930s – even if technological limitations greatly reduced their accessibility and reusability. This historical comparison demonstrates how the government has held a longstanding interest in accumulating, sharing, and repurposing individuals' voter registration data.

---

[61] "Findings: Voter Privacy in the Digital Age," The California Voter Foundation, June 9, 2004, accessed September 17, 2011, http://www.calvoter.org/issues/votprivacy/pub/voterprivacy/findings.html.
[62] Ibid.
[63] For example, the study cites how the National Abortion and Reproductive Rights Action League ("NARAL") used voter registration lists to target two million Republican and Independent women whose media preferences matched those of NARAL's existing membership. Ibid.

The government's participation in repurposing and reusing citizens' data in 2010 has dramatically proliferated since 2002. While the government's interest to re-use citizens' data may be traced back to the early 1920s and the enabling technologies to the 1970s and 1980s, a unique ingredient to the government's escalation in buying and selling citizens' data hinges on individuals' intensified sharing practices in the twenty-first century. While the Data Cycle's emergence stems from an inseparable combination of technological and non-technological drivers, the penultimate catalyst in its materialization is the rise in individuals' online sharing habits.

Voter Registration Data in the Data Cycle

To participate in the voting process in the twenty-first century, of-age citizens must provide specific pieces of personal information to their local and/or state election offices. Since the early 2000s, this information is then often given or sold to private data companies. These companies subsequently enhance the government's voter registration lists and often sell back augmented versions of the lists to MCs who purchase this data. Pursuant to the Member's Handbook, MCs may only buy information on their constituents; however, many MCs have begun doing so using taxpayer dollars under the pretenses of the franking privilege ("the frank").

To vote in an election, citizens must complete voter registration forms collected by their state of residence (i.e., Figure 1, Node 1). This voter registration form asks citizens to provide personally identifiable information (i.e., name, home address, mailing address, date of birth, phone number, party affiliation, and race/ethnicity).[64] As of November

[64] U.S. Election Assistance Commission, *Online Voter Registration Form* (Washington, DC, March 1, 2006), accessed October 29, 2011, http://www.eac.gov/assets/1/Documents/national%20mail%20voter%20registration%20form%20english%20February%2015%202011.pdf.

2011, voter data existed in the following two formats: (1) the actual registration forms (or "affidavits") and, (2) voter lists compiled by election agencies, consisting of the information supplied by voters on their registration forms.[65] In general, voter lists are created from the data collected on voter registration forms.[66]

After completing his/her registration form or affidavit, a citizen may then participate in the voting process, while his/her election office enters his/her data into a computerized voter list. At this point, the citizen has only engaged in the Data Cycle to the extent to which he/she has complied with providing the government information required to vote. However, many citizens do not realize that although the data gathering transaction appears complete, it does not necessarily stop there.

Since HAVA and increases in connectivity, storage space, and processing power over the first decade of the twenty-first century, state governmental offices have been selling this information to various companies (i.e., Figure 1, Node 2). In the first decade of the twenty-first century, it has become increasingly common for private data firms to acquire these lists. Specifically, pursuant to HAVA, once a citizen provides his/her information to the state, that information is compiled into a federally required, centralized database.[67] States, however, do not closely police the distribution of voter lists to third parties. In fact, the ways in which local election offices may engage with the private sector is not clearly articulated in HAVA or the law. Thus, citizens have remained largely unaware of the secondary and tertiary uses of their personal information, which superficially appears as an exchange that ends after the primary point of collection of

---

[65] "Voter Privacy in the Digital Age," The California Voter Foundation.
[66] Ibid.
[67] Pursuant to the HAVA, citizens' state election officials must maintain statewide personal information "computerized list[s]." "Help America Vote Act of 2002." (P.L. 107-252), *United States Statutes at Large.* 2002. 166 Stat. 1666. 107th Cong.

their information.[68]

In digital format, voter lists are easy to duplicate at the push of a button, transfer in bulk, and associate with other lists of information.[69] According to a PEW Research Center study conducted in 2002, "Almost every state that produces a statewide list also sells that list."[70]

The Election Assistance Commission ("EAC"), an independent, bipartisan agency created by the HAVA in 2002, regularly conducts audits of various secretaries of states ("SOS") to continue to improve the program. According to West Virginia's May 2011 EAC audit, the state amassed a net income of $219,410 from the sale of voter registration lists in 2011.[71] The EAC's March 2011 audit of Alabama, revealed that the state netted $88,000 from the sale of voter registration lists in 2011.[72]

Since the early 2000s, commercial data brokers increasingly seek to obtain personal information about American citizens from county and state election offices around the nation as a means to combine this information (with other collected personal information) to build reliable, accurate, and powerful databases. The information that

---

[68] In regards to the notion of "practical obscurity," this is a key example of how the Internet and the computerization of government records have transformed public documents in the case of voter lists into actually very public documents, baring unprecedented types of implications and uses. "Voter Privacy in the Digital Age," The California Voter Foundation.

[69] Ibid.

[70] Such lists are considered public record "because they are government documents created by government agencies." In fact, if a state does not distinctly address how its voter lists may be redistributed, public record laws are generally used to defend its distribution. "Voter Privacy in the Digital Age," The California Voter Foundation; "Statewide Voter Registration Database," PEW Center on the States, 7.

[71] U.S. Election Assistance Commission, "Report No. E-HP-WV-04-09," (Washington, DC, May 5 2011), accessed September 29, 2011, http://www.eac.gov/assets/1/Documents/FINAL%20EAC%20Management%20Decision%20West%20VA%20E-HP-WV-04-09.pdf.

[72] Both of the EAC's audits of the Secretaries of States ("SOS") of Alabama and West Virginia found that neither SOS deposited his/her income netted from the sale of the voter legislation lists into the HAVA election fund. U.S. Election Assistance Commission, "Report No. E-HP-AL-06-1," (Washington, DC, March 28, 2011), accessed September 28, 2011, http://www.eac.gov/assets/1/Documents/Final%20EAC%20Management%20Decision%20Alabama%20E-HP-AL-06-10.pdf.

commercial data brokers receive from state election offices serves as useful data points in creating data products that may be sliced and diced depending on any one client's needs, such as for election organizations, Members of Congress, commercial retailers, insurance agencies, and credit card companies (i.e., Figure 1, Node 3). Once commercial data brokers obtain citizens' voter registration data, the brokers augment this information with other online and offline private and public information. Examples include magazine subscriptions, credit card records, consumer purchasing behaviors, participation in certain types of events, census data, religious affiliations, and the list goes on.[73] In particular, private political data aggregation and management companies sell enhanced versions of citizens' voter registration information to government officials (who buy back this data using citizens' tax dollars).[74] Specifically, political data firms, such as Lockheed Martin, ChoicePoint, Dun & Bradstreet, Aristotle International, Catalist, Target Point, and Franking Grid are often among the list of private parties that boast maintenance of American voter databases.[75]

It should be noted that individuals' online sharing and communication habits, which have been increasing since the introduction of the Web in the 1990s, are fundamental to the Cycle's existence. Connecting the three nodes of the Cycle, private sector companies obtain datasets from the government, augment the datasets with additional information, and then sell the enhanced versions of the datasets back to the government.[76] As discussed at greater detail in Chapter Three and the Conclusion, the

---

[73] "Around The Clock Access To Voter Data: VoterListsOnline.com," Aristotle, accessed November 29, 2011, http://www.aristotle.com/content/view/35/119/.
[74] Kim Zetter, "For Sale: The American Voter," *Wired*, December 12, 2003, accessed November 29, 2011, http://www.wired.com/politics/security/news/2003/12/61543?currentPage=all.
[75] Laura Quinn (CEO of Catalist), interview by Melissa Oppenheim, November 9, 2011. 6:30pm EST. Omni Parker Hotel, Boston, MA, transcript in *Appendix B*, "Oral History Interviews."
[76] Often the government entity providing the data is not the same government entity purchasing the data.

American public is increasingly growing neurologically dependent on the Internet and communication technologies.[77] Through regularly connecting online, the human brain stimulates higher levels of neurological chemicals that humans become habitually trained to desire.[78] As we continue to feel like we need to be connected, and therefore, continue to share increasing amounts of personal information with third party services, such as social media platforms, e-mail websites, shopping websites, and gaming websites, an entire industry has emerged to track individuals online and buy and sell gleaned information – often with the government. Without individuals' heightened usage of the Web, private sector companies would not be able to offer the government much information and correspondingly, the government would likely not feel as compelled to hire private sector companies as middlemen to enhance its datasets. Thus, the Data Cycle is dependent on individuals' increased usage of online services.

Most Americans do not realize that when they provide the state with their personal information in order to vote, they are implicitly consenting to allow their information to enter the endless Data Cycle. It is even more doubtful that voters know they are funding their elected officials' ability to purchase their voter registration information – in an enriched form – from political data brokers.

For this case study, I chose to focus on the activities of the following two political data firms: Aristotle International and Catalist. I chose these two firms not because they are the biggest or most frequently employed firms, although they are popularly used. Instead, I chose to focus on these two firms because they both clearly publicize their

[77] Paul Zak (Professor of Economics and Department Chair and Founding Director of the Center for Neuroeconomics Studies at Claremont University), phone interview by Melissa Oppenheim, July 16, 2010. 9:45am EST, transcript in *Appendix B*, "Oral History Interviews."
[78] Ibid.

companies' ability to offer a unique, comprehensive database of voting-age Americans that combines standard demographic and high-quality political information with various types of commercial data. Additionally, while those in this industry argue that Lockheed Martin and Dun & Bradstreet[79] are actually much more heavily entrenched in this arena, I chose to focus my study on Catalist and Aristotle because the elected officials who use these two firms disclose the ways in which they employ these firms (for "publication/reference materials") in a much more explicit way.[80] Moreover, since Aristotle's legislative clients are majorly Republican, while Catalist is a service that is only available for Democrats and/or progressives (CEO Laura Quinn emphasizes "progressives with a small p"), this makes for an interesting usage comparison.[81] Through analyzing the Quarter 1, 2010 – Quarter 1, 2011 of the Members of the U.S. House of Representatives' expense disbursement records, these two firms were the most commonly used for the specified reason of data consumption.

---

[79] Dun & Bradstreet has issued and maintained Data Universal Numbering System ("DUNS") numbers since the 1960s. DUNS numbers are unique identifiers for businesses. In October 1994, the DUNS number was widely adopted by federal and commercial entities around the world as a standard business identification number for federal electronic commerce. In April 1999 the Federal government formally recognized the DUNS number system as the government's contractor identification code for all procurement-related activities. Today, many U.S. government agencies will only do business with firms that have DUNS numbers. Many other smaller, newer firms feel that Dun & Bradstreet is monopolizing the industry. "About the D-U-N-S Number," Dun & Bradstreet, accessed October 29, 2011, http://fedgov.dnb.com/webform/pages/dunsnumber.jsp; Quinn, interview.

[80] The other firms also offer a wider selection of services and products that are not solely data-related. Daniel P. Beard, comp. U.S. Congress. Statement of Disbursements of the House from January 1, 2010 to March 31, 2010. 111th Cong., 2d sess., April 13, 2010. Washington, DC; Daniel P. Beard, comp. U.S. Congress. Statement of Disbursements of the House from April 1, 2010 to June 30, 2010. 111th Cong., 2d sess., July 27, 2010. Washington, DC; Daniel J. Strodel, comp. U.S. Congress. Statement of Disbursements of the House from July 1, 2010 to September 30, 2010. 111th Cong., 2d sess., November 15, 2010. Washington, DC; Daniel J. Strodel, comp. U.S. Congress. Statement of Disbursements of the House from October 1, 2010 to December 31, 2010. 112th Cong., 1st sess., January 11, 2011. Washington, DC; Daniel J. Strodel, comp. U.S. Congress. Statement of Disbursements of the House from January 1, 2011 to March 31, 2011. 112th Cong., 1st sess., May 23, 2011. Washington, DC.

[81] "Progressive" may entail democratic political candidates, elected officials, and political organizations. However, Quinn explained how Catalist also provides services to other types of "progressive" efforts so long as they are not commercially selling the data. For instance, Harvard University has a contract with Catalist to grant researchers access to the database to conduct studies aimed at providing insight into some aspect of the political process. Quinn, interview.

Although these two firms both provide enhanced constituent data to MCs, these two firms are quite different from each other. Aristotle International is a political data firm, started in 1983, that develops databases for political professionals and elected officials. Aristotle represents itself as non-partisan. However, from 2006 to 2011, the majority of Aristotle's clients were Republican; in particular, around 88% of Aristotle's House of Representative clients in the 2010 election cycle were Republican.[82] Comparatively, Catalist is a trust-owned Democratic political operative, started by Harold M. Ickes, that specializes in providing data for scores of liberal groups supporting the Democratic and/or progressive tickets. Catalist remains a much younger and smaller operation than Aristotle, with only forty-five people working for the company, but has enjoyed strong support by large, political donors such as George Soros. Both of these firms widely advertise a customized legislative product for Congressional members in their elected capacities.[83]

Aristotle advertises on its website that it obtains and catalogues the original files from more than 4,000 election boards, county clerks and Board of Registrars and then "appends telephone numbers for phone banks and cleanses the addresses through the United States postal service's National Change of Address program and the Social Security Administration file, flagging deceased voters."[84] Similarly, Catalist asserts, "[We] maintain and constantly update a complete national (fifty states, plus the District of Columbia) database of over 265 million persons (more than 180 million registered voters

---

[82] Aristotle International, Inc. v. NGP Software, Inc. (U.S. District Court for D.C. 2011), accessed October 18, 2011, http://www.campaignsandelections.com/article_assets/articledir_519/259782/Aristotle_NGP_Opinion.pdf.

[83] Quinn, interview.

[84] "Political Data," Aristotle, accessed October 2, 2011, http://www.aristotle.com/content/blogcategory/22/45/.

and 85 million unregistered adults). This sets us apart from other voter file vendors who assemble their database only immediately prior to election cycles or when customer demand for the data is high."[85] Additionally, these two firms offer special services for MCs in particular. Aristotle offers "Constituent Service!" and "Hill Data," services directly targeted for legislative offices.[86] Catalist, too, offers "Catalist Congressional," a product made available to MCs.[87] The website reads that Catalist Congressional provides, "[D]ata exports, and unlimited matching and uploads," to allow MCs to get their messages across with "unprecedented precision and effectiveness."[88]

Catalist CEO, Laura Quinn, stated in October 2008, "We're trying to build a complete record of every American over the age of 18… We aspire to be much more than just a database provider…"[89] In the same article Vinay Ravindran, then Chief Technology Officer of Catalist echoed, "We aspire… to build an ecosystem."[90]

In November 2011, I spoke with CEO Laura Quinn and asked her if her opinion and/or goals for the company have changed at all. She offered that they had ever so slightly. Today, Catalist is in the business of helping provide a more complete picture of individuals in order to project their anticipated civic engagement. "We are trying to help

---

[85] "Products," Catalist, accessed October 4, 2011, http://catalist.us/product.

[86] As Aristotle describes on its website, "In Constituent Service! every registered voter in the district is automatically assigned his or her own record. Each record may contain… numerous enhancements, such as phone number, estimated income, homeowner status or presence of children…" "Hill Data," Aristotle, accessed October 2, 2011, http://www.aristotle.com/content/view/92/117/.

[87] Catalist states that its congressional subscriptions are specifically only available for "the official use of Democratic Members of the House of Representatives for constituent communications only, as governed by House rules, and may not be used for fundraising, voter registration, or other electoral work." However, the extent to which this is enforceable akin to traditional franking abuses that occur with snail mail is hazy at best. "Products," Catalist; Garret Graff, "They Have Your Number," *The Washingtonian*, October 1, 2008, accessed June 20, 2011, http://www.washingtonian.com/print/articles/6/171/9627.html.

[88] "Hill Data," Aristotle.

[89] The founder of Catalist hopes that the service becomes the living record of nearly every political action ever undertaken by an American. Catalist highlights that its database builds on voter lists with more than 450 commercially available data layers. Graff, "They Have Your Number;" Quinn, interview.

[90] Graff, "They Have Your Number."

progressive organizations see the civic personalities of the people they are interacting with."[91]

Without the technical transformation of a paper-based recordkeeping system to a digital one, the creation of such an operation would be economically and physically impossible. Laura Quinn offers, "A couple of years ago only Fortune 500 companies could afford the names and addresses of their customers and build models to attempt to predict their behaviors… Not too long ago, the majority of people doing these types of activities had to rely on a crude set of tools, such as word of mouth and the phonebook."[92] However, with the computational resources that have been made widely available since the early 2000s, data aggregation services have become straightforward. "Now it is cheap enough," says Quinn, "that people can afford to maintain large datasets, append large datasets, and build mathematical and statistical models to predict behaviors."[93] Quinn points to her and Ickes' disappointments with the 2000 and 2004 elections as the catalyst for realizing the need to create the company Catalist. She felt that at the time, the overwhelming sentiment among Democrats was that if the campaign had just organized itself a bit smarter and had a real utility to employ, they could have made a huge political difference.[94] Representatives from Aristotle remained unresponsive for an interview.

[91] Quinn, interview.
[92] Ibid.
[93] Ibid.
[94] Ibid.

Quantitatively, in September 2011, I analyzed the most recently published five quarters (Quarter 1, 2010 – Quarter 1, 2011) of the Members of the House of Representatives' official statement of disbursements ("SOD").[95] Intriguingly, members of the House of Representatives have evidenced using their MRAs to purchase the enhanced versions of their constituents' voter registration data. By analyzing the House of Representatives' SOD and looking for outgoing payments to Aristotle and Catalist, I have drawn the conclusion that MCs often pay for detailed voter data with taxpayer dollars.

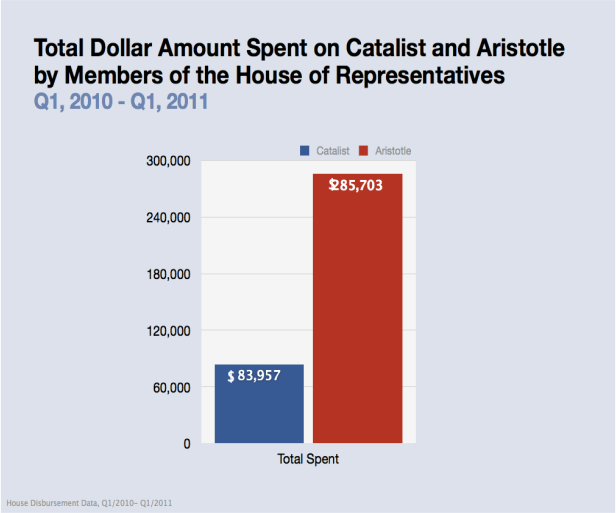

Figure 2: Melissa, Oppenheim, "MRA Expenditures on Catalist and Aristotle 2010-2011," August 2011.

Specifically, I have been able to identify both the amounts of money MCs have spent with these two firms and the frequency in which MCs have used them over the past year (*see Figures 2, 3, and 4*). From 2010 - 2011, fifty-one different House offices purchased data from Aristotle, and forty House

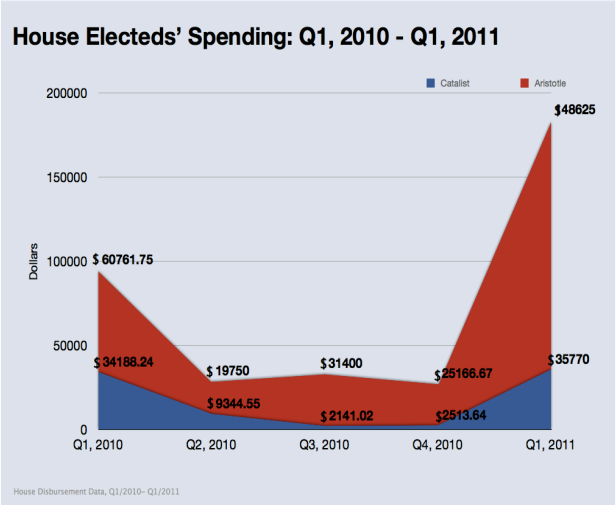

Figure 3: Melissa Oppenheim, "MCs Official Spending on Catalist and Aristotle per Quarter, 2010-2011," August 2011.

[95] The Senate does not have the same disbursement disclosure requirements as the Members of the House. While the House provides delineated expense reporting, the Senate does not break down how each senator spends his/her franking budget in the Secretary of the Senate quarterly report. Even after acquiring the reports in hard copy from the Library of Congress, Washington, DC, I have not been able to find an answer as to where the public can access its senators' delineated disbursements.

offices purchased data from Catalist. For instance, Rep. Henry Waxman (D-CA), a generally staunch privacy advocate and Chairman of the House Committee on Energy and Commerce, the committee which oversees areas of consumer protection and telecommunications, spends a portion of his publicly allocated budget on Catalist services that provide his office with personal data on the voters in his district. Further, over Quarter 1, 2011, $35,000 in taxpayer funds was spent on data from Catalist, among twenty-six different offices.[96] Familiar names such as Speaker of the House Nancy Pelosi (D-CA), member of the Ways and Means Committee Xavier Becerra (D-CA), member of the Judiciary Committee Zoe Lofgren (D-CA), and member of the Oversight and Government Reform Committee Jackie Speier (D-CA), each spent taxpayer dollars to purchase data on their constituents.[97]

Additionally, Representatives

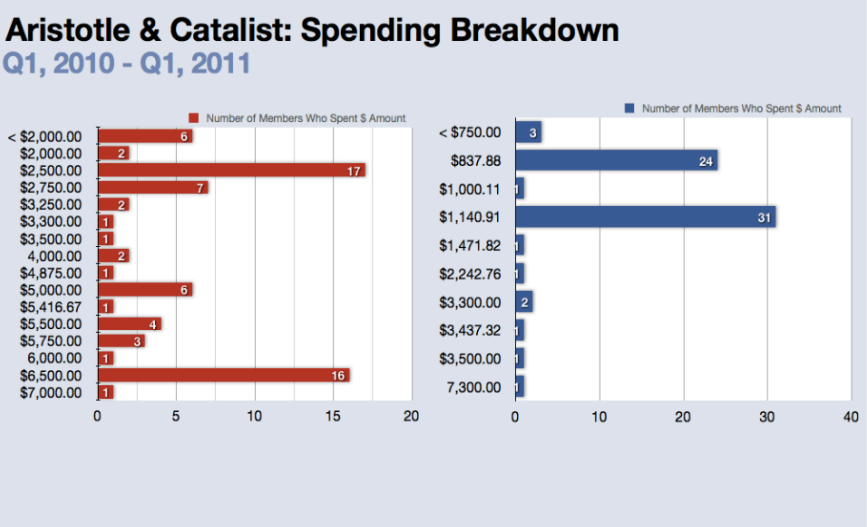

Figure 4: Melissa Oppenheim, "Spending Breakdown per Office on Catalist and Aristotle 2010-2011," August 2011.

[96] While at first blush it may appear that this sum of money is insignificant, its insignificance is what is so significant. The fact that one can spend so little in order to access such detailed data suggests how inexpensive it now is to collect and maintain such repositories. In years past, collecting and maintaining such data would have been in all likelihood cost prohibitive due to the general unavailability of such technology to the companies that now provide these data services. Beard, Statement of Disbursements of the House, April 13, 2010; Beard, Statement of Disbursements of the House, July 27, 2010; Strodel, Statement of Disbursements of the House, November 15, 2010; Strodel, Statement of Disbursements of the House, January 11, 2011; Strodel, Statement of Disbursements of the House, May 23, 2011; Quinn, interview.

[97] Beard, Statement of Disbursements of the House, April 13, 2010; Beard, Statement of Disbursements of the House, July 27, 2010; Strodel, Statement of Disbursements of the House, November 15, 2010; Strodel, Statement of Disbursements of the House, January 11, 2011; Strodel, Statement of Disbursements of the House, May 23, 2011.

Michele Bachmann (R-MN), formerly a 2012 presidential candidate, John Conyers (D-MI), ranking committee member of the Energy and Commerce Committee, Bob Goodlatte (R-CA), member of the Judiciary Committee, and Kevin McCarthy (R-CA), majority whip, have each purchased data from Aristotle over the 2010 – 2011 fiscal year.[98] In fact, over the course of Quarter 1, 2011, $148,000 was spent on Aristotle services between twenty-seven offices.[99]

MCs are able to legally pay for official, informational (non-promotional) mail and e-mail under the franking privilege. MCs have sent newsletters with free postage under the frank since it was written into law in 1789 during the First Congress.[100] Often taken advantage of by incumbents, the frank has a controversial history of usage where for almost twenty years during the latter half of the nineteenth century the privilege was eliminated.[101] Restored in 1891, the frank has traditionally provided a way for MCs to inform their constituents about official legislative business. In the twenty-first century, however, the use of the frank, a historically maintained congressional privilege, with new technological advancements such as the Internet is posing new issues.[102]

According to the Franking Rules in the Member's Handbook, the only major rule limiting the way MCs may purchase mailing lists addresses political party and constituent information. Specifically, mailing lists purchased under the frank may not contain "any campaign, campaign related, or political party information" and must only "contain

---

[98] Ibid.
[99] Ibid.
[100] U.S. Congress, Senate, "January 22, 1873 Senate Ends Franked Mail Privilege," accessed October 3, 2011, http://www.senate.gov/artandhistory/history/minute/Senate_Ends_Franked_Mail_Priviledge.htm.
[101] Ibid.
[102] Daniel Newhauser, "Franking on Facebook May Run Afoul of Founders," *Roll Call*, May 2, 2011, accessed June 5, 2011, http://www.rollcall.com/issues/56_114/Franking-on-Facebook-May-Run-Afoul-of-Founders-205207-1.html.

information about individuals within the MC's district."[103] In other words, House Members cannot prepare their mailing lists solely based on political party and cannot use their official resources to send mass mailings to constituents outside of their district. While this is easy to control via snail mail, in the online world, geographic boundaries are a vague notion. For instance, on the social networking platform Facebook, House Members may use their franking budget to run advertisements for their constituents about official, informational events (e.g., an upcoming tele-town hall). Because Facebook users do not always list a location, or often list their current town instead of hometown, MCs may be targeting individuals who are not actual constituents of their district.[104] This serves as just one example of how technological and social developments of the twenty-first century are intersecting with traditional political practices, resulting in new types of consequences.

As of January 2012, very few regulations limit how MCs may strategically target their constituents.[105] Besides the rule that forbids MCs from sending or addressing mail with regard to voter preference, offices may purchase data lists that target constituents in other ways.[106] When looking at the quarterly SODs, all that is reported in a lump-sum number is the spending on franked communications (which includes hard-copy mail, online adverts, e-mail newsletters, and other types of mass mail).

Additionally, according to the study conducted by the California Voting Foundation, most states do not warn voters that their registration data may be sold to

---

[103] U.S. Congress, House, Committee on House Administration. *The Member's Handbook*, 111th Cong., 2d sess. (Washington, DC, 2011), 40.
[104] Newhauser, "Franking on Facebook May Run Afoul of Founders."
[105] With regards to online adverts, the House Administration Committee "requires Members to report how many times an online ad was clicked, referred to as a click-through, and how many times the ad appeared on a website, which is called an impression." Newhauser, "Franking on Facebook May Run Afoul of Founders."
[106] Danielle Kurtzleben, "More Taxpayer Money Spent on Mail During Election Years," *US News,* April 8, 2011, accessed September 5, 2011, http://www.usnews.com/news/articles/2011/04/08/more-taxpayer-money-spent-on-mail-during-election-years.

third parties, such as political data firms.[107] Researcher at the California Voter Foundation Keith Mills said, "The simple act of signing up to register to vote leads to this chain of data that goes to a lot of places you're not aware of."[108] Thus, official MCs' offices may purchase data similar to almost any other third party buyer in the industry in order to target their advertisements, mail, and e-mail messages to specific groups of constituents.[109]

Discussion

Through the lens of this case study, we clearly understand how the government has dramatically altered its collection and handling practices of voter registration data. In 2002, the federal government mandated the creation of computerized databases in each state to store (once manually collected) voter registration data. However, very few limitations were placed on how this information may be shared with other parties. Under the guise of the timeworn franking privilege, the frank has become a mechanism through which MCs access otherwise unattainable data on their constituents at no additional cost to their offices. In the voter registration case, an explicit marriage exists between the private and public sectors whereby the franking privilege has enabled new grey-area interactions to regularly take place without the public's knowledge or consent. This case study demonstrates how unintended consequences emerge when processes become

[107] Very few states explicitly reveal how the data voters provide on their voter registration forms may be used for secondary or tertiary purposes. For instance, the CalVoter project found that in 2002 only New Mexico, Texas, Tennessee, and Iowa included language referring to the fact that voters' personal information would be treated as a public record. "Findings: Voter Privacy in the Digital Age," The California Voter Foundation.

[108] Zetter, "For Sale: The American Voter."

[109] In 2003, *The New York Times* published an article on e-mail spam and congressional mailings, "… [A]t least 40 House members have bought… email address lists..." Private companies acquire voter registration files from a member's district and cross matches such lists "with large databases of names and email addresses…" For further expansion, see Jennifer 8 Lee, "We Hate Spam, Congress Says (Except When It's Sent by Us)" *The New York Times*, December 28, 2003, accessed September 4, 2011, http://www.nytimes.com/2003/12/28/us/we-hate-spam-congress-says-except-when-it-s-sent-by-us.html?scp=1&sq=franking+and+internet&st=cse&pagewanted=all.

ingrained without a transparent understanding by those initiating the data collection (government) or by those who are providing the data (individuals).[110] Additionally, a paradox exists whereby Americans are unknowingly funding their own augmented surveillance both literally, through their tax dollars, and metaphorically, through their personal information, which is captured by various third parties that share that information with the government.

With specific regard to voter registration data, Kim Alexander, President of the California Voting Foundation states, "[There] is [a] need to establish a national dialogue about how to protect voter privacy in the digital age and ensure that voter data practices are not a deterrent to voter participation."[111] Jim Harper, policy expert at the think tank CATO, adds, "[P]rotecting privacy from government is often impossible. When Americans apply for licenses or permits… [t]hey must submit information that the government requires… [P]ower to control personal information – is totally absent in the governmental context."[112]

The Data Cycle is especially interesting to consider in regards to the Privacy Act's "routine use" clause.[113] Although the Act outlines that personal information provided by individuals should only be used by the government for specific purposes, in the twenty-first century, when various third parties obtain control of voter registration lists no one has any real idea of how that information is being used and re-used.[114]

---

[110] On this note, often MCs themselves do not even realize how the data, their offices use, is obtained.
[111] Alexander and Mills, "New Study Examines Voter Data Privacy in the US."
[112] "Findings: Voter Privacy in the Digital Age," The California Voter Foundation.
[113] After conversations with representatives from various data firms that do business with the federal government, it is clear that this Act is not taken into consideration from the private sector side of the transaction.
[114] For instance, as early as 2000, *The New York Times* ran article in September reporting on how voter registration lists were being collected, enhanced, and sold by commercial firms such as Aristotle International. The article stated, "Voters may well provide personal information for voter registration

David Holtzman, former cryptographic analyst for the U.S. government and author of *Privacy Lost*, argues, "In practice, it's become difficult to use the Privacy Act for protection… The act contains major… provisions allowing government agencies to obtain a variety of information from private corporations…"[115] Holtzman argues that the "unfortunate reality" is that information collected for one purpose is almost always re-used for multiple others.[116] Today, it is almost impossible to quantify just how many hands have access to one's personal information because so many transactions are taking place between the public and private sector and within parties in each respective sector.

Privacy advocates argue that the U.S. Federal Government currently circumvents the Privacy Act by employing private sector companies as middlemen.[117] UC Berkeley Law Professor, Chris Hoofnagle articulates, "Even though existing laws strictly limit the government's ability to conduct surveillance on U.S. citizens, those limitations don't apply to corporations."[118] In other words, the government provides private commercial data brokers with its publicly collected data in order to later acquire enhanced versions of those datasets that would otherwise be legally inaccessible to the government.[119] Even Laura Quinn admits, "If you are willing to pay the price, it seems as though there is almost nothing that is not for sale [in terms of data]."[120]

---

without knowing that such information may be used for other purposes." Leslie Wayne, "The 2000 Campaign: The Internet; Voter Profiles Selling Briskly As Privacy Issues Are Raised," *The New York Times*, September 9, 2000, accessed November 29, 2011, http://www.nytimes.com/2000/09/09/us/2000-campaign-internet-voter-profiles-selling-briskly-privacy-issues-are-raised.html?pagewanted=all.

[115] Holtzman, *Privacy Lost,* 110.

[116] Ibid., 112.

[117] Trevor Potter, a former Federal Election Commissioner, argues, " 'Federal election data cannot be used for commercial purposes, but no one has ever challenged it... These are some fuzzy areas, and no one has really pushed the edges of it.'" Wayne, "The 2000 Campaign."

[118] Shane Harris, "FBI, Pentagon Pay for Access to Trove of Public Records," *Government Executive,* November 11, 2005, accessed June 18, 2011, http://www.govexec.com/dailyfed/1105/111105nj1.htm.

[119] Hoofnagle, "Big Brother's Little Helpers,"16.

[120] Quinn, interview.

While the federal government is not directly repurposing the data itself (in agreement with the Privacy Act), it is providing a means for others to do so on an immeasurably large scale, so long as the government gets to stick its fingers back in the honey jar when the honey is ready. For instance, in the Data.gov case detailed at the onset of this chapter, we are able to gain a significant taste of how the government's collection and handling of individuals' data is changing in light of both new technologies and antiquated legislation. In fact, Data.gov may be the ultimate example of the Data Cycle occurring in U.S. society today. The government is itself incentivizing the repurposing of data on an unprecedented, massive scale. Federal agencies collect data for specific purposes. After the data is collated it ends up as a large file sitting on a government computer, ideally used for the purpose in which it was collected. Pursuant to the Data.gov project, if this file meets certain criterion deeming it safe, reliable, and threat-free, the agency housing this file is then strongly encouraged to post it online.

While Data.gov promotes transparency by revealing what government collected data entails, one cannot ignore the fact that the federal government is spurring federal agencies to release collected data (in aggregate form) for public use. Once in the public domain, the number of ways in which the published data could be used and re-used by other parties is difficult to estimate. It does not seem farfetched to hypothesize that often much of the data that is made public on Data.gov is enhanced with other datasets, repackaged, and sold back to the government (perhaps to a different agency) for a completely different use. In this respect, the government is naively championing "transparency." The government is supporting the repurposing of big government data.

On January 21, 2009, President Barack Obama welcomed his senior staff to the White House and stated, "The way to make government responsible is to hold it

accountable. And the way to [do this]… is to make it transparent so that the American people can know exactly what decisions are being made, how they're being made, and whether their interests are being well served."[121] However, this mantra is not applied uniformly across the government and begs the question as to why only some taxpayer-funded data is made available to the public (e.g., certain data sets available on Data.gov) and why the practice of such collection is kept quiet (e.g., voter list data used by MCs while in office). Regardless of the type of record, information collected by the government, funded by taxpayer dollars, has conventionally been understood to be a public good. As Diane Smith, well-known historian and economist states, "If the government is a government of the people, then all Americans own that which was gathered by their taxes."[122]

While the messy, siloed government files characteristic of the years prior to the Internet and computerization of public records certainly left much to be desired, the ways in which the government has taken advantage of technological advancements and individuals' sharing habits, is giving way to the development of a new norm. Specifically, government mandated information is today being used and re-used by the government and/or other third parties in an endless number of ways that may not be known or condoned by the public or even the government supplying the information. Laura Quinn believes that as a society "we have not sufficiently grappled with this [the issue of data privacy] at all."[123] While she stated that Catalist limits itself to the types of information it stores on an individual, the waters remain extremely murky.[124] After I explained the Data

[121] Lakhani, "Data.gov."
[122] Diane Smith, "The Commercialization and Privatization of Government Information," *Government Publications Review,* 12 (1985): 47.
[123] Quinn, interview.
[124] Ibid.

Cycle to Jim St. George of NGPVAN, a political data software firm that contracts with the Democratic National Party and ran the campaign data operations for the Obama campaign in 2008, he responded, "I'm not one of those people who is worried about my name in a database – get over it; your name is in a lot of databases."[125]

The Data Cycle reveals how the government and other entities are currently able to take advantage of the secondary and tertiary uses of government collected data, originally required of Americans for a specific purpose. The Data Cycle showcases how the government frequently gives or sells individuals' information to various private parties; such parties then repurpose the government collected information in ways that make the data useful for their clients. The evidence suggests that the government benefits from the repurposing of such information, as the government itself is not only a provider of information in the data marketplace, but also a client of many of these private companies. In spite of the outdated Privacy Act of 1974, the government has allowed its political and economic interests to guide the employment of modern computing and networking technologies. The following chapter specifically explores why the government has entered the Data Cycle and how its participation in the Cycle has revealed a slew of unchecked, questionable activities. While the Data Cycle is often championed as part of new government transparency efforts, the Cycle directly ignores the stated intentions of the Act. The remainder of this thesis discusses the government and individuals' roles in the Cycle and uncovers how and why the Cycle has been able to emerge in the twenty-first century.

[125] Mr. St. George's job relies on building and selling access to large, personalized databases. St. George's sentiments are reminiscent of then CEO of Sun MicroSystems Scott McNealy. In 1999 Scott McNealy allegedly told the public "privacy is dead, deal with it." St. George, interview.

CHAPTER TWO

The Government Has Gone Rogue

Figure 5: "… In a bureaucracy, the filed cards are reality. Punching new holes, we recreate the world." *Figure 5:* Alan Moore, David Lloyd, Steve Whitaker, and Siobhan Dodds, *V for Vendetta* (New York: DC Comics, 1989), 218.

The government continually sidesteps the Privacy Act of 1974, taking advantage of individuals' privacy. Unbeknownst to the public, the government currently participates in the Data Cycle, buying and selling Americans' data with very little restraint, serving as both a major provider and beneficiary of individuals' information. By actively participating in the Data Cycle and deliberately circumventing the Privacy Act, the government is functioning in an unregulated manner, buying and selling individuals' data with whomever, whenever it chooses. The government's role in the Data Cycle sounds alarms of waning privacy norms in the twenty-first century. In this chapter, I discuss why the government has entered the Data Cycle in the twenty-first century and why its involvement in the Data Cycle has rendered such actions rogue.

In this chapter, I argue that the government entered the Cycle due to an inseparable amalgamation of technological, socioeconomic and legal events that catalyzed a paradigm shift. Beginning in the 1980s, the American public began to adopt and widely

use modern computing technologies and Internet connectivity, initially invented for government use. This phenomenon gave rise to new types of private companies that profit off of gathering and packaging massive amounts of personal information, commonly referred to as the "data economy." The government, once supreme in gathering data about its constituents, has taken a back seat since the early 2000s to such companies. Unlike private companies, the government is restricted from directly amassing troves of individuals' personalized information for no specific purpose. However, the government is able to ultimately acquire such information by hiring private data companies to do what it cannot, circumventing existing legislation. In this chapter, I investigate how major technological advancements since the mid-twentieth century, the creation of the data economy, and an outdated legal infrastructure have allowed the government to unrestrictedly participate in the Data Cycle.

Laying the Foundation for a Future Shift in Government Data Collection Practices

In the twenty-first century, most Americans use personal computers and the Internet on a daily – if not hourly – basis.[126] In this section, I demonstrate that many of the technologies that have become crucial enabling elements of the Data Cycle were created originally for governmental use in the mid-twentieth century. The long-term effects of such breakthroughs have not only directly propelled the fields of computer science and engineering in ways key to the Cycle's emergence, but also have dramatically altered the established notion of the government's informational dominance.

[126] As of February 2011, more than 90% of Americans were reported to own some type of computerized gadget. Amy Gahran, "Report: 90% of Americans Own a Computerized Gadget," *CNN*, February 3, 2011, accessed February 15, 2012, http://articles.cnn.com/2011-02-03/tech/texting.photos.gahran_1_cell-phone-landline-tech-gadget?_s=PM:TECH.

During World War II ("WWII"), the United States led a "scientific war," boasting technology projects such as radar and the Manhattan Project. Although analog-computing machines had been developed in the 1930s, historians of technology Martin Campbell-Kelly and William Aspray detail that "the need for [quick and accurate] mathematical computation" remained a major theme throughout WWII, as wartime scientific contributions such as the atomic bomb and ballistic calculations relied on computing.[127]

Although most computing machines and computing projects during WWII were used as a "means to an end" to formulate complex calculations for wartime efforts, the existing analog technologies remained extremely primitive.[128] Even the most advanced defense department computer took ten to twenty minutes to solve one complex trajectory equation.[129] Often there were over 3,000 trajectories lined up for processing. During this time, the types of computing machines available were based on various types of analog technologies, such as differential analyzers,[130] punched card technologies, and "teams of human computers equipped with desk calculating machines."[131] The government sought

---

[127] There was also an existing effort by the British during this time period to use computers for what would later become cryptanalysis. Martin Campbell-Kelly and William Aspray, *Computer: a History of the Information Machine* (New York: Basic Books, 1996), 79-81.

[128] Ibid.

[129] Such calculations were often aimed at figuring out the single trajectory of a shell fired from a particular type of a gun at a given angle and would factor in seven types of variables, such as wind conditions. Ibid., 81.

[130] Vannevar Bush, an electrical engineer and later Chair of the National Defense Research Committee, and his students developed the Differential Analyzer at MIT in the early 1930s. The Differential Analyzer was composed of shafts, wires, wheels, pulleys and 1,000 gears, and was able to find the solution to differential equations. "Predecessors: The Differential Analyzer," in *The Case Files: John W. Mauchly and J Presper Eckert*, the Franklin Institute, accessed January 7, 2012, http://www.fi.edu/learn/case-files/eckertmauchly/analyzer.html

[131] The Moore School of Electrical Engineering at the University of Pennsylvania, a hot bed for computing projects during the 1940s, began to admit women in 1942 in order to train them to operate calculating machines. An undergraduate at Moore recollected, "The girls sat at their desks for hours, pounding away at the calculators of metallic digits. While it would take the most advanced defense department computer about thirty days to complete 3,000 trajectory calculations, Campbell-Kelly and Aspray note that such a scenario would equivalently take two rooms full of "calculating girls" a full month to complete – not a

desperately to optimize the ways in which digital computing machines could assist in such calculations. Nevertheless, throughout WWII, computers remained relatively scarce, large, and expensive.[132]

After WWII ended, the government still possessed an interest in building faster, more reliable computing machines to help perform calculations needed for future wartime efforts such as gun and missile trajectories, artillery firing tables, bomb construction, and piloting systems. Such computing machines were popularly described as "giant brains" since they were usually used to complete calculations, bookkeeping, and/or math.[133]

Throughout the 1940s and early 1950s, significant breakthroughs in computing hardware and design occurred specifically for governmental use. In hindsight, these breakthroughs were forerunners not only for the personal computers later widely used by the American public, but also as the symbols behind an informational power shift between the public and private sectors in the first decade of the twenty-first century.

Specifically, in the early 1940s, the notable Moore School Physics Professor John Mauchly teamed up with his research associate John Presper Eckert to devise a new electronic computing machine.[134] Prior to working with Eckert, Mauchly in 1942 published a research paper entitled, *The Use of High Speed Vacuum Tubes for Calculating.* In his renowned paper, Mauchly encouraged the use of vacuum tubes, which functioned akin to on-off switches and therefore allowed for easier control of electricity, in the

---

hugely improved time difference from using the single computer. Campbell-Kelly and Aspray, *Computer: a History of the Information Machine*, 79-82.

[132] Tracy Kidder, *The Soul of a New Machine* (Boston: Little, Brown and Company, 1981), 11.

[133] Ibid.

[134] Campbell-Kelly and Aspray, *Computer: a History of the Information Machine*, 86.

creation of new computing machines.[135] Mauchly believed that by employing electronic means, rather than mechanical ones, the speeds of calculating differential equations would greatly increase.[136] Sponsored by the American Army Ordinance Department, in 1943 Mauchly began working with Eckert to build a computing machine using these electronic components.[137] While Mauchly and Eckert were clearly working on a government project to improve automated calculations, their creation would subsequently serve as a working model from which other researchers would learn, iterate, and build new computing technologies.

Ultimately, Mauchly, Eckert, and a team of engineers built the first mainframe computer using vacuum tubes. From this point forward, computers developed using vacuum tubes became known as "first generation" computing machines.[138] Named the

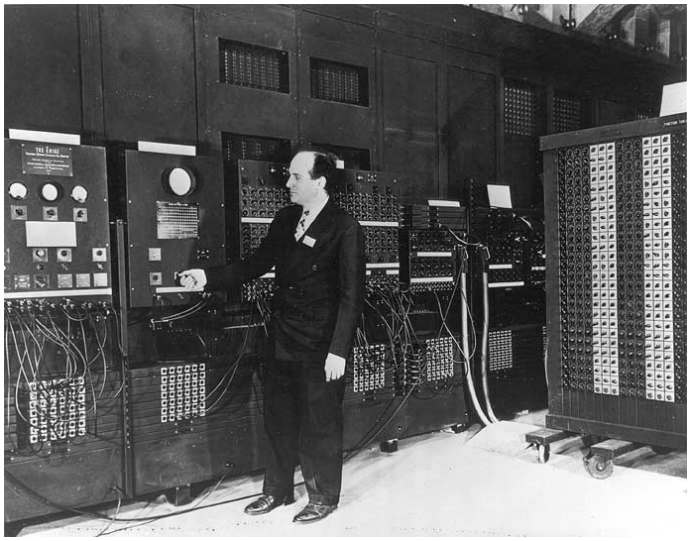

Figure 6: "The ENIAC Computer," The Franklin Institute, accessed January 2, 2012, http://www.fi.edu/learn/case-files/eckertmauchly/medium/eckertmauchly_photo_4.jpg.

[135] "Vacuum Tubes and Flip-Flops," in *The Case Files: John W. Mauchly and J Presper Eckert*, the Franklin Institute, accessed January 7, 2012, http://www.fi.edu/learn/case-files/eckertmauchly/vacuum.html.
[136] Mauchly wrote, "It is the purpose of this discussion to consider the speed of calculation and the advantages which may be obtained by the use of electronic circuits which are interconnected in such a way as to perform a number of multiplications, additions, subtractions or divisions in sequence, and which can therefore be used for the solution of difference equations." John Mauchly, "The Use of High Speed Vacuum Tube Devices," in *The Origins of Digital Computers,* ed. Brian Randell and David Gries. (Berlin: Springer-Verlag Berlin Heidelberg, 1975), 355, accessed January 7, 2012, http://books.google.com/books?id=Dwj4RmcZ1AoC.
[137] John Mauchly and J. Presper Eckert. "The Electronic Numerical Integrator and Computer (ENIAC)," in *The Origins of Digital Computers,* ed. Brian Randell and David Gries. (Berlin: Springer-Verlag Berlin Heidelberg, 1975), 359, accessed January 7, 2012, http://books.google.com/books?id=Dwj4RmcZ1AoC.
[138] A.P. Godse, and D.A. Godse, "Basic Structure of Computer" in *Computer Organization and Architecture*. (India: Technical Publications Pune, 2008), 1-18.

ENIAC, Mauchly and Eckert unveiled the machine on February 15, 1946 (*see Figure 6*). The machine proved to be "more complex than any previous electronic system."[139]

With the machine's début, ideas for how such a machine might influence or benefit other areas of society, other than the military, were immediately articulated. On the day of the ENIAC's unveiling, *The New York Times* ran a piece entitled, "Electronic Computer Flashes Answers, May Speed Engineering," proclaiming the ENIAC's completion and groundbreaking potential.[140] The reporter exulted, "Leaders who saw the device in action for the first time heralded it as a tool with which to begin to rebuild scientific affairs on new foundations. Such instruments… could revolutionize modern engineering…"[141] At this time, Mauchly and Eckert's use of vacuum tubes transformed the way electronic engineers thought about building computing machines and the understood potential of computing power.[142]

Soon after its inauguration, though, the ENIAC proved to be an extremely clunky machine (and not just by modern day standards).[143] Nevertheless, many in the field of computing considered Mauchly and Eckert's development as "the real starting point" for

---

[139] The full name of the mainframe was the Electronic Numerical Integrator and Computer. Campbell-Kelly and Aspray, *Computer: a History of the Information Machine*, 81.

[140] T.R. Kennedy Jr., "Electronic Computers Flash Answers, May Speed Engineering," *New York Times*, February 15, 1946, accessed January 2, 2012, http://www.fi.edu/learn/case-files/eckertmauchly/full/460215_eckertmauchly_article_1of2.jpg.

[141] Ibid.

[142] As subsequently discussed in this section, John von Neumann, renowned mathematician of the Institute for Advanced Study at Princeton, in 1945 wrote *A First Draft of a Report on the EDVAC,* detailing what would later become the foundation of modern computing architecture. In his paper, von Neumann claimed, "It is clear that a very high speed computing device should ideally have vacuum tube elements..." John Von Neumann, "The First Draft of a Report on the EDVAC" in *The Origins of Digital Computers,* ed. Brian Randell and David Gries (Berlin: Springer-Verlag Berlin Heidelberg, 1975), 5, accessed January 7, 2012, http://books.google.com/books?id=Dwj4RmcZ1AoC.

[143] Physically, the ENIAC possessed around 18,000 vacuum tubes, 70,000 resistors, 10,000 capacitors, 6,000 switches, 1,500 relays, and exposed wiring. The ENIAC soon revealed that vacuum tubes actually had a very short life span – there would be some sort of tube failure in the ENIAC every ten minutes (in addition to all of the other variable hardware and exposed wiring). Campbell-Kelly and Aspray, *Computer: a History of the Information Machine*, 81, 87-88.

the first electronic, large-scale, general-purpose computer.[144] While the ENIAC was built for the military, the machine possessed the capability to do much more than solely solve specific ballistic equations. Mauchly and Eckert elaborated, "The machine was developed primarily for the use of calculating firing tables for the armed forces. Its design is, however, sufficiently general to permit solution of a large class of numerical problems, which could hardly be attempted by more conventional computing tools."[145] Mainframe computing developments made throughout the 1940s and 1950s were generally funded and used by the military. However, as the ENIAC demonstrated, with the development of new mainframe computing technologies for one purpose came other newly envisioned uses and applications.

Contemporaneously, in 1945 the Hungarian and American mathematician John Von Neumann wrote the seminal paper, *A First Draft of a Report on the EDVAC,* which laid out the basis of modern computing architecture. Von Neumann outlined the designs for a new computing machine, the Electronic Discrete Variable Automatic Computer ("EDVAC"). Von Neumann described for the first time how the EDVAC's storage device would be able to "hold the instructions of a program and the numbers on which it operates."[146] Interested in von Neumann's work, the government funded the machine's development.[147] Retrospectively, scholars overwhelmingly point to von Neumann's

---

[144] "Validation," in *The Case Files: John W. Mauchly and J. Presper Eckert*, the Franklin Institute, accessed January 7, 2012, http://www.fi.edu/learn/case-files/eckertmauchly/valid.html.
[145] Mauchly and Eckert, "The Electronic Numerical Integrator and Computer (ENIAC)," 359.
[146] Von Neumann presented this as the "stored-program concept" or an architecture in which data and program are stored in the same memory in the computer. Campbell-Kelly and Aspray, *Computer: a History of the Information Machine*, 92-94.
[147] However, due to the ending of the war and wartime funding, the EDVAC was not finished until 1952. In 1949 Maurice Wilkes of Cambridge University actually developed the world's first operational, modern computing machine based on von Neumann's designs, known as the Electronic Delay Storage Automatic Calculator ("EDSAC"). Ibid., 104.

computing architecture as a significant marker of the beginning of the computer age.[148] To this end, von Neumann's architecture eventually offered the basic architectural concept from which the modern day general purpose computer was built.

As the ENIAC and EDVAC were introduced to the American public, the prospects of computer use by other parts of American society expanded. In the mid-1940s, the idea that computers could handle more diverse responsibilities than mere computing calculations began to mobilize – and with it, so did the idea that the computer could be used by masses of people for different purposes. For instance, in 1945 Vannevar Bush, an electrical engineer and later Chair of the National Defense Research Committee, proposed the idea of "memex." In his article "As We May Think," which ran in the *Atlantic Monthly* in July 1945, Bush proposed the idea of an adjustable microfilm-viewer, possessing qualities that are somewhat analogous to the structure of the twenty-first century's World Wide Web.[149] Bush's memex, which he envisioned would serve individuals as a memory bank to organize and retrieve data, was one of the first times in history that the computing machine had been envisioned as useful and necessary for everyday individuals.[150]

Government Support Increases Technologies' Accessibility

While Mauchly and von Neumann's contributions catalyzed further advancements made in computing hardware and storage design, the U.S. Defense Department continued to push for progress in computing hardware throughout the 1950s

---

[148] Ibid.

[149] Vannevar Bush, "As We May Think," *The Atlantic Monthly*, July 1945, accessed February 8, 2012, http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/3881/.

[150] In his piece, Bush offered, "The world has arrived at an age of cheap complex devices of great reliability; and something is bound to come of it... Consider a future device for individual use, which is a sort of mechanized private file and library… A memex is a device in which an individual stores all his books, records, and communications… It is an enlarged intimate supplement to his memory." Ibid.

and 1960s. Such progress made computing machines only slightly less expensive and therefore, accessible to few sectors of the public. During this time the semiconductor industry began to develop mass produced silicon-based components, such as high-performing silicon transistors, to create reliable, faster computing devices for the U.S. Air Force.[151] The transition from vacuum tubes to the transistor marked one of the major changes that gave birth to "second generation" computing machines characteristic of the 1950s and 1960s.[152] Soon thereafter, integrated circuits, characteristic of "third generation" computing machines,[153] came onto the market in the early 1960s and greatly aided the computer's functionality, reliability, speed, and cost.[154]

While computers in the 1950s, 1960s, and even early 1970s, remained gargantuan in size and still very expensive, entities other than the military began purchasing and employing the machines in unforeseen ways. During this time, big companies and universities began purchasing computing machines. By the mid-1960s businesses regularly

---

[151] Specifically, silicon transistors allowed for a more efficient way to alter and control the flow of electricity in circuits. The United States Air Force played a major role in propelling the semiconductor industry forward in the 1950s. In 1945, the Air Force began to encourage the electronics industry to create more advanced silicon transistors because the Air Force wanted to convert its existing flight control systems to digital computers. This provoked the electronics industry to meet the Air Force's demands. Many large companies transformed their development techniques into assembly-line fabrication processes. Beginning in 1958, Fairchild Semiconductor, the forerunner company, dominated the market for about a year and a half, as they were really the only firm that could produce reliable, high performance silicon transistors on a large scale. Fairchild witnessed its sales grow from $500,000 in 1958 to $21 million in 1960. By 1960 other firms such as Motorola, Texas Instruments, and Hoffman Electronics were in the business of semiconductor manufacturing, demonstrating the growing need for such electronic components. Christophe Lécuyer, "Revolution in Silicon," in *Making Silicon Valley: Innovation and the Growth of High Tech, 1930-1970* (Cambridge: MIT Press, 2006), 141-148, 156-162.

[152] Godse and Godse, "Basic Structure of Computer," 1-18.

[153] Ibid.

[154] Further, in 1959 Jean Hoerni, a physicist at Fairchild Semiconductor, patented a new development process (which was later licensed to the industry) to create more reliable transistors. Known as the "planar process," Fairchild created a manufacturing process to produce transistors with optimized characteristics for computer use. This process produced integrated circuits or society's modern day "computer chips." The planar process allowed for the packing of many transistor components to be built onto complex, single units, named integrated circuits. Lécuyer, "Revolution in Silicon," 129-130, 143.

installed large mainframe machines for data-processing purposes.[155] IBM and a dozen of other computer manufacturing companies in the mid-1960s sold large computing mainframes to commercial entities, such as airline companies that wanted to operate centralized reservation systems and banks that wanted to operate ATM machines.[156]

Perceived as an enormous resource, the government, large companies and universities, were the only entities that could afford to purchase and employ people who knew how to use such technology.[157] As historian of science Joe November describes, even in the early 1960s, mainframe computers were revered as "demigods," rather than used as convenient tools. In particular, operating a mainframe computer required significant technical know-how and even then, was often unpredictable.[158] November details how entire university departments would race to sign up to reserve their own sliver of computing time.[159] In fact, in order to transport computing machinery inside a building, the building would have to have an especially large elevator, which was often built specifically to transport a mainframe.[160] Unlike the twenty-first century's sleek, small gadgets and peripherals, the mainframe's internal components were often exposed, such as its wiring, as was certainly true of the ENIAC.[161] Generally, computing up until the late 1960s remained foreign to most people, rather inaccessible, challenging to use, and associated with large, mechanical and electrical machinery.

---

[155] Campbell-Kelly and Aspray, *Computer: a History of the Information Machine*, 222.
[156] Ibid.
[157] Humans were needed to manually program the mainframes by tediously moving circuits and inputting data via punched cards or magnetic tape. Dick, "History of Computing."
[158] Joe November, "LINC: Biology's Revolutionary Little Computer," *Endeavour* 28, no 3. (2004), 125-131.
[159] Ibid.
[160] Dick, "History of Computing."
[161] Ibid.

Technology Invented for the Government Enters Mainstream Society

By the mid-1970s computer hardware costs had fallen dramatically. As the prices of technology continued to significantly decrease over the course of the 1980s, a much larger slice of society could access more efficient and dynamic tools. With greater access though, new uses and consequences followed.

In 1959 Richard Feynman, physicist and Nobel Prize Winner for his work in quantum mechanics, gave a presentation at CalTech entitled, "There's Plenty of Room at the Bottom" where he encouraged a room full of students to investigate computing on the atomic scale.[162] Although his recommendation was not immediately pursued, six years later, during the height of transistor and integrated circuit development, Intel's co-founder Gordon Moore popularized the power of miniaturization in the 1960s. In his 1965 article, "Cramming More Components onto Integrated Circuits," Moore asserted that every eighteen months, the size of transistors will decrease by half, doubling computing power and storage capacity.[163]

Pressured to meet this goal, the electronics industry competed with each other to produce smaller, less expensive devices (e.g., pocket-sized calculators and digital watches) that continued to meet large consumer demand, further propelling the market.[164] This trajectory quickly transformed into a famous self-fulfilling prophecy; in the 1960s and 1970s furious innovation in semiconductors occurred. Moore wrote, "The complexity for minimum component costs has increased at a rate of roughly a factor of two per year…

---

[162] Richard Feynman, "There's Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics," *Engineering and Science* (1960), accessed September 3, 2011, <http://www.zyvex.com/nanotech/feynman.html>.
[163] In other words, the same amount of power could be obtained for half of the price every eighteen months. Hallam Stevens, "Nanocultures" (lecture, Sever Hall 306, Harvard University, Cambridge, MA, February 23, 2011).
[164] Ibid.

this rate can be expected to continue, if not to increase."[165] This prophecy has remained true since its introduction in 1965, as the miniaturization of electronic components has been inextricably linked with increasing processing power and storage capacity. Commonly known as "Moore's Law," this prediction has served as an invisible hand bringing smaller, cheaper, and more powerful components and devices to a larger group of the American public.

Historians Campbell-Kelly and Aspray detail that there were less than 1,000 minicomputers distributed worldwide in 1961.[166] Throughout the 1960s and 1970s, growth and adoption of the minicomputer increased at a rapid pace. In accordance with Moore's prediction, in the 1970s prices of computing devices steadily decreased, becoming increasingly accessible to the public. By the late 1970s the personal computer was developed and used by over 150,000 people for various purposes.[167] Describing this growth, Campbell-Kelly and Aspray offer, "By 1972 there were 65,000 minicomputers in use, and 150,000 by 1974."[168] In the late 1970s historians describe how the computer was growing more and more ubiquitous. "Practically every organization in America had come to rely upon computers, and ordinary citizens were buying them for their homes…"[169]

In the 1990s, computer processing-speeds were measured by how many *millions* of instructions per second the processor could read. In 2012, the average computer's

---

[165] Gordon Moore, "Cramming More Components onto Integrated Circuits," *Electronics* 38, no. 8 (April 9, 1965): 82.

[166] Minicomputers or personal computers were not developed overnight. It was not until the early 1970s with the widespread introduction of the minicomputer that the notion of computing would become more accessibly understood and forever change. In particular, in 1970 minicomputers were developed and priced around $20,000. In 2011, this would be a very expensive price to pay for a modern minicomputer or personal laptop, which are usually around priced around $1,000. Comparatively, in 1965 the price for a mainframe computer with the same amount of power as a minicomputer in 1970 cost ten times as much. Campbell-Kelly and Aspray, *Computer: a History of the Information Machine*, 229.

[167] Ibid., 207-228.

[168] Ibid.

[169] Kidder, *The Soul of a New Machine*, 12.

processing-speeds are measured by how many *billions* of instructions per second the machine can process. In agreement with Moore's law, computers continue to grow faster while prices continue to decrease. The first modern personal computer had less than ten megabytes of storage memory; smart phones in 2012 have around sixty-four gigabytes of storage memory.[170] A new Mac desktop computer with 500 gigabytes of memory in 2006 was double the cost of a new Mac desktop in 2012 with a terabyte of memory. Processing power and storage capacity have grown faster and larger, as the prices of such devices continue to decrease.

The Internet and World Wide Web is Invented

As more and more Americans began to obtain personal computers, the Internet became ubiquitous. Originally developed for governmental and university use, the Internet moved out of the government's domain and into mainstream society, as the growth of personal computers continued to expand. Today, Americans use the Internet on their personal devices, laptops, and phones as a primary means of communication. However, as Internet penetration increased throughout the late 1990s and early 2000s, an entirely new industry of online services and data scrapers emerged to profit off of individuals' interactions on the Web.

The start of what became the Internet began in 1957 as a military response to the USSR's launch of Sputnik. Under President Dwight Eisenhower, the government created the Advanced Research Projects Agency ("ARPA") as a means to gain a technological

[170] Random access memory (or "RAM") is a different form of memory from storage memory. RAM, akin to human's short-term memory, is where the computer stores information that can be made quickly accessible. Similar to the trajectory of storage memory, RAM has gone from around sixty-four kilobytes of memory in the early personal computers, such as the "Commodore 64," from the early 1980s to multiple gigabytes in some smart phones in 2012. "Timeline of Computer History," Computer History Museum (2006), accessed November 7, 2011, http://www.computerhistory.org/timeline/?category=cmptr; "Highest Ram Amount, Cell Phones Top List," PhoneEgg, accessed January 5, 2012, http://www.phonegg.com/Top/Ram-Amount-Cell-Phones.html.

edge over other countries. J.C.R. Licklider, appointed head of the Information Processing Techniques Office at ARPA in 1962, paved the way for the development of a network, ARPANET, to connect computers. In his paper, *Man-Computer Symbiosis* written in 1960, Licklider described his idea of how computing could be used to augment the human intellect and enhance human's everyday work. In 1962, Licklider wrote an article "Online Man-Computer Communication," outlining the ways in which humans and computers could work together to solve problems.[171] Licklider wanted to create an "Intergalactic Network," which was the first conception of the modern day Internet.[172] In 1965, Licklider left ARPA and Robert Taylor took over Licklider's project of connecting computers together. ARPA's goal was to increase overall computer power and decentralize information storage in case of an attack.

During WWII, the Defense Department channeled unprecedented funds into academic research in order to bring university scientists together to tackle the growing "technology gap" between the United States and its global competitors. According to *Fortune Magazine's* October 1948 issue, during WWII, about $500 million of governmental research and development money went to universities each year, accounting for

Figure 7: A cartoon of the ARPANET as it existed in the early 1970s, demonstrating how all of the users were universities, think tanks, and government agencies. "ARPA has a Network of Supercomputers," *Science Magazine* iv (March 10, 1972), 175.

[171] J. C. R Licklider, "Man-Computer Symbiosis," *IRE Transactions on Human Factors in Electronics* (1960).
[172] Ibid.

the majority of universities' research budgets.[173] In 1968, ARPA contracted Bolt, Beranek, and Newman ("BBN") to connect computers in different locations, namely computers at UCLA, Stanford Research Institute, University of California, and University of Utah.[174] This was called the ARPANET (*see Figure 7*). The ARPANET proved to work in 1969 when the first message was sent between two nodes.

After the ARPANET's 1972 public presentation at the first International Conference on Computer Communications, hobbyists, technologists, and scientists began building their own similar, personalized networks in the 1970s and 1980s for various purposes.[175] In 1974, Vint Cerf and Robert Elliot Kahn, then computer scientists, developed a standardized mode of connection, TCP/IP Internet Protocol Suite (known as "TCP/IP"), which established the protocol to create a network of networks or the modern day Internet. Throughout the 1970s and early 1980s different research groups experimented with the protocol, creating different systems. Finally, in 1980 the U.S. Defense Department set TCP/IP as the standard protocol to be used for all military computer connections.[176] By 1985, TCP/IP was being showcased around the world at different workshops and conventions, publicizing its reliability and speed.

Soon thereafter, the World Wide Web, a system of linked hypertext accessed through a Web browser, was built to organize the Internet and make it more accessible. Building from Vannevar Bush's 1945 conception of the memex, in 1965, Ted Nelson

---

[173] Steven Shapin, "Science and Business in Modern America" (lecture, Science Center, Harvard, University, Cambridge, MA, March 30, 2009).
[174] Barry Leiner et al., "A Brief History of the Internet," last modified January 23, 1999, accessed November 2, 2011, http://arxiv.org/html/cs/9901011v1/#Origins.
[175] Different groups, hobbyists, companies, and governments began creating their own networks for specific purposes. For instance, IBM created BITNET in 1979 as a means to conduct e-mail, while the National Science Foundation created NSFNET in 1984 to create faster and more efficient connections. Graduate students from Duke University launched USENET in 1980 order to create a global discussion group system. Ibid.
[176] However by 1985 security became a concern since many users began connecting to the ARPANET on TCP/IP. The U.S. Defense Department split off of the network and created MILNET to use instead. Ibid.

presented the idea of Xanadu in his "A File Structure for the Complex." This was the first time hypertext (or the linking of text with each other) was described.[177] While working as a consultant at CERN, Tim Berners-Lee launched the modern day Web on the Internet in 1991. This greatly popularized the usability and utility of the Internet. In 2011, Internet penetration globally has grown approximately 528% since 2000.[178]

Technological Advancements Shape the Government's Information Collection Practices

The methods of government data collection have radically changed in highly visible ways due to these advances in processing power, storage capacity, and connectivity. Computerized records enable government operations to run much more efficiently; details are no longer lost in stacks of paper, the same information does not have to be asked for repeatedly, and information may be easily shared and accessed. Formerly, government data maintenance was a process that depended on a physical organization system. The sharing of information with external entities (i.e., other governmental agencies, individuals, commercial parties) was challenging, as the files were easily misplaced, lost, too bulky to carry, or too sacred to transport.

However, the Office of Technology Assessment ("OTA"), a bi-partisan agency created during the Nixon administration to offer Congress technical expertise,[179] cautioned Congress in 1986 of the serious consequences that may emerge from digitally

---

[177] Ted Nelson, "A File Structure for the Complex, the Changing and the Indeterminate," *Association for Computing Machinery: Proceedings of the ACM 20th National Conference* (1965), 84-100.
[178] Globally, in 2000 there were approximately 360,985,492 Internet users; in 2011, there were approximately 2,267,233,742. "World Internet Usage Statistics News and World Population Stats," Internet World Stats, accessed February 15, 2012, http://www.internetworldstats.com/stats.htm.
[179] The Office of Technology Assessment was created in 1972 as an office of U.S. Congress. In 1995, however, it was shut down. *The New York Times* reports, "The agency, one of Congress' smallest… fell victim to budget cutting by the Republican majority and, its supporters say, shortsightedness about its value in providing unbiased, understandable advice on complex issues…" Warren Leary, "Congress's Science Agency Prepares to Close Its Doors," *New York Times*, September 24, 1995, accessed December 8, 2011, http://www.nytimes.com/1995/09/24/us/congress-s-science-agency-prepares-to-close-its-doors.html?scp=1.

storing individuals' personal information.[180] It is clear that the powerful technological advancements made since the 1940s and 1950s have helped facilitate robust transfers of digitally recorded information between parties. Digital records are now easily transferred, shared, copied, sliced and diced at the click of a few buttons. Harvard Law Professor John Palfrey offers, "Both private and public entities are increasingly encouraged to retain more data as a result of legal and market pressures."[181] Characteristic of the Data Cycle, public and private entities are not just retaining but also sharing more data with each other.

The Emergence of a Data Economy

The explosion of the use of the Internet and personal computing devices by the American public has reshaped life in the twenty-first century, creating undeniable convenience and improvements; however, such usage has consequences. As major technological advancements have moved out of the silo of government and into mainstream American society, new industries have emerged, taking advantage of the recent diffusion and novel uses of such technologies.

In particular, individuals share more information about themselves than ever before through the Internet. "The Digital Life," a website that showcases data from one of the largest-ever online research projects (covering 90% of the world's online population in forty-six countries), found that 83% of online Americans participate in social networking and that 50% of them participate daily.[182] EMC Corporation, "the world leader in information infrastructure solutions," announced in May 2009 that the "Digital

[180] U.S. Congress, Office of Technology Assessment, "Federal Government Information Technology," 11.
[181] Palfrey contends that data shared between such entities is either extremely effective or intrusive depending on one's perspective. Palfrey, "The Public and the Private at the United States Border with Cyberspace," 242.
[182] "The Digital Life: Study," TNS, accessed September 18, 2011, http://discoverdigitallife.com.

Universe is expected to double in size every eighteen months. In 2012, five times as much digital information will be created versus 2008."[183] Finally, in July 2010 PEW Research Center reported, "In a survey about the future impact of the internet, a solid majority of technology experts and stakeholders said the Millennial generation [aged 18-29] will lead society into a new world of personal disclosure and information-sharing using new media."[184]

The information that individuals volunteer online may be viewed as "a currency of value" where an entire industry has developed to capitalize off of it.[185] By employing tracking tools such as cookies, many Internet services such as Google, Facebook, Amazon, Yahoo!, and EBay, quietly gather information about their consumers to sell to advertisers and/or other third parties while offering individuals "free" services in return. This practice has largely spawned the Internet tracking industry.

"Cookies" were first incorporated into the early Web browser, Netscape Navigator, in 1994 to make online shopping easier.[186] *Wall Street Journal* reporter Julia Angwin discusses how it was not until the late 1990s during the "dot-com boom" that online ads became widely employed by advertisers.[187] In the wake of the dot-com crash, however, a "power shift [occurred], away from websites and toward advertisers" as

---

[183] Dave Farmer, "Digital Information Growth Outpaces Projections, Despite Down Economy," *EMC*, May 18, 2009, accessed November 19, 2010, http://www.emc.com/about/news/press/2009/20090518-01.htm.
[184] Janna Anderson and Lee Rainie, "Millennials Will Make Online Sharing in Networks a Lifelong Habit," Pew Research Center's Internet & American Life Project, July 9, 2010, accessed November 18, 2010, http://www.pewinternet.org/Reports/2010/Future-of-Millennials/Overview.aspx.
[185] Schonfeld, "The Privacy Dilemma."
[186] Julia Angwin, "The Web's New Gold Mine: Your Secrets," *The Wall Street Journal*, July 30, 2010, accessed August 2, 2010, http://online.wsj.com/article/SB10001424052748703940904575395073512989404.html.
[187] Ibid.

advertisers began paying for ads only if someone clicked on them."[188] In the late 1990s, advertising networks were motivated to routinely deploy cookies in order to get paid.

In 2012, various tracking tools, such as persistent cookies and beacons, are placed on users' computers whenever they visit different websites across the Internet, allowing advertisers and other entities to make a profit from collecting and analyzing consumers' preferences and behaviors. Regardless of the type of Web tracker, information about an individual's behavior on a particular website or across many websites, is stored in a specific file that identifies that person's computer with a unique code placed in such deployed tools. As author and reporter Ken Auletta puts it, "While it is true… 'on the Internet, nobody knows if you're a dog,' it is also true that your cookies know what that dog does online" (*see Figure 8*).[189] While online tracking has been in existence for approximately ten years, the technology has become not only much more sophisticated, but also much more routine, as corporations seek to produce as high a return of capital as possible.[190]

Figure 8: Steiner, Peter. *The New Yorker.* 5 July 1993. Vol.69 (LXIX). P 61.

[188] Ibid.
[189] Ken Auletta, "Media Maxims," Scribd, 21, accessed November 19, 2010, http://www.scribd.com/doc/22564045/Ken-Auletta-Media-Maxims.
[190] However, tracking technologies do pose some advantages to consider. In an article written by Jim Harper of the CATO Institute, Harper argues that online tracking is not a bad thing. Harper argues that the reason Internet users are able to benefit from "free" services, such as Facebook, Yahoo, MSN, and Google, is because personalized advertising is more profitable than "advertising aimed at just anyone." Plainly, consumer tracking is the way that "free" online services are able to remain (at least superficially) "free" to their users. "Free" online services are capable of appealing to a larger consumer base than normal because they are easily accessible and require no out-of-pocket costs. Consequently, by appealing to more people, advertisers' messages have a higher impact on a larger audience, and they can strategically target individual consumers within that audience. Jim Harper, "Why Online Tracking Isn't Bad," *The Wall Street Journal*, August 7, 2010, accessed November 9, 2010, http://online.wsj.com/article/SB10001424052748703748904575411530096840958.html.

For instance, on July 30, 2010, *The Wall Street Journal* published "The Web's New Gold Mine: Your Secrets," the first article of an ongoing, investigative series entitled "What They Know," which aimed to expose the hushed, complex practices and controversial implications of modern Internet-tracking technology to the American public. Written by Angwin, this article reveals how Internet-tracking companies have begun eroding Internet users' privacy. Angwin's article profiles how "a normal" Internet user navigates her way through the Web, performing humdrum activities such as watching online videos and taking quizzes. By performing such trivial activities, Angwin demonstrates how consumers tacitly and unknowingly invite hundreds of tracking companies to follow their online habits in order to turn a profit.[191] As Angwin states, "One of the fastest-growing businesses on the Internet… is the business of spying on Internet users."[192]

The mood conveyed in Angwin's piece is one of immediacy; the title of the series "What They Know" ignites a sense of eeriness that is mirrored in Angwin's article by her use of anecdotes and studies that highlight complicated and thorny privacy issues. Angwin explains technical issues in layman's terms, attempting to empower her readership to better understand the current landscape and not feel unknowingly victimized.

As the consumer tracking industry has become more effective and profitable, the use of its products has intensified. Angwin points out that in 2010 researchers found

---

[191] Chris Cox, Facebook's vice president of product, commented on the use of the information that Facebook collects from its over 800 million users. He indicated that the "information hoard" is both a blessing and a curse for the company. Cox stated, "The challenge of the information age is what to do with it," he said. Somini Sengupta and Evelyn Rusli, "Personal Data's Value? Facebook is Set to Find Out," *New York Times*, January 31, 2012, accessed January 31, 2012, http://www.nytimes.com/2012/02/01/technology/riding-personal-data-facebook-is-going-public.html?pagewanted=1&_r=1&emc=eta1; Angwin, "The Web's New Gold Mine: Your Secrets."
[192] Angwin, "The Web's New Gold Mine: Your Secrets."

tracking technology on 80% of 1,000 popular sites, up from 40% of those sites in 2005.[193] The consumer tracking industry produced more than $23 billion in ad spending just in 2009-2010. Featured in Angwin's article is a study conducted by the *Wall Street Journal*, testing the "fifty most popular U.S. websites, which account for about 40% of the websites viewed by Americans " to observe how the most widely visited websites employ tracking devices and programs.[194] The study found that on average, each website deployed about sixty-four pieces of tracking technology without notice onto the visitor's machine. Combined, the fifty websites installed "3,180 tracking files" on *The Journal's* test computer.[195] These "questionable files" were installed by over 131 distinct companies in the business of pervasive consumer tracking to create profitable databases. Similar to ads placed on television or in a newspaper, advertisers once paid for "dumb" ads to appear on different websites. Now advertisers pay a premium for ads that are perhaps too "smart" for their own good. The trackers that Angwin describes can reproduce themselves "even after users try to delete them."[196]

While the U.S. courts have legally approved the use of cookies, the simplest tracker, they have not yet considered the newer, more intrusive tools such as beacons or flash cookies (also known as "evercookies").[197] This has led to a series of court cases in 2010, where an increasing number of consumers are taking legal action against various

[193] Ibid.
[194] The top five websites that deploy the most number of trackers are the following: dictionary.com, merrium-webster.com, comcast.net, careerbuilder.com, and photobucket.com. Ibid.
[195] A deeper look into what kind of tracking files were deployed, divulged that approximately one-third of the files were harmless (i.e., just password-saving), while the other two-thirds of them (or 2,224 files) were more controversial. Ibid.
[196] Ibid.
[197] Ibid.

tracking companies.[198] Scott A. Kamber, a privacy and technology lawyer with Kamber Law, explains, "What these cases are about is the right of a computer users to dictate the terms by which their personal information is harvested and shared. This is all about user control."[199] Even one of the world's largest communications service groups, WPP, publicly conceded to a *Journal* reporter, "…[A]dvertisers must do better to inform customers about the tracking and mapping of online behavior."[200] Such articles may easily plant seeds of reservation in the minds of Internet users who were previously oblivious to such practices. While scientists and engineers continue to rapidly develop new technologies, Angwin's article demonstrates how new practices involving science and technology, if prematurely and widely adopted, may later cause confusion or fear especially when the implications of such practices are not fully analyzed or publicized.[201]

The Government's Entrance into the Data Economy

The online data industry, operating largely free from regulation, has distorted the traditional informational balance between individuals, private sector companies, and the government. The millions of Americans who access and contribute information over the Internet are largely unaware that their data is gathered by the tracking industry and often sold to the government. The informational landscape in the twenty-first century has empowered private companies – rather than the government – to amass copious amounts

---

[198] Wendy Davis, "The Forever Cookie: New Tracking Technologies Continue To Threaten Privacy," *MediaPost,* October 11, 2010, accessed November 19, 2010, http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=137459.
[199] Tazina Vega, "Code That Tracks Users' Browsing Prompts Lawsuits," *New York Times*, September 20, 2010, accessed November 19, 2010, http://www.nytimes.com/2010/09/21/technology/21cookie.html?pagewanted=1&_r=1.
[200] Angwin, "The Web's New Gold Mine: Your Secrets."
[201] It is worth noting that on September 27, 2011, *The Wall Street Journal* revised its own privacy policy; the policy now allows for the *Journal* to combine any user's personally identifiable information with his/her Web browsing data without consent, which had formerly been required. Julia Angwin, "Wall Street Journal Revises Its Privacy Policy," *The Wall Street Journal*, September 27, 2011, accessed October 8, 2011, http://blogs.wsj.com/digits/2011/09/27/wall-street-journal-revises-its-privacy-policy/.

of individuals' information gleaned from the Web. Because corporations routinely gather information both free from legal restraint and with an efficient use of resources, the government is playing "catch up," recalibrating its data collection practices in order to not fall behind. In an effort to hold onto its informational edge, the government has ignored existing legislation and entered the data economy as a major stakeholder. In effect, the government has gone rogue.

Corporations that collect information as their main source of revenue have grown hugely powerful – in some respects even more powerful than the government itself. This shift in informational power has triggered the government, which has held a longstanding political interest in learning about its constituents' interests, preferences, behaviors, and locations, as detailed in Chapter Three, to enter the Data Cycle. By employing digital technologies, Solove highlights, "Small details that were once captured in dim memories or fading scraps of paper are now preserved *forever* in the digital minds of computers, in vast databases with fertile fields of personal data…"[202] Private companies that profit solely from collecting and analyzing a deluge of individuals' personal information are operating free from legal safeguards, resulting in business relationships with the government and other public entities.[203]

A major issue surrounding the Data Cycle is that the growing amounts and types of information made available through the Internet are being collected from an

---

[202] Daniel Solove, *The Digital Person: Technology and Privacy in the Information Age* (New York: NYU Press, 2004), 1-2.

[203] Currently, there is no comprehensive U.S. law that protects consumers' privacy online. In fact, the Federal Trade Commission can only police online privacy intrusions *after* an entity has performed a "deceptive" and/or "unfair" practice . If the entity has breached one's privacy but not in a "deceptive" or "unfair" manner, the Federal Trade Commission is powerless. Julia Angwin, "Obama Administration Seeks Internet Privacy Protections, New Policy Office," *The Wall Street Journal*, November 11, 2010, accessed November 23, 2010, http://online.wsj.com/article/SB10001424052748703848204575608970171176014.html.

unsuspecting population base.[204] Although individuals share information online "publicly," most are unaware that the government may eventually access such information. Harvard Law Professors Lawrence Lessig and John Palfrey agree that individuals and experts alike are unlikely to know what information is being collected about them and how it is being used and/or re-used.[205] Interestingly, on October 25, 2010, the *Wall Street Journal* published another piece as part of the "What They Know" series, touching on this exact issue. The article, written by reporter Emily Steel, exposed the ways in which public and private entities currently interact with each other to obtain personal information from Americans.[206] Steel articulated how data tracking companies are increasingly being tapped by political campaigns to help target likely voters.[207] Specifically, Steel profiled how Republican Senator hopeful Jim Bender of New Hampshire employed an online tracking company, RapLeaf, to shower individual Internet users with targeted advertisements.[208]

---

[204] Having a significant, uninformed consumer base could easily become a colossal issue for businesses should a major problem arise. Akin to doing business with aliens who would have no understanding of the significance of a contract, modern Internet-based companies are serving a consumer base that similarly has no recollection of consenting to any sort of meaningful agreement between them and the company. From a business perspective, this issue could gravely threaten many businesses' existence and wholly affect future industry practices. Melissa Oppenheim, "From a Business Perspective: Terms of Service" (research paper, Independent Study on Social Media and Politics with Professor Nicco Mele, Harvard University, Cambridge, MA, April 2010).

[205] For instance, information that an individual volunteers on a private company's website may not necessarily be "private information," but is not intended to end up in the hands of the government either. Websites that possess large amounts of personal information such as Facebook are increasingly being cited as evidence in arrests and issuances of search warrants. See Palfrey, "The Public and the Private at the United States Border with Cyberspace," 243; CNN Wire Staff, "Brash Facebook Posts Lead to Bank Heist Arrests," *CNN,* April 22, 2011, accessed April 23, 2011, http://www.cnn.com/2011/CRIME/04/22/texas.heist.facebook/.

[206] The publishing of this article was bolstered by a flurry of *Wall Street Journal* "Technology" section articles that provided advice on how to evade data tracking companies' spying eyes. The articles further reported on the aftermath of both individuals' and politicians' reactions to this growing phenomenon.

[207] Emily Steel, "A Web Pioneer Profiles Users by Name," *The Wall Street Journal*, October 25, 2010, accessed October 25, 2010, http://online.wsj.com/article/SB10001424052702304410504575560243259416072.html?KEYWORDS=rapleaf.

[208] Ibid.

Americans tend to view corporations as entities that they do business with whereby personal information has become another commodity to trade for access to goods and services. However, individuals do not realize that when they regularly provide substantial amounts of personal information to various third party entities, including email hosting services, social networking sites, banking sites, and news media websites, they are actually exchanging their privacy for enhanced convenience. As Palfrey contends, "The tradeoffs involved are rarely conscious ones."[209]

Even though government entities cannot collect individuals' personal information directly, both economic and political pressures have incentivized the government to act as a data broker, selling information to and buying information from private companies. While some corporations may not sell their users' data to the government (regardless of the price) in order to (at least publicly) maintain trust with their consumers, corporations may be convinced to sell their data to the government, if the government agrees to keep the relations private.[210] Further, as discussed in Chapter Three, the American public has grown neurologically reliant on the very technologies that empower the Cycle. Thus, although individuals remain unaware of how their information is being used and re-used by various public and private parties, they continue to share information online because it is convenient and neurologically satisfying.

---

[209] Palfrey, "The Public and the Private at the United States Border with Cyberspace," 243.

[210] However, WikiLeaks, an anonymous whistleblower website, has demonstrated that keeping information "private" is becoming less of a tangible reality. To this point, Jim Harper from the CATO Institute argues, "Advancing technologies and the Internet will continue to…undercut secrecy… in all large organizations." Jim Harper, "It's a WikiLeaks World, Get Used to It," CATO Institute, August 5, 2010, accessed April 20, 2011,http://www.cato.org/pub_display.php?pub_id=12035&utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+CatoHomepageHeadlines+(Cato+Headlines).

The Privacy Act is Ignored and Sidestepped

In order to grasp how the Data Cycle has fully emerged in the first decade of the twenty-first century, it is crucial to consider the previously discussed technological advancements, informational changes, and economic developments, in conjunction with the existing obsolete legal infrastructure that has allowed for various interests to guide the technology's use. Branded as "toothless,"[211] the Privacy Act has collapsed under unwavering political and economic interests of those in power to employ digital technologies to acquire constituents' personal information in order to enhance their knowledge or influence.[212]

The Privacy Act of 1974, which was created when the government began implementing digital computing systems in the 1970s, was hastily passed in 1974 after the Watergate Scandal. However, nearly a decade after its enactment, the Act grew obsolete and has been successfully disregarded ever since.[213] Even in the late 1970s, the government continued to employ advancing information technologies discounting the Act's policies in order to increase administrative convenience and efficacy.

In the late 1960s and early 1970s a number of congressional hearings were held to discuss the implications of a created National Data Center, federal databanks, and automated personal data systems.[214] The Privacy Act, often colored as an attempt of "so-called Watergate reform," was initially created to provide an affirmative way for

---

[211] See Anna Kimbol, "The Privacy Act May be Toothless," *Health Law Perspectives* (September 2008), accessed December 2, 2011. http://www.law.uh.edu/healthlaw/perspectives/homepage.asp; Alex Kardon, "Damages Under the Privacy Act: Sovereign Immunity and a Call for Legislative Reform," *Harvard Journal of Law & Public Policy* 34, no. 2. (2011): 705-767.
[212] Waldo, Lin, and Millet, *Engaging Privacy,* 349-50.
[213] I asked Laura Quinn, the CEO of Catalist, what she and the company thought about the Privacy Act. She asked me if I could first explain the Act to her. A similar series of events occurred with a representative from Google's legal team. The fact that I had to explain what the Privacy Act was to two firms that are actively engaged in trading individuals' information with the government speaks to the notion that the law has been successfully ignored. Quinn, interview.
[214] Westin and Baker, *Databanks in a Free Society,* 101.

individuals to protect information about themselves from particular federal government uses.[215] Ripe for legislation, the political climate post-Watergate understood that the public held a strong distrust for the government's informational activities. Quickly passed under President Gerald Ford after Nixon's resignation, the Act sought to empower individuals to learn which federal agencies were keeping files on them, review such records, make corrections, and inquire how they were being used.[216] The Act, at the time, aimed to limit the government's potential involvement in inappropriately and/or illegally using advanced computing technology to access and store personal information.

However, shortly after the passage of the Privacy Act, the Act received little attention and was barely enforced, let alone implemented. The Office of Management and Budget ("OMB") which was appointed to oversee the Act's implementation, typically created annual reports for the President on the "Implementation of the Privacy Act." As documented by the OTA's 1986 report, entitled "Federal Government Information Technology: Electronic Record Systems and Individual Privacy" ("the report"), the OMB's presidential reports were well executed during the years directly following the Act's passage, 1975-1978.[217] However, in 1981 and 1982, the annual reports were poorly conducted.[218] In 1983 the OMB testified that it had not taken much initiative to oversee how the Act was being adhered to in practice.[219] According to the OTA's 1986 report, each federal agency was supposed to have selected a representative to deal with Privacy Act matters. In many cases, this official was either found to have many responsibilities or

---

[215] Harold Relyea, "The Privacy Act: Emerging Issues and Related Legislation," *Congressional Research Service and the Library of Congress' Report for Congress* (February 26, 2002), 2.
[216] Ibid.
[217] U.S. Congress, Office of Technology Assessment, "Federal Government Information Technology," 117.
[218] Ibid.
[219] In particular, in 1980 other pieces of legislation such as the Paperwork Reduction Act, weakened the OMB's role in overseeing how the federal agencies were implementing the Privacy Act. Ibid.

did not exist.[220] During the twenty years following the passage of the Privacy Act, the Act was left without adequate support, funding, or staff to be meaningfully executed.

Further, the OTA's 1986 report to Congress, which focuses on the federal agencies' digitization and storage of individuals' personal information, offers insight into the Act's implementation during the 1980s. The report highlights how in 1985, the number of computerized records for agency use had significantly increased.[221] The report details the results of a survey that asked agencies to disclose their largest Privacy Act record systems in order to gauge the computerization penetration rates of Americans' records in 1985. Of the responding thirteen agencies and their corresponding 3.5 billion records, 60% were either totally or partially computerized while 40% were still wholly manual. However, of the large systems of records (i.e., over 500,000 persons) 78% were either totally computerized or partially computerized while 22% were wholly manual.[222]

The computerization of governmental records enhanced governmental agencies' efficiency in terms of detecting fraud and abuse and decreasing wasted time and energy, but also made it "nearly impossible" for individuals to use the remedies available through the Act.[223] Since the time of the Privacy Act's creation and enactment, new information technologies have emerged that are clearly not addressed by the Act. Although the Act has been amended six times, instead of being meaningfully updated or interpreted more broadly, the Act has been left outdated and irrelevant.[224]

For instance, as early as 1986 real concern was expressed by the OTA that the holes in the 1974 Act were allowing "agencies to collect, use, store, exchange, and

---

[220] A survey conducted by the OTA exposed that 67 of 100 responding agencies admitted that they had "one full time equivalent staff person *or less* assigned to Privacy Act matters." Ibid.
[221] Ibid., 22.
[222] Ibid.
[223] Ibid., 12.
[224] Relyea, "The Privacy Act: Emerging Issues and Related Legislation," 5.

manipulate individual records, as well as entire record systems, in electronic form."[225] In 2011, Alex Kardon, graduate of Yale Law School, wrote a white paper discussing the potential damages under the Privacy Act in the twenty-first century. Kardon argues that the government's possible mistreatment of personal information "is a real harm."[226] He observes, "[Today, I]n the Internet age… information spreads more quickly and reaches the wrong hands more swiftly than the Privacy Act's drafters could possibly have imagined."[227]

The Act's largest loophole concerning its "routine use" clause still exists.[228] The Act's "routine use" clause indicates that agencies must state how they will make of use of individuals' collected personally identifiable information in order to curtail the discretion of agency officials to use and/or re-use such information in any way they deem appropriate.[229] Further, such information is not to be shared with other agencies.[230] The Privacy Act requires each agency engaging in the collection and maintenance of personal information to publish how the agency plans to use such information in the Federal Register each year. However, the majority of individuals are oblivious to these notices and their rights.

The Congressional Research Service ("CRS"), which conducted an updated study for Congress on the Privacy Act's implications in 2002, quoted a policy analyst who

---

[225] U.S. Congress, Office of Technology Assessment, "Federal Government Information Technology," 12.
[226] Kardon argues that individuals should be allowed to claim damages related to harm caused by the mistreatment of personal information such as identity theft and widespread reputational harm; as the Privacy Act exists now, such harms are non-pecuniary and thus, not interpreted as actual damages. See Kardon, "Damages Under the Privacy Act," 767.
[227] Ibid., 715-767.
[228] "(a)(7)…[T]he term 'routine use' means, with respect to the disclosure of a record, the use of such record for a purpose which is compatible with the purpose for which it was collected..." See "U.S. Privacy Act of 1974 (5 U.S.C. § 552a as amended)," *The U.S. Justice Department*, accessed November 2, 2011, http://www.justice.gov/opcl/privstat.htm.
[229] Relyea, "The Privacy Act: Emerging Issues and Related Legislation," 9.
[230] Ibid.

argued, "[A]gency officials have interpreted the routine use clause broadly and have created almost unlimited ability to move data among federal agencies."[231] Further, the CRS report listed seven emerging issues with the Privacy Act to be considered by Congress. One of these issues called for "better enforcement or overhaul."[232] Another proposal made by the CRS called for Congress to expand the Act to manage federal agencies' deployment of browser cookies, a twenty-first century, commonly used tracking technology. The report reveals that in 2000, twenty-three of seventy federal agencies had disclosed the personal information of individuals visiting their websites to third parties, including other federal agencies, banks, retailers, distributors, product manufacturers, and trade organizations.[233] The OMB responded to these findings in 2000 by issuing a memorandum to every federal agency stating that browser cookies should not be used on Federal websites (or by their contractors managing such websites) unless certain consent and conditions are met.[234] Four months later a follow-up study was conducted that revealed at least thirteen federal agencies had ignored the memorandum and were still tracking their websites' visitors.[235]

As technology has continued to advance since the late 1970s, when the Privacy Act was enacted and had already begun to prove itself weak, the government has been able to repurpose Americans' information without much objection. The law has not been updated to address technological implications since the 1970s with respect to the

[231] Ibid.
[232] Ibid., 5.
[233] Ibid.
[234] The OMB included in the memorandum that "persistent cookies" or a type of tracker that remains on a visitor's machine for long periods of time should have been banned altogether. Ibid.
[235] Further, in 2001 another study was conducted and found sixteen agencies still employing persistent cookies on their websites, which were strictly prohibited in the OMB's memorandum. Ibid.

government's use or misuse of individuals' personal information.[236] Thus, the Privacy Act provides the government with the necessary freedom to engage in the Data Cycle.

A Shift in Information Supremacy

Once referred to as the "nation's chief generator of knowledge in just about every field," the federal government no longer holds this title.[237] Although the federal government used to be the only entity that could afford the physical collection of data and garner wide participation of individuals (by instating a direct mandate), private data companies in the twenty-first century collect much more information than the government ever has.[238] In addition to the government's participation in the Data Cycle, the government has enacted legislation such as the Patriot Act, the Foreign Intelligence Surveillance Act ("FISA") and the Electronic Communications Privacy Act ("ECPA") to enhance its data gathering power. Further, by technically de-identifying its datasets, the government has broadened its ability to access and repurpose individuals' information. Through these measures, the U.S. government aims to ensure that it possesses the freedom to acquire comparable amounts of information as the private sector.

Legislatively, after September 11, 2001 ("9/11") the ways in which the government understood its need to access information on its constituents dramatically changed and set the precedent for how the government would acquire information on individuals for the first decade of the twenty-first century.[239] Congress enacted the Patriot

---

[236] Kardon, "Damages Under the Privacy Act," 767.
[237] Smith, "The Commercialization and Privatization," 47.
[238] Ibid.
[239] A survey conducted right after 9/11 by Harris Interactive & Alan Westin evidenced very high public approval ratings of expanded governmental investigative powers (i.e., "93% approved expanded under-cover activities in suspected groups…"). However, the same survey recorded very high majority percentages of the public's concern for the way such expanded powers may be used in practice (i.e., "Judges will not look closely enough at the justifications for surveillance (79%), Congress will not include adequate safeguards in its authorizations (78%),… communications of innocent people will be checked (72%), …and

Act shortly after 9/11 to increase the federal government's ability to collect more information from a greater number of sources than had previously been allowed in criminal or foreign intelligence investigations.[240] The government believes that such expanded authority to collect personal information is necessary in order to prevent future terrorist attacks and respond to the new technologies and signals of modern day threats. Additionally, under the amendments to FISA and ECPA, the federal government gained easier access to new powers to collect intelligence information from foreigners, intercept one's e-mail and telephone conversations, search homes and businesses, and infiltrate one's spending and communication patterns.[241]

The federal government is not just acquiring information through its own means as permitted by these expanded laws. Instead, as the Data Cycle illustrates, the government enters relationships with private sector entities to acquire information that the government is legally restricted from directly collecting. Research conducted by Chris Hoofnagle demonstrates how the shift from the analog to digital world has altered the ways in which the federal government collects and handles information on individuals. "Through just visiting a single website," Hoofnagle indicates, "law enforcement can obtain a comprehensive dossier on almost any individual… custom-tailored…" by a commercial data broker ("CDB").[242] The issue of interest is not necessarily that the

---

new surveillance powers will be used to investigate crimes other than terrorism (67%)"). Alan Westin, "Social and Political Dimensions," *Journal of Social Issues* 59, no. 2 (2003): 27-28.

[240] Anna Henning, "Government Collection of Private Information: Background and Issues Related to the USA PATRIOT Act Reauthorization," *Congressional Research Service*, March 2, 2010.

[241] The CRS Report to Congress on the Patriot Act elaborated, "In some cases, evidentiary standards required to obtain court approval for the collection of information were lowered. Other approaches included expanding the scope of information subject to search, adding flexibility to the methods by which information could be collected, and broadening the purposes for which information may be sought." Ibid., 2-9.

[242] "ChoicePoint, Inc. is a company based in Alpharetta, Georgia, that concentrates on selling information and data services to insurers, businesses, government, and direct marketers." In 2004, ChoicePoint amassed

government has access to a CDB's database, but rather that there are hardly any meaningful rules guiding the use and re-use of such information.[243] Increasingly, information that the federal government legally cannot collect is collected by CDBs. Because the data originated from the CDB, the Privacy Act is not triggered to guide the federal government's or the CDB's use of such information.[244]

Although Congress passed the Privacy Act in 1974 with the intent to place limitations on the Executive branch and federal agencies' ability to handle Americans' personal information, the government has rendered these protections irrelevant in the twenty-first century. Alex Kardon claims, "The world has not changed so much since the days of Watergate that the government should once again be given free rein to use personal information with no fear of repercussions..."[245] Similarly, Palfrey has commented, "[There are no] constitutional protections for the re-use of privately collected data by state actors."[246] The private sector has grown to serve as a middleman, allowing the federal government to circumvent the Privacy Act's intentions and mandates. The law has been left, perhaps intentionally, outdated, allowing the government to participate in the Data Cycle.

Additionally, the government takes technical measures to sidestep the Privacy Act and repurpose individuals' information. In regards to Data.gov, the federal government incentivizes the re-using and repurposing of Americans' information in aggregated, anonymized form. However, in response to how Data.gov deals with complying with the

---

revenues of over $795,700,000, largely from government contracts. See Hoofnagle, "Big Brother's Little Helpers."

[243] Ibid.

[244] However, if the information originates from the federal government and is shared with a CDB, the Privacy Act is applicable. Ibid.

[245] Kardon, "Damages Under the Privacy Act," 767.

[246] Palfrey, "The Public and the Private at the United States Border with Cyberspace," 242.

Privacy Act, Vivek Kundra, who spearheaded the Data.gov initiative while serving as President Obama's Chief Information Officer (2009-2011), responded, "We de-identify the data in order to comply with the Privacy Act…"[247] Kundra continued, "There are actually a number of ways to get around 'privacy issues.'"[248]

However, Carnegie Mellon University Computer Scientist Latanya Sweeney has evidenced that one only needs access to a few characteristics about individuals in order to uniquely identify "anonymized" data points. For example, Sweeney conducted a study using American individuals' 5-digit zip codes, birth dates, and genders and demonstrated how this combination of data can effectively de-anonymize previously anonymized datasets.[249] Sweeney found that although individuals often do not perceive these three pieces of information to be particularly privacy-invasive, and therefore readily provide them (and often, provide even more identifying pieces of information to all types of third parties online) the mere three-piece combination of zip codes, birth dates, and genders is actually uniquely identifying for approximately 87% of the entire U.S. population.[250] Sweeney illustrated how data that is often believed to de-identified, such as some datasets made available on Data.gov, is re-identifiable.[251] The re-identification process is easier

---

[247] In a roundtable discussion with Vivek Kundra about the limitations of public services versus private ones, Kundra offered, "The government does not have the luxury of trading individuals' privacy for more convenience [like private companies do]." Kundra is currently working on a white paper at Harvard's Kennedy School of Government that proposes a private sector-like incubator program for the public sector in order to encourage the development of innovative applications and services that access government databases. Vivek Kundra (Former Chief Information Officer in the Obama Administration), questions asked by Melissa Oppenheim at a roundtable discussion at Harvard University's Institute of Politics, November 29, 2011, 6pm EST, Harvard University, Boston, MA, transcript in *Appendix B*, "Oral History Interviews."

[248] Ibid.

[249] By using the 1990 U.S. Census summary data, which is publicly available online, Sweeney determined "how many individuals within geographically situated populations had combinations of demographic values that occurred infrequently." See Latanya Sweeney, "Simple Demographics Often Identify People Uniquely," (Pittsburgh: Carnegie Mellon University, 2000).

[250] 87% of the U.S. population reflects approximately 216 million people out of 248 million people. Ibid.

[251] Sweeney demonstrated how information across datasets can be linked and easily re-identified. By purchasing the voter registration list for Cambridge, MA, for $20 Sweeney associated this information with

than one may think. Sweeney offers, "The computational and informational ability to re-identify data is no longer limited to directly linking… a single dataset or using demographics alone… [A]s time has passed, more data has become available about individuals, providing more ways to re-identify de-identified data."[252]

For instance, in addition to one's demographic information, other unique identifiers may stem from the increasing amounts of information individuals share online, such as one's search terms, purchase habits, browsing preferences, and the structure of one's social network.[253] The enhanced ease in identifying a previously de-identified dataset is important insofar as the federal government continues to operate under the premise that anonymized datasets are forever anonymous and are therefore publishable and repurposable.

Prior to the existence of the Internet, individuals relied much more heavily on communicating with one another through the U.S. Postal Office. However, in order to mail a letter, each individual had to pay for the amount of postage necessary to transport his/her message. In return, the government blindly facilitated the carrying of letters. Online, the transportation of mail is free. Thus, it may seem apropos that individuals pay a different price. However, what this price *should* be remains questionable.

---

publicly available medical, anonymized data such as hospital discharge data (forty-four of fifty states collect this information) and ambulatory care data (seventeen out of fifty states collect this information). "Many of these states have subsequently… sold copies to industry and made versions publicly available..." Ibid.

[252] Sweeney linked individuals' voter information with medical information by using ZIP code, birth date and gender, pieces of information that both datasets provided. Once linked, Sweeney was able to uniquely identify individuals to match their medical diagnoses, procedures, medication information, ethnicity, visit dates, and total charges. Latanya Sweeney, "Patient Privacy Risks in U.S. Supreme Court Case Sorrell v. IMS Health Inc.: Response to Amici Brief of El Emam and Yakowitz," Data Privacy Lab Working Paper 1027-1015B, Cambridge, 2011.

[253] Seth Schoen, "What Information is 'Personally Identifiable'?" *The Electronic Frontier Foundation*, September 11, 2009, accessed November 5, 2011, https://www.eff.org/deeplinks/2009/09/what-information-personally-identifiable

Discussion

The government's participation in the Data Cycle and its continual repurposement of individuals' data exemplifies how the government has seemingly gone rogue in the twenty-first century. As argued in this chapter, the government entered the Data Cycle due to an entwined interaction of trends. From the American public's adoption and use of the modern computer and Internet, new types of private companies emerged, creating the data economy. As private companies have grown experienced and powerful in gathering individuals' data, a major shift in informational power between the public and private sectors has occurred. This shift has motivated the government to circumvent existing legislation and hire private companies to access data that the government cannot. This chapter proffered that major technological advancements since the mid-twentieth century, the creation of the data economy, and an outdated legal infrastructure, encouraged and permitted the government to participate in the Cycle.

Why are Americans not aware of this phenomenon? And if they are, why are they not doing anything to change it? In the following chapter, I discuss how privacy expectations in America have remained an evolving notion shaped by sociopolitical, economic, and legal factors. I argue that individuals continue to share information with private companies, mistaking them for wholly separate entities from the government. I suggest that this misunderstanding is further clouded by evidence demonstrating humans' growing neurological dependence on their wired devices. I offer how such trends have influenced Americans' acceptable notions of privacy and contributed to the Data Cycle's proliferation.

CHAPTER THREE

Where is the Public's Concern for Privacy?

Historically, the legal relationship in the United States between privacy expectations and actual privacy protections has been complicated and confusing.[254] Perhaps better viewed as a moving target, "privacy" has meant different things to Americans throughout time, as even our most widely used dictionaries offer varying definitions of the term.[255] Early treatises on privacy in the United States, such as Boston lawyers Samuel Warren and Louis Brandeis' seminal piece, "The Right to Privacy," date back to the 1890s when Americans began to question the encroachment of privacy in relation to print newspapers and instantaneous photographs.[256] Such technological innovations allowed individuals to propagate gossip in a more concrete way than ever previously experienced. Many individuals felt that the adoption of such technologies had invaded their boundaries of personal privacy.[257]

Since its founding, the U.S. government has always held an interest in learning information about its constituents. This chapter traces the impetuses driving the American public's participation in the Data Cycle. Why do Americans (knowingly or unknowingly) contribute to the Data Cycle? How are their contributions affecting their own privacy? And, if individuals knew the consequences of their actions online, would they still knowingly act the same?

---

[254] Holtzman discusses that most Americans think privacy is a right. He continues, "A Gallup Poll conducted in February 1999 found that 70% of respondents believed that the Constitution guaranteed citizens the right to privacy… [but] the word privacy doesn't event appear in the Constitution – not once." Holtzman, *Privacy Lost,* 93.
[255] Ibid., 1.
[256] Solove, *The Digital Person*, 57.
[257] Brandeis called privacy, "the right most valued by civilized men," exemplifying how Americans have historically been deeply concerned with protecting their right to privacy. Jon Mills, *Privacy The Lost Right* (New York: Oxford University Press, 2008), 25.

This chapter specifically investigates why the American public's historically held concern for privacy has not come to bear on its actions as reflected in the Data Cycle in the twenty-first century. I will refer to "privacy" as one's ability to maintain personal information control. I argue that although individuals' privacy is declining in light of the Cycle, individuals are nevertheless fueling the Cycle's existence. In light of Chapter Two's discussion of rapidly advancing and subsequently adopted technological innovations, I examine how notions of individuals' privacy since the mid-twentieth century in the U.S. have continuously waxed and waned. I contrast the public's concern for privacy in the mid-twentieth century with those currently held in the face of the Data Cycle. By discussing how political interests have encouraged the government to obtain Americans' personal information, I suggest that political interests in the twenty-first century are able to achieve this goal due to both social and neurological factors. Socially, I propose that the structure of the Data Cycle supports individuals' participation in that Cycle. Neurologically, recent scientific studies have evidenced that humans are growing addicted to remaining continuously connected.

Throughout this chapter, I argue that the notion of privacy in America is alive but at a crossroads. While individuals claim they are interested in maintaining personal privacy in the Information Age, their actions, influenced by ignorance of the Data Cycle and/or developing neurological addictions, speak differently.[258] Nevertheless, as humans

[258] A number of studies conducted in the first decade of the twenty-first century reveals that Americans are concerned about their privacy online. For instance, 91% of Americans who participated in a Zogby Interactive survey in 2007 said they were concerned that their identities might be stolen and used to make unauthorized purchases. Further, in March 2000 BusinessWeek/Harris Poll found that 89% of its poll's participants were "uncomfortable with Web tracking schemes where data was combined with an individual's identity." Lastly, in April 2001 a study conducted by the American Society of Newspaper Editors found that 81% of its participants were either "very concerned" or "somewhat concerned" that a company might violate their personal privacy. "EPIC: Public Opinion on Privacy," Electronic Privacy Information Center, accessed February 2, 2012, http://epic.org/privacy/survey/.

grow increasingly comfortable with sharing information with private companies and dependent on the very services that collect our personal information, individuals are not only empowering the Data Cycle, but are also tacitly eroding their own understood notion of privacy.

Privacy Concerns in America Since the Mid-Twentieth Century

In the twenty-first century, articles published by news organizations such as *The Wall Street Journal* regularly invoke alarm by exposing new types of privacy intrusions online. In fact, over the first decade of the twenty-first century, disturbing headlines such as *The Wall Street Journal's* "Sites Feed Personal Details to New Tracking Industry,"[259] and *The New York Times'* "Little Brother is Watching"[260] have become increasingly commonplace.

Such articles though rarely place today's alleged alarm in a larger historical context, giving the impression to readers that today's concerns are novel. Harvard Law Professor Lawrence Lessig recognizes that individual privacy has become a major theme in modern legal history, yet posits, "It is not as if we have never faced this question before."[261] Privacy experts highlight that privacy debates *must* be considered within historical, technological contexts.

In agreement with Lessig, contemporary authors Jon Mills and Daniel Solove emphasize how privacy has been part of an ongoing societal discussion for hundreds of years. Mills and Solove detail how the patterned emergence of new, "disruptive"

---

[259] Julia Angwin and Tom McGinty, "Sites Feed Personal Details to New Tracking Industry," *The Wall Street Journal*, July 30, 2010, accessed November 2, 2011, http://online.wsj.com/article/SB10001424052748703977004575393173432219064.html.
[260] Walter Kirn, "Little Brother is Watching," *The New York Times*, October 15, 2010, accessed November 3, 2011, http://www.nytimes.com/2010/10/17/magazine/17FOB-WWLN-t.html.
[261] Lawrence Lessig, "The Architecture of Privacy *Draft 2*," Taiwan Net Conference (1998), 13, accessed November 4, 2010. http://cyber.law.harvard.edu/works/lessig/architecture_priv.pdf.

technologies, such as the computer in the 1940s and new information technologies and record keeping systems in the 1960s and 1970s, have continuously required Americans to reevaluate their perceived right to privacy.[262] In particular, throughout the second half of the twentieth century, various sectors of American society voiced fears surrounding the waning of individuals' privacy, as the government adopted new uses of technological innovations such as computerized databases and connectivity. Exploring such conversations and cultural movements provide us with a useful context from which we may compare the privacy concerns held by the American public regarding the Data Cycle.

Privacy at the Center of the 1960s Counterculture

Columbia University Professor and privacy scholar Dr. Alan Westin posits that by the end of World War II the public had placed a high level of trust in government, businesses, and the non-profit sector, and therefore "the information collection and use activities of those organizations."[263]

However, the public's level of trust soon changed.

From 1945 to 1960 the government engaged in many personal privacy intrusions, such as the government's loyalty-security programs, blacklists, and McCarthy-era allegations. Throughout the 1950s and 1960s, then F.B.I. Director J. Edgar Hoover infamously executed his secret investigations on "suspicious" residents and public figures throughout the U.S. Dr. Westin argues though that government surveillance did not

[262] Mills, *Privacy The Lost Right*, 25.
[263] Westin, "Social and Political Dimensions," 8.

actually reach peak levels until the 1960s and 1970s, "when hundreds of organizations and hundreds of thousands if not millions of individuals were investigated."[264]

The writing of contemporary authors Stephen Goode and Vance Packard demonstrate Americans' strong privacy concerns in the 1960s, 1970s, and throughout the 1980s due to the rise of computerization and the digitization of individuals' information. In particular, Packard cautioned in his 1964 novel, *The Naked Society*, that privacy was under attack due to the government's increasing computerization and ability to collect and store data in newly enhanced ways.[265] By exploring contemporary written works, one finds that an "anti-technocracy" counterculture movement began to mobilize during the 1960s.

Amidst a general climate of anxiety and social unrest in America in the 1960s, as epitomized by the anti-Vietnam War movement, the civil rights movement, the feminist movement, the sexual revolution, and student protests and demonstrations, little information was really known about databanks or how those in charge used them, adding uneasiness to an already rocky social environment.[266] The anti-technocratic movement held that society was employing technology too robustly and hastily, producing an increasingly conformist, militaristic, and authoritarian society. The movement perceived technology as the center enabler of all of these problems.

Throughout the 1960s and 1970s, Dr. Westin details how information privacy emerged as a major issue.[267] As advances in physical, psychological, and data surveillance technologies were made and embraced by government and the private sector in the early

---

[264] Philippa Strum, Gerald Nash, and Richard Etulain, *Privacy, the Debate in the United States since 1945* (Fort Worth: Harcourt Brace College, 1998), 149.
[265] Vance Packard, *The Naked Society* (New York: D. McKay, 1964), 13-30.
[266] Westin and Baker, *Databanks in a Free Society,* xiv.
[267] Ibid., 8.

1960s, all levels of government began implementing such projects.[268] Although most major media outlets in the early 1960s were excited by the improvements that these technologies and projects promised, popular commentators, such as Packard and Westin, along with government officials, especially those who worked for the OTA, began to warn Congress of the potentially dark implications of many of these technological changes.[269]

Notably, Mario Savio, student and member of the Free Speech Movement, gave a speech at U.C. Berkeley in 1964 in front of more than 5,000 students, voicing many of the concerns held by the anti-technocratic movement. Savio referred to established institutions as "machines." He proclaimed, "There's a time where the operation of the machine [the institution] becomes so odious... [Y]ou've got to put your bodies upon the gears and upon the wheels, upon the levers, upon all the apparatus, and you've got to make it stop..."[270] Savio likened that U.C. Berkeley had begun to treat its students as nothing more than "IBM card[s,]…subjected to all the techniques of factory methods: tight scheduling, speedups, rules of conduct..."[271]

Further, in author Fred Turner's book, *From Counterculture to Cyberculture*, he details how during the 1960s the counterculture viewed "the corporate world, the university, the military and the punch-card universe of information … [as] mirrors of one another."[272] Turner offers, "Each presented the otherwise whole and authentic individual with a world in which he or she must pare away some part of his or her self in order to participate... threaten[ing] to alienate the individual from her or his own lived

---

[268] Ibid., 9.
[269] Westin and Baker, *Databanks in a Free Society*, 9.
[270] Savio continued, "And you've got to indicate to the people who run it… that unless you're free, the machine will be prevented from working at all…" See Fred Turner, *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism* (Chicago: University of Chicago Press, 2008), 11.
[271] Ibid., 12.
[272] Ibid.

experience."[273] Additionally, in 1969, Theodore Roszak published the nonfiction work, *The Making of a Counter Culture*, describing how the counterculture of the 1960s blamed society's technocratic, dehumanizing mindset for the disasters of Vietnam and the military-industrial complex. Roszak explains how a part of society viewed science and technology – and especially the computer – as the symbol of bombs, the arms race, and deterrence occurring at the time.[274] Thus, as exemplified through nonfiction works, speeches, and demonstrations, the 1960s counterculture passionately promoted that technology could be antiseptically dangerous if not carefully and strategically employed.

In addition to the countercultures' skeptical view of technological advancements in general during the 1960s, Dr. Westin detailed how wide public anxiety existed over the extent to which various organizations were particularly sharing, using, and re-using electronic records. Westin called this "the databank issue."[275]

Although computers were not widely distributed throughout the federal government in the 1960s, Congress and the executive branch began to examine the effects of government information collection on individuals' privacy and agencies' accountability during this time.[276] Such reports were conducted regularly due to an

---

[273] Ibid.

[274] In the 1960s, other types of less pessimistic, countercultural movements concerning technology existed as well. In 1968, Stuart Brand published the Whole Earth Catalog, a collection of items including books, outdoor supplies, and artifacts that could be useful for the anti-technocratic countercultural revolution. Around the same time of the Catalog's publishing, a subset of the counterculture movement became interested in how computers could be used as a liberating technology. Throughout the 1960s and 1970s, MIT's Tech Model Railroad Club and the Homebrew Computer Club played a major role in propelling the ethos of tinkering. Both of these groups served as loci for computer enthusiasts to gather, discuss how electronics worked, and how computers could be used for productive purposes. Ibid.

[275] Westin and Baker detail, "By the late 1960s, the 'databank issue' had become one of the most widely discussed and emotionally laden civil-liberties questions facing American society." See Westin and Baker, *Databanks in a Free Society*, xiv.

[276] Federal agencies only began gradually adopting mainframe computers for information collection and maintenance during the 1960s. U.S. Congress, Office of Technology Assessment, "Federal Government Information Technology," 22.

"explosion in information activities" catalyzed by "Great Society programs."[277] Additionally, in 1966 and 1967, Congress held two large hearings on using mainframe computers to create a National Data Center.[278] Throughout the early 1970s various Congressional committees held hearings debating the possibilities and/or implications of creating and pursuing federal databank projects.[279] Many individuals objected to such projects, as they feared that those in power could easily abuse their access to sensitive information.

To respond to this growing anxiety, in 1972 Dr. Westin and Michael Baker, then a PhD student, published the first nation-wide, government-sponsored study on computer databanks. The study sought to reveal how organizations and governmental entities were actually using their record-keeping systems.[280] Westin and Baker concluded that by 1972, "The United States ha[d] become a records-oriented society… [whereby,] [i]n each major area of personal and civic life formal, cumulative records are assembled about each of us by many private and government recordkeeping organizations."[281] They concluded, "It is feared that far more personal data might be assembled about the individual than it had been feasible to collect before; that much greater sharing of confidential information might take place among the computerized record holders; and that there might be a

[277] After the assassination of President JFK, President Johnson enacted many programs under what he called the "Great Society" to combat poverty, tackle racial injustice, and create "model cities." Johnson signed into law the Social Security Act on 1965, which added health care programs to the Social Security Act from 1935 (originally signed into law by President Roosevelt). Specifically, in 1965, Medicare and Medicaid were created. Ibid., 22, 100.
[278] Ibid., 100.
[279] Ibid., 101.
[280] Westin and Baker, *Databanks in a Free Society,* xiv.
[281] Ibid., 25.

lessening of the individual's ability to know what records have been created about him, and to challenge their accuracy or completeness."[282]

Dr. Westin described how the computer in the 1960s and 1970s was actually viewed by parts of the American public as an instrument for continuing discrimination. Westin details, "The common element among such groups [blacks, other non-whites, women, and political dissidents] is their claim that they should have an equal competitive opportunity with others in the award of credit, employment, housing, education, licensing, and government benefits."[283] He explains that these groups viewed computerization as a means to accelerate discrimination since "computerized data systems usually involved more systemic collection, more extensive recording, more centralization, and easier dissemination."[284] Additionally, there existed a growing skepticism about the government's "data-based," social programs. For instance, the government frequently asked individuals to disclose personal information under the premise of assisting the creation of certain social programs – such was the case with the computer records kept on a Washington D.C. slum clearance project.[285] Despite the fact that such programs often fell flat, the government still kept the collected personal information.[286]

---

[282] Ibid.

[283] Alan Westin, *Computers, Health Records, and Citizen Rights* (Washington, DC: U.S. Government Printing Office, 1976), 212.

[284] Lori Andrews, Law Professor at Chicago-Kent Law School, wrote an article for the *New York Times* in February 2012, referring to the fear held in the 1960s and 1970s about discrimination drawn from data analysis. In particular, she highlights Northwestern Professor of Communication Studies John McKnight who described how in the 1970s banks and insurance companies often would not invest in inner-city neighborhoods or provide loans to individuals based on race or ethnicity (even if they were well-off). In 2012, Andrews contends that similar practices are occurring, whereby judgments are being made based off of individuals' online activity. She offers, "Institutions are denying people opportunities based on their digital selves." For full discussion, see Lori Andrews, "Facebook Is Using You." *The New York Times*, February 4, 2012, accessed February 10, 2012, http://www.nytimes.com/2012/02/05/opinion/sunday/facebook-is-using-you.html?_r=1&ref=opinion&pagewanted=all; Westin, *Computers, Health Records, and Citizen Rights,* 213.

[285] Westin, *Computers, Health Records, and Citizen Rights,* 213-214.

[286] Ibid.

An important outcome of these studies appeared in the late 1970s, when Dr. Westin produced a monograph for the U.S. National Bureau of Standards describing how popular conceptions of "the computer" held by Americans in the 1960s and early 1970s were actually quite misinformed.[287] Even in the mid-1960s when the government began using computers to automate personal data files, Westin believed that the majority of individuals remained "fundamentally ignorant of what computers did and how."[288] Westin clarified that computers were not the "unforgiving, unforgetting" automatons generally described by anti-computer enthusiasts in the 1960s.[289]

Watergate and the Creation of the Privacy Act

Nevertheless, in the early 1970s the public's concerns regarding the government and private sectors' computerization of records continued to mount. The federal government began to take regulatory action to assuage some of the public's worries.[290] For example, in 1970 Congress passed the Fair Credit Reporting Act to respond to the consumer reporting industry's computerization of its 110 million consumer files.[291] Additionally, in the 1970s the National Academy of Sciences and an Advisory Committee to the U.S. Department of Health, Education, and Welfare conducted studies investigating changing privacy patterns; in 1973 Congress developed a Fair Information Practice framework to help guide information privacy standards.[292]

---

[287] Ibid., 217.
[288] Ibid.
[289] Ibid.
[290] Then Director of the Institute for Computer Science and Technology Dr. Ruth M. Davis detailed, "The use of computers to automate the information handling and record-keeping activities of government and private organizations has brought… concerns for privacy stemming from the desire of individuals to control the collection of information about themselves and to exercise some measure of control over the use of that information." Ruth M. Davis, "Foreword" in Westin, *Computers, Health Records, and Citizen Rights,* iii.
[291] However, the Act did not really set privacy standards for consumer reporting. Westin, "Social and Political Dimensions," 10.
[292] Westin, *Computers, Health Records, and Citizen Rights,* 10.

Although Congress had begun taking action, privacy intrusions continued to take place throughout the 1970s. Many federal agencies in the 1970s employed "computer matching" or the comparison of machine-readable records containing individuals' used to detect "cases of interest."[293] Towards the latter half of the decade more federal agencies implemented computer systems, and instances of computer matching greatly increased. Serving as the last straw, the Watergate Scandal embarrassingly unfurled on the world's stage in 1974. The scandal frustrated Americans who began demanding that the opaque doors, behind which the government operated, be made more transparent. President Nixon, who resigned on August 9, 1974, epitomized privacy advocates' worst fears. Specifically, Watergate demonstrated how top government officials could easily abuse their informational power. In 1974, feared situations once only dreamt by the counterculture movement of the 1960s suddenly became a harsh reality.[294] Thus, at a time of high political pressure, the federal government passed the Privacy Act in December of 1974.[295]

Although the Privacy Act sought to ameliorate Americans' privacy concerns by establishing "rights and remedies for individuals who are the subjects of agency recordkeeping," the Act soon became but a "cute," distant idea.[296] Just three years later, in 1977, the Privacy Protection Study Commission warned Congress of the "gradual erosion of individual liberties through the automation, integration, and interconnection of

---

[293] U.S. Congress, Office of Technology Assessment, "Federal Government Information Technology,"101.
[294] Westin, *Computers, Health Records, and Citizen Rights,* 215.
[295] In the 1970s, Congress passed other privacy legislation including the Family and Education Rights in 1974 and the Right to Financial Privacy Act of 1978. Additionally, the Privacy Act of 1974 created the U.S. Privacy Protection Study Commission. Westin, "Social and Political Dimensions,"10.
[296] U.S. Congress, Office of Technology Assessment, "Federal Government Information Technology," 101.

many small, separate recordkeeping systems, each of which alone may seem innocuous, even benevolent, and wholly justifiable."[297]

A decade later, the OTA conducted the largely ignored 1986 report, which resounded similar concerns to the Privacy Commission's study in 1977.[298] The OTA's report stated, "Technology has now altered that balance in favor of the agencies. Computers and telecommunications capabilities have expanded the opportunities for federal agencies to use and manipulate personal information."[299] The report explicitly warned Congress of the Privacy Act's incompatibility and irrelevance with the new, advancing digital technologies being employed by the government.

Specifically, the OTA's report alerted Congress that the Privacy Act, which was created and debated in a paper-based working environment, no longer adequately applied to the technology used in the mid-1980s.[300] The report addressed how the federal agencies' shift to using primarily electronic record-keeping systems posed serious privacy implications for Americans.[301] Even after the Privacy Act passed in 1974, computer matching by federal agencies tripled from 1980 to 1984.[302] By 1986, the OTA's report declared that computer matching had become routine in a number of programs.[303] Additionally, the OTA's report highlights the potential consequences of the "direct linkage of computer records via telecommunication systems" – a primitive version of the

[297] Ibid., 23, 108.
[298] For first reference of the Office of Technology Assessment's report, entitled "Federal Government Information Technology: Electronic Record Systems and Individual Privacy," see Chapter Two, 69.
[299] Ibid., 11.
[300] Ibid.
[301] The OTA warned that the government's use of "computer and telecommunications… to process detailed information on millions of citizens," could pose dark consequences. Ibid.
[302] Ibid., 11, 101.
[303] Ibid., 11.

Internet.[304] Although federal agencies enjoyed the "ease and efficiency" of "communicating directly with the computer as information [was]… collected or needed…" the OTA warned that such widespread adoption of direct linkages would allow for "easy disclosure and exchange of information."[305] Thus, beginning in the late 1970s and increasingly in the 1980s, a recognizable grey area started to emerge whereby technologies began to outpace the ability of individuals to meaningfully protect their information through mechanisms made available under the Privacy Act.[306]

The concerns voiced by the authors of the 1986 OTA study are analogous to those voiced by the Privacy Commission in 1977, demonstrating the uneasiness felt by Americans regarding their changing right to privacy. As the OTA's report claims, "[In 1986] … large volumes of …[personal] information are collected, retrieved, disclosed, disseminated, manipulated, and disposed of by computers. Moreover, direct on-line linkages now make it possible to compare individual [pieces of] information with a host of public and private agencies."[307] As early as 1977 and again in 1986, members of the Privacy Study Commission and OTA stressed how the Privacy Act had already been rendered irrelevant. Additionally, both reports point to potentially worrisome implications of the ways in which the government had begun employing digital record-keeping systems and networking technologies to maintain individuals' personal information.

---

[304] Ibid., 11, 23.
[305] The report describes how the adoption of "direct linkages" by federal information systems likely occurred due to decreasing costs and increasing conveniences. For instance, in 1982 the cost of "a typical network interface" was around $500 but "by 1987 the cost [was] expected to drop to around $50." Ibid.
[306] Ibid., 11.
[307] The report continued, "Computer tapes, software, and networking also make it possible to compare personal information stored in different record systems." Ibid., 23.

Affecting Individuals' Privacy Concerns in 2012: Political, Social, and Neurological Factors

Prior to the 1970s, computers were too expensive, big, and difficult to use for individuals to even contemplate the Privacy Act's importance. For instance, in the early 1960s, federal agencies still largely operated in "paper environments."[308] In the mid-1970s, when the Privacy Act of 1974 was debated and passed, the "vast majority of federal record systems [during this time] were manual."[309] In fact, up until the mid-1980s, the government ran a primarily paper-based information collection system. A survey in 1980 revealed that federal agencies possessed only "a few thousand microcomputers."[310] By 1985 though, the same agencies suddenly possessed around 100,000 microcomputers.[311] Thus, by the time individuals began to understand the valuable protections granted by the Privacy Act, the mechanisms through which individuals sought protection had already been rendered irrelevant, as the technology had advanced beyond the Act's guarantees.

This is certainly the case in 2012 – except that the gap concerning advancing information technology, individuals' understanding of how that information technology works, and the façade created by the Privacy Act has only grown in scope and size. While the public's concern for privacy may exist, these concerns are not holding individuals back from sharing online without a full understanding of how their personal information is used and re-used by countless entities. The stakes have grown higher.[312]

---

[308] U.S. Congress, Office of Technology Assessment, "Federal Government Information Technology,"12, 22, 104.
[309] Ibid.
[310] Ibid.
[311] Ibid.
[312] While most online companies and services advocate that consumers have a choice between services and/or using a service at all, in developed societies that choice is, in reality, an expectation. Telling consumers that they do not *have* to use a service, such as e-mail, is futile in a society where people depend on others to use that service. Opting-out of using such services would prove extremely challenging in today's

Because individuals enjoy sharing personal information online, data companies enjoy making a profit, and the government enjoys obtaining more data on its constituents, a vicious cycle of underlying short-term benefits incentivizes the Data Cycle.[313] The public cares about privacy but is currently blinded by the perceived instant gains and benefits of the very technologies that exploit its information. Meanwhile, companies, which thrive on driving their bottom line, continue to collect and sell individuals' personal information to the government. The government pretends the Privacy Act is providing adequate protection to individuals so that it may gain influence, insight, and power. In the next section, I argue that presently, the government is able to successfully access unprecedented amounts of individuals' information for two primary reasons. First, individuals mistakenly share information with private companies without realizing that private companies actually share such information with the government. Second, we humans are growing increasingly neurologically addicted to information technologies. As such, our understanding of the long-term privacy implications of our actions is greatly obscured.

The Government's Longstanding Political Interest in Collecting Individuals' Information

The actual practice of data collection by governments is an ancient phenomenon.[314] Dating back to primitive Egyptian civilizations, rulers of societies kept

---

developed world where we are essentially technologically locked into using such convenient means to accomplish everyday tasks and effectively participate in society.

[313] 70% of participants in a poll conducted by the *Economist* on February 14, 2012, indicated that they believe "society benefits when we share personal information online." "Debate: Social Networking," *The Economist*, February 8, 2012, accessed February 5, 2012, http://www.economist.com/debate/debates/overview/222.

[314] I refer to government data collection in this section as an *indirect* surveillance activity. Unlike deliberate types of surveillance such as wiretapping, spying, and other types of espionage, indirect surveillance activities may be defined as the collecting individuals' humdrum, personal information. Americans remain unaware of how this information is later used and re-used.

population records in order to conduct taxation, immigration and military checks.[315] Additionally, as early as the fifteenth century BC, the Book of Numbers provides records detailing how the migrating Israelites conducted censuses in order to record population details.[316]

In the case of the United States, where the Data Cycle has been specifically identified, even as early as in the colonial era organizations were eager to know the details of the colonists' daily lives.[317] America's own Founding Fathers included a mandate in the U.S. Constitution (Article I, Section 2) that the federal government conduct a decennial census in order to reapportion districts based on changing populations.[318] Since 1790 the government has continued to conduct the decennial survey in addition to the "hundreds of surveys" of the U.S. Census Bureau and labor statistics bureaus (*see Appendix C)*, as well as ethnographic surveys conducted by the Smithsonian.[319] All of these surveys are conducted every year in an attempt to help the government learn more about the people they serve.[320]

Throughout time, the government has maintained that all of its collected information assists the government, politicians, and legislators to better serve their

---

[315] David Lyon, *The Electronic Eye: The Rise of Surveillance Society* (Minneapolis: University of Minnesota Press, 1994), 23.
[316] Ibid.
[317] Waldo, Lin, and Millet, *Engaging Privacy,* 349.
[318] This section reads, "Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective Numbers . . . The actual Enumeration shall be made within three Years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such Manner as they shall by Law direct." U.S. Census Bureau, *History: Decennial Census* (Washington, DC: Bureau of the Census), accessed November 2, 2011, http://www.census.gov/history/www/programs/demographic/decennial_census.html.
[319] Waldo, Lin, and Millet, *Engaging Privacy,* 349.
[320] U.S. Census Bureau, "1890: Fast Facts," The U.S. Census Bureau (Washington, DC: Bureau of the Census), accessed November 2, 2011, http://www.census.gov/history/www/through_the_decades/fast_facts/1890_fast_facts.html.

constituents.[321] As David Lyon, author of *The Electronic Eye* puts it, "Surveillance is not new. Since time immemorial, people have 'watched over' others to check what they are up to, to monitor their progress, to organize them or to care for them."[322] Indeed, according to author James Scott in *Seeing Like a State*, data collection made possible modern statehood. He offers, "The techniques devised to enhance the legibility of a society to its rulers [such as censuses, surnames, straight roads, tax records…] have become vastly more sophisticated, but the political motives driving them have changed little."[323]

Dr. Westin acknowledged in 2003 that democracies, which inherently have a strong commitment to individualism and freedom of association, have gathered information on their people in order to most efficiently govern and organize their populations.[324] Scott echoes, "Appropriation, control, and manipulation (in the non-pejorative sense) remain the most prominent [legibility techniques]. If we imagine a state that has no reliable means of enumerating and locating its population, gauging its wealth, and mapping its land… we are imagining a state whose interventions in that society are necessarily crude..."[325] Assuming that Scott's claim is true, most would agree that in order for democratic governments to "appropriate," "control," and "manipulate" society (in the non-pejorative sense) – to efficiently provide various services, benefits, and protections to its citizens – the government needs information about those individuals. Only after

---

[321] Paul Starr, *The Creation of the Media: Political Origins of Modern Communications* (New York: Basic Books, 2004), 98.
[322] Lyon, *The Electronic Eye,* 23.
[323] James C. Scott, *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed* (New Haven, CT: Yale University Press, 1998), 77.
[324] Westin, "Social and Political Dimensions," 2-3.
[325] Scott, *Seeing like a State,* 77.

collecting data on its people may a government then effectively provide its public with such deliverables.

Scott makes an excellent point that the government's perceived informational need exists in order to best serve its constituents. Nevertheless, there exists a fine line between collecting and using critical information on individuals to achieve a specific, articulated purpose versus collecting and using information for that purpose and then sharing, using, and/or re-using that same information for other purposes – no matter how benign or beneficial. The Data Cycle represents a slippery slope whereby the government is collecting data under the guise of one purpose and then distributing and re-accumulating the data at a later point in time. Existing laws are either obtrusive (yet still allowing these circuitous transactions to take place in a highly inefficient manner) or not specific enough to stop these transactions from occurring.

In order to better understand why the government pretends the Privacy Act is satisfactory in the twenty-first century, we may consider the political factors driving MCs' desires to obtain information on their constituents. In particular, the remainder of this section considers political theorists' proposals as to why MCs have maintained an interest in not only collecting information on their constituents, but also in learning more about them through partnering with private sector participants, as characteristic of the Data Cycle.[326]

Independent of technical advancements, MCs have always had an interest in learning more about their constituents in order to produce better policy, satisfy more needs, and get reelected. Generally speaking, while MCs possess the desire to achieve

[326] We focus on MCs since they govern the legislative branch of our society. If we understand MCs' motives, we may better understand why certain laws do or do not exist.

multiple goals in office, in order to keep their jobs for a wide array of reasons (i.e., prestige, income, power), winning reelections is undoubtedly extremely important. By achieving this primary goal, MCs are then granted the opportunity to pursue their other goals. As political scientist Steven Smith states, "…[I]f a legislator is to continue pursuing other goals in public office, it is not too surprising that reelection dominates all other considerations."[327] In order for MCs to increase their likelihood of staying in office, MCs must ascertain the beliefs and interests of their constituents; they must learn the needs and preferences of the people they represent.[328] By collecting information about their constituents, MCs may garner valuable feedback, politically useful information (e.g., how to vote on salient issues), and input on how to most effectively lead and therefore hold onto their jobs.[329] As political scientist Richard Fenno states, "When a Congressman describes his seat as 'safe,' he implicitly adds: 'because, and so long as, I work actively to keep it so.'"[330] Thus, as more legislators seek to convert their positions into long-term careers, MCs have a clear motive to learn more about their constituents, as their jobs and opportunities to accomplish other secondary and tertiary goals depend on learning and appropriately acting on such relevant pieces of information.

---

[327] Steven Smith, Jason Roberts, and Ryan Vander Wielen, *The American Congress* (New York: Cambridge University Press, 2009), 6th edition, 90.

[328] A study published in 1973 by political scientist John Kingdon demonstrates how MCs, when deciding to vote on salient, controversial issues most heavily factor in the views of their constituencies. Conducted via extensive interviews with a sample of MCs in the House of Representatives, Kingdon concluded that the more salient the issue, the more legislators weighed constituents' preferences because they understood how important it is to align themselves with their voters' views in order to stay in office. Ibid., 112.

[329] Additionally, author and political scientist Steven Smith describes how since the early 2000s, MCs have increasingly sought to transform their positions into long-term careers. Statistics from the first decade of the twenty-first century suggest, the number of MCs aiming to turn their positions into careers is at an all time high of 90-95%, whereas in the 1940s it was only around 80%. Because of this trend, legislators today arguably must cater to their constituents to a higher degree than their predecessors did should they wish to stay in office longer. Ibid., 87.

[330] Richard Fenno, Jr., "U.S. House Members in Their Constituencies: An Exploration," *The American Political Science Review* 71, no. 3 (1977): 914.

Engaging the Private Sector Veils the Government's Orwellian Incentives

While state actors have seemingly always possessed an unswerving political interest to collect constituents' data, what makes the Data Cycle more insidious than any other similar endeavor?

In order to circumvent the Privacy Act, the government employs private sector companies to collect and provide massive amounts of individuals' information. In other words, private companies act as secret middlemen to facilitate a concealed one-way connection between the government and individuals. Socially, individuals would never voluntarily provide the government with the same amounts and types of data (e.g., shopping lists, music selections, browsing preferences, and personal communications) that they share with private companies. It is through the weak interpretation and enforcement of the Privacy Act that private sector companies have emerged as disguised arms of the government, duping individuals into believing that sharing with them is entirely different than sharing with the government itself.

Over the first decade of the twenty-first century, Americans have grown increasingly comfortable with sharing vast amounts of private information with private companies. As discussed in Chapter Two, entire company business models are based off of monetizing individuals' shared personal data. On February 1, 2012, Facebook filed for its initial public offering ("IPO"), estimations of which have predicted a capitalization of at least $80 billion.[331] Facebook however, does not profit by selling a tangible product to consumers; instead, its success hinges upon individuals' continued sharing of personal information on the platform so that the company may meaningfully analyze and exploit

[331] Andrews, "Facebook Is Using You."

such data.[332] Similarly, Google's business model relies on people using its services: search, e-mail, and others. Both companies analyze and sell the data that individuals provide while using their services. In light of Facebook's IPO, individuals have begun realizing their importance to Facebook's business model, calling for CEO Mark Zuckerberg to share some of the company's profits, which are generated from individuals' use of his site.[333]

Nevertheless, individuals continue to share with private companies like Facebook and Google as if they are entities completely separate from the government. Would individuals feel comfortable with the Internal Revenue Service ("IRS") knowing what type of cupcake they had with lunch? Probably not. However, as reported in *The New York Times* in February 2012, "The Internal Revenue Service searches Facebook and MySpace for evidence of tax evaders' income and whereabouts."[334] Further, the United States Citizenship and Immigration Services "has been known to scrutinize photos and posts to confirm family relationships or weed out sham marriages."[335] The article exposes how LexisNexis offers a product, "Accurint for Law Enforcement," which provides government agents with information about what people do on social networking sites.[336] Information shared with private companies has reportedly been used by the government as evidence against people fighting for child custody.[337] Lastly, as exemplified in Chapter One, MCs buy personalized profiles of their constituents from political data companies.

---

[332] Ibid.
[333] David Lazarus, "Facebook Users Want a Piece of the Action," *LA Times*, February 2, 2012, accessed February 18, 2012, http://www.latimes.com/business/money/la-fi-mo-money-minute-20120202,0,1944277.story.
[334] Andrews, "Facebook Is Using You."
[335] Ibid.
[336] Ibid.
[337] Ibid.

As Harvard Law Professor Jonathan Zittrain argues, "If you're not paying for something, you're the product – not the customer."[338] This is one of the major problems with how individuals, U.S. corporations, and increasingly the government currently interact. The ubiquity of the Internet complicates many of the traditional distinctions between privacy protections of U.S. individuals, corporations, and the government. Today, corporations routinely collect information on individuals without those individuals knowing that information is being gathered, how that information will be used, and/or later re-used. Further, the government, which has an interest in learning about its population, has entered the data economy as a major player. Such interactions, however, have led to a symbiotic relationship between corporations and the government, where data collection and retention practices are becoming intertwined and increasingly ingrained without individuals' consent or awareness.

Thus, the structure of the Data Cycle, resulting from limitations of the Privacy Act, has uniquely situated state actors to indirectly achieve access to more information than ever previously possible. The Act's vague language and outdated guidelines permit data sharing, using, and re-using between public and private entities. As Professor John Palfrey states, "In the twenty-first century, a state can come to know more about each of its citizens via surveillance than ever before in human history… [through] data … held in private hands as well as public [ones]…"[339] The Data Cycle veils the fact that sharing with the Googles and Facebooks of the world is actually not too far off from directly sharing with the government itself. Nevertheless, individuals continue to fuel the Data

[338] Jonathan Zittrain, "Ideas for a Better Internet" (class discussion, Areeda Hall, Harvard University, Cambridge, MA, April 24, 2011).
[339] Palfrey, "The Public and the Private at the United States Border with Cyberspace," 242.

Cycle, as they remain largely oblivious to how the distinctions between the public and private sectors are growing increasingly ambiguous.

Humans are Growing Neurologically Addicted to our Information and Communication Technologies

While individuals inaccurately believe that sharing information with private companies is distinctively different from sharing such information directly with the government, individuals are additionally empowering the Cycle, as they grow neurologically dependent on using communication and information technologies.

Even though privacy concerns certainly exist in the twenty-first century, they are not strongly defended or practiced. Why? Other than the clear answers of enhanced access to information and communication, the public's conception of privacy in America has been greatly influenced by our growing inability to decouple our human selves from our virtual selves. Intriguingly, human beings are changing both neurologically and habitually through our use of new digital technologies. In comparing the works of Nicholas Carr, Clay Shirky, and William Powers, three modern, contemporary authors who discuss technology's effects on human behavior, we may better understand how American culture in the twenty-first century is being shaped through our uses of modern technologies. Further, this discussion suggests why individuals – even if they are aware of the Data Cycle – may not want to or cannot stop using the technologies that power its existence.

First, in Nicholas Carr's *The Shallows*, Carr expresses concern over how his computer use is biologically altering the electrical wiring in his brain and changing the way he thinks. "It seemed ludicrous to think that fiddling with a computer, a mere tool, could alter in any deep or lasting way what was going on inside my head. But I was

wrong."[340] Carr discusses his self-diagnosed "addiction to being connected" in relation to neural plasticity, or the capability of the human brain – even in adult form – to change in response to repeated experiences and behaviors.[341]

Carr's "connectivity concerns" find supportive evidence in recently conducted biological and neurological studies. For example, developmental doctor and Harvard Professor of Pediatrics Charles Nelson and his lab have conducted numerous studies in developmental cognitive neuroscience. Nelson offers, "Collectively, we now know that the adult brain is quite capable of being molded by experience… [L]earning and memory represent the quintessential examples of adult plasticity…"[342] Although the impact of experience on one's brain is not constant throughout one's life (early experiences in childhood development usually have a greater impact), studies have proved that both the developing and developed brain are capable of being molded by experience.[343] Relatedly, Neuroeconomist Professor at the University of Claremont Paul Zak conducted a study featured in the July 2010 issue of *Fast Company* that investigated social networking's effect on the brain. Professor Zak explained that when you interact online with other people your brain releases oxytocin, a type of neurotransmitter that makes humans feel connected to others, empathetic, and less stressed.[344] Zak indicated in an interview, "The more you have oxytocin released, the more your brain wants it. The brain becomes used to this, as it sets down these pathways and trains itself to release oxytocin from online

---

[340] Nicholas Carr, *The Shallows: What the Internet Is Doing to Our Brains* (New York: W.W. Norton, 2010), 38.
[341] Ibid., 31.
[342] Charles Nelson and S. Jeste, "Neurobiological Perspectives on Developmental Psychopathology," in *Rutter's Child and Adolescent Psychiatry* (Oxford: Blackwell Publishing Ltd., 2009), 5th edition, 152.
[343] It was once thought that the adult brain had little plasticity. However, studies have proved this incorrect. Research demonstrates that the part of the brain controlling the playing hand of violinists is much more developed than the other hand. Additionally, London cab drivers who must acquire significant knowledge of the city in order to perform their job demonstrate an overdevelopment of the hippocampus. Ibid., 152.
[344] Zak, phone interview.

social interactions..."[345] Evidence shows that humans today who habitually use the Internet and social networking are becoming neurologically addicted to such practices.[346] Thus, when Carr states, "We like to feel connected – and we hate to feel disconnected," he makes a convincing argument for how technology and habit are working together to gradually affect human behavior.[347]

In his book *Hamlet's Blackberry*, William Powers adds, "Today the stimuli we receive from our environment are different [from the primitive ones, such as]… wild animals lurking in the trees."[348] Powers contends that instead, "We're on alert for ringtones and new messages—but the biochemical effect is hypothetically the same."[349] Research conducted by Harvard psychiatry experts Dr. Hallowell and John Ratey conclude that when you receive a text or chat with a friend online, the stimulation "provokes excitement" or a "dopamine squirt."[350] These constant chemical squirts, however, are "rewiring our brains," as Director of the National Institute of Drug Abuse Nora Volkow states.[351] Without constant stimulation, people may feel bored, withdrawn, or isolated, as constant connectivity neurologically stimulates the same wires of our brains as drugs of abuse and pleasure.[352]

Powers elucidates, "From the earliest days of computers, there have been worries about the effects… on the human mind." Citing Alvin Toffler, who coined the term "information overload" in the 1970s, Powers notes how this term has become increasingly

---

[345] Ibid.
[346] Ibid.
[347] Carr adds, "The Internet doesn't change our intellectual habits against our will. But change them it does…" Carr, *The Shallows,* 91-92.
[348] William Powers, *Hamlet's Blackberry: a Practical Philosophy for Building a Good Life in the Digital Age* (New York: Harper, 2010), 74.
[349] Ibid.
[350] Ibid.
[351] Matt Richtel, "Attached to Technology and Paying a Price," *New York Times*, June 6, 2010, aaccessed August 23, 2010, http://www.nytimes.com/2010/06/07/technology/07brain.html?pagewanted=all.
[352] Ibid.

popular in recent time as exposure to using our digital devices has grown. [353] Powers describes how "connected" humans are becoming busier and busier managing our connectedness.[354] He recognizes that this could just be a "transitional issue," yet echoes Carr's fears.

Clay Shirky directly disagrees with Carr's views, which Carr reiterated in an *Atlantic Monthly* article, entitled "Is Google Making Us Stupid?" In response to Carr, Shirky wrote an article for *The Wall Street Journal*, "Does the Internet Make You Smarter?" In his article, Shirky purports that the ability to connect with billions of other people through the Internet is allowing us to tap into "cognitive surplus," or participate in thought-provoking activities, rather than merely consume information.[355] Shirky argues against Carr's case for "digitally-driven stupidity" by positing that we have yet to fail to integrate our new digital freedoms into society.[356] In accordance with Shirky's views, Vint Cerf, the creator of TCP/IP standard, wrote in the 1980s with regard to the ARPANET, "It is not that the human is being replaced by machines (except for the most menial of functions); rather, human capability is being greatly extended by the leverage of automation."[357] A contemporary of Cerf's, J.C.R. Licklider similarly viewed computer networking as a means to improve human capabilities. In Licklider's "Man-Computer Symbiosis" Licklider articulated his hope that "… [i]n not too many years, human brains and computing machines will be coupled together very tightly…"[358]

---

[353] Powers, *Hamlet's Blackberry,* 50.
[354] Ibid., 2.
[355] Richtel, "Attached to Technology and Paying a Price."
[356] Clay Shirky, "Does the Internet Make You Smarter?" *The Wall Street Journal*, June 4, 2010, accessed October 24, 2010, http://online.wsj.com/article/SB10001424052748704025304575284973472694334.html.
[357] Vint Cerf, "Military Requirements for Packet-switched Networks and Their Implications for Protocol Standardization," *Computer Networks* 7, no. 5. (1983): 296.
[358] Licklider, "Man-Computer Symbiosis," 5.

In 2012, Licklider's envisioned coupling is becoming a reality. *New York Times* reporter Matt Richtel has taken a closer look at the way constant connectivity and the data deluge is affecting the way humans' function in his modern series "Your Brain on Computers."[359] One of Richtel's most startling articles focuses on how high school students are "growing up digital." Richtel conducted a four-month study at Woodside High School, a high school in Silicon Valley, California, interviewing students, parents, teachers, the school's principal, and scientists to expose how profound the distractions of computers and cellphones are on ordinary Americans high school students.[360] Since our devices offer constant stimuli, younger adults – whose brains are still developing – are training themselves to constantly change their focus, making once-typical school tasks like reading an increasingly difficult task.[361] Richtel concludes, "Just as food nourishes us and we need it for life, so too — in the twenty-first century and the modern age — we need technology. You cannot survive without the communication tools."[362] Additional scientific studies on the issue confirm Professor Zak's findings – people who regularly switch their focus are grooming their brains to habitually favor distraction and necessitate a continual stream of stimuli.[363]

Powers proposes a solution that speaks to Carr's concerns. Powers offers a "new digital philosophy" that aims to balance "screen time" and "time apart."[364] Nevertheless, Powers' "solution," especially when considered in tandem with *The New York Times'* study

---

[359] Matt Richtel, "Your Brain on Computers: A Series," *New York Times*, June 7, 2010, aaccessed February 3, 2012, http://topics.nytimes.com/top/features/timestopics/series/your_brain_on_computers/index.html.
[360] Matt Richtel, "Growing up Digital, Wired for Distraction," *New York Times*, November 21, 2010, accessed November 23, 2010, http://www.nytimes.com/2010/11/21/technology/21brain.html?src=me&ref=general.
[361] Ibid.
[362] "Digital Overload: Your Brain On Gadgets," *NPR*, August 24, 2010, accessed February 15, 2012, http://www.npr.org/templates/story/story.php?storyId=129384107.
[363] Richtel, Matt. "Growing up Digital."
[364] Powers, *Hamlet's Blackberry,* 4.

of Woodside High School (and its "Unplugged Challenge"), recognizes that humans have become so reliant on our technologies that some of us are worried we cannot live without them.[365]

Regardless of individuals' waning rights to informational privacy, the idea that humans cannot – or really, do not want to – separate ourselves from our gadgets speaks to why privacy concerns are not going to stop the Data Cycle any time soon.

Discussion

As Dr. Westin wrote in 2003, "A loss of overall citizen privacy in America post 9/11… will surely be the judgment of twenty-second century historians. Whether the new privacy balances will be seen as a necessary and justified shrinkage or a disastrous and authority-abused decline remains to be determined."[366] Unbeknownst to most Americans, the U.S. government is often using their tax dollars to acquire personal information about them from private companies. However, while individuals remain unaware of the Cycle, they participate as active enablers of this phenomenon. By sharing online without understanding the consequences and/or demanding more protections, individuals are tacitly consenting to perpetuating this pattern.

From his book *The Digital Person,* Daniel Solove states, "We are currently confronting the rise of what I refer to as 'digital dossiers,'" where private companies, government agencies, and other parties continually collect data from unsuspecting users in order to make judgments.[367] However, the government and private companies are not merely creating dossiers on individuals, as they are also acting as brokering agents, selling

---

[365] *The New York Times* "dared" its readership to go "unplugged" for 24 hours (and submit their observations) in 2011. "Unplugged: Take the Challenge," *New York Times*, June 7, 2010, accessed October 2, 2011, http://bits.blogs.nytimes.com/2010/06/07/unplugged-take-the-challenge/; Powers, *Hamlet's Blackberry,* 4.
[366] Westin, "Social and Political Dimensions," 27.
[367] Solove, *The Digital Person,* 1-2.

and trading information that individuals do not know about and that arguably, does not belong to either of the entities trading such information. Operating within an outdated, weak legal environment, individuals, the government, and private companies participate in the Data Cycle without any oversight. The government, which has an interest in preserving its participation in the Data Cycle, has chosen to ignore the Privacy Act and instead, exploit individuals' ignorance of the Data Cycle and their addiction to the Web. Addressing this concern, Palfrey has written extensively about the dire need to rethink legal protections as we continue to drive full speed ahead into the hazy issue of what constitutes "private" versus "public" realms.[368] Thus, the information toll Americans pay in the twenty-first century has expanded. Americans may no longer truly know how the information they initially provide to the government for official use is being used, re-used, by whom, when, and for what purpose(s).

Clearly articulated by the OTA's 1986 report, a grey area began to emerge in the 1980s when technological advances began to quickly outpace the Act's outlined legal protections. As Senator Susan Collins (R-ME) argues, "[W]e will need to be more vigilant in ensuring that the wheels of progress are not inadvertently running over our basic privacy rights."[369] Although various government officials and privacy experts voice the need for Congress to alter the Privacy Act so that it may keep up with new technologies, since 1974 little has changed.[370] Society continues to produce more technical

---

[368] Palfrey warns that there are basically no clear, legal protections prohibiting a private company from turning over an individual's or group's data to the state, nor are there protections against state entities using data in more than one way at different periods in time. Palfrey, "The Public and the Private at the United States Border with Cyberspace," 242.

[369] Grant Gross, "U.S. Privacy Act Outdated, Hasn't Kept up With Technology, Experts Say," *Computer World,* June 18, 2008, accessed April 24, 2011, http://www.computerworld.com/s/article/9100258/U.S._Privacy_Act_outdated_hasn_t_kept_up_with_technology_experts_say_.

[370] Ibid.

advancements and develop new uses while the legal limitations of such uses have remained static, stuck in a different era.

CONCLUSION

This thesis began as an effort to uncover an observed pattern among individuals, the government, and private companies. After following a trail of transactional breadcrumbs that led from official state election voter records to political data companies to MCs' disbursements, I uncovered the Data Cycle. By exploring the Cycle, I revealed how MCs and other government officials are able to purchase augmented versions of publicly collected personal information with public funds. Substantiated throughout this thesis, the Data Cycle connects the public and private sectors together through a voracious desire to acquire individuals' personal information.

Historically, this thesis demonstrated how the public and private sectors have actually engaged in similar data sharing practices since as early as the 1930s when the sectors began exchanging data originally gathered by the quasi-public entities of political campaigns. However, this thesis shows how both the Data Cycle and its implications in the twenty-first century are completely different from any comparable construct we have observed before.

An informational shift has occurred between the public and private sectors. Currently, the public sector heavily leans on the private sector to acquire information that the government is legally prohibited from directly collecting. The government and private companies continue to trade individuals' personal information through existing legal loopholes. The government's active participation in the Data Cycle and circumvention of the Privacy Act reveals how the privacy of individuals in the twenty-first century is objectively waning.

Although the public remains victim to the government and private sectors' veiled collecting and trading practices, the public's naïve and/or apathetic sharing of vast

amounts of personal information online directly contributes to the Cycle's power. Individuals share personal information online with private companies because they either mistakenly believe that sharing with private companies is dissimilar from sharing directly with the government and/or are growing addicted to the heightened levels of stimulation produced by connecting online. In both cases, the Data Cycle benefits.

In particular, the Cycle thrives on the manifestation of two symbiotic relationships. In the first, individuals depend on the same technologies and online services that have spawned and support the data economy, enabling private companies to enter business relationships with the government. In the second, the government and private sectors rely on one another for different kinds of datasets (i.e., the private sector wants the government's official public records and the government wants the private sector's comprised consumer profiles). This "cycle" we are headed down disproportionately favors short-term gains and completely disregards potentially harmful, long-term repercussions. Individuals who are aware of Data Cycle may not change their online sharing habits since those habits produce immediately realizable rewards, such as ease, convenience, and the release of feel-good hormones. Meanwhile, private companies gather, process, and later exploit this information. Regardless of whether an individual ever experiences an unfortunate consequence from this interaction, individuals recognize they must surrender their personal information in order to effectively function in certain parts of American society in 2012.[371]

[371] Modern society heavily relies upon electronic communication to function. On Harvard's campus in 2012, electronic dependence is intertwined with the school's innovative nature and embrace of information technology. In fact, Harvard daily life without electronic communication is essentially impossible, considering not only the way students communicate with each other, professors, and employers, but also the students' reliance on technology to access information and conduct research. Even "old-fashioned" systems have been revamped to use electronic means, such as locating library books or turning in course

Thus, paradoxically, the victims of the Data Cycle are fueling its strength. The manifestation of the Data Cycle in the first decade of the twenty-first century raises red flags concerning the unchecked powers of political and economic interests in the Information Age. With multiple interests and dependencies at stake, the development of the Data Cycle symbolizes how as a society we tacitly do not prioritize individuals' privacy.

The Unintended Adoption of Social Security Numbers as Unique Identifiers

The government's creation of social security numbers ("SSNs") in the 1930s serves as a useful precedent to better understand how the Data Cycle produces unintended consequences from the widespread adoption of new technologies and practices. Although the government did not intend for SSNs to be used as unique identifiers, they have been used as such since the 1960s. Akin to the Data Cycle, SSNs offered the government, individuals, and companies alike short-term organizational benefits and conveniences. The government's practice of employing SSNs as unique identifiers grew popular in the mid-twentieth century and is now seemingly permanent. After fifty years of using SSNs as identifiers, the American public is only beginning to realize many of their counterproductive implications. As to the Data Cycle, we may likewise expect the government to alter its handling of personal information, producing serious, unforeseen consequences.

Prior to the 1930s, punched cards were generally regarded as mere processing tools.[372] However, with the passage of the Social Security Act, a key component of

---

assignments online. Modern society prefers efficiency and we, therefore, constantly rely on technology to function.

[372] After the processing was complete, the cards were usually discarded. See Lars Heide, *Punched-card Systems and the Early Information Explosion 1880-1945* (Baltimore: Johns Hopkins University Press, 2009), 212.

President Franklin D. Roosevelt's New Deal, punched cards and supporting technologies (e.g., collators and tabulators) were used to monitor the wages of tens of millions of people.[373] The Social Security Administration ("SSA") sought to maintain records on workers' contributions so that when those workers retired or turned age sixty-five, the government could provide workers with their entitled benefits. Lars Heide, author of *Punched-Card Systems and the Early Information Explosion* comments, "Administering the Social Security program introduced for the first time a register on a national scale, giving it unprecedented size and scope."[374] The Social Security register became one of the first of many large register-based systems in the United States.

The SSA created SSNs in the U.S. initially as a means to universally register, organize, and identify people's true identities. Because this decision was controversial in the 1930s, the Social Security board in 1937 required that all records remain confidential and that all numbers only be used for the administration of the program.[375] In fact, between 1930 and 1960, social security cards explicitly read, "*Not for identification*."[376]

---

[373] The Social Security Board, a newly created federal organization to run the program, initially envisioned manual processing of the reports. However, as the Board continued to design the program's architecture, punched cards triumphed manual tabulating. It is important to note that punched cards in and of themselves would have certainly not been an effective means to tabulate information. IBM provided the government with two other technologies that eased its operation: namely, the collator and a new posting attachment to the tabulator. Ibid., 214.

[374] By employing such a large-scale system, the Social Security Board greatly catalyzed the technology industry's development at the time – similar to how the U.S. military spurred computing hardware and design development throughout the early to mid twentieth century. Heide argues that the wide employment of punched cards through the Social Security Act provided "the essential impetus for further development of the established closure of punched cards for bookkeeping." Ibid., 212, 220.

[375] Ibid., 224.

[376] This statement was meant to notify anyone to whom a social security card might be presented "that it cannot be relied upon, by itself, as evidence of the identity of the person presenting it." U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, *Records, Computers and the Rights of Citizens: Report of the Secretary's Advisory Committee on Automated Personal Data Systems* (Washington, DC: Government Printing Office, July 1973), accessed December 8, 2011, http://aspe.hhs.gov/datacncl/1973privacy/c7.htm; Strum, Nash, and Etulain, *Privacy: The Debate in the United States Since 1945,* 46.

Initially, only the SSA and the IRS used SSNs, and even then for limited purposes.[377] However, in the early 1940s the government grew to rely on SSNs for additional uses.[378] Other countries such as France, Germany, and the U.K had contemporaneously built large, register-based systems that closely resembled that of the United States' Social Security register.[379] During a time of utter chaos in the 1940s, one regrettable result of punched card systems was the Nazis' ability to methodically locate, deport, and isolate Jews throughout WWII.[380]

Besides the increasing popularity in punched card technology by other countries as an efficient means to track and organize substantial numbers of people, various administrations and government agencies, particularly the IRS, expanded the use of the SSNs.[381] In 1961, the IRS decided to require individuals to use SSNs for taxpayer identification, allowing the government to track and serve citizens with improved efficiency.[382] In 1973, the Report of the Secretary's Advisory Committee alerted the government that since the 1960s the federal government had continued to expand the

---

[377] As articulated in the 1973 Report of the Secretary's Advisory Committee on Automated Personal Data Systems, "[In 1936 and 1937] most Americans had not been issued a number, and few organizations felt the need of a numeric identifier for purposes of data processing… In fact, for some years after its inception in 1936, there was no substantial use of the SSN..." See U.S. Department of Health and Human Services, "Records, Computers and the Rights of Citizens."

[378] Strum, Nash, and Etulain, *Privacy: The Debate in the United States Since 1945,* 47.

[379] Heide, *Punched-card Systems,* 222.

[380] In 1942, the Germans invaded and controlled unoccupied portions of France and began using their national registers to track Jews, who were already being registered separately. The French register had assigned each citizen a 13-digit "national identification number" (today's French social security number) and recorded his/her personal information. Ibid.

[381] In 1943 President Roosevelt inadvertently encouraged federal agencies to use SSNs as unique identifiers by passing Executive Order 9397, which prompted federal agencies to use SSNs to manage records. The Order stated, "Hereafter any Federal department, establishment, or agency shall, whenever the head thereof finds it advisable to establish a new system of permanent account numbers pertaining to individual persons, utilize exclusively the Social Security account numbers…" U.S. Department of Health and Human Services, "Records, Computers and the Rights of Citizens."

[382] While in hindsight Roosevelt's Executive Order 9397 in 1943 may have "provided the basis for a change in conception of the role of the SSN," the 1973 Report of the Secretary's Advisory Committee details that it was not until "after the 1961 decision to use the SSN as an individual identifier for federal tax purposes," did the order have any real significance. Ibid.

uses of the SSN "beyond its original purposes."[383] Nevertheless, since the 1960s SSNs have been used as personal identification numbers.

In the twenty-first century, SSNs have morphed into a major liability for tax refund fraud.[384] In order to commit refund fraud, according to the Government Accounting Office, an identity thief just needs a taxpayer's name and SSN.[385] Finding these pieces of information has become progressively easier to do on a large scale since many electronic databases do not adequately defend phishing or hacking attempts. Over the last three years, the number of identified fraudulent tax return cases has more than tripled in size.[386] According to the U.S. Treasury Department, "As of March 4, 2011, the IRS had identified 335,341 tax returns with $1.9 billion claimed in fraudulent refunds," a 181% increase in the number of fraudulent tax returns identified during this same period from the preceding year.[387] Additionally, this statistic only encompasses the number of fraudulent claims the IRS was capable of detecting and does not account for the huge

---

[383] The Committee advised against "the adoption of any nationwide, standard, personal identification format, with or without the SSN, that would enhance the likelihood of arbitrary or uncontrolled linkage of records about people, particularly between government or government-supported automated personal data systems." In 1973, the Committee argued, "… [T]he legislative history clearly indicates that such universal enumeration was not intended." Ibid.

[384] Besides tax refund fraud, other forms of fraud including credit card fraud also require the use of a person's SSN. U.S. Government Accountability Office, *Taxes and Identity Theft: Status of IRS Initiatives to Help Victimized Taxpayers* (Washington, DC: Government Printing Office, 2011), accessed November 10, 2011, http://www.gao.gov/new.items/d11721t.pdf; U.S. Government Accountability Office. *2010 Tax Filing Season: IRS' Performance Improved in Some Key Areas, but Efficiency Gains are Possible in Others.* Washington, DC: Government Printing Office, 2010. Accessed November 2, 2011. http://www.gao.gov/new.items/d11111.pdf.

[385] Melissa Oppenheim, "Why No Real Free and/or Open Source Software Movement Has Developed Around Filing Income Tax Returns: An Exploratory Analysis" (research paper, Independent Study on Internet Policy with Professor Jonathan Zittrain, Harvard University, Cambridge, MA, December, 2011); U.S. Government Accountability Office, *2010 Tax Filing Season;* U.S. Government Accountability Office, *Taxes and Identity Theft*.

[386] In 2008 there were 380,656 reported tax refund fraud incidents, 457,369 in 2009, and 975,511 in 2010. U.S. Government Accountability Office, *2010 Tax Filing Season;* U.S. Government Accountability Office, *Taxes and Identity Theft*.

[387] While the Treasury Department reports that it "prevent[ed] the issuance of $1.8 billion (97%)" of the fraudulent returns, one may question the credibility of the government's statistic due to self-reporting bias. U.S. Department of the Treasury. *Interim Results of the 2011 Filing Season*. Washington, DC: Government Printing Office, March 31, 2011. Accessed January 20, 2012. http://www.treasury.gov/tigta/auditreports/2011reports/201140032fr.pdf.

government bureaucracy that has been specifically created to monitor this issue. In fact, recovery of the lost returns cost taxpayers hundreds of millions of additional dollars to fund an entire operation of government investigators, accountants, and prosecutors dedicated to recovering the lost revenue. Thus, from the development of punched card technologies to the government's changed uses of SSNs, it is clear that unintended consequences arise when technologies that produce short-term gains are hastily adopted. After nearly fifty years of widespread implementation of SSNs, the public is only now realizing their serious consequences.

Unintended Consequences of the Data Cycle

Analogous to the unintended adoption of SSNs as a means for identification, the manifestation of the Data Cycle was certainly not anticipated, as existing law specifically sought to restrict the repurposing of individuals' collected data. Inevitably, substantial, unforeseen, and counterproductive consequences may result from the Data Cycle. What will these unintended consequences look like? What will happen?

The federal government's participation in the Data Cycle is disconcerting. Individuals' sense of privacy is diminishing as more of their information is being collected, often without their awareness, and used in ways by different parties they do not know.[388] If such status quo continues, individuals may lose full control of their personal information. As private companies and the government continue to build robust digital dossiers on individuals from personal information gathered online, one's sense of privacy in and outside of the home becomes the same.

---

[388] In 2011, the *Wall Street Journal* conducted an online "Privacy Poll," asking its readership how concerned they were about being tracked online by private companies. 84.7% of the readership voted either "somewhat concerned" or "very alarmed." "Privacy Poll," *The Wall Street Journal,* 2011, accessed February 2, 2012, http://online.wsj.com/community/groups/media-marketing-267/topics/how-concerned-you-about-advertisers; Mills, *Privacy The Lost Right,* 32-34.

Akin to how the government altered its use of SSNs, the government may change the way it handles individuals' personal information in the Data Cycle. If the government was to release more publicly collected data to private companies (and/or third parties) or release the enhanced datasets it acquires from such private companies, worrisome consequences would result. Once publicly collected information is associated with private companies' data points, the resulting dataset is much more powerful than disparate information. The government's possession of robust, detailed datasets on its residents may morph into a huge national security liability if accidentally leaked. One could imagine that an obtained dataset on individuals' health records detailed complete with names, addresses, phone numbers, and birthdays could greatly aid a terrorist organization looking to unleash a virus on a certain population. Such possibilities are endless. However, until a catastrophic event that relies on the harnessing of individuals' personal information occurs, most people will likely continue to not actively guard their personal information.

Technological Determinism and the Future of the Data Cycle

Modern American society has begun to fuse humanity with technology. We depend on and are not willing to live without digital technologies and the Internet. In February 2012, a Stanford University researcher uncovered that Google and other advertising networks were secretly collecting information from iPhone users' Internet activity. IPhone users, while enraged for a day or so, did not relinquish their iPhones or even change their behaviors on their iPhones to avoid this privacy intrusion.[389] As society

[389] On February 17, 2012, reporters from the *Wall Street Journal* revealed, "Google Inc. and other advertising companies have been bypassing the privacy settings of millions of people using Apple Inc.'s Web browser on their iPhones and computers—tracking the Web-browsing habits of people who intended for that kind of monitoring to be blocked." See Julia Angwin and Jennifer Valentino-Devries, "Google's

becomes increasingly dependent and even addicted to the neural effects of constant communication and instant accessibility to information, humans may not insist on maintaining privacy if doing so means surrendering the benefits made possible by these tools.

As individuals we must remember that the Internet and digital technologies are constituted through our use. The ways in which we decide to employ different technologies shape their function. Within the History of Science community, controversy exists surrounding the theory of "technological determinism," or the belief that changes in technology are the greatest drivers in the way a society, its processes, and its culture develop.[390] Given the scientific evidence of humans' growing addiction to connectivity, technological determinism may be directly applicable to the Data Cycle in the twenty-first century.

In general, many historians of science such as John McKay and George Daniels dismiss technological determinism as an outlandish, narrow theory. This sentiment is clearly articulated in historian Claude S. Ficsher's book *America Calling* where Fischer quotes historian George Daniels, "No single invention – and no group of them taken together in isolation from non-technological elements – ever changed the direction in which a society was going..."[391] Nevertheless, there certainly exist many historians of science who defend technological determinism and currently debate its relevance to

---

iPhone Tracking," *The Wall Street Journal*, February 17, 2012, accessed February 17, 2012, http://online.wsj.com/article/SB10001424052970204880404577225380456599176.html.

[390] Leo Marx and Merritt Roe Smith eds., *Does Technology Drive History?: the Dilemma of Technological Determinism* (Cambridge: MIT Press, 1995), 2.

[391] John McKay similarly postulates that consequences are largely socially conditioned by non-technological factors. Claude Fischer, *America Calling: A Social History of the Telephone to 1940* (Berkeley and Los Angeles: University of California Press, 1992), 9.

society.[392] Langdon Winner and mid-twentieth century philosophers Lewis Mumford, Allen Ginsberg, and Jacques Ellul each hold slightly different conceptions of the theory but still defend technology as the catalyst of cultural change.[393] For instance, in Langdon Winner's 1977 book *Autonomous Technology*, Winner offers that technology in its most basic sense is about "determining" or providing direction for a set of materials to accomplish a certain human activity.[394] Winner posits that as technologies "… become woven into the texture of everyday existence… [and] shed their tool-like qualities… [they] become part of our very humanity."[395]

Winner's description accurately describes life in the first decade of the twenty-first century. Laws against texting on cell phones have been instated because people cannot even resist connecting while driving. Historian Merrit Roe Smith's question, "Does technology drive culture?" is difficult to concretely answer in light of the ways human beings are changing both neurologically and habitually through their constant use of new technologies. Regardless of individuals' privacy concerns, the idea that we humans cannot

---

[392] A modern debate exists in the twenty-first century between historians Adrian Johns and Elizabeth Eisenstein over how the rise of print culture relates to technology's impact in society. In her *The Printing Press as an Agent of Change*, Eisenstein believes that the technology of printing transformed culture, while Johns holds in his *The Nature of the Book* that communities of people interacted with the technology in a more porous way. John Sommerville, "Review," *American Historical Review* 104, no. 5. (December 1999): 1751, accessed November 20, 2011, http://www.jstor.org/stable/2649491; Adrian Johns, *The Nature of the Book* (Chicago: University of Chicago Press,1998), 30, accessed October 24, 2010, http://books.google.com/books?id=ERpBdEUdhz8C&lpg=PP1&dq=nature%20of%20the%20book&pg=PP1#v=onepage&q=nature%20of%20the%20book&f=false.

[393] In the early 1950s, Lewis Mumford, Allen Ginsberg, and Jacques Ellul articulated their own nuanced versions of technological determinism. Mumford maintained that American society had fallen victim to "the myth of the machine," and warned that if Western civilization is to progress, it must give "…the human personality… the precedence it now gives to machines and computers." Similarly, author of the Beat Generation in the 1950s, Allen Ginsberg cautioned society from placing too much emphasis on technological capacities rather than humans. In 1954, French Philosopher Jacques Ellul discusses the disappearance of "nature" in conjunction with humans' increasing dependence on machines. Marx and Smith, *Does Technology Drive History?*, 29-32; Langdon Winner, *Autonomous Technology: Technics out of Control* (Cambridge: MIT Press, 1977), 75, accessed October 24, 2010, http://books.google.com/books?id=uNIG0gi4b40C&printsec=frontcover&dq=langdon winner&hl=en&ei=hebCTIKvF4L_8AbkveDXBg&sa=X&oi=book_result&ct=result&resnum=1&ved=0CCgQ6AEwAA#v=snippet&q=neutral&f=false., 3-5.

[394] Winner, *Autonomous Technology,* 75.

[395] Ibid., 12

separate ourselves from our digital gadgets and the Internet favors the technological determinists' view of the world.[396] As humans grow increasingly dependent on the very technologies that power the Data Cycle, privacy in 2012 is at a crossroads. While we can still tease apart humanity from technology, this distinction is growing noticeably more difficult.

Concluding Thoughts

Just this month, President Obama publicly announced his decision to see through the adoption of "new Internet privacy laws to protect consumers."[397] In a press release on February 23, 2012, Obama stated, "American consumers can't wait any longer for clear rules of the road to ensure their personal information is safe online."[398]

But what about the rules of the road protecting individuals from the ways the government repurposes their personal information? What about the rules of the road protecting individuals' publicly collected data from how the government may later trade this information with private companies?

While consumers need to be educated about how their personal information online is collected, used, and re-used by multiple entities, the government cannot point fingers at everyone else but itself. Additionally, while Obama's proposed "do-not-track" mechanism is a good start, this type of tool is hardly enforceable. There is nothing truly stopping companies whose business models depend on gathering personal information (to often do business with the government) from continuing to track individuals online. Even the government itself agreed in 2000 to stop employing trackers on its own webpages and

---

[396] Leo Marx and Merritt Roe Smith, *Does Technology Drive History?,* 1-10.
[397] Edward Wyatt, "White House, Consumers in Mind, Offers Online Privacy Guidelines," *The New York Times*, February 23, 2012, accessed February 23, 2012, http://www.nytimes.com/2012/02/23/business/white-house-outlines-online-privacy-guidelines.html?_r=1.
[398] Ibid.

yet, a couple months later some agencies were found to still be tracking their visitors.[399]

Notably, Internet privacy issues have brewed on Capitol Hill before. The Federal Trade Commission released a Privacy Report on December 1, 2010, which proposed a similar universal "do-not-track" mechanism for consumers to opt-out of online tracking.[400] This plan faced difficulty in its implementation, as "universal opt-out" tools ignore complex technological and enforcement challenges.

Even if the government required private companies to honor the voluntary tracking preferences of their online consumers, private companies and the government would still be able to engage in the Data Cycle. Obama's measures do not limit the number or types of entities that may purchase data from the private companies that "fairly" track their consumers. Thus, the Data Cycle would presumably still operate under the status quo, as the government remains able to obtain individuals' personal information from private companies.

The Privacy Act's "routine use" clause needs to be updated. Should the government update the Privacy Act by abolishing the Act's "routine use" clause to actually allow the repurposing and sharing of Americans' data with commercial parties and other federal agencies, the government could enhance its datasets more efficiently without tiptoeing around existing loopholes. Federal agencies could share data amongst each other, eliminating duplicative costs and efforts.

Yet, if the government were to decide to meaningfully revise and enforce the Privacy Act, an updated Act might directly restrict the government's ability to freely buy

---

[399] Relyea, "The Privacy Act," 5-7. For first reference, see Chapter Two, page 70.
[400] Edward Wyatt, "F.T.C. Backs Plan to Honor Privacy of Online Users," *The New York Times*, December 1, 2010, accessed December 2, 2010, http://www.nytimes.com/2010/12/02/business/media/02privacy.html?_r=1&nl=todaysheadlines&emc=a2.

and sell personal information with private companies. The Act could force governmental entities to serve notifications or consent forms to the individuals whose data they provide to third parties. The Act could also limit private companies from selling individuals' personal information to the government. A modernized Privacy Act could push the government to engage in the same exercise they wish to require of private companies – that is, provide individuals with more awareness and choice regarding their information control. Such a twenty-first century Privacy Act would directly address the Data Cycle and state which uses are permissible. An independent organization would require ample resources to then oversee the Act's implementation and enforcement.

The Data Cycle will continue unless public awareness is raised and either MCs' reelections are threatened or an agency such as the Federal Trade Commission is convinced private companies are engaging in unfair and deceptive trade practices with the government. Until that day, the government has gone rogue – employing whomever, whenever, to accomplish activities that are prohibited by law.

The tradeoffs involved in using such technologies are rarely obvious. Instead of inquiring about the cost, individuals regularly and blindly exchange personal privacy for convenience with the click of a button. Having explored the Data Cycle, its development, and strong influences such as economic, sociopolitical, legal, and neurological factors, we may arrive at an understanding of how individuals – even those who fall victim to the Cycle – are cementing its existence by clinging onto its immediate benefits. While individuals remain concerned about losing their rights to informational privacy, it is worrisome that soon they may not have a convincing argument *for* privacy, as they continue to use and rely on the very technology and interactions that are eroding their own sense of the word.

APPENDIX

A. Examples of the Data Cycle Occurring in Other Sectors

The Cycle is occurring in a number of different activities. This appendix will illuminate how the Cycle is present in the two following areas of citizen data: (1) Foreclosure Data, and (2) Pharmaceutical and Prescription Drug Data.

1. Foreclosure Data

As an unfortunate response to the "largest crash in the history of the real estate market," between 2007 and 2011 foreclosures, which deal exclusively with state laws, have become increasingly common in the current economic climate.[401]

Foreclosure data is first, recorded with state courts. As observed over the last decade, courts often give or sell this information to data analytic services and/or private companies such as, CoreLogic.

For example, as CoreLogic touts on its website, "CoreLogic has the most comprehensive repository of public, contributory and proprietary data in the United States which combines property and mortgage information; legal parcel and geospatial data; motor vehicle records, criminal background records, and tax records…"[402] Once CoreLogic obtains the foreclosure information, it generally appends this information with other privately acquired data, advertises it to targeted clients, and ultimately sells or gives it back to the government in its robust form.[403]

2. Pharmaceutical and Prescription Drug Data

Additionally, a slightly different version of the Data Cycle is occurring with regards to pharmaceutical and prescription drug information.[404]

Specifically, the government requires that citizens provide certain pieces of information in order to acquire prescription drugs. However, the government itself does not initially collect this data. Instead, the information is collected by a third party rather than by the government itself. This type of example still fits into the Data Cycle because the information mandated by the government (but initially collected by a third party) is still collected for one purpose, yet appended, shared with other parties, and repurposed.

---

[401] Roy Oppenheim and Jacquelyn K. Trask, "Deconstructing the Black Magic of Securitized Trusts: How the Mortgage-Backed Securitization Process Is Hurting the Banking Industry's Ability to Foreclose and Proving the Best Offense for a Foreclosure Defense," *Stetson Law Review* (Forthcoming Spring 2012): 2.
[402] "About Us: Our Company," CoreLogic, accessed September 29, 2011, http://www.corelogic.com/about-us/our-company.aspx.
[403] As CoreLogic's website confirms, "Clients of CoreLogic include government agencies that require real estate data." "About Us," CoreLogic.
[404] Scott Gottlieb, "Physician Prescribing Data and the Public Health: How Efforts to Thwart the Collection of Prescribing Data Undermine Drug Safety," *The Legal Pulse*, June 22, 2011, accessed October 20, 2011, http://wlflegalpulse.com/2011/06/22/physician-prescribing-data-and-the-public-health-how-efforts-to-thwart-the-collection-of-prescribing-data-undermine-drug-safety/.

For example, in its simplest form, a citizen must obtain a prescription from his or her doctor in order to obtain certain pharmaceutical drugs. IMS Health, "a data miner," buys prescription drug data in bulk from pharmacies around the country. Information about the patients is removed to preserve patient privacy (and to comply with complex laws that safeguard patient medical data).[405] Once the patients' prescription data is aggregated, "the information forms a valuable record of the drugs that are prescribed worldwide and the clinical conditions under which they are used." Federal agencies such as, the Food and Drug Administration ("FDA") either buys or is sometimes given this IMS data. In particular, the FDA uses "prescriber-level data" in three major ways: (1) measure incidence of drug side effects and prevalence, (2) target drug safety updates to doctors, (3) help the FDA mandate post-market safety programs.[406]

B. Oral History Interviews

Below are the transcripts of the professionals with which I spoke and included in this thesis.

1. Clay Johnson (Author, Technology and Transparency Advocate), phone interview by Melissa Oppenheim, November 4, 2011, 11am EST.

Introductions and Salutations.

Melissa: I'm curious as a transparency advocate, what you think about Data.Gov? What type of uses has it provided to people?

Clay: Well, one success story is the federal register, where all regulations and/or posts for comments get posted into… In 2009, the federal register's xml feeds were put on the internet on Data.Gov. The Sunlight Foundation was holding a contest "Apps for America." The Foundation said they would give $50,000 for the best app developed using Data.Gov's data. The people who came in second place, a team called Govpulse.us, took the federal register's xml feeds and made it something nice, presentable, and usable. The administration actually contacted them and said that they would like to use that software to power the official FederalRegister.Gov site. This resulted in free software getting developed for the federal government. Of course, it wasn't "free" in the end, but it was much, much cheaper for the federal government than what they would've done on their own.

Melissa: What do you think though of the government using Data.Gov to repurpose citizen data? For example, if the government collected my voter registration information for one purpose and then anonymized that data and put it online in a big dataset for anyone to download. Do you think this is an issue that people don't know this is

[405] However, the identities of individual doctors remain attached to the prescription data. Thus, once aggregated, this information is sold to drug companies. "The drug makers use the details on specific doctors, their locations, and the drugs they prescribe to target education programs, conduct studies to evaluate the safety and effectiveness of their medicines, and monitor for drug side effects. Ibid.
[406] Ibid.

happening? Especially since "anonymized data sets" are now, fairly easy to de-anonymize with a couple of data points.

Clay: If you look critically at Data.Gov, the number of actual data sets is far fewer from what Data.Gov advertises. For example, the last time I checked out the site, at least 50% of all the data hosted on there, is from one data set – the EPA's toxic release inventory. The EPA has broken one huge dataset down in hundreds – they have a dataset for every state from every year. There definitely exists a problem with incentives… Agencies give as many datasets as possible by breaking down the same dataset into multiple categories. So I'm more concerned they aren't sharing enough – not too much. All of the parsed datasets makes it very difficult for a developer because now that person has to download each set one by one. I think there is a massive misalignment of incentives between the n-user and bureaucrat inside the government. However, there are some good examples where multiple people were submitting FOIA requests for the same types of information (such as certain datasets from the Department of Labor) and now that information is hosted on Data.Gov.

I think there is a lot of promise behind Data.Gov to provide an opportunity for government transparency and utility to the public but a lot of work needs to be done. If I had a say in the matter, I'd say first and foremost incentives between the data providers and consumers need to be aligned. Second, why focus on providing data to everyone? Data.Gov should focus 100% on the technology developer community. The success of the website is not to teach people how to deal with government data, but rather to talk with developers who already have these skills to innovate using the data.

Melissa: What do you think though of the government incentivizing the repurposement of data that isn't theirs? Especially with regard to the Privacy Act of 1974, which forbids this. (i.e. foreclosures; IMS health; voter registration data)

Clay: There certainly exists a delicate balance between privacy and transparency, but right now I firmly believe that we protect privacy too much and not nearly transparent enough. Right now we lack very good redaction technologies; in fact, this is where most costs of FOIA come from. On Data.Gov, datasets do exist with black lines through some of the data entries – why post this online then? For a pat on the back? Sure, I wouldn't want to share my salary information at the dinner table, but this information for the government is useful: it helps prevent scams, fraud, and it protects ourselves.

Melissa: How do you personally think that we have historically arrived at the current imbalance you are describing between government transparency and privacy?

Clay: Well it is always a combination of things. Often cited as the two "big wins" of open government data are GPS and weather data. Without the national weather service we would have no weather data or the 1000s of apps that use this information. Without GPS we would have no Google Maps. However, how the weather data became "open" was accidental. Somehow it always seems as though the wins or success stories are never really planned. For example, the new iPhone's "Siri" is actually named after a huge government

contractor. We for sure have arrived here accidentally due to a hodgepodge of strange efforts.

Going forward I don't think we have the luxury of being so accidental. It was really bizarre that the FAA put some datasets on Data.Gov right when the site launched; no one cared about it. But, then a plane got hit by a bird, and emergency landed in the Hudson. All of a sudden everyone cared about the FAA's datasets – the bird database, and it was available online for free. This speaks to the accidentalness of this space – you never know what you want until you want it.

We have to be proactive and think critically about how this stuff works and what is successful and what isn't. For instance, the EPA's toxic release inventory data is likely not as useful as other types of data that are not being shared…

Melissa: Do you view the government's publicly collected information as a public good belonging ultimately to the public?

Clay: The Federalist Papers No. 42 has some interesting information about this. The government during that time was debating to establish post offices or not. Madison, in particular, argues that this is important. Back then, Congress was looking at post office as something that was hugely important and new – the government would ensure that people could communicate with one another.

Clay quotes, "Nothing which tends to facilitate the intercourse between the states can be deemed unworthy of the public care." In other words, if a government service facilitates transactions between nodes of people in the country, then it ought to be a public good.

Melissa: Thank you.

2. Vivek Kundra (Former Chief Information Officer in the Obama Administration), questions asked by Melissa Oppenheim in a roundtable discussion, November 29, 2011, 6pm EST, at Harvard University's Institute of Politics, Cambridge, MA.

Vivek: Private data services offer more conveniences than public ones. Right now the government doesn't have the luxury of trading privacy for better services. We need to make services more social; this is why services like Facebook are so hot right now; imagine taxes served through Facebook.

The government's NIH data helped fuel the human genome project; the government's satellite data birthed GPS. We should create a public sector "Y Combinator" [a private incubator, startup company] to leverage public sector data. Businesses doing smart things with public data should be able to profit from those activities; they are also serving the public with greater utility. Can you just imagine if all data was "open" by default? What if this was the presumption?

Melissa: How does your Y Combinator thesis deal though with the Privacy Act and individual protections from the government repurposing data?

Vivek: The government de-identifies data to comply with privacy act. There are actually a number of ways to get around privacy issues...

Melissa: Thank you.

3. Laura Quinn (CEO of Catalist), interview by Melissa Oppenheim, November 9, 2011, 6:30pm EST at the Omni Parker Hotel, Boston, MA.

Introductions and Salutations.

Melissa: Regarding the founding of the company in 2005, what were the greatest needs that the company tried to fulfill? And, how did these needs present themselves at the time and how did the founders of Catalist recognize them?

Laura: Sure. So political information starts from a list; a voter registration list. "People that are for me" – I want to find them if I am running for a position in office. To do this most effectively and efficiently, I really need to find a theory of the types of people I should spend time, energy, and funds persuading (i.e. east side versus west side populations; business associates versus senior citizens, etc.)
The way most people politically organized in the past was by collecting citizens' voter registration information. Although states vary the information they collect, they usually collect each citizen's birthday, age, party, address (but not always) – a core set of basic voter information. This information often included one's registrar records – whether or not you voted; over time you could potentially build up a voter's history and create a profile. In the past people had this set of information to work with in order to best target their resources. This was the case until 2005.

Additionally, there is the mass media. You couldn't buy a TV commercial to target just a specific person; political campaigns are always dealing with general demographics. So the campaign would use this information to buy TV space.

This is essentially how it worked since the 1950s. However, people got better and better each year at organizing and using this type of data to target demographics more effectively.

Today though, we are at a completely different place in the political world – there have been massive changes in the types and amounts of data that is available and the types and numbers of channels through which you can communicate. Changes in both of these areas have almost created a "wild west."

Further, with decreasing costs of large scale data infrastructures, almost anyone can get into this business. Before technology was affordable, really only "fortune 500" companies could afford the names and addresses of their customers to build models to attempt to predict how they can shop.

Now though, this type of targeting and modeling is affordable for civic organizing. At Catalist we don't just serve political parties; we also serve grass roots advocacy groups,

NGOs, supporting service providers, petition groups, and universities (they come in under the guise of the small "p" for "progressive"). We do, do some work with universities for academic research to make elections better. In the past, people doing these activities had to rely on early on a crude set of tools (word of mouth, polling, phone book). Now because it is cheap enough, people can afford to maintain large datasets, append them, and build mathematical and statistical models off of them. (Once a dataset is large enough, statistics are extremely useful in predicting behavior.)

Shopping and online commerce (via cookies) figured this out about 15 years ago. Regarding, Catalist's start… We were basically a bunch of progressive individuals getting together and agreeing slash complaining how each year we spend the same money over and over again to collect the same updated lists, ask similar questions, learn about you in "different way" from the campaign before, and then we end up throwing that all away at the end of the campaign (a lot of these paper datasets would end up stored in a wet basement or something). Then, two years or four years would go by, and we'd have to start all over. We wanted to stop paying the same money over and over again and eliminate the duplicative efforts.

Harold was involved because he felt the pains of the 2000 and 2004 elections – and people felt that if we just did our work a little bit better, a little bit smarter this might have made a difference and had a huge impact on the future of our country.

Melissa: In your own words how you describe what Catalist does or aims to do now?

Laura: When we started Catalist, we had the idea that we wanted to collect "civic transactions" – not just what you shop for, but also what you care about. We wanted to know how intensely and where you express what you care about through civic participation. I.e. What groups have you joined? All of these things could become a way of predicting your civic behavior.

Catalist was put together to create a utility that provides some basic data services to our clients and stores this really important information year to year. We seek to raise the quality of the conversation we are having with voters and citizens by better allocating our resources and building upon learned information.

Catalist viewed these resources as a utility that we can provide to our clients. We made Catalist for profit – owned by a trust – and wanted to be a store-house for important information. The "republican version of us" is a company called "Data Trust."
The costs of the services we offer from say, a competitor commercial provider are much higher than many of the organizations we serve could ever afford. We sell access at a flat annual fee to our clients; most of our competitors have a pricing structure that rents or re-sells the same lists. We wanted our organizations to not be penalized for using more data and becoming smarter. Our structure incentivizes people to not wait until the last possible moment to buy data: if you buy earlier you actually get a better deal – like a rental car where you get unlimited mileage and drive from Boston to Argentina – more for your buck if you do that.

Melissa: How do you try to record and/or quantify political action?

Laura: We understand that there are two parts to political action: 1. Organic action – genuine sentiment, bubbling up [i.e. the occupy movements]. 2. Harnessing power [i.e. how to harness the occupy sentiments and leverage them effectively?]

Melissa: In your opinion, before the recent rise of the "consumer data culture," how did the federal government handle its need for personal information? Was there even a (known) need for instant access to tons of personal information before the Internet or was the government's response more reactionary (everyone's doing it, we need to be involved)?

Laura: Regarding appending personal information to these basic voter information datasets, this is an extension that definitely was not contemplated but in hindsight, was inevitable. Commercial side of the world changed people's expectations of how we should similarly, communicate political information.
Knocking on all doors versus knocking on the right doors is a big deal.
Where do you draw the line though? I have no idea. As a society we haven't gripped this at all. Catalist has put rules on itself. For example, we do not store medial history data, legal record, financial information, credit history, or social security numbers. In fact, we usually don't touch these types of information at all.

Nevertheless, a judgment does need to be made to help facilitate effective political organizing. There is certainly a difference though between a data service remembering that you showed up and participated in an Avon walk versus remembering that I have cancer. People care about things that have a private expectation; things that are usually discussed behind closed doors, I believe should not be accessible.
However, this is all very murky on the commercial side.

Melissa: Doesn't Catalist buy data though from the commercial side? How do you decide what you will buy and append and what you "shouldn't" acquire?

Laura: Catalist buys from commercial vendors. We buy all of your volunteer actions, shopping actions, census demographics – long form (is about as detailed we get). It is definitely a grey area – if you are wiling to pay the price, there seems like almost nothing that isn't for sale.

Melissa: Could you explain the product of "Catalist Congressional" and how/when that particular service was started? What services and information do MCs receive for paying Catalist the standard prices of "Catalist congressional"?

Laura: This service was started c. 2008. Members of Congress ("MCs") can't access political party information. But other questions exist that get at this (i.e. environmental work, birth control advocacy, etc.). In general, MCs don't have a lot in their tool chest to buy a lot of anything… Their franked budgets are tiny compared to news coverage and media outlets. They can access all basic contact information and issue related propensities. They can tell who is registered who isn't.

Melissa: How did her company develop the relationship they have now with elected officials to provide this service? Is that service growing more popular since its introduction?

Laura: So, Lockheed Martin does the biggest work in this area because they do good lobbying. Additionally, large contractors exist in this space. Dun & Bradstreet provides personal level data services to a whole load of governmental agencies. I mean, they are so burrowed in with the government, that when large governmental agencies, such as the DHS, put out RFPs one of the requirements in the RFP is that if you are going to provide data services with us you need to use Dun Numbers. So they definitely have a pretty big monopoly over this area.

Additionally, Target Smart, Labels and Lists, VCS, PDI (California), SAGE (MA) provide a lot of data to governmental entities.

Aristotle International mostly sells to Republicans – and smaller candidates. They make most of their money by how they sell their data outside of political arena. They also sue a lot of people.

Melissa: What major laws (if any?) does her company and the "Catalist Congressional" service in particular, have to pay attention to? Does Catalist pay much attention to the Franking restrictions?

Laura: The data that we collect cannot be used for commercial purposes; our clients cannot sell things with the data we provide. Additionally, each state has a different set of laws – so we adhere to those. We think we are more scrupulous than our competitors. We spend a lot of effort making our data secure.

Melissa: What does Laura think of the Privacy Act of 1974 in regards to Catalist? How does she believe Catalist relates to the Privacy Act –if at all?
[Didn't know what it was… explained it to her.]

Laura: I think this is pretty well observed. I have seen various federal agencies buying the same lists, which is an enormous cost. For example, the IRS doesn't share with DHS.

In regards to voter registration information, the purpose that it can be used for is electioneering. We think that Catalist makes sure all its clients use it the data for this purpose.

I think that barriers have been pretty well observed – even to the detriment of the efficiency of some things that could work better and more efficiently without them.

There certainly exists an enormous amount of pressure for government to get more organized and entrenched in data field as we become more and more connected.
The biggest amounts of information actually come from yourself: self-published data. The stuff people put on Facebook (the amount people are sharing on Facebook doubles each

year). However, right now it is true that people only feeling benefit – no one has been seriously burned yet; remains to be seen if there are more nefarious sides to this industry.

Melissa: What does Laura think of her quote today that appeared in the *Washingtonian* (back in October 2008) – Does she still agree with these goals? Or has her opinion/hopes for Catalist changed at all?
[As a reference the quote from *Washingtonian*, October 2008, stated "Our hope is that Catalist becomes the living record of nearly every political action ever undertaken by an American… We're trying to build a complete record of every American over the age of 18… We aspire to be much more than just a database provider—we're looking to build an ecosystem…"]

Laura: We are truly trying to help progressive organizations see the civic personalities of the people with which they are interacting. We are not interested in sharing or showing the private or shopping personalities, but more so their civic engagement level so that the organizations we serve can make decisions or try to effect literally how these people participate in this arena.

Right now we are in a very shopping-focused paradigm. People are selling things and they don't care where they make the sell. However, political power is distributed by geography… The apportionment of power is handed out by geography. So in politics, to figure out how to change power and affect decision making is very location based… In a way that the shopping world is not – maybe that is why we like to Occupy places. What we really want is that civic picture of you.

Melissa: Thank you.

4. Jim St. George (Co-owner of NGPVAN), interview by Melissa Oppenheim, November 11, 2011, 11:30am EST, at NGPVAN Offices, Somerville, MA.

Introductions and Salutations.

Melissa: How did VAN get started and what was Jim's role?

Jim: So my partner, Mark, started the company… The idea arose out of our experience from working on political campaigns. I started working for Tom Harkin's 1992 presidential campaign. From there, I worked with the senate staff, the senate campaign, and again the senate staff.

As brief background, in the world of politics there are some people you work for and you respect what they do, but your life stinks because they are a miserable person. Then some are ordinary, and then some are actually just great people. Tom is one of those great people – he's smart, talented, successful; only person in American history to have defeated 5 sitting MCs; he beat 5 incumbents. You actually have to be a really good person and good at what you do to do that. So people who like him stick with him; he has a very loyal following.

During his 2002 reelection campaign, Harkin called Mark, the campaign manager at the time, and said that he had this vision where his campaign should collect data on palm pilots. They could use early voter data and harness this information to leverage their campaign's resources. Mark, the campaign manager, thought about it and decided to essentially go out and built it for Harkin… Mark did not have any real intention of starting a company; he was just building a tool for the campaign.

However, in Missouri, Jean Caranahan was running for reelection (her husband won even though he died a week earlier in a plane crash) – so she decided she wanted to run and that she wanted to take advantage of what was going on with this data stuff too. In order to provide her with this service, Mark had to turn his informal operation into a company and sell her the service. The interesting part of this model was that no one else was really doing anything as effective at the time…

Campaigns used to go to board of elections and/or secretary of states, get a data file, and then someone on the campaign would build a system around it and use it; this was very ad-hoc for each race and each election at every level of the spectrum (local to presidential).

A vision came out of our experiences in Iowa, Missouri, and Wisconsin – there existed an ethos that we are all democrats and that we should all cooperate and collaborate to make a difference. So we tried to build one, large database and make it available to all Democrats in one state; it was expensive to build, but cheap to share broadly. So this cut costs, time, and energy across the board once we started sharing. Our system proved genuine efficiencies and cost savings; now we had built something that was once difficult to acquire and made it easy to acquire.

Between 2001 (start of company by Mark) and 2006 we had expanded to work with 25 different state parties. The DNC stepped in and said they needed a natinoal program for Obama's presidential election in 2008. It was super interesting because in 2004 and 2006 the DNC used companies that used to be our competitors – they now no longer exist. Their systems crashed; so having a hugely reliable, flexible system is imperative to these highly important races. The DNC wanted a national, dependable database – no more crashing. They put out a RFP in 2006 – most people thought we would win the contract; we had proved we could scale, were serving a lot of states already… but it wasn't guaranteed. Nevertheless, we obtained it.

Melissa: So how exactly do you work with the DNC now?

Jim: So the DNC licenses their contract; 49 of state democratic parties distribute to democratic campaigns. (California does not…) Other entities though, like AFL CIO, America Votes (coalition of 501(c)4 groups), State Voices (coalition of 501(c)3 groups), all use our services. We now also do work with the liberal party of Canada, liberal democrats in the U.K., among some other smaller international work. Effectively, if you are a democrat or progressive you have access to our software to run your campaign through a license somewhere in the pipeline.

Melissa: What distinguishes your firm from say, Catalist, which also serves progressives and democrat groups?

Jim: Catalist is data firm. They acquire data and sell data – they write a narrative that they are the most important data firm in the world for democrats. However, they have nothing really to do with the DNC. They were modestly involved with Obama campaign… They provide voter file data to AFL, America Votes, State voices – but not to the DNC.

Their data goes into our [NGPVAN] software sometimes. Target smart, a competitor of Catalist – is actually growing bigger than Catalist; they seem to have better quality and better customer service in all 50 states…

The biggest difference though, is that NPGVAN is a software firm. We are mission based to serve the progressive community. Catalist would claim to be the same but because they get a significant amount of revenue from billionaires (George Soros) who want to engage in the political world and want to fill gaps, they don't have as much of a bottom line at actually succeeding on economic grounds…

The DNC's data comes from the DNC (however they get it) and is transferred to NGPVAN for us to process into our software. The DNC does their own hygiene checks: SSNs, address checks, living/deceased, etc. When NPGVAN receives the data from the DNC, we then do more hygiene and standardization checks.

Melissa: Regarding the founding of the company in 2002, what were the greatest needs that the company tried to fulfill and how did Jim recognize them?

Jim: One lesson that I learned studying the history of technology around 20 years ago, which I think is true and I mean I act as though it is true, is that when a new technology is developed and distributed, people use it initially to do the same thing that they were doing before that technology. It is only over time that people learn there are new things you couldn't do before with a technology. This seems to be the case with what we provide to campaigns. Initially we needed databases to create the same old stuff we always created: canvas lists, phone bank lists, etc. Now though, campaigns have better access to data and thus, are running much more data-centric campaigns – using the data in new types of ways.

NPGVAN does not guarantee you are going win if you use our software services –we don't guarantee you will even win a close election. We do believe though that at the margins you will get more votes by using our software. You will be *more likely* to win; better access to data provides you with more information about voters. You are then much more better off than having less access and less data.

We have seen great change in the industry. Many more campaigns are significantly more data focused - demographics, advertising, addresses, etc. Without giving away any proprietary information, the 2012 Obama campaign is going to have more data staff than any campaign has ever had before in an election. They believe that understanding as

much as they possibly can about the enormous amount of data we are collecting is helpful and effective.

My dad was an electrician and while growing up, if someone borrowed his tools, those tools degraded over time (for example the sharpness would dull). Our tool, NPGVAN, though works in the opposite way. The more you use it, the better it becomes with no end in sight – you literally cannot tap out of value. Every piece of information that someone collects is useful in a bigger context.

When you register to vote, you have a choice in some states (i.e. in MA) to register with a party (some states there is no party affiliation field, i.e. in Minnesota) So how then in Minnesota can we tell who the Dems are? We can tell who Dems are by canvases and getting at controversial issues (i.e. abortion rights, environmental issues, etc), coupled with demographics and geography. We then put all this together in a system called "Likely Party" and we calculate answers to guess what party these people are. All of the canvassing information – phone banks, etc. – are done by the DNC party.

Melissa: Explained the Data Cycle… Do you have any comments?

Jim: Campaigns really like knowing who has guns in the house. In some cases, you could typically go to the agency and buy a list of hunting license. We experience the value added from linking a lot of different datasets. A lot of our clients have interests in knowing commercial data, government data, etc.

Like Catalist, NPGVAN will not store social security numbers as a field [they might store it as a note next to someone's name but they won't create a field for it]. We also have to be careful when we take credit card information. We also, certainly have to adhere to state laws; some states have significant restrictions on how their voter files can be used – i.e. CA. Also it is never *our* data; we don't own the data. Data providers send us the data that we host on our services/software. We do though need to be careful how we go about doing that. We collect voter files and also voting age population files (for those individuals who are not registered but are of age).

Regarding privacy, I have a business interest in this I suppose… My opinion is that I think that campaigns and issue campaigns have a legitimate interest in data. I'm not one of those people who is worried about my name in a database – get over it; your name is in a lot of databases.

Voter registration is information that is required in order to vote. Campaigns have a legitimate case to make to access/use this data. In a democracy, we are trying to engage the public. We have to know where they are and who they are in order to effectively engage them… Should we sell this same information to Protctor and Gamble? I don't know.

Melissa: Thank you.

5. Paul Zak (Professor of Economics and Department Chair and Founding Director of the Center for Neuroeconomics Studies at Claremont University), phone interview by Melissa Oppenheim, July 16, 2010. 9:45am EST.

Introductions and Salutations

Melissa: You recently looked at social media's affects on the brain. Could you tell us about the experiment you conducted and what inspired it?

Paul: So there is this question of whether social media is good for us or not. Whether we are replacing in person interactions with online interactions… In this experiment we took blood before and after someone interacted with others through social media… What we found is that when you interact online your brain processes this interaction as if the person is in the room right next to you. The brain doesn't clearly differentiate between actual social interactions and online interactions…

We particularly looked at a neurotransmitter that connects us to other people called oxytocin… When you interact online with other people your brain releases oxytocin, it connects you to them, makes you feel empathy towards them and also reduces your stress hormones.

This is really big news because online interactions are a good substitute for in person interactions. And obviously it's much easier and faster to connect to people online than it is to meet your friends at a bar or meet your friends at a restaurant.

While this study does not suggest that we shouldn't also cultivate interactions in the real world, it does say that from the brain's perspective, online interactions through social media are almost as good.

Melissa: So from your study you were able to conclude that the brain processes e-connections that we make online using social networks the same way that the brain also processes in person connections, such as when we're having a conversation with our parents or friends. Did this shock you?

Paul: One thing that was shocking was the strength of this connection. We found in our laboratory studies that the more positive an interaction, the more someone shows they're in tune with what you're doing, the more they trust you, and the more your brain releases oxytocin. In the study that I conducted the amount of oxytocin released after using social networking was equivalent to the levels we found when a man says, "I do," at his wedding. So these are pretty strong bonding effects and I think that's valuable to know.

Again, we don't want to spend our whole lives connecting through the Internet, but given that we spend a lot of time on the Internet anyway, it's good to know that the evidence we have so far, suggests its effects are actually positive. Social networking reduces cardiovascular stress, for example.

One of the implications is that the "richness" of social interactions is likely to increase the positive impact on our brains and bodies. As we move to having not just a text, but more photos, and more online videos, the more interactive these interactions become, the more the brain will process them as if the person is right in the room with us… Our brains want to have as full of an interaction as possible…We want the richest possible experience or the highest bandwidth, if you will.

So some of the things that we don't have yet, for example, are very high resolution images so we can see clear expressions. We don't have things like smell; we don't have a good sense of body language. As these other senses are drawn in, the online experience will only get more powerful from the individual's perspective.

Melissa: I actually took a course during my sophomore year in college where we had to do the very challenge that *The New York Times* recently "dared" their readership to try, which was to go "unplugged" for 24 hours. Placing the obvious obstacles aside, do you think that these types of challenges are becoming increasingly difficult because our brains aren't necessarily used to functioning without a certain level of oxytocin for even 24 hours?

Paul: That's exactly the issue. The larger brain circuit that is activated when the brain releases oxytocin is a circuit called "HOME", Human Oxytocin Mediated Empathy circuit. The HOME circuit essentially makes social interactions rewarding. So this is the basis for a large number of "moral behaviors" or the reason why we behave in a moral way even when no one is looking. For instance, it feels "right" to pick up the wallet you found on the ground or the thousands of other things that people "don't have to do" but do because we're all connected as a big human family.

The more you have this oxytocin release, the more your brain wants it. The brain becomes used to this, as it sets down these pathways and trains itself to release oxytocin from online social interactions, as demonstrated by our study.

This can produce withdrawal symptoms when you are offline if you cannot replace all the online interactions with real world interactions, which may be difficult since we're so used to getting immediate gratification and feedback with our online interactions.

So it is not surprising that there is this withdrawal and that the HOME circuit uses the similar neurotransmitters that are released when people use drugs of abuse, like cocaine and methamphetamine. It would not surprise me if we have an actual withdrawal when we are separated from our friends by not being able to use social media.

I think it is really interesting that now all of a sudden we have trained ourselves to be hyper connected, and Facebook is certainly one of the places where that happens.

Melissa: So in some sense, we are addicted to the oxytocin that we have trained our brains to release, but on the other side, oxytocin also makes us healthier human beings. How exactly can social networking increase users' health?

Paul: Oxytocin has a positive effect on the cardiovascular system. Our study showed a reduction in stress hormones from the use of social media and that's very important to maintain for heart health and reduce the risk of stroke.

On the first level, connections of any type are healthy from the cardiovascular perspective. From the oxytocin perspective, oxytocin also reduces anxiety. It gives us a sense of wellbeing; that we're connected or part of a larger group or family. Human beings are herd creatures. We need to be part of a social group.

For example, we punish prisoners in jail by putting them in isolation. That's very stressful for humans. As we interact more on social media, we want less and less to be isolated. The downside of that is if we actually are isolated then we have a sense of withdrawal, which is the stress response.

I think the key issue if you're a Facebook user is you need to continue to be a Facebook user because your brain is used to that level of interaction. I know for myself, when I travel overseas and do not get good wireless connection, I feel isolated or unconnected to my people and my groups. There's a sense of loneliness that I don't have if I'm traveling in the U.S. or a place where I can interact in person or online with the people around me.

There's this small risk that as we are used to being in the online world that we need that online world. From the brain's perspective that's healthy as long as it's not crowding out your in person interactions. But it also means that you need to stay online.

Melissa: So, if you could convey one thing about your research to the social media community what would you like them to know?

Paul: I think the most important thing about oxytocin is what I call the "moral molecule," which motivates human beings to engage in a variety of helping behaviors, like giving to charity, sharing money with strangers, or just sacrificing something to help others.

So in a very real sense, interacting on Facebook and other social media can make us better people. I think the key issue is that carryover, the kind of good feeling you have when you connect online to the outside world. My research suggests that there's about a half an hour or so carryover time from interacting on social media where your oxytocin level is elevated.

If you want to do something fun, interact on Facebook and then go and see how you feel when you interact with the next person you see. My research suggests that, that's going to be a much more positive interaction. You'll feel more emphatic and more connected to people right in the room after you interact on Facebook. All Facebook users should participate in this challenge: interact on Facebook and then go do something good for other people or for the world. Maybe they can make the world just a teeny bit better by using social media.

Melissa: Thank you.

C. Case Study: Punched Cards and the U.S. Census

Soon after the founding of the United States, the federal government conducted the first decennial census of 1790 to get a better picture of "information about its citizens."[407] Specifically, in 1790 and 1800, the censuses were conducted more as mere "head counts" of the population, demonstrating how non-technological factors may contribute to the government's desire to collect information.[408] While there is reason to believe that the U.S. decennial census may have expanded over time due to the government's growing needs to keep increased tabs on a growing country's various interests and beliefs, it is useful to consider how the survey has been able to expand both independently of and in conjunction with the impact of the computing technologies.

As Princeton University Professor and Pulitzer Prize winning author Paul Starr argues in his 2004 book, *Creation of the Modern Media*, the "creation of the modern census" did not depend on technological innovation.[409] The census, which possesses an "ancient lineage," was a tool characteristic of state control – not public knowledge.[410] Interestingly, in ancient Rome, Starr explains how the "census" referred to "a register of adult male citizens and their property for purposes of taxation, the distribution of military obligations and the determination of political status."[411] Moreover, Starr offers how in ancient Rome the Roman censor was also in charge of controlling peoples' manners, relating the two terms "census" with "censorship."[412]

While in 1790 the U.S. became the first country to instill "a periodic census" with published statistical findings, Congress began to contemplate whether to expand the robustness of the census.[413] Congress found itself split and did not automatically support the idea of expanding the census. Specifically, dissenters questioned the government's right to extend statistical inquiries beyond mere enumeration; supporters, however, argued that the country was composed of diverse interests and groups and thus, those interests and groups needed to be better accounted.[414] The latter camp was able to persuade Congress, and thus, later censuses transformed into more comprehensive tools that sought to "better understand the nation's inhabitants, their pursuits and activities, and needs."[415] The gradual expansion of the census during this time stemmed from Congress' changing political interests, independent of the technology available.[416]

By examining an increase in census questions during periods of time without major technological advancements demonstrates how non-technological factors

---

[407] Waldo, Lin, Millet, *Engaging Privacy,* 354.
[408] U.S. Census Bureau, "Index of Questions," (Washington, DC: Bureau of the Census), accessed November 2, 2011, http://www.census.gov/history/www/through_the_decades/index_of_questions/.
[409] Starr, *The Creation of the Media,* 109.
[410] Ibid., 97.
[411] Ibid.
[412] Ibid., 97-99.
[413] Specifically, Congress debated whether or not to expand the census in regards to the questions asked and statistical inquiries. Ibid., 98
[414] Ibid.
[415] U.S. Census Bureau, "Index of Questions."
[416] Starr, *The Creation of the Media,* 98.

contributed to the government's desire to collect information. For example, the 1830 census unprecedentedly asked citizens about physical defects, while the 1840 census asked citizens about insanity.[417] The increase in such sensitive questions on the census began to cause some discomfort. Starr comments, "…[A]s more questions were added to the census, it began to represent more of an intrusion into private life."[418] By the 1840 census, "the government began to pledge that those involved in the enterprise would keep [such private] information confidential," promising that the information would not be used for surveillance.[419] Thus, the censuses continued to prompt citizens to fill out personal information. The census of 1880 asked citizens a wide array of personal questions including age, birthdate, gender, marital status, birthplace, and literacy status(es) of house members.[420]

Strong political interests in collecting more information on more citizens was certainly amplified in the late 1880s when technological advancements began to make such goals much easier to accomplish. Specifically, in the late 1880s Herman Hollerith's punched cards and tabulating machines greatly assisted the U.S. government to carry out its census surveys.[421] Hollerith's tabulation system was first employed during the 1890 census survey. His machines streamlined processing census data by aggregating information based on the patterns on the punched cards.[422] The system shortened the traditional turn around time (between the collection of information and the analysis of results) from eighteen weeks to six weeks.[423] Thus, by the late nineteenth century, the Hollerith punched card system revolutionized the government's ability to not only process census data, but also expand the survey. Those in positions of power began to increasingly employ available technology within their legal capacities to assist in carrying out specific interests – in this case, finding out more information about the American population.

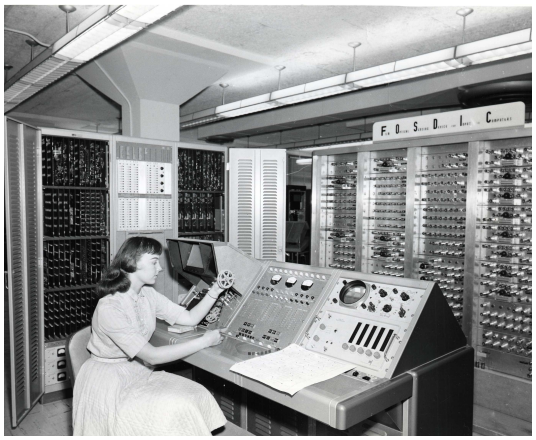

Figure 9: U.S. Census Bureau, "Image of FOSDIC," accessed November 2, 2011, http://www.census.gov/multimedia/www/photos/census_history/early_census_machines_2/machines_1960_08012.jpg.

By looking at the decennial census questions asked in the 1800s, thirty -- the most number of questions asked during this century --

[417] Ibid., 99.
[418] Ibid., 98.
[419] Ibid., 99.
[420] Waldo, Lin, Millet, *Engaging Privacy,* 357.
[421] The first punched card system grew out of problems with census processing in the U.S. in the late 19th century. Since its invention, Author Lars Heide argues that punched card technology experienced four distinct closures: 1. Counting-based census statistics (1880s-1910); 2. Processing general statistics (1894-1930s); 3. Bookkeeping (1906-1960s); 4. Operating large registers of people (1935-1960s). Heide, *Punched-card Systems*, 252.
[422] Waldo, Lin, Millet, *Engaging Privacy,* 356.
[423] Ibid.

occurred on the 1890 census.[424] Perhaps this had something to do with employing Hollerith's new technology. Although data is missing from the Census.gov Archive between the 1940 to 1970 censuses, during the twentieth century, the census was broken down into two components beginning in 1940. These components included a "short form" and "long form."[425]

For the first time in history, large mainframe machines processed almost all of the data from the 1960 census.[426] The Bureau employed a "film optical sensing device" ("FOSDIC") to convert the completed questionnaires into data on magnetic tapes for the computer then to easily read (*See Figure 9)*. This process eliminated time and expenses involved in manual entry and tabulation. The FOSDIC was able to transfer the data on the punched cards into magnetic, readable tape for the computer to process at a speed of 3,000 items per minute. Soon thereafter, subsequent versions of the FOSDIC device were able to process up to 70,000 items per minute.[427] Consequently, the number of questions since the 1960 census increased. For instance, in 1980 the census asked nineteen questions on its short form and fifty-two questions on its long form.[428]

Similarly in 1990, the census asked seventy-one questions but distributed them differently: thirteen questions on its short form and fifty-eight questions on its long form.[429] In the latest decennial census, conducted in 2010, only ten questions were asked on the short form and the long form was abolished. Instead of the Census Bureau, the annual American Community Survey now conducts the "long form" questions.[430] Thus, we can safely conclude that the mass collection and standardization of personal information by governments is not a new practice in and of itself and that non-technological factors such as strong political interests drive the government's desire to collect information.

---

[424] U.S. Census Bureau, "1890: Fast Facts. Washington," (DC: Bureau of the Census), accessed November 2, 2011, http://www.census.gov/history/www/through_the_decades/fast_facts/1890_fast_facts.html.
[425] U.S. Census Bureau, "Index of Questions."
[426] U.S. Census Bureau, "Overview: 1960," (Washington, DC: Bureau of the Census), accessed November 2, 2011, http://www.census.gov/history/www/through_the_decades/overview/1960.html.
[427] Ibid.
[428] U.S. Census Bureau, "Overview: 1980," (Washington, DC: Bureau of the Census), accessed November 2, 2011, http://www.census.gov/history/www/through_the_decades/fast_facts/1980_new.html.
[429] U.S. Census Bureau, "Overview: 1990," (Washington, DC: Bureau of the Census), accessed November 2, 2011, http://www.census.gov/history/www/through_the_decades/fast_facts/1990_new.html.
[430] U.S. Census Bureau, "Index of Questions."

BIBLIOGRAPHY

SECONDARY SOURCES

The Data Cycle

1. "A Review of Federal and State Privacy Laws." BBOnline, Inc. and the Council of Better Business Bureaus, Inc., 2006. Accessed June 23, 2011. http://www.bbbonline.org/understandingprivacy/library/fed_stateprivlaws.pdf.
   This documents describes major federal laws protecting the American's right to privacy.

2. Elliott, Justin. "GOP Rep Runs Taxpayer-Funded Campaign-Style Ads On 'Out Of Control Spending.'" *Talking Points Memo,* March 5, 2010. Accessed June 8, 2011. http://tpmmuckraker.talkingpointsmemo.com/2010/03/roskam_runs_taxpayer-funded_campaign-style_ads_on.php.
   This article describes how Rep. Peter Roskam (R-IL) ran web advertisements for an officially held "tele-town hall" financed by taxpayer dollars.

3. Graff, Garret. "They Have Your Number." *The Washingtonian,* October 1, 2008. Accessed June 20, 2011. http://www.washingtonian.com/print/articles/6/171/9627.html.
   In his article, Garret describes how campaigns and candidates are collecting and buying Americans' personal information in order to target potential voters.

4. Harris, Shane. "FBI, Pentagon Pay for Access to Trove of Public Records." *Government Executive,* November 11, 2005. Accessed June 18, 2011. http://www.govexec.com/dailyfed/1105/111105nj1.htm.
   Author Harris discusses how services like ChoicePoint, a large aggregator of public records on American citizens, is providing new types of "exclusive" searching capabilities to law enforcement and intelligence agencies. Harris details, "According to an outside expert on ChoicePoint… the exclusive service looks like something ChoicePoint built specifically for federal agencies, and the arrangement raises questions about whether the company is effectively becoming an arm of the federal government."

5. Kurtzleben, Danielle. "More Taxpayer Money Spent on Mail During Election Years." *US News,* April 8, 2011. Accessed September 5, 2011 http://www.usnews.com/news/articles/2011/04/08/more-taxpayer-money-spent-on-mail-during-election-years.
   This article describes how Congressional members often try to circumvent the limitations placed on the franking privilege in order to use their official budgets for purposes outside of publicizing their official business.

6. Lakhani, Karim, Robert Austin, and Yum Yi. "Data.gov." *Harvard Business School Case Study,* May 23, 2010. Accessed September 8, 2011. http://www.data.gov/documents/hbs_datagov_case_study.pdf.
   This case study examines how the government's online public data initiative, Data.Gov came into fruition.

7. Oppenheim, Roy and Jacquelyn K. Trask. "Deconstructing the Black Magic of Securitized Trusts: How the Mortgage-Backed Securitization Process Is Hurting the Banking Industry's Ability to Foreclose and Proving the Best Offense for a Foreclosure Defense." *Stetson Law Review* (forthcoming): 2-30.
This white paper describes how the economic crisis of 2008 gravely affected the real estate industry. The paper particularly focuses on securitized trusts.

8. Shirky, Clay. "Epilogue: Open Source Outside the Domain of Software." In *Perspectives on Free and Open Source Software,* edited by Joe Feller, Brian Fitzgerald, Scott Hissam and Karim Lakhani, 483-488. Cambridge: MIT Press, 2005.
Although Shirky is writing for a book focused on the free and/or open source software movement, Shirky's chapter touches on how data accessed by a diffuse population of individuals allows multiple people to work on projects together. Shirky comments on how digitized data is crucial to building successful, open source software.

9. Smith, Diane. "The Commercialization and Privatization of Government Information." *Government Publications Review* 12 (1985): 45-63.
In her journal article, Diane Smith identifies how the government has published more data in the public domain during 1975-1985 than in prior decades. She investigates the trend of increased government publishing from an economic and social policy perspective.

10. Solove, Daniel. "Access and Aggregation: Public Records, Privacy, and the Constitution." *Minnesota Law Review* 86 (2002): 1137-1184.
In this white paper, author Daniel Solove describes how governmental entities have begun sharing individuals' private data without their knowledge or consent.

11. "State Laws Related to Internet Privacy." National Conference of State Legislatures, March 17, 2011. Accessed April 21 2011.
http://www.ncsl.org/default.aspx?tabid=13463#isp.
This website highlights how various state legislatures have responded to privacy issues by issuing their own types of laws and mandates.

Case Study: Voter Registration Data in the Data Cycle

12. Alexander, Kim and Saskia Mills. "New Study Examines Voter Data Privacy in the US." The California Voter Foundation, May 28, 2004. Accessed September 29, 2011.
http://www.calvoter.org/news/releases/052804release.html.
Authors Kim and Mills summarize the California Voter Foundaiton's "Voter Data Privacy Study" in a press release, highlighting key statistics and conclusions.

13. "Congressional Franking Privilege, Background and Recent Legislation." *Congressional Research Service*, August 20, 2010. Accessed September 3, 2010.
http://opencrs.com/document/RS22771/.
This report describes how federal law, House and Senate rules, and committee regulations limit Congressional members' use of the franking privilege.

14. Lee, Jennifer 8. "We Hate Spam, Congress Says (Except When It's Sent by Us)." *The New York Times*, December 28, 2003. Accessed September 4, 2011. http://www.nytimes.com/2003/12/28/us/we-hate-spam-congress-says-except-when-it-s-sent-by-us.html?scp=1&sq=franking+and+internet&st=cse&pagewanted=all.
This article describes a significant loophole plauging the franking privilege policy. The loophole specifically refers to Congressional members' electronic communications and campaign-related blackout periods. The article argues that political e-mails and official business e-mails are increasingly looking more similar.

15. "Members of House Spent $45M on Outreach." iConstituent, May 5, 2010. Accessed September 20, 2011. http://www.iconstituent.com/?sectionid=26&parentid=5&sectiontree=5,26&itemid=124.
On political data company, iConstituent's website, the company discusses Congressional members' recent spending trends on franked communications. The site describes, "[L]awmakers spent the most money and sent the most communications during the fourth quarter (Oct. 1 - Dec. 31). Historically, the last quarter of a non-election year is the busiest franking period, according to the Congressional Research Service."

16. Newhauser, Daniel. "Franking on Facebook May Run Afoul of Founders." *Roll Call*, May 2, 2011. Accessed June 5, 2011. http://www.rollcall.com/issues/56_114/Franking-on-Facebook-May-Run-Afoul-of-Founders-205207-1.html.
This article is about how Congressmen of major metropolitan areas may run the risk of violating the Constitution in terms of their recent use of Facebook to run political advertisements. On Facebook, many individuals do not list their actual state of residence, but rather their current locales.

17. ———. "Facebook Raises Franking Questions." Congress.org, June 3, 2010. Accessed July 2, 2011. http://www.congress.org/news/2010/06/03/facebook_raises_franking_questions.
This article discusses how politicians' campaign Facebook pages and Congressional members' official Facebook pages are growing dangerously similar looking. The article concludes that members need to be careful about how they use their online social media presences.

18. Shesgreen, Deidre. "GOP Censors Courtney's Franked Mail on Medicare as 'Political Propaganda.'" *The Connecticut Mirror*, June 15, 2011. Accessed June 18, 2011. http://www.ctmirror.org/print/12955.
By looking at Rep. Joe Courtney's controversial use of the frank, reporter Deidre discusses the "blurriness" involved in propagating political versus official messages.

19. Verini, James. "Big Brother Inc." *Vanity Fair*, December 13, 2007. Accessed June 20, 2011. http://www.vanityfair.com/politics/features/2007/12/aristotle200712.
This article describes political data firm Aristotle's role in the voter data marketplace and provides insight into how the political data company started. Reporter Verini details, "Knowing your business is big business for Aristotle Inc., whose Orwellian database of

voter records has been an essential campaign tool for every president since Ronald Reagan."

20. Wayne, Leslie. "The 2000 Campaign: The Internet; Voter Profiles Selling Briskly As Privacy Issues Are Raised." *The New York Times*, September 9, 2000. Accessed November 29, 2011. http://www.nytimes.com/2000/09/09/us/2000-campaign-internet-voter-profiles-selling-briskly-privacy-issues-are-raised.html?pagewanted=all.
In Wayne's article, she describes how political data companies, such as Aristotle International, gather voters' information in order to compile large databases and sell such information to their clients.

21. "Web Franking 101: Overview of the Rules." Art Stiefel, 2010. Accessed July 20, 2011. http://www.slideshare.net/ArtStiefel/web-franking-101.
This slideshow, created by the private company Art Stiefel outlines the franking rules as they relate to electronic communications.

22. Wonderlich, John. "Web-Use Reform Happy Ending." The Sunlight Foundation, October 3, 2008. Accessed July 26, 2011. http://sunlightfoundation.com/blog/2008/10/03/web-use-reform-happy-ending/.
This blog post entry discusses how the House and Senate have updated their guidelines in order to cover issues of how they should use the Internet in their official capacities.

History of Government Data Collection

23. Heide, Lars. *Punched-card Systems and the Early Information Explosion 1880-1945*. Baltimore: Johns Hopkins University Press, 2009.
Lars Heide charts the historical development, implementation, and implications of punched card systems employed globally. From the U.S. Census to tracking Jews and political dissidents in Europe during World War II, Heide discusses how punched cards were employed by different cultures in often unintended ways.

24. Lockwood, Brad. *Domestic Spying and Wiretapping.* New York: The Rosen Publishing Group, 2007. Accessed April 24, 2011. http://books.google.com/books?id=WoiNocUQRNkC&pg=PA16&dq=domestic+spying+to+new+heights&hl=en&ei=eZavTcKmKYP30gGS9oCgCQ&sa=X&oi=book_result&ct=result&resnum=1&ved=0CD4Q6AEwAA#v=onepage&q=domestic%20spying%20to%20new%20heights&f=false.
In his book on modern government surveillance, Lockwood describes how the U.S. government has historically possessed the most information concerning its citizens.

25. Lyon, David. *The Electronic Eye: The Rise of Surveillance Society*. Minneapolis: University of Minnesota Press, 1994.
David Lyon discusses in *The Electronic Eye* how government surveillance is not a new phenomenon. He explores the increasingly fuzzy distinction between banal and/or necessary surveillance and intrusive surveillance in the 1990s.

26. Starr, Paul. *The Creation of the Media: Political Origins of Modern Communications*. New York: Basic Books, 2004.
Starr discusses the history of global modern communication and global media with respect to non-technological factors, such as the political decisions being made during the twentieth century.

History of Computing Technologies and the Internet

27. Campbell-Kelly, Martin, and William Aspray. *Computer: a History of the Information Machine*. New York: Basic Books, 1996.
Historians Campbell-Kelly and Aspray provide a thorough recounting of how the modern computer was developed and distributed. They highlight how the computer continuously provoked innovation over the course of the twentieth century. The "Introduction" as well as the chapters, "Inventing the Computer," and "New Modes of Computing" are especially insightful concerning the development of computing technologies and cultural understandings.

28. Dick, Stephanie. "History of Computing" (lecture, Sever Hall 306, Harvard University, Cambridge, MA, February 9, 2011).
In a lecture given by guest speaker, Stephanie Dick, Dick discussed how mainframe computers were created, how they worked, how they were being employed by certain parts of society that could afford them in the mid-1990s.

29. Godse, A.P., and D.A. Godse. "Basic Structure of Computer." In *Computer Organization and Architecture*, 1-18. India: Technical Publications Pune, 2008.
In this book, the authors give a technical description of how the architecture between first, second, and third generation computers differ.

30. "Highest Ram Amount, Cell Phones Top List." PhoneEgg. Accessed January 5, 2012. http://www.phonegg.com/Top/Ram-Amount-Cell-Phones.html.
This website lists modern day cell phones and ranks them according to how much RAM (or memory) they each have. This is useful in comparing today's average cell phone memory capacities to mainframe computers' memory capacities from the mid-twentieth century.

31. Lécuyer, Christophe. "Revolution in Silicon." In *Making Silicon Valley: Innovation and the Growth of High Tech, 1930-1970*, 129-168. Cambridge: MIT Press, 2006.
In the chapter "Revolution in Silicon," Lécuyer describes how the semiconductor industry, with Fairchild Semiconductor at its forefront, took off in the 1950s and 1960s due in part to support by the U.S. military.

32. Mowery, David, and Timothy Simcoe. "Is the Internet a U.S. invention? An Economic and Technological History of Computer Networking, Research Policy." *Science Direct* 31 (December 2002): 1369-1387. Accessed October 24, 2011. http://www.sciencedirect.com/science/article/pii/S0048733302000690.
This white paper provides a useful account of how the Internet was created in the United States.

33. "Predecessors: The Differential Analyzer." In *The Case Files: John W. Mauchly and J Presper Eckert*. The Franklin Institute. Accessed January 7, 2012. http://www.fi.edu/learn/case-files/eckertmauchly/analyzer.html
On this website, the Franklin Institute details Vannevar Bush's development of the Differential Analyzer and notes how it was "essential to the solving of differential equations prior to the advent of ENIAC." The website makes available correspondence from May 1948 between Dr. John von Neumann to Mr. Frazer, asking for opinions about the ENIAC.

34. Shapin, Steven. "Science and Business in Modern America" (lecture, Science Center, Harvard, University, Cambridge, MA, March 30, 2009).
In Professor Steven Shapin's lecture on March 30, 2009, Professor Shapin discussed how the U.S. government played a major role in propelling research and development in the computing hardware and design industries.

35. Stevens, Hallam. "Nanocultures" (lecture, Sever Hall 306, Harvard University, Cambridge, MA, February 23, 2011).
In this lecture of Professor Stevens' course, Nanocultures, Stevens discussed the creation and mobilization of the electronics industry.

36. "Timeline of Computer History." Computer History Museum, 2006. Accessed November 7, 2011. http://www.computerhistory.org/timeline/?category=cmptr.
Hosted by the Computer History Museum, this website offers an interactive timeline of major developments in computing.

37. "Vacuum Tubes and Flip-Flops." In *The Case Files: John W. Mauchly and J Presper Eckert*. The Franklin Institute. Accessed January 7, 2012. http://www.fi.edu/learn/case-files/eckertmauchly/vacuum.html.
The Franklin Institute describes how Mauchly and Eckert used vacuum tubes in the ENIAC machine. Of interest, the website includes original images of the ENIAC and describes how vacuum tubes worked.

38. "Validation." In *The Case Files: John W. Mauchly and J. Presper Eckert*. The Franklin Institute. Accessed January 7, 2012. http://www.fi.edu/learn/case-files/eckertmauchly/valid.html.
On this website, the Franklin Institute makes available images of various letters, medals, and awards that Mauchly and Eckert received after having completed the ENIAC.

The Privacy Act of 1974

39. Henning, Anna. "Government Collection of Private Information: Background and Issues Related to the USA PATRIOT Act Reauthorization." *Congressional Research Service*, March 2, 2010.
This document produced by the Congressional Research Service, an organization that works for the U.S. Congress, describes how the Patriot Act changed the legal and political landscape in the twenty-first century.

40. Kardon, Alex. "Damages Under the Privacy Act: Sovereign Immunity and a Call for Legislative Reform." *Harvard Journal of Law & Public Policy* 34, no. 2. (2011): 705-767.
In his white paper, graduate of Yale Law School Alex Kardon discusses how the Privacy Act of 1974 has become antiquated. He argues that individuals should be allowed to claim damages related to harms caused by the mistreatment of personal information, such as identity theft and widespread reputational harm.

41. Kimbol, Anna. "The Privacy Act May be Toothless." *Health Law Perspectives* (September 2008). Accessed December 2, 2011. http://www.law.uh.edu/healthlaw/perspectives/homepage.asp.
Anne S. Kimbol exemplifies how the Privacy Act of 1974 has been rendered toothless through considering the case *Cooper v. Federal Aviation Association.* After examining this case, Kimbol highlights that one must have more than "emotional distress" to successfully prove that the government violated the Privacy Act.

42. Leary, Warren. "Congress's Science Agency Prepares to Close Its Doors." *The New York Times*, September 24, 1995. Accessed December 8, 2011. http://www.nytimes.com/1995/09/24/us/congress-s-science-agency-prepares-to-close-its-doors.html?scp=1.
This *New York Times* article describes why the Office of Technology Assessment was shut down in 1995. The office was closed mainly due to budget cuts to eliminate duplicative efforts.

43. "Updating the Privacy Act of 1974." The Center for Democracy and Technology, June 5, 2009. Accessed September 10, 2011. http://www.cdt.org/policy/updating-privacy-act-1974.
This website features a policy analysis by the Center for Democracy and Technology on how the Privacy Act of 1974 should be updated and made relevant to the twenty-first century.

44. Westin, Alan. "Social and Political Dimensions." *Journal of Social Issues* 59, no. 2 (2003): 1-35.
In this article, Alan Westin defines "information privacy" and analyzes how this notion has been thought of differently by modern American society during each decade since the 1960s. Westin examines the relationships between Americans and the U.S. Government, consumers and businesses, and employees and employers.

The Data Economy

45. Andrews, Lori. "Facebook Is Using You." *The New York Times,* February 4, 2012. Accessed February 10, 2012. http://www.nytimes.com/2012/02/05/opinion/sunday/facebook-is-using-you.html?_r=1&ref=opinion&pagewanted=all.
In this opinion piece, Law School Professor Lori Andrews discusses how governmental entities use individuals' personal information to make judgments that could be considered discriminatory. She discusses how Americans were skeptical of this practice in the 1960s and 1970s and argues in the twenty-first century that this practice still exists.

46. Angwin, Julia. "The Web's New Gold Mine: Your Secrets." *The Wall Street Journal*, July 30, 2010. Accessed August 2, 2010. http://online.wsj.com/article/SB10001424052748703940904575395073512989404.html.
Angwin's article demonstrates how data mining is becoming an increasingly popular and lucrative business practice even though most consumers do not know it is happening.

47. Harper, Jim. "Why Online Tracking Isn't Bad." *The Wall Street Journal*, August 7, 2010. Accessed November 9, 2010. http://online.wsj.com/article/SB10001424052748703748904575411530096840958.html.
Director of Information Policy Studies at CATO, Jim Harper posits in his article that online tracking is useful and beneficial to users. Harper does not believe that Web tracking is necessarily privacy intrusive.

48. Farmer, Dave. "Digital Information Growth Outpaces Projections, Despite Down Economy." *EMC*, May 18, 2009. Accessed November 19, 2010. http://www.emc.com/about/news/press/2009/20090518-01.htm.
This report discusses the rate of generated information online and how its trend continues to increase.

49. Lazarus, David. "Facebook Users Want a Piece of the Action." *LA Times,* February 2, 2012. Accessed February 18, 2012. http://www.latimes.com/business/money/la-fi-mo-money-minute-20120202,0,1944277.story.
In this article (and video), reporter David Lazarus describes Facebook users' desire to acquire a piece of Facebook's lucrative initial public offering. Because Facebook profits off of individuals' continued use of the platform and the information individuals share with each other, Facebook users have voiced that they want a piece of the company's revenue.

50. Oppenheim, Melissa. "From a Business Perspective: Terms of Service." Research Paper, Independent Study on Social Media and Politics with Professor Nicco Mele, Harvard University, Cambridge, MA, April 2010.
I wrote a paper in the spring of 2010 looking at "Terms of Service" contracts between online services and consumers. I tackled the issue from a business perspective, attempting to identify the main problems facing the industry and offered solutions as to how the industry could improve.

51. Schonfeld, Erick. "The Privacy Dilemma." *TechCrunch*, January 28, 2009. Accessed November 18, 2010. http://techcrunch.com/2009/01/28/the-privacy-dilemma/.
This article on *TechCrunch* discusses the value inherent in consumer-produced information. It also discusses the danger of sharing online. Schonfeld writes, "The more of our lives that we put online, the less privacy we have."

52. Sengupta, Somini, and Evelyn Rusli. "Personal Data's Value? Facebook is Set to Find Out." *The New York Times*, January 31, 2012. Accessed January 31, 2012.

http://www.nytimes.com/2012/02/01/technology/riding-personal-data-facebook-is-going-public.html?pagewanted=1&_r=1&emc=eta1.
This article discusses how Facebook's initial public offering will be one of the biggest in history – notwithstanding the fact that the company will be profiting from their users' data shared on the site. The article details how Facebook makes money from targeted adverts and plug-ins. The reporters state, "More than the world's largest social network, it is a fast-churning data machine that captures and processes every click and interaction on its platform."

Individuals' Online Sharing Habits

53. Anderson, Janna, and Lee Rainie. "Millennials Will Make Online Sharing in Networks a Lifelong Habit." Pew Research Center's Internet & American Life Project, July 9, 2010. Accessed November 18, 2010. http://www.pewinternet.org/Reports/2010/Future-of-Millennials/Overview.aspx.
This study concludes that "millennials" view online sharing as an integral part of their lives and will most likely never stop communicating with each other online.

54. Auletta, Ken. "Media Maxims." Scribd. Accessed November 19, 2010. http://www.scribd.com/doc/22564045/Ken-Auletta-Media-Maxims.
In his paper, Ken Auletta outlines "media maxims" or pieces of advice on how successful, twenty-first century business have been created. He compares the rapid adoption of the Internet to predecessor technologies.

55. "Debate: Social Networking." *The Economist*, February 8, 2012. Accessed February 5, 2012. http://www.economist.com/debate/debates/overview/222.
On February 8, 2012, *The Economist* launched a poll among its readership asking if they believed sharing information online benefits society. The poll overwhelmingly demonstrates that the participants do.

56. Gahran, Amy. "Report: 90% of Americans Own a Computerized Gadget." *CNN*, February 3, 2011. Accessed February 15, 2012. http://articles.cnn.com/2011-02-03/tech/texting.photos.gahran_1_cell-phone-landline-tech-gadget?_s=PM:TECH.
In this article, reporter Amy Gahran announces that 90% of Americans own a computerized gadget of some kind as of February 2011. She goes on to detail the types of most popularly owned devices by specific demographics.

57. Magid, Larry. "Study: 92% of U.S. 2-year-olds have online record." *CNET News*, October 6, 2010. Accessed October 8, 2010. http://news.cnet.com/8301-19518_3-20018728-238.html.
This article describes how Americans love to share information online, including photos of their infants.

58. Mollman, Steve. "How Can We Cope with Information Overload?" *CNN*, February 3, 2010. Accessed November 18, 2010. http://articles.cnn.com/2010-02-03/tech/content.overload_1_content-rss-algorithm?_s=PM:TECH.
Mollman's *CNN* article discusses how Americans today often feel overwhelmed by the

large amounts of content on the Web.

59. "More Teens Are Creating and Sharing Material on the Internet." Pew Research Center's Internet & American Life Project, December 19, 2007. Accessed November 30, 2010. http://www.pewinternet.org/Press-Releases/2007/More-teens-are-creating-and-sharing-material-on-the-internet.aspx.
This study demonstrates how American teens are sharing and interacting online.

60. Saul, Hansell. "Zuckerberg's Law of Information Sharing." *The New York Times*, November 6, 2008. Accessed February 10, 2012. http://bits.blogs.nytimes.com/2008/11/06/zuckerbergs-law-of-information-sharing/.
In this article, Saul Hansell describes Facebook's CEO Mark Zuckerberg's projection about the future of individuals' online sharing. Zuckerberg concludes that on average people are sharing twice as much than the year before, every year.

61. Shirky, Clay. "It's Not Information Overload. It's Filter Failure." Web 2.0 Expo, September 16-19, 2010. Accessed November 18, 2010. http://web2expo.blip.tv/file/1277460/.
This video excerpt is of Clay Shirky giving a speech at the Web 2.0 Expo, a conference held on Internet innovations and implications, on information overload and privacy in the Information Age.

62. "The Digital Life: Study." TNS. Accessed September 18, Sept. 2011. http://discoverdigitallife.com.
This project, hosted online, offers an abundance of data on individuals' global and domestic online sharing habits. The studies expose how humans all over the globe are using, sharing, and interacting with one another on the Internet.

63. "The Future of Online Socializing." Pew Research Center's Internet & American Life Project, July 2, 2010. Accessed October 10, 2011. http://pewresearch.org/pubs/1652/social-relations-online-experts-predict-future.
This piece of research suggests how Americans will share over the Internet in the future.

64. "World Internet Usage Statistics News and World Population Stats." Internet World Stats. Accessed February 15, 2012. http://www.internetworldstats.com/stats.htm.
This website provides up-to-date statistics on the world's usage of the Internet. The site offers Internet penetration rates by region, year, and country.

Online Tracking

65. Angwin, Julia. "Online Interview with Julia Angwin." *The Wall Street Journal*, November 11, 2010. Accessed November 19, 2010. http://online.wsj.com/article/SB10001424052748703848204575608970171176014.html?mod=djemalertNEWS#video=49C70BB6-E963-4994-802B-065655247A67&articleTabs=video.
This video-interview highlights how website publishers are growing skeptical of granting third-party data trackers access to individuals' data on their sites.

66. Davis, Wendy. "The Forever Cookie: New Tracking Technologies Continue To Threaten Privacy." *MediaPost,* October 11, 2010. Accessed November 19, 2010. http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=137459.
This article discusses various cookie technologies, highlighting the "evercookie," which is more powerful than other types of cookie trackers.

67. Richmond, Riva. "Resisting the Online Tracking Programs." *The New York Times*, November 10, 2010. Accessed November 15, 2011. http://www.nytimes.com/2010/11/11/technology/personaltech/11basics.html?src=un&feedurl=http%3A%2F%2Fjson8.nytimes.com%2Fpages%2Ftechnology%2Fpersonaltech%2Findex.jsonp.
Richmond's *New York Times* article discusses the advertising industry's aggressive efforts to track consumers online. The article highlights how it has become difficult to remove cookies and other trackers permanently from one's computer.

68. Tynan, Dan. "21 Things 'They' Don't Want You to Know." *PCWorld*, March 29, 2010. Accessed November 19, 2010. http://www.pcworld.com/article/191312-6/tech_secrets_21_things_they_dont_want_you_to_know.html.
This report highlights twenty-one revealing facts that consumers may not have known about the technology they use everyday.

69. Vega, Tazina. "Code That Tracks Users' Browsing Prompts Lawsuits." *The New York Times*, September 20, 2010. Accessed November 19, 2010. http://www.nytimes.com/2010/09/21/technology/21cookie.html?pagewanted=1&_r=1.
Tazina's article describes the increasing trend of upset, disturbed consumers suing online tracking companies for deceptively collecting their data.

70. "What Are Computer Cookies?" *WiseGEEK.* Accessed November 9, 2010. <http://www.wisegeek.com/what-are-computer-cookies.htm>.
This website explains how different types of popular Web trackers, such as cookies, work. The site lists and differentiates between the various kinds of trackers.

Online Privacy

71. Angwin, Julia. "Obama Administration Seeks Internet Privacy Protections, New Policy Office." *The Wall Street Journal*, November 11, 2010. Accessed November 18, 2010. http://online.wsj.com/article/SB10001424052748703848204575608970171176014.html?mod=djemalertNEWS.
In this article, reporter Julia Angwin discusses the White House's increased efforts and new strategy to protect consumer privacy online.

72. Angwin, Julia and Tom McGinty. "Sites Feed Personal Details to New Tracking Industry." *The Wall Street Journal,* July 30, 2010. Accessed November 2, 2011. http://online.wsj.com/article/SB10001424052748703977004575393173432219064.htm

l.
In this article, reporters Angwin and McGinty expose how online websites are harvesting and selling individuals' shared information.

73. Auletta, Ken. "Authors@Google: Ken Auletta." YouTube, November 20, 2009. Accessed November 18, 2010. <http://www.youtube.com/watch?v=qVMFdsfTzb0>.
This video is a recording of Ken Auletta's talk at Google about his new book's debut, *Googled.* He speaks about issues of consumer trust, privacy, and corporate responsibility.

74. Bennett, Jessica. "Online Privacy: What the Internet Knows About You." *Newsweek*, October 22, 2010. Accessed November 18, 2010. http://www.newsweek.com/2010/10/22/forget-privacy-what-the-internet-knows-about-you.html.
This article discusses the public's eroding sense of privacy online, quoting Nicholas Carr, author of *The Shallows*.

75. Cameron, Chris. "Pew: 71% of Young Adults Change Online Privacy Settings." *ReadWriteWeb*, May 26, 2010. Accessed June 18, 2011. http://www.readwriteweb.com/archives/pew_71_of_young_adults_change_online_privacy_settings.php.
Cameron's article highlights findings from the PEW Research Center's Study on how often young adults change their privacy settings online.

76. Harper, Jim. "It's a WikiLeaks World, Get Used to It." CATO Institute, August 5, 2010. Accessed April 20, 2011. http://www.cato.org/pub_display.php?pub_id=12035&utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+CatoHomepageHeadlines+(Cato+Headlines).
In this article Jim Harper discusses how secrecy is becoming a growing weakness for organizations, such as the military. He proffers that controlling information has become harder and harder to do.

77. Holtzman, David. *Privacy Lost: How Technology Is Endangering Your Privacy*. San Francisco: Jossey-Bass, 2006.
This is a book about what privacy is, how it is changing, and what risks certain technologies pose. In particular, this book discusses how privacy is hard to actually define and that advancing technologies are altering how privacy is experienced.

78. Kerry, Cameron and Christopher Schroeder. "White House Council Launches Interagency Subcommittee on Privacy & Internet Policy." TheWhiteHouse.Gov, October 24, 2010. Accessed November 18, 2010. http://www.whitehouse.gov/blog/2010/10/24/white-house-council-launches-interagency-subcommittee-privacy-internet-policy.
This press announcement discusses the Obama Administration's creation of a new Subcommittee on Privacy and Internet Policy under the National Science and Technology Council.

79. King, Cecilia. "Commerce Dept. Weighs Privacy Policy Guidelines." *Washington Post*, November 16, 2010. Accessed November 18, 2010. http://voices.washingtonpost.com/posttech/2010/11/the_commerce_department_which.html.
In this article, reporter Cecilia King discusses the U.S. Commerce Department's draft of broad, online privacy recommendations.

80. ———. "Is Internet Privacy Dead?" *Washington Post*, March 15, 2010. Accessed October 10, 2011. http://voices.washingtonpost.com/posttech/2010/03/is_internet_privacy_dead_no_ju.html.
This article discusses the ironic phenomenon that a lot of people use the Internet to share personal information but then are upset when they find that information is being used for other purposes. The article highlights a PEW study demonstrating that "most people said they cared greatly about online privacy but they didn't do much about it."

81. Kirn, Walter. "Little Brother is Watching." *The New York Times*, October 15, 2010. Accessed November 3, 2011. http://www.nytimes.com/2010/10/17/magazine/17FOB-WWLN-t.html.
In Kirn's article, he writes about how unlike George Orwell's *1984* conception of "Big Brother," in the twenty-first century there exists many "little brothers." Kirn is referring to ease and ability for almost anyone online to invade another individual's sense of privacy.

82. Lewis, Harry, Hal Abelson, and Ken Ledeen. *Blown to Bits: Your Life, Liberty and the Pursuit of Happiness after the Digital Explosion*. Crawfordsville: Addison-Wesley Professional, 2008. Accessed November 18, 2011. http://www.bitsbook.com/excerpts/.
This book describes the "digital revolution" and how technology is changing the way humans interact with each other and in society in general. The book has an in-depth review of how accessibility alters our conception of "public information." It discusses the "digital revolution" and how privacy is changing. The book is licensed under the "Creative Commons" so it is available for free PDF download.

83. Madden, Mary, and Aaron Smith. "Reputation Management and Social Media." PEW Internet Reports, May 26, 2011. Accessed August 20, 2011. http://pewinternet.org/Reports/2010/Reputation-Management/Summary-of-Findings.aspx?r=1.
The PEW Research Center conducted a study on how Internet users in the U.S. view their reputations online. The report also highlights social norms concerning Americans' social media use.

84. Rainie, Lee. "Trust and Privacy Online: What the Public Really Wants." PEW Research Center, December 14, 2002. Accessed October 2, 2011. http://www.pewinternet.org/Presentations/2002/Trust-and-Privacy-Online.aspx.
This study looks at how Americans' view disclosing different types of information online.

85. Siegel, Robert, and Melissa Block. "Cybersecurity Expert On China Net Hijacking." *NPR*, November 18, 2010. Accessed November 21, 2010. http://www.npr.org/templates/story/story.php?storyId=131423973.
This report describes China's hijacking of Internet traffic that occurred on April 8, 2010, for a little less than 20 minutes. The ability to re-route traffic on the Internet and wiretap sensitive information poses a huge security issue to the Internet's fundamental design. Further, China's hijacking of the Web poses additional privacy issues as to how individuals openly share information online.

86. Solove, Daniel. *The Digital Person: Technology and Privacy in the Information Age*. New York: NYU Press, 2004.
This book describes how Americans share an increasing amount of information online, allowing various parties to collect this information. Solove argues that this information is used to create "digital dossiers" on individuals. He discusses many privacy-related issues that the everyday person may not know about while surfing the Web and offers some solutions.

87. Sprenger, Polly. "Sun on Privacy: 'Get Over It.'" *Wired*, January 26, 1999. Accessed October 24, 2010. http://www.wired.com/politics/law/news/1999/01/17538#ixzz1380e4KYp.
This piece in *Wired* describes how in 1999, then CEO of Sun Microsystems Scott McNealy publicly stated that consumers have "zero privacy" and "to get over it." Allegedly, McNealy made the comment in response to fielding a question about what privacy safeguards the company was considering for a new technology product that would allow different types of consumer devices to communicate and share processing resources with one another.

88. Stone, Brad. "Amazon Erases Orwell Books From Kindle." *The New York Times*, July 17, 2009. Accessed April 20, 2011. http://www.nytimes.com/2009/07/18/technology/companies/18amazon.html.
Brad Stone discusses in this *New York Times* article how Kindle readers who had purchased *1984*, was, paradoxically, removed from their Kindle devices due to licensing issues. Although Amazon admitted it was a bad idea to reclaim the book, the company erased the book (and refunded the customers) without many individuals' consent or knowledge. For example, Stone describes, "In a move that angered customers and generated waves of online pique, Amazon remotely deleted some digital editions of the books from the Kindle devices of readers who had bought them."

89. Waldo, James, Herbert S. Lin, and Lynette I. Millet, eds. *Engaging Privacy and Information Technology in a Digital Age*. Washington, DC: The National Academies Press, 2007. Accessed May 20, 2011. http://books.nap.edu/openbook.php?record_id=11896&page=R1.
The National Academy of Sciences, Academy of Engineering, and the Institute of Medicine, supported by the U.S. government, worked to put together a book on privacy and information issues in the twenty-first century. This book thoroughly explains many of the privacy tradeoffs individuals make while accessing information and communicating

online. The book analyzes prior legislative measures aimed at addressing privacy and security issues in the Information Age.

90. Wyatt, Edward. "F.T.C. Backs Plan to Honor Privacy of Online Users." *The New York Times*, December 2, 2010. Accessed December 3, 2010. http://www.nytimes.com/2010/12/02/business/media/02privacy.html?_r=1&nl=todaysheadlines&emc=a2.
This article discusses the Federal Trade Commission's Privacy Report, released on December 1, 2010. The Report overwhelming sides with consumers, calling for companies that employ Web tracking tools to offer an opt-out mechanism.

91. ———. "White House, Consumers in Mind, Offers Online Privacy Guidelines." *The New York Times*, February 23, 2012. Accessed February 23, 2012. http://www.nytimes.com/2012/02/23/business/white-house-outlines-online-privacy-guidelines.html?_r=1.
In this article, reporter Wyatt summarizes President Obama's announcement to offer consumers more privacy controls online. Specifically, the White House intends to compel private companies to provide individuals online with an "opt-out" tracking mechanism.

Countercultural Perspectives

92. Mills, Jon. *Privacy The Lost Right*. New York: Oxford University Press, 2008.
This book describes the privacy debate in the U.S. that has been occurring since the 1890s. Mills highlights Americans' changing privacy concerns over the last century.

93. Strum, Philippa, Gerald Nash, and Richard Etulain. *Privacy, the Debate in the United States since 1945*. Fort Worth: Harcourt Brace College, 1998.
This book offers a historical account of privacy issues in the United States during the mid- to late- 1990s. The book possesses a particularly useful chapter on personal information stored in government databases.

94. Turner, Fred. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. Chicago: University of Chicago Press, 2008.
In Turner's book, he explains how in the 1960s a certain sector of society feared computers, as they considered to have produced the military industrial complex. Turner contrasts the existing fears held in the 1960s with the held beliefs of a "digital utopia" that later emerged in the twentieth century.

Political Interests in Data Collection

95. Fenno, Richard Jr. "U.S. House Members in Their Constituencies: An Exploration" *The American Political Science Review* 71, no. 3 (1977): 883-917. Accessed October 7, 2010. http://links.jstor.org/sici?sici=00030554%28197709%2971%3A3%3C883%3AUHMITC%3E2.0.CO%3B2-1.
In this paper, Richard Fenno explores the questions, "What does a an elected representative see when he or she sees a constituency?" and, "What consequences do these perceptions have for his or her behavior?" Fenno purports that Members of

Congress seek to pursue goals such as good public policy and influence, but that they are also preoccupied with issues of reelection.

96. Scott, James. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press, 1998.
In Scott's book *Seeing like a State*, he considers a wide arrange of large-scale, big-government regimes and analyzes their failed outcomes. I used this source as a means to investigate the political motives driving government data collection projects.

97. Smith, Steven, Jason Roberts, and Ryan Vander Wielen. *The American Congress.* New York: Cambridge University Press, 2009. 6th edition.
This book provides a survey of the history of the U.S. Congress and how it currently functions. The book discusses contemporary studies that have focused on better understanding the motives driving Members of Congress to take certain actions while in office.

The Neurological Effects of Digital and Communication Technologies

98. Carr, Nicholas. "Is Google Making Us Stupid?" *The Atlantic Monthly*, July 2008. Accessed November 2, 2010. http://www.theatlantic.com/magazine/archive/2008/07/is-google-making-us-stupid/6868/.
This article discusses Carr's theory of how the Internet is teaching our brains to become accustomed to a certain practice of constantly wanting and finding information.

99. ———. *The Shallows: What the Internet Is Doing to Our Brains*. New York: W.W. Norton, 2010.
Carr discusses how he believes that our brains are neurologically changing due to our Internet usage.

100. "Digital Overload: Your Brain On Gadgets." *NPR*, August 24, 2010. Accessed February 15, 2012. http://www.npr.org/templates/story/story.php?storyId=129384107.
This *NPR* article describes recent scientific studies on how the Internet and constant connectivity is effecting the brain. The article comments on the work produced by Matt Richtel for *The New York Times.*

101. Powers, William. *Hamlet's Blackberry: a Practical Philosophy for Building a Good Life in the Digital Age*. New York: Harper, 2010.
This book written by Powers strives to offer a healthy philosophy for balancing screen time and real life. Powers describes the power of the crowd and how it seems often hard to turn it off.

102. Shirky, Clay. "Does the Internet Make You Smarter?" *The Wall Street Journal*, June 4, 2010. Accessed October 24, 2010.
http://online.wsj.com/article/SB10001424052748704025304575284973472694334.html.
In this article, Shirky counters Carr's argument by advocating that the Internet actually makes mankind smarter via the "cognitive surplus." Instead of merely consuming media,

Shirky argues that we now have the power to create it ourselves.

The Blurring of "Private" and "Public" Conceptions

103. Angwin, Julia. "Privacy Advocate Withdraws from RapLeaf Advisory Board." *The Wall Street Journal*, October 24, 2010. Accessed November 6, 2010. http://blogs.wsj.com/digits/2010/10/24/privacy-advocate-withdraws-from-rapleaf-advisory-board/?KEYWORDS=rapleaf.
This article announces how a board member of RapLeaf, a consumer-tracking firm, left the company after RapLeaf's controversial data collecting practices were publicly exposed.

104. CNN Wire Staff. "Brash Facebook Posts Lead to Bank Heist Arrests." *CNN,* April 22, 2011. Accessed April 23, 2011. http://www.cnn.com/2011/CRIME/04/22/texas.heist.facebook/.
This article describes how individuals' Facebook status updates were used by law enforcement to arrest two individuals.

105. Hafner, Katie, and Matthew Richtel. "Google Resists U.S. Subpoena of Search Data." *The New York Times*, January 20, 2006. Accessed November 14, 2010. http://www.nytimes.com/2006/01/20/technology/20google.html?_r=1.
This article describes the tension between the U.S. government's desire for information and private data companies' wish to please their consumers. Specifically, this article highlights a case where the government wished to subpoena large amounts of information from Google.

106. Lessig, Lawrence. "The Architecture of Privacy *Draft 2.*" Taiwan Net Conference, 1998. Accessed November 4, 2010. http://cyber.law.harvard.edu/works/lessig/architecture_priv.pdf
Lessig posits how the public and private sectors are growing increasingly interconnected and difficult to tease apart.

107. Liptak, Adam. "Court to Rule on Right to Privacy for Referendum Petition Signers." *The New York Times*, January 16, 2010. Accessed January 30, 2011. http://www.nytimes.com/2010/01/16/us/politics/16scotus.html?_r=1&pagewanted=print.
This article demonstrates how traditionally labeled pubic records were not always intended to actually be public.

108. Palfrey, John. "The Public and the Private at the United States Border with Cyberspace." *Mississippi Law Journal* 78 (2008): 241-294. Accessed October 24, 2010. http://nrs.harvard.edu/urn-3:HUL.InstRepos:2554026.
Palfrey argues that the public and private sectors need federal regulations clarifying their relationship. He contends that currently, the government enjoys unregulated access to private companies' databases of individuals' personal information.

109. "Rapleaf: Interactive Graphics." *The Wall Street Journal*, 2010. Accessed November 6, 2010. http://s.wsj.net/public/resources/documents/st_RAPLEAF_20101018.html.
This website presents an interactive graphic of a tracked consumer's data. Code is translated into layman's terms so that *Wall Street Journal* visitors may understand how alphanumeric characters are used to signify different pieces of profiling information.

110. Schoen, Seth. "What Information is 'Personally Identifiable'?" The Electronic Frontier Foundation, September 11, 2009. Accessed November 5, 2011. https://www.eff.org/deeplinks/2009/09/what-information-personally-identifiable.
In this article, Schoen refers to Professor Latanya Sweeney's studies to exemplify how easy it has become to to de-identify previously anonymized datasets. He states, "Gender, zip code, and birth date feel anonymous…," but this combination is generally all it takes to uniquely identify an individual (as it does for 87% of the U.S. population).

111. Weintraub, Jeff Alan, and Krishan Kumar. *Public and Private in Thought and Practice: Perspectives on a Grand Dichotomy*. Chicago: University of Chicago, 1997.
This book offers a historical definition of what the "public" and "private" spheres meant to Aristotle.

Technological Determinism

112. Fischer, Claude. *America Calling: A Social History of the Telephone to 1940.* Berkeley and Los Angeles: University of California Press. 1992.
In his book, Claude Fischer details how the telephone and its underlying technology was introduced to society and ultimately adopted. In his book, Fischer discusses technological determinism and the differences between various nuanced theories.

113. Johns, Adrian. *The Nature of the Book*. Chicago: University of Chicago Press, 1998. Accessed October 24, 2010. http://books.google.com/books?id=ERpBdEUdhz8C&lpg=PP1&dq=nature%20of%20the%20book&pg=PP1#v=onepage&q=nature%20of%20the%20book&f=false.
Adrian Johns argues in his book that he does not believe fixity or technological determinism propelled the printing revolution; instead he believes that society and technology interacted with one another.

114. Marx, Leo, and Merritt Roe Smith, eds. *Does Technology Drive History?: the Dilemma of Technological Determinism*. Cambridge: MIT Press, 1995.
This anthology of work describes how different philosophers and historians of science interpret the theory of technological determinism.

115. Sommerville, John. "Review." *American Historical Review* 104, no. 5. (December 1999): 1751-1752. Accessed November 20, 2011. http://www.jstor.org/stable/2649491.
Sommerville reviews Elizabeth Eisenstein's book, *The Printing Press as an Agent of Change*, which discusses the rise of print culture and its effects on society. Eisenstein argued in her novel that the idea of fixity – or the idea that once something is printed it remains the same – was intrinsic to print technology.

116. Winner, Langdon. *Autonomous Technology: Technics out of Control*. Cambridge: MIT Press, 1977. Accessed October 24, 2010. http://books.google.com/books?id=uNIG0gi4b40C&printsec=frontcover&dq=langdon winner&hl=en&ei=hebCTIKvF4L_8AbkveDXBg&sa=X&oi=book_result&ct=result&re snum=1&ved=0CCgQ6AEwAA#v=snippet&q=neutral&f=false.
In his book, Winner argues that technology is not value-neutral. Winner defends technological determinism and comments how culture is changing based on individuals' increased uses of technology.

The Linear Model and Rise of Applied Science

117. Devettere, Raymond. *Introduction to Virtue Ethics: Insights of the Ancient Greeks*. Washington, DC: Georgetown University Press, 2002.
In his book, Devettere discusses the Ancient Greek origins of many ethical distinctions that we use in modern society. In particular, Devettere comments on how the Greeks defined "pure" versus "applied" science. He introduces the ideas of "episteme" and "phronesis" and explains their differences.

118. Hounshell, David. "Evolution of Industrial Research in the United States." In *American System to Mass Production, 1800-1932: The Development of Manufacturing Technology in the United States*, 13-85. Baltimore: Johns Hopkins University Press, 1984.
Hounshell offers insight into the rise of industrial and government support of research and development over the twentieth century. In particular, he focuses in the first chapter on the rise of the linear model in the 1940s and early 1950s.

PRIMARY SOURCES

The Data Cycle

119. "About the D-U-N-S Number." Dun & Bradstreet. Accessed October 29, 2011. http://fedgov.dnb.com/webform/pages/dunsnumber.jsp.
On Dun & Bradstreet's "About Us" website, the organization explains how the DUNS number works and what purpose it serves.

120. "Data.gov." Data.gov, 2009. Accessed December 8, 2011. http://www.data.gov/.
Data.gov is the website hosted by the U.S. Government that seeks to provide any online visitor access to freely download certain collections of previously vetted, public datasets.

121. Hoofnagle, Chris. "Big Brother's Little Helpers: How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement." The Electronic Privacy Information Center, Summer 2004. Accessed August 20, 2011. http://epic.org/privacy/choicepoint/cp_article.pdf.
In this research paper by Chris Hoofnagle, Hoofnagle demonstrates how ChoicePoint and other private data companies have entered into government contracts in order to provide the government with individuals' data that the government cannot acquire itself.

122. Robinson, David, Harlan Yu, William Zeller, and Felten, Edward W. "Government Data and the Invisible Hand." *Yale Journal of Law & Technology* 11 (Fall 2009): 160-175. Accessed October 2, 2011. http://ssrn.com/abstract=1138083.
This white paper argues that publicly distributed government data should be done through private means. Specifically, the authors posit that private parties can build better interactive designs and user interfaces needed to encourage people to use government data. The authors argue that the government should provide the data, but the private sector should work on presenting it.

Case Study: Voter Registration Data in the Data Cycle

123. Aristotle International, Inc. v. NGP Software, Inc. (U.S. District Court for DC 2011). Accessed October 18, 2011. http://www.campaignsandelections.com/article_assets/articledir_519/259782/Aristotle_NGP_Opinion.pdf.
In this Washington, DC district court conclusion document, political data firm Aristotle claims that software company NGP has engaged in false and misleading advertising in violation of the law. The document provides insight into background information on Aristotle and the types of clients they typically serve.

124. Committee on State Voter Registration Databases, Computer Science and Telecommunications Board, Division on Engineering and Physical Sciences, and the National Research Council, *Improving State Voter Registration Databases Final Report*. Washington, DC: The National Academies Press, 2010.

This report produced by the National Academies of Science, Engineering, Medicine, and the National Research Council, describes how state voter registration databases and systems may be improved.

125. "Findings: Voter Privacy in the Digital Age." The California Voter Foundation, June 9, 2004. Accessed September 17, 2011. http://www.calvoter.org/issues/votprivacy/pub/voterprivacy/findings.html.
This website presents the "Findings" to the study conducted by the California Voter Foundation. In particular, this website offers statistics on how voter data is shared with third parties.

126. Gottlieb, Scott. "Physician Prescribing Data and the Public Health: How Efforts to Thwart the Collection of Prescribing Data Undermine Drug Safety." *The Legal Pulse*, June 22, 2011. Accessed October 20, 2011. http://wlflegalpulse.com/2011/06/22/physician-prescribing-data-and-the-public-health-how-efforts-to-thwart-the-collection-of-prescribing-data-undermine-drug-safety/.
This blog post, written by Dr. Scott Gottlieb comments on the Supreme Court case *Sorrell v. IMS Health.* Gottlieb argues that for public health reasons, pharmaceutical companies should have access to anonymized prescription information.

127. Harris, Joseph P. "Election Administration in the United States." The Brookings Institution and Institute for Government Research Studies in Administration, no. 27 (1934). Accessed September 23, 2011. http://www.nist.gov/itl/vote/upload/chapter1.pdf.
Joseph Harris describes in this study, which was supported by the Brookings Institution, how the electoral process functioned throughout the United States during 1929 and 1930. Harris describes the impetus for assuming this research project "due to the present backward and generally unsatisfactory administration of elections." In particular, Harris draws our attention to issues of voter fraud, data collection, and tabulating issues.

128. "Statewide Voter Registration Databases." PEW Center on the States, Electonline.org, and the Constitution Project (March 2002): 1-16. Accessed September 24, 2011. http://www.pewcenteronthestates.org/uploadedFiles/Statewide%20Voter%20Registraion%20DB.pdf.
This study conducted by the PEW Research Center provides useful information on how states and other third parties in the early 2000s are using voter registration databases. The study touches on how private companies are often hired to intervene in updating states' voter registration data.

129. "Voter Privacy in the Digital Age." The California Voter Foundation, June 9, 2004. Accessed September 29, 2011. http://www.calvoter.org/issues/votprivacy/pub/voterprivacy/introduction.html.
This website provides the "Introduction" to the California Voter Foundation's study on "Voter Privacy in the Digital Age." This study investigates how voter registration data is often shared with various parties.

130. Zetter, Kim. "For Sale: The American Voter." *Wired,* December 12, 2003. Accessed November 29, 2011.
http://www.wired.com/politics/security/news/2003/12/61543?currentPage=all.
This article describes the growing trend of how commercial data companies are selling enhanced voter lists with personally indefinable information.

Raw Data & Government Documents

131. Beard, Daniel P., comp. U.S. Congress. Statement of Disbursements of the House from January 1, 2010 to March 31, 2010. 111th Cong., 2d sess., April 13, 2010. Washington, DC.
This government document lists each Member of the House of Representative's official expenses during Quarter 1 of 2010. I used this document and the subsequent documents in this section to search Members' payments made to political data companies Catalist and Aristotle.

132. ———. Statement of Disbursements of the House from April 1, 2010 to June 30, 2010. 111th Cong., 2d sess., July 27, 2010. Washington, DC.
This government document lists each Member of the House of Representative's official expenses during Quarter 2 of 2010.

133. "Help America Vote Act of 2002." (P.L. 107-252), *United States Statutes at Large.* 2002. 166 Stat. 1666. 107th Cong.
The full text of the Help American Vote Act of 2002 assisted my analysis in understanding recent business trends tied to the computerization of voter registration data.

134. Strodel, Daniel J., comp. U.S. Congress. Statement of Disbursements of the House from July 1, 2010 to September 30, 2010. 111th Cong., 2d sess., November 15, 2010. Washington, DC.
This government document lists each Member of the House of Representative's official expenses during Quarter 3 of 2010.

135. ———. Statement of Disbursements of the House from October 1, 2010 to December 31, 2010. 112th Cong., 1st sess., January 11, 2011. Washington, DC.
This government document lists each Member of the House of Representative's official expenses during Quarter 4 of 2010.

136. ———. Statement of Disbursements of the House from January 1, 2011 to March 31, 2011. 112th Cong., 1st sess., May 23, 2011. Washington, DC.
This government document lists each Member of the House of Representative's official expenses during Quarter 1 of 2011.

137. U.S. Congress. House. Committee on House Administration. *The Member's Handbook.* 111th Cong., 2d sess. Washington, DC, 2011.

The Member's Handbook provides the rules and regulations that all Congressmen and Congresswomen must follow while serving in office. Of particular interest, the handbook lays out the rules for how such officials may and may not spend their franking budget.

138. U.S. Congress. Office of Technology Assessment. "Federal Government Information Technology: Electronic Record Systems and Individual Privacy." Washington, DC: Diane Publishing, 1986. Accessed October 24, 2011. http://books.google.com/books?id=jrP7XBOZUK4C.
Nearly a decade after the Privacy Act of 1974's enactment, the Office of Technology Assessment wrote a report to Congress in 1986 identifying problems and weaknesses with the Act.

139. U.S. Congress. Senate. "January 22, 1873 Senate Ends Franked Mail Privilege." Accessed October 3, 2011. http://www.senate.gov/artandhistory/history/minute/Senate_Ends_Franked_Mail_Priviledge.htm.
This website, hosted by the United States Sente, offers insight into the history of the franking privilege.

140. U.S. Congress. Senate. Select Committee on Ethics. *Senate Code of Official Conduct*. 110th Cong., 2d sess., April 2008. Accessed June 12, 2011. <http://ethics.senate.gov/public/index.cfm/files/serve?File_id=efa7bf74-4a50-46a5-bb6f-b8d26b9755bf>.
The Senate Code of Official Conduct outlines the rules and regulations governing the U.S. Senate.

141. U.S. Election Assistance Commission. "Report No. E-HP-AL-06-1." Washington, DC, March 28, 2011. Accessed September 28, 2011. http://www.eac.gov/assets/1/Documents/Final%20EAC%20Management%20Decision%20Alabama%20E-HP-AL-06-10.pdf.
The Election Assistance Commission conducted an audit (from April 30, 2003 – January 31, 2010) of the state of Alabama's payments received under the Help America Vote Act.

142. ———. *Online Voter Registration Form*. Washington, DC, March 1, 2006. Accessed October 29, 2011. http://www.eac.gov/assets/1/Documents/national%20mail%20voter%20registration%20form%20english%20February%2015%202011.pdf.
This is a copy of the U.S. government's federal voter registration applications for U.S. citizens.

143. ———. "Report No. E-HP-WV-04-09." Washington, DC, May 5 2011. Accessed September 29, 2011. http://www.eac.gov/assets/1/Documents/FINAL%20EAC%20Management%20Decision%20West%20VA%20E-HP-WV-04-09.pdf.
The Election Assistance Commission conducted an audit (from April 28, 2003 – August 31, 2009) of the state of West Virginia's payments received under the Help America Vote Act.

144. ———."SUBCHAPTER I–H—NATIONAL VOTER REGISTRATION." Accessed September 20, 2011. http://www.eac.gov/assets/1/workflow_staging/Page/27.PDF.
As set forth by the U.S. government, this document provides the federal laws governing national voter registration.

145. U.S. Securities and Exchange Commission. "Form S-1 Registration Statement Under the Securities Act of 1933: Facebook Inc." Washington, DC, February 2012. Accessed February 12, 2012. http://s3.documentcloud.org/documents/288894/facebook-prospectus.pdf.
This is a copy of Facebook's official S-1 Registration form for the U.S. Security and Exchange Commission. The document provides a prospectus detailing official facts and statistics about the company.

Private Companies' Data Products

146. "About Us: Our Company." CoreLogic. Accessed September 29, 2011. http://www.corelogic.com/about-us/our-company.aspx.
On CoreLogic's "Abut Us" website, the company details what type of data it collects and what type of services the company offers its clients.

147. "Around The Clock Access To Voter Data: VoterListsOnline.com." Aristotle. Accessed November 29, 2011. http://www.aristotle.com/content/view/35/119/
This Aristotle website describes the type of voter data packages available for purchase.

148. "Hill Data." Aristotle. Accessed October 2, 2011. http://www.aristotle.com/content/view/92/117/.
This Aristotle website describes its "Hill Data" product, which was created specifically for elected officials.

149. "Political Data." Aristotle. Accessed October 2, 2011. http://www.aristotle.com/content/blogcategory/22/45/.
On this website, Aristotle describes its "Political Data" package and the types of information made available in this product.

150. "Products." Catalist. Accessed October 4, 2011. http://catalist.us/product.
On Catalist's "Product's" webpage, Catalist lists and describes the types of data packages that it makes available to its consumers.

History of Government Data Collection

151. Westin, Alan, and Michael Baker. *Databanks in a Free Society: Computers, Record-Keeping and Privacy*. New York: Quadrangle Books, 1972.
In this report, sponsored by the Project on Computer Databanks of the Computer Science and Engineering Board and the National Academy of Sciences, Alan Westin and Michael Baker conduct a study on how different sectors of American society in the early 1970s were using databanks. The study investigates what types of information were being

collected, stored, and shared. Westin and Baker explain that project was undertaken due to the concern held by a large majority of the American public about government databanks.

152. Westin, Alan. *Computers, Health Records, and Citizen Rights.* Washington, DC: U.S. Government Printing Office, 1976.
In this book, sponsored by the Institute for Computer Sciences and Technology and the National Bureau of Standards, Alan Westin reveals the findings of his 1970s study on how computers had been employed to digitize, organize, store, and share personal, medical information. The study specifically maps out how different parties were accessing individauls' health records.

History of Computing Technologies and the Internet

153. Bush, Vannevar. "As We May Think." *The Atlantic Monthly*, July 1945. Accessed February 8, 2012. http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/3881/.
In this article, Bush describes his idea of "memex," or a device that individuals could use to compress and store all of their books, records, and communications. Through his introduction of "memex," Bush advocates for a new type of relationship between human, machines, and knowledge.

154. Cerf, Vint. "Military Requirements for Packet-switched Networks and Their Implications for Protocol Standardization." *Computer Networks* 7, no. 5. (1983): 293-306.
In this journal article Vint Cerf, known as one of the major creators of the Internet, describes how packet-switched networks work. He discusses precautions the military should take in order to ensure security if it wished to adopt such a network.

155. Feynman, Richard. "There's Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics." *Engineering and Science* (1960). Accessed September 3, 2011. http://www.zyvex.com/nanotech/feynman.html.
In his speech, Richard Fenyman advocates for the exploration of the science of small. In front of a room full of students, Feynman describes the great potential that he believes exists on the micro- and nano- scales for computing and storing information.

156. Kennedy Jr, T.R. "Electronic Computers Flash Answers, May Speed Engineering." *The New York Times*, February 15, 1946. Accessed January 2, 2012. http://www.fi.edu/learn/case-files/eckertmauchly/full/460215_eckertmauchly_article_1of2.jpg.
In this 1946 *New York Times* article, the reporter announces the introduction of the ENIAC and celebrates the potentials that such technology holds for revolutionizing the fields of engineering and science.

157. Leiner, Barry, Vint Cerf, David Clark, Robert Kahn, Leonard Kleinrock, Daniel Lynch, Jon Postel, Larry Roberts, and Stephen Wolff. "A Brief History of the Internet." Last modified January 23, 1999. Accessed November 2, 2011. http://arxiv.org/html/cs/9901011v1/#Origins.

This website, written by the "fathers of the Internet" themselves, offers a historical account of the Internet's development. The authors describe their respective roles in detailing the history of the Internet.

158. Mauchly, John. "The Use of High Speed Vacuum Tube Devices." In *The Origins of Digital Computers,* edited by Brian Randell and David Gries, 355-358. Berlin: Springer-Verlag Berlin Heidelberg, 1975. Accessed January 7, 2012. http://books.google.com/books?id=Dwj4RmcZ1AoC.
In this paper, Mauchly describes why high speed vacuum tubes may offer computers greater calculation speeds and performances. He goes on to discuss the benefits of employing electronic circuits.

159. Mauchly, John, and J. Presper Eckert. "The Electronic Numerical Integrator and Computer (ENIAC)" In *The Origins of Digital Computers,* edited by Brian Randell and David Gries, 359-374. Berlin: Springer-Verlag Berlin Heidelberg, 1975. Accessed January 7, 2012. http://books.google.com/books?id=Dwj4RmcZ1AoC.
In their report, engineers Mauchly and Eckert give an overview of how the ENIAC functions and how it may be programmed.

160. Moore, Gordon. "Cramming More Components onto Integrated Circuits." *Electronics* 38, no. 8 (April 9, 1965): 82-85.
In this journal article, then Director of Research and Development at Fairchild Semiconductor Gordon Moore posits the following self-fulfilling prophecy: every eighteen months the number of transistors on a chip should double. This idea later became known as "Moore's Law" and still holds true today.

161. Nelson, Ted. "A File Structure for the Complex, the Changing and the Indeterminate." *Association for Computing Machinery: Proceedings of the ACM 20th National Conference* (1965): 84-100.
In Ted Nelson's piece "A File Structure for the Complex, the Changing and the Indeterminate," he describes a primitive version of hyperlinking. Building on Nelson's idea, Tim Berners-Lee eventually designed the architecture behind the World Wide Web in the early 1990s.

162. November, Joe. "LINC: Biology's Revolutionary Little Computer." *Endeavour* 28, no 3. (2004): 125-131.
November describes the 1963 Laboratory Instrument Computer ("LINC") as the computer that was designed to meet the failure of biologists to use computer technology. November discusses the computerization of a biologist's laboratory and how mainframe computers during the second half the twentieth century were extremely hard to access and use.

163. Von Neumann, John. "The First Draft of a Report on the EDVAC" In *The Origins of Digital Computers,* edited by Brian Randell and David Gries, 383-392. Berlin: Springer-Verlag Berlin Heidelberg, 1975. Accessed January 7, 2012. http://books.google.com/books?id=Dwj4RmcZ1AoC.
This is a very technical report on the functioning of the the Electronic Discrete Variable

Automatic Computer ("EDVAC"). The EDVAC, produced in the mid-twentieth century, was one of the earliest electronic computers. Von Neumann wrote this widely distributed paper that described a computer architecture whereby data and program information could be stored in the same place. This became the basis of modern computer design.

The Privacy Act of 1974

164. Gross, Grant. "U.S. Privacy Act Outdated, Hasn't Kept up With Technology, Experts Say." *Computer World*, June 18, 2008. Accessed April 24, 2011. <http://www.computerworld.com/s/article/9100258/U.S._Privacy_Act_outdated_hasn_t_kept_up_with_technology_experts_say_>
In 2008, *Computer World* ran an article on how the Privacy Act of 1974 has not kept up with twenty-first century technologies.

165. Relyea, Harold. "The Privacy Act: Emerging Issues and Related Legislation." *Congressional Research Service* and the Library of Congress' Report for Congress, February 26, 2002.
This report produced by the Congressional Research Service summarizes the history of the Privacy Act of 1974 and identifies newly emerging issues in the twenty-first century. Specifically, the report highlights the law's out-dated nature.

166. Sweeney, Latanya. "*Simple Demographics Often Identify People Uniquely."* Pittsburgh: Carnegie Mellon University, Data Privacy Working Paper 3, 2000.
In Latanya Sweeney's research paper she demonstrates how one may identify previously anonymized datasets by using the following three simple pieces of information: gender, birth date, and zip code.

167. ———."Patient Privacy Risks in U.S. Supreme Court Case Sorrell v. IMS Health Inc.: Response to Amici Brief of El Emam and Yakowitz." Data Privacy Lab Working Paper 1027-1015B. Cambridge, 2011.
In this paper, Sweeney details how she was able to link individuals' hospital discharge data with their voter registration information. These findings are interesting to consider in light of the U.S. government's support of publishing "anonymized" datasets online.

168. "U.S. Privacy Act of 1974 (5 U.S.C. § 552a as amended)." The U.S. Justice Department. Accessed November 2, 2011. http://www.justice.gov/opcl/privstat.htm.
This website provides the full text of the Privacy Act of 1974.

Online Tracking

169. Arthur, Charles. "iPhone Keeps Record of Everywhere You Go." *The Guardian,* April 20, 2011. Accessed January 1, 2012. http://www.guardian.co.uk/technology/2011/apr/20/iphone-tracking-prompts-privacy-fears.
This article exposes how Apple's iPhone keeps track of users' location by timestamping and recording individuals' location coordinates.

170. Cohen, Noam. "It's Tracking Your Every Move and You May Not Even Know." *The New York Times,* March 26, 2011. Accessed December 14, 2011. http://www.nytimes.com/2011/03/26/business/media/26privacy.html.
In this article, Cohen describes how cell phone users are regularly tracked by their devices whether they are cognizant or not. In the United States, cell phone companies do not need to disclose the specific types of information that they record.

Online Privacy

171. Angwin, Julia. "Wall Street Journal Revises Its Privacy Policy." *The Wall Street Journal*, September 27, 2011. Accessed October 8, 2011. http://blogs.wsj.com/digits/2011/09/27/wall-street-journal-revises-its-privacy-policy/.
Julia Angwin discusses the *Wall Street Journal's* latest privacy policy changes to its own website. The change allows the *Journal* to share users' data more broadly. This is somewhat ironic when juxtaposed next to the *Journal's* privacy investigative series, "What They Know," which seeks to uncover how various entities are using individuals' information without their knowledge and/or consent.

172. ———. "Obama Administration Seeks Internet Privacy Protections, New Policy Office." *The Wall Street Journal*, November 11, 2010. Accessed November 23, 2010. http://online.wsj.com/article/SB10001424052748703848204575608970171176014.html
In this article, Angwin describes how in the fall of 2010 the Obama administration concerned itself with "policing Internet privacy." She reports that the administration seeks to update and/or create new laws.

173. Angwin, Julia and Jennifer Valentino-Devries, "Google's iPhone Tracking," *The Wall Street Journal,* February 17, 2012, accessed February 17, 2012, http://online.wsj.com/article/SB10001424052970204880404577225380456599176.html.
In February 2012, a Stanford University researcher uncovered that Google and other advertising networks were secretly collecting iPhone users' online browsing activity. IPhone users, while upset, did not change their behaviors or stop using this gadget.

174. "EPIC: Public Opinion on Privacy." Electronic Privacy Information Center. Accessed February 2, 2012. http://epic.org/privacy/survey/.
On this website, the Electronic Privacy Information Center provides contemporary surveys and poll results related to the topic of online privacy.

175. "Privacy Poll." *The Wall Street Journal,* 2011. Accessed February 2, 2012. http://online.wsj.com/community/groups/media-marketing-267/topics/how-concerned-you-about-advertisers.
In 2011, *The Wall Street Journal* conducted a poll of its online readership asking how concerned they felt as to private companies' ability to track them online. Over 85% of survey participants (11,800 people) revealed that they are concerned.

176. Steel, Emily. "A Web Pioneer Profiles Users by Name." *The Wall Street Journal,* October 25, 2010. Accessed October 25, 2010. http://online.wsj.com/article/SB10001424052702304410504575560243259416072.html?KEYWORDS=rapleaf.
This article discusses how an everyday Internet user, Ms. Twombly, is targeted with political endorsement ads through the Web tracking firm, RapLeaf, which collects individuals' personal data.

177. ———. "Candidate Apologizes For Using RapLeaf To Target Ads." *The Wall Street Journal,* October 27, 2010. Accessed October 27, 2010. http://blogs.wsj.com/digits/2010/10/27/candidate-apologizes-for-using-rapleaf-to-target-ads/?KEYWORDS=jim+bender+apologizes.
New Hampshire Republican hopeful for Senate Jim Bender publicly apologizes for employing RapLeaf and admits to not knowing how intrusive it was until the *Wall Street Journal* article exposed its practices.

178. ———. "Thousands of Web Users Delete Profiles From RapLeaf." *The Wall Street Journal*, October 26, 2010. Accessed October 27, 2010. http://online.wsj.com/article/SB10001424052702304248704575574653801361746.html?KEYWORDS=rapleaf.
This article describes how thousands of users deleted their RapLeaf profiles. *The Wall Street Journal* provides step-by-step instructions on how to view, change, and/or delete one's personal profile on RapLeaf's website.

179. Steel, Emily, and Peter Wallsten. "Politicians Tap Sophisticated Online Tracking Tools." *The Wall Street Journal,* October 24, 2010. Accessed October 30, 2010. http://online.wsj.com/article/SB10001424052702304915104575572173837099644.html?KEYWORDS=rapleaf.
This article discusses how Web tracking technology is becoming more popular among political campaigns. Political campaigns are using tracking tools to gain information on potential voters.

180. Zittrain, Jonathan. "Ideas for a Better Internet" (class discussion, Areeda Hall, Harvard University, Cambridge, MA, April 24, 2011).
In this class, Professor Zittrain discussed privacy issues related to the Internet. He questioned whether anyone even really cares about Internet privacy or rather, if they even has a say in the matter. Professor Zittrain commented on how "free services," such as Facebook and Google, are not really serving "clients" in the traditional sense, since such clients are really part of the product.

Countercultural Perspectives

181. Goode, Stephen. *The Right to Privacy*. New York: Franklin Watts, 1983.
This book from the 1980s discusses various, historical privacy debates in the United States. Goode highlights how different people continue to redefine what privacy means. Goode discusses how privacy is at risk and how the technologies of the time are posing credible threats.

182. Kidder, Tracy. *The Soul of a New Machine*. Boston: Little, Brown and Company, 1981.
Kidder's book describes how computers were perceived and used throughout the 1970s and 1980s. In particular, he describes the high-tech industry boom and how people began realizing they could make money off of computer companies.

183. Licklider, J. C. R. "Man-Computer Symbiosis." *IRE Transactions on Human Factors in Electronics* (1960).
This paper describes how humans and computers may positively enhance one another's capabilities. Licklider proposes that the computer will lead in data crunching while humans should direct rational, decision making.

184. Packard, Vance. *The Naked Society*. New York: D. McKay, 1964.
In Packard's novel, he describes the privacy worries held by American citizens in the 1950s and 1960s. His book is particularly useful in understanding some of the major concerns held by the public during the mid-twentieth cenutry.

185. Roszak, Theodore. *The Making of a Counter Culture; Reflections on the Technocratic Society and Its Youthful Opposition.* Garden City: Doubleday, 1969.
In this book, Roszak describes the counterculture that existed against computers and technology throughout the mid-twentieth century.

The Neurological Effects of Digital and Communication Tech

186. Nelson, Charles, and S. Jeste. "Neurobiological Perspectives on Developmental Psychopathology." In *Rutter's Child and Adolescent Psychiatry*. Oxford: Blackwell Publishing Ltd., 2009.
This scientific article on neural plasticity shows how the human brain's circuitry is capable of manipulation even into adulthood.

187. Richtel, Matt. "Growing up Digital, Wired for Distraction." *The New York Times*, November 21, 2010. Accessed November 23, 2010.
http://www.nytimes.com/2010/11/21/technology/21brain.html?src=me&ref=general.
In this article, reporter Matt Richtel does an in-depth study of Woodside High School and the students' addiction to technology. He interviews various students, their parents, school faculty members, and scientists on the issue.

188. ———. "Attached to Technology and Paying a Price." *The New York Times*, June 6, 2010. Accessed August 23,
2010. .http://www.nytimes.com/2010/06/07/technology/07brain.html?pagewanted=all.
This article written by Matt Richtel follows the story of a businessman who cannot "disconnect" even when he is at home spending time with his family. Richtel contextualizes this anecdote within a larger body of scientific studies discussing humans' addiction to technologies.

189. ———. "Your Brain on Computers: A Series." *The New York Times*, June 7, 2010. Accessed February 3, 2012. http://topics.nytimes.com/top/features/timestopics/series/your_brain_on_computers/index.html.
This page on the *New York Times'* website provides information about a series of articles written on the topic of "Your Brain on Gadgets." The website is comprised of articles, interactive games, and related links to the topic.

190. "Unplugged: Take the Challenge." *The New York Times*, June 7, 2010. Accessed October 2, 2011. http://bits.blogs.nytimes.com/2010/06/07/unplugged-take-the-challenge/.
This article poses a "challenge" to its readers to go without technology for twenty-four hours. This challenge is indicative of how difficult it is for individuals to function in society today without their gadgets.

Conducted Interviews
(See Appendix B "Oral History Interviews" for Transcripts)

191. Johnson, Clay (Author, Technology and Transparency Advocate). Phone interview by Melissa Oppenheim. November 4, 2011, 11am EST.
I spoke to Clay Johnson about his views on public data and government transparency. Although Clay recognizes the trade off between privacy and transparency, he believes that more data made publicly available by the U.S. government is better for everyone.

192. Kundra, Vivek (Former Chief Information Officer in the Obama Administration). Questions asked by Melissa Oppenheim in a roundtable discussion. November 29, 2011, 6pm EST, at Harvard University's Institute of Politics, Cambridge, MA.
In a roundtable discussion at Harvard's Institute of Politics, I asked Vivek Kundra about "open government initiatives" like Data.gov and how they function with respect to the Privacy Act of 1974.

193. Quinn, Laura (CEO of Catalist). Interview by Melissa Oppenheim. November 9, 2011, 6:30pm EST at the Omni Parker Hotel, Boston, MA.
I spoke to Laura Quinn about her personal career and experience at Catalist. We discussed the perceived need for political data firms and how Catalist reconciles their business practices with privacy issues.

194. St. George, Jim (Co-owner of NGPVAN). Interview by Melissa Oppenheim. November 11, 2011, 11:30am EST, at NGPVAN Offices, Somerville, MA.
Jim St. George and I discussed how NGPVAN, a political data firm, handles its individuals' information in regards to the unregulated legal landscape.

195. Sweeney, Latanya (Professor of Computer Science and Privacy at Harvard University). In person conversation with Melissa Oppenheim. October 4, 2011, 4pm EST, at Harvard University, Boston, MA.
From a discussion with Professor Sweeney, I was able to flesh out the idea of the "Data Cycle."

196. Zak, Paul (Professor of Economics and Department Chair and Founding Director of the Center for Neuroeconomics Studies at Claremont University). Phone interview by Melissa Oppenheim. July 16, 2010. 9:45am EST.
Over the phone, Professor Paul Zak and I discussed some studies that he conducted on how interaction via social media, such as Facebook and Twitter, affect the human brain.

Implications of Social Security Numbers

197. Oppenheim, Melissa. "Why No Real Free and/or Open Source Software Movement Has Developed Around Filing Income Tax Returns: An Exploratory Analysis." Research Paper, Independent Study on Internet Policy with Professor Jonathan Zittrain, Harvard University, Cambridge, MA, December, 2011.
In this paper, I explored the current trend of filing online tax returns and questioned why no open source and/or free software solution existed. The paper discusses issues of voluntary compliance.

198. U.S. Department of Health and Human Services. Office of the Assistant Secretary for Planning and Evaluation. *Records, Computers and the Rights of Citizens: Report of the Secretary's Advisory Committee on Automated Personal Data Systems*. Washington, DC: Government Printing Office, July 1973. Accessed December 8, 2011.
http://aspe.hhs.gov/datacncl/1973privacy/c7.htm
This document details how the the Secretary of Health, Education, and Welfare in the early 1970s was concerned about the consequences of using social security numbers as unique identifiers. The report concludes that the "net effect of computerization is that it is becoming much easier for record-keeping systems to affect people than for people to affect record-keeping systems."

199. U.S. Department of the Treasury. *Interim Results of the 2011 Filing Season*. Washington, DC: Government Printing Office, March 31, 2011. Accessed January 20, 2012.
http://www.treasury.gov/tigta/auditreports/2011reports/201140032fr.pdf.
This report reveals the interim results of the 2011 tax filing season. The Treasury Department lists statistics regarding tax return fraud.

200. U.S. Government Accountability Office. *2010 Tax Filing Season: IRS' Performance Improved in Some Key Areas, but Efficiency Gains are Possible in Others.* Washington, DC: Government Printing Office, 2010. Accessed November 2, 2011.
http://www.gao.gov/new.items/d11111.pdf.
This document produced by the U.S. GAO identifies key problems with the current practice of tax return filing and proposes solutions to these issues.

201. ———. *Taxes and Identity Theft: Status of IRS Initiatives to Help Victimized Taxpayers.* Washington, DC: Government Printing Office, 2011. Accessed November 10, 2011. http://www.gao.gov/new.items/d11721t.pdf.
This documented produced by the U.S. GAO describes how online tax return filing has spawned large amounts of tax fraud and identity theft. The document sites how individuals' social security numbers are an enabling piece of information to this recent trend.

The Linear Model and Rise of Applied Science

202. Bush, Vannevar. "Science – The Endless Frontier: A Report to the President on a Program for Postwar Scientific Research." Washington, DC: Government Printing Office, 1945.
In this report to the President, Vannevar Bush highlights the success of the linear model and calls for the creation of a federal science research foundation.

203. Cohen, I. Bernard. "Chapter 1" and "Chapter 5" In *Science, Servant of Man*. Boston: Little, Brown and Company, 1948.
In 1948, I. Bernard Cohen discussed how American society was about to embark on "a program of unprecedented large-scale expenditure of public funds for the support of scientific resarch." He details how during this time it was believed that scientific research "pay[ed] dividends in a very practical way."

204. Hoover, Herbert. "The Vital Need for Greater Financial Support of Pure-Science Research." *National Research Council* 65 (1926): 6-7.
In this article, Herbert Hoover advocated for the need to fund pure scientific research in order to produce useful byproducts.

205. Rowland, Henry A. "The Highest Aim of the Physicist." *American Journal of Science* 8, no. 43 (1899): 401-412. Accessed November 29, 2009.
http://books.google.com/books?id=bUs7AQAAIAAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.
Henry Rowland's article from 1899 discussed his beliefs that at the turn of century he did not think that America had been paying enough attention to scientists and researchers. He argued that even though physicists strove to unlock the universe's secrets, an ordained pursuit, American society remained oblivious to the types of value that could come from pure scientific discoveries. He hopes this will change with the dawning of the twentieth century.

206. Weinberg, Alvin M. "Impact of Large-Scale Science on the United States." *Science* 134, no. 3473 (1961): 161-164.
In this article, Weinberg describes the legacy that he believes twentieth century society will leave on the world. Unlike societies' "landmarks" of the past, Weinberg argues that twentieth century society's landmarks will be measured in the technologies and tools that were produced.

Punched Cards and the U.S. Census
(See Appendix C)

207. U.S. Census Bureau. *1890: Fast Facts*. Washington, DC: Bureau of the Census. Accessed November 2, 2011.
http://www.census.gov/history/www/through_the_decades/fast_facts/1890_fast_facts.html.
This government website provides official statistical information about the 1890 census survey.

208. ———. *History: Decennial Census.* Washington, DC: Bureau of the Census. Accessed November 2, 2011.
http://www.census.gov/history/www/programs/demographic/decennial_census.html.
On this website, the U.S. Census Bureau provides historical information about the creation and implementation of the decennial census.

209. ———. *Index of Questions.* Washington, DC: Bureau of the Census. Accessed November 2, 2011.
http://www.census.gov/history/www/through_the_decades/index_of_questions/.
On this website the U.S. Census Bureau provides a lists of questions that its surveys have asked citizens to answer between 1790 and 2010.

210. ———. *Overview: 1960*. Washington, DC: Bureau of the Census. Accessed November 2, 2011. http://www.census.gov/history/www/through_the_decades/overview/1960.html.
This government website provides official statistical information about the 1960 census survey.

211. ———. *Overview: 1980.* Washington, DC: Bureau of the Census. Accessed November 2, 2011.
http://www.census.gov/history/www/through_the_decades/fast_facts/1980_new.html.
This government website provides official statistical information about the 1980 census survey.

212. ———. *Overview: 1990*. Washington, DC: Bureau of the Census. Accessed November 2, 2011.
http://www.census.gov/history/www/through_the_decades/fast_facts/1990_new.html.
This government website provides official statistical information about the 1990 census survey.

Definitions

213. *Merriam-Webster Dictionary and Thesaurus Online,* s.v. "Aggregate." Accessed December 8, 2011,
http://www.merriam-webster.com/dictionary/aggregate.
This website provides the definition used for "aggregate" in this thesis.

214. ———. s.v. "Repurpose," Accessed December 8, 2011, http://www.merriam-webster.com/dictionary/repurpose.
This website provides the definition used for "repurpose" in this thesis.

215. "Definitions: Routine Use." In the Privacy Act 1974 (5 U.S.C. § 552a). Electronic Privacy Information Center, last modified January 3, 2005. Accessed October 15, 2011, http://epic.org/privacy/laws/privacy_act.html
This website clearly defines and elucidates the "routine use" clause, pursuant to the Privacy Act of 1974.

216. The Glossary of Terms of the Statement of Disbursements. U.S. House of Representatives, s.v. "Member Representational Allowance," Disbursements.House.Gov. Accessed October 1, 2011, http://disbursements.house.gov/glossary.shtml.
This website provides the official glossary of terms as defined by the federal government. In particular, this site offers the definition for "member representational allowance" used in this paper's voter registration data case study.