Big Data and the Bigger Picture

Morgan E. Kaptur
Krannert School of Management
Purdue University

Steven D. Fisher
Krannert School of Management
Purdue University

Delores Robinson
Krannert School of Management
Purdue University

Clifford D. Fisher
Krannert School of Management
Purdue University

Abstract

In 1972, New York passed the Controlled Substances Act which allowed the state to collect personal information about which individuals were prescribed drugs that were deemed high priority on the illegal market. The case Whalen v. Roe was one of the first recorded instances to investigate this issue and truly analyze the legality of mass data collection. Fast forward 40 years and data has only become a larger part of everyday life for anyone uses a computer, cell phone, or even visits their local pharmacy. Companies love collecting high volumes of data because it allows them to analyze human behaviors and trends to optimize their businesses. Fashion companies, grocery stores, government agencies, insurance companies, and many more organizations desire information about their constituents for varying reasons. This mass collection of data is termed "Big Data". Big Data has become a popular and controversial topic in the news. The two real-life stories below clarify a valid argument for each side.

Elderly Richard Guthrie answered his phone to an insurance company to discuss ways to save on his life insurance plan. The next day when Richard went to the bank, he discovered his money was gone. What happened was that Richard's frequent entry into online sweepstakes gave criminals access to his information, allowing them to pose as life insurance agents and hack his accounts. In opposition, IBM is creating a Big Data Analytics Platform to enhance government agencies ability to predict and prevent real time crime. Their plan suggests uncovering suspicious relationships between people, events, and transactions to tenfold increase general safety.

The case and examples above bring up an interesting question to the benefit of collecting data versus the potential privacy and security risks. This paper defines Big Data, references current matters in society surrounding Big Data, and examines existing legal statutes and regulations. These legal issues are used to help predict the future of Big Data and set expectations for society and the government as time progresses. As technology is changing, it is critical the courts continually analyze and question the precedents set for Big Data collection. The paper presents the Big Picture in regard to Big Data.

Keywords: Big Data, Privacy, Data Collection, Technology, Corporations, Hacking.

Introduction

A user's online experience while scrolling through any webpage or application will begin with a multitude of advertisements relevant to their interests. Past searches made through search engines like Google or through online browsing will be tracked in an online database to provide a series of product suggestions. A shopper scrolling through Amazon will encounter an abundance of features that Amazon uses to enhance the customers' shopping experience. Amazon's Personalized Recommendation System, Book Recommendations from Kindle Highlighting and One-Click Ordering are some of the key features that customers interact with that increase Amazon's annual revenue around 30%. [1]Anticipatory Shipping Model, Supply Chain Optimization and Price Optimization are features that consumers interact with indirectly, but that are very much apart of Amazon's operations and aid in their growth. The ability to know which products to recommend to specific customers or which prices to display at certain times takes a talent and resources to do so effectively. Information is one of those resources and a specific form of information called big data provides programmers behind websites like Amazon the loads of information related to user actions to develop the shopping experience that they have created for their consumers.

Information is everywhere. What makes information valuable is aggregating those bits and pieces of information to paint a picture with which one can draw conclusions for decision making, as Amazon does. Amazon is one of many entities that utilizes big data to make decisions to help reach their objectives. While the Amazon has used big data to increase revenues with creative and successful implementation such powerful data is subject to being used to cause harm to those who the data is about actually about.

The term "Big Data" refers to any sizable amount of facts or statistics that upon analysis reveals an abundance of information surrounding human behaviors and trends. Organizations across all industries and development stages use these analytics to understand consumer preferences and create the most value for their firm. Firm applications of big data include but are not limited to, developing products appropriate for their target audience, forecasting customer demand, and creating effective marketing. Big Data uses in various organizations across industries span from minimizing risks of deception for bank loans, identifying at risk students in the school systems, preventing crime through the government, optimal treatment plans in health care organizations, to even how retailers organize their stores.[2]

The power of Big Data naturally raises concerns about security. Customers want to know when and where their information is collected and any potential threats because of it. Corporations want to know that they can make money by creating the best experience to keep their consumers coming back. The nation wants to know that Big Data will prevent and not cause any threats to homeland security. The term "Big

[1] Wills, Jennifer. 7 Ways Amazon Uses Big Data to Stalk You. Investopedia.7 Sept. 2016. https://www.investo pedia.com/articles/insights/090716/7-ways-amazon-uses-big-data-stalk-you-amzn.asp. Accessed 22 Feb. 2018.

[2] "What is Big Data and why it matters." What Is Big Data? | SAS US, SAS Institute, www.sas.com/en_us/insig hts/big-dat a/what-is-big-data.html#. Accessed 20 Feb. 2018.

Data" is associated with many assumptions and stereotypes therefore, it is critical to understand the basic principles of Big Data collection. Given the relatively new nature of Big Data, the pre-existing laws must be monitored to ensure they protect everyone involved.

Given the power of Big Data, it is critical that the laws surrounding Big Data are fitting to protect everyone. This paper analyzes what current legal statutes and regulations exist with regard to Big Data to help examine its future. In addition, it presents opposing views to the necessity of big data by analyzing the rights of both users and corporations. Understanding the scope of Big Data is critical as technology progresses and this paper presents the big picture and key legality issues of mass data collection.

Defining big data

1. What is Big Data?

N.J. Stat. § 52:17C-3.4 states "'Big Data' means high volume information assets, high velocity information assets, high variety information assets, or all three, that require new forms of processing to enable enhanced decision making, insight discovery, and process optimization.[3] Big data is essentially a sizable amount of facts or statistics (data) that upon analysis reveals an abundance of information surrounding human behaviors and trends. Big Data's foundation lies in the history of data itself. Perhaps the most fascinating aspect of data is that the concept itself is its appearance in ancient history as tally marks on a rock used by the Paleolithic tribespeople to record things such as trading goods and keeping track of time.[4] The term Big Data manifested itself in the late 1990's when computers could not store all the information that was needed within its local memory, that meant they had to acquire more resources.[5] It was around the mid 2000's that computer storage technology became capable of storing large amounts of data, leading to a boom in the size of the internet as well as the data that was collected from it. The collection of personal data was not an important thing for the average person until the early 2010's. The first data collection agencies began appearing due to cheap data storage solutions for consumers and the idea of collecting people's behavior to analyze what products or services people want when accessing the internet.

2. Three Parties of Data

Many internet users have had the experience of a Google search turning into a stream of advertisements. Web browsing is one of the most notorious sources to obtain information about individuals. Information on what sites are visited, what is searched for within the site, what is bought, and even what is posted about a particular item. Not only are these records online, but store owners have been found

[3] N.J. Stat. § 52:17C-3.4.

[4] Marr, B. A Brief History of Big Data Everyone Should. LinkedIn. 24 Feb. 2015 Read" https://www.linkedin.com /pulse/brief-history-big-data-everyone-should-read-bernard-marr. Accessed 21 March 2018.

[5] Houpert, J.F. "What You Need To Know About 1st, 2nd and 3rd Party Data." Intent-Based retargeting, Datacratic, 20 June 2017, www.datacratic.com/blog/first-second-third-party- data. Accessed 20 Feb. 2018.

selling their sales information. This consumer data may contain names, addresses, and other personal information. Sensitive information such as online health records and conditions may be tracked online. A top broker, Acxiom, has developed a website called AboutTheData.com to try and ease tension about their data collect. Users can log in and see the information the company has about them, after, of course, entering more personal data. [6]A recent New York Times article claims that although it shows what data is being recorded it leaves out many major elements that they market to their corporate clients.[7] As unnerving as this can sound, the actual collection of data is legal. The specifics of legality will be discussed later, but it is important to understand that data collection is a recognized practice within business and government agencies. The specifics of data brokering can be broken into three types of data called parties. First party data is information the first party gathers from their own audiences. For example, if a cell phone store wanted to know how users felt about their coverage they could use data they collected through online surveys, contact centers, or in-store beacons. In this type, customers typically can assume their data is being used by the company since the customers provided it. Unfortunately for the users of data, first party data provides a limited scope of information, which creates the need for second party data. This type of data is easily defined as "someone else's first party data". Organizations can receive this data through their business partnerships with other organizations. Second party data is more distant from the actual users limiting its accuracy and users may not be aware their data has moved to a second party source. Furthest from the customer is third party data. A company that specializes in data analytics buys and sells the data between varying organizations.[8] By understanding the behind the scenes of Big Data, users can make more educated decisions on where they want to share their data.

3. Disposal of Data

Currently data storage systems are not advanced enough to have kept all the data they obtained when they started collecting therefore, to make more room for the most current and useful data, collections of data must be rid of all the less useful data. In a 2012 article, the Director of Information Lifecycle Governance Solutions at IBM, Deidre Paknad, stated that 90% of the data in the world was created in the last two years.[9] With data production as rapid as it is, storage technology needs to advance as well for industries to keep informatics for extended periods of time. Ms. Paknad recognized

[6] Boutin, Paul. "Theres very little oversight in the industry of data brokers." The Secretive World of Selling Data About You, Newsweek, 16 June 2016, www.newsweek.com /secretive-world-selling-data-about-you-464789. Accessed 21 Feb. 2018.

[7] Larson, Selena. "Every single Yahoo account was hacked." CNNMoney, Cable News Network, 14 Oct. 2017, money.cnn.com/2017/10/03/technology/business/yahoo-breach-3-billion-accounts/index.html. Accessed 22 Feb. 2018.

[8] Houpert, J.F. "What You Need To Know About 1st, 2nd and 3rd Party Data." Intent-Based retargeting, Datacratic,
20 June 2017, www.datacratic.com/blog/first-second-third-party-data. Accessed 20 Feb. 2018.

[9] Savitz, Eric. Defensive Disposal: You Can't Keep All Your Data Together, Forbes Magazine, 18 July 2012, www.forbes.com/sites/ciocentral/2012/07/17/defensible-disposal-you-cant-keep-all-your-dataforever/#fca0fbc6bb3f. Accessed 21 Feb. 2018.

this as a problem and founded *The Compliance, Governance, and Oversight Counsel*. This organization and similar ones are developing internationally to make save disposal of data as protected as possible. They work to combine legal, compliance, record, and IT departments to develop a system that makes online users more protected and saves the organizations from spending millions on storage systems. Unfortunately, current disposal mechanisms and related information is limited as no organization has mastered the disposal process, or is not willing to share.

Since major concerns have been raised about the current data disposal mechanism, major data brokers claim they have 'opt out' options to halt data collection. Currently, StopDataMining.me is a source that provides links to opt out of data collection from the top 50 brokers.[10] Questions have risen in the legitimacy of this 'opt out' system, but legally there is no way to ensure the data is correctly being disposed. In addition, it is impossible to opt out of the potential 4,000 data broker sites that are operating in the United States. Although opting out is a good starting strategy, there is still a great deal to understand when it comes to damage control over a consumer's private information.

Current matters surrounding big data

1. *Negatives of Big Data*

In the current news cycle, Big Data has been an extremely prevalent topic. Most often publications have portrayed the collection of data from users in a negative light. A looming concern for many people nationally is how their information is going to be used against them. Users had a sense of inflated security in the scenarios above, by believing their passwords protected their information.[11] A current topic of concern is known as "consumer scoring". This scoring system uses data collected and inputs into a program that attends to give each consumer number based on varying factors. These scores can be used to predict a person's likelihood to get sick or even pay off a debt. Factors such as purchasing preferences are used to decide if someone is more likely to be diagnosed with a life threatening illness. Looking forward, this numeric system could be sold to healthcare providers to determine individual pricings for life insurance.[12] A more terrifying notion is that consumers may never know when their number is sold from a data broker, effectively changing their lives.

This information is uneasy for consumers, but even breaches in data security are causing concerns for data brokers. Large data breaches have been advertised greatly in the past few years, justifiably raising concerns for users. It is becoming easier for hackers to create and spread malware data on large company software systems. In 2017, a prominent software firm, Bitdefender, discovered that total ransom where payments reached $2 billion.[13] Technology firms are projecting growth in these

[10] Boutin, Paul, supra, note 7.

[11] Larson, Selena. "10 biggest hacks of 2017." CNNMoney, Cable News Network, 20 Dec. 2017, money.cnn.com/20 17/12/18/technology/biggest-cyberattacks-of-the year/index.html.Accessed 21 Feb. 2018.

[12] Boutin, Paul, supra, note 7.

[13] Larson, Selena, supra, note 12.

breaches as data grows in 2018. Determined as "among the worst breaches of all time", the attack on Equifax endangered 145 million people's personal data. Extremely sensitive information being stolen, including social security numbers, exposed thousands of people identity theft.

In another scenario, Yahoo suffered from a significant breach in account information in the year 2013. It was announced that every single Yahoo account (approximately 3 million) was hacked, but Yahoo was not aware of the breach until 2016 former CEO Marissa Mayer told congress. Luckily in this case, only names and addresses were collected and no financial information was collected[14]. Yahoo was also faced with a smaller scale hack in 2014 which affected about 500 million people. Four people were found connected with the 2014 hack, two Russian spies and two hackers. However, no parties were found guilty for the 2013 hack.[15]

Forty-one million customer payment records were lost in a data breach in Target breach in 2013. In addition, contact information was stolen for 60 million customers. The breach caused Target to agree a comprehensive security program to prevent similar instances from happening again.[16]The examples from above show some of the most extreme data breaches that existed, but this continually happens on a smaller scale. Companies need to continue to following in Target's footsteps enhancing protocol to prevent breaches, hopefully instead preventing incidents.

The breaches above occurred without company and broker knowledge, but these are not the only way that data gets into the wrong hands of users. Data scams occur when data brokers knowingly sell information they collect to criminals. Back in 2007, a top data broker, InfoUSA, sold data that led to the scam of approximately 19,000 elderly persons. By releasing information on frequent sweepstakes players, scammers were able to steal $100 million by impersonating government and insurance workers.[17] One of the elderly men scammed out of his life savings was quoted saying, " "I loved getting those calls…. I don't have many people to talk with. I didn't even know they were stealing from me, until everything was gone".[18] In less than a year after that case, United States agencies have filed lawsuits or injunctions against 68 broker firm. The frequency of brokers knowingly assisting criminals has only increased in the 10 years since then. These scams cause questioning of the integrity of the entire industry.

Safety issues have been joined with privacy concerns have around the massive use data tracking. Earlier this year a Belgian court addressed Facebook for breaking privacy laws by tracking user's activity through third-party websites. The court agreed and ruled that Facebook must delete all data and will be fined 250,000 euros each day it does not comply with the court's judgement. Facebook's vice president of public

[14] Larson, Selena, supra, note 8.

[15] Larson, Selena, supra, note 12.

[16] McCoy, Kevin. "Target to pay $18.5M for 2013 data breach that affected 41 million consumers." USA Today, Gannett Satellite Information Network, 23 May 2017, www.usatoday.com/story/money/2017/05/23/target-pay-185m-2013-data-breach-affected-consumers/102063932/. Accessed 21 Feb. 2018.

[17] Boutin, Paul, supra, note 7.

[18] Duhigg, Charles. "Firms sell elderly Americans data to telemarketing con artists." The New York Times, The New York Times, 21 May 2007, www.nytimes.com/2007/05/21/world/americas/21iht-data.1.5803543.html?_r=0. Accessed 21 Feb. 2018.

affairs has agreed to following the ruling.[19] This privacy issue is also being recognized in the United States government. Five years ago, President Obama published what he called the Consumer Privacy Bill of Rights. This document pointed out the lack of legislation for privacy in regards to data. Three years later, when this Bill of Rights was officially proposed to legislation, it was attacked by both sides and therefore never made it to a vote.[20] Although no agreement has been met since, these attempted actions from both governments lay out immediate privacy concerns.

Not only do these threats pose privacy issues, but the mass collection of data sets easy targets for dangerous organizations. In 2016, a hacker pleaded guilty to providing material to the US enemy after hacking United States corporate computers. The Islamic State terror group he provided the information to leaked the private information of 1,351 military and US government officers.[21] These threats to privacy and data security pose physical, monetary, and emotional concerns as the situation progress. The light that has been shed on the potential detrimental effect of data breaches have made both businesses and people more aware of the posed risk. Current issues are inevitably changing consumer behavior.[22]

2. *Positives of Big Data*

Although the current news issues tends to focus on threatening aspects of big data, many data brokers are working to point out the positives of their work. In a simple sense, mass data collection creates a convenience from custom advertising and store browsing experiences. The ability to use big data to enhance user experiences was one of the first reasons big data was collected. This allowed customers to have an easier time going to a clothing store, buying groceries, or even minimizing the wait time for their favorite amusement park ride. When browsing the internet for a product, the data you are searching for is collected and brings you thousands of items you may want. By allowing data analysts to understand your habits, they are able to add convenience to your life in ways you didn't know were possible.

The positives of data collection span to more important topics such as national security. The issue that occurred in April 2017 scared many citizens when an anonymous group leaded a variety of hacking tools allegedly used by the National Security Agency. This can be viewed as a terrifying breach of privacy, but way to understand the issue viewing it at a safety from terrorism. These systems that are developed can help the United States and other foreign partners safe from breaches that can occur from extremely dangerous organizations.[23] The NSA's data collection tools have most likely protected United States citizens on multiple occasions without anyone outside

[19] Armerding, Taylor. "The 5 worst big data privacy risks (and how to guard against them)." CSO Online, CSO, 14 July 2017,www.csoonline.com/article/2855641/privacy/the-5-worst-big-data-privacy-risks-and-how-to-guard-against-them.html.

[20] Stevens, Micheal. "Can Big Data Help Keep Us Safe?" Can Big Data Help Keep Us Safe? | IBM Big Data & Analytics Hub, www.ibmbigdatahub.com/blog/can-big-data-help-keep-us-safe.

[21] Kravets, David. "ISIS hacker pleads guilty to giving terrorists US military kill list." Ars Technica, 15 June 2016, arstechnica.com/tech-policy/2016/06/isis-hacker-pleads-guilty-to-giving-terrorists-us-military-kill-list/. Accessed 22 Feb. 2018.

[22] Larson, Selena, supra, note 12.

[23] Larson, Selena, supra, note 12.

the organization finding out.

Companies such as IBM are noticing how useful this data can be in detecting threatening activity. IBM is creating a Big Data Analytics Platform to enhance government agencies ability to predict and prevent real time.[24] Part of their plan is to incorporate real-time social media into threat analysis. Their article describing their idea is titled "Threat Prediction and Prevention Solution Framework from IBM". It suggests what sort of information could be used in safety precautions. For example, it suggests uncovering suspicious relationships between people, events, and transactions. Analyzing big data in a new way can tenfold increase general safety. Police departments all over the United States have begun using IBM's ideas to prevent crimes on a smaller scale. Big Data has even been referred to as "the largest natural resource of law enforcement" by a senior consultant at IBM.[25]

Rights and laws associated with big data

1. Obligations of Corporations

With the power of Big Data, and the potential issues that could arise from, corporations have obligations to their customers to protect them. There are two major users of big data, corporations and the government, and they both use the collected data for different purposes.[26] Whether it is other companies that must be kept at bay or our own government, the companies must try to protect their customers. Corporations having access to your personal data may not seem like a risk when they use it mostly for advertising, the most worrying is the government. They can use the collected data to see if people are engaging in suspicious behavior. To help protect United States citizens, laws have been used to protect privacy, meaning the government must have a legal basis for mass data collection. This results in some protection for Big Data and the companies that collect and store it, but sometimes they must still fight to keep their information private from the governmental agencies as well as competitors.

2. Rights of Users

The biggest concerns with the rights of the users are if the data that is being collected is being shared and the invasion of privacy that can result. There has been a lot of inquiry as to what specific rights are protected or not by the current laws and regulations that can be carried over to Big Data. Corporations are not as prevalent in this area because they have contracts with the customers that say what the company can, or can not, do with the data they collect. Governmental agencies, such as Homeland Security, have had issues with what rights are and are not protected which has lead to research in both civil and criminal justice systems. Since Big Data breaches and investigations using those tools has become more prevalent in court cases in the United States, they

[24] Stevens, Micheal. "Can Big Data Help Keep Us Safe?" Can Big Data Help Keep Us Safe? | IBM Big Data & Analytics Hub, www.ibmbigdatahub.com/blog/can-big-data-help-keep-us-safe.

[25] Wyllie, Doug. "How Big Data is helping law enforcement." PoliceOne, 20 Aug. 2013, www.policeone.com/p olice-products/software/Data-Information-Sharing Software/articles/6396543-How-Big-Data-is-helping-law-enforcement/.

[26] ARTICLE: Small Data Surveillance v. Big Data Cybersurveillance, + Copyright (c) 2015 Margaret Hu., 42 Pepp. L. Rev. 773.

are leading the courts to recognize the need to establish rules and boundaries for use of Big Data without legal cause. For Homeland Security, there has to be a balance between the use of the information for national safety and the data they can use without a legal standing.

Every time a person clicks the "accept terms and conditions" when using new software or creating an account, they are waiving all their rights to whatever has been listed in that agreement. Normally there are rules and regulations protecting consumers, however when they waive their rights to gain access, the company can do whatever they like with the data they collect from the user. Corporations that collect data are restricted by privacy laws, until the terms and conditions are accepted. However other entities, such as the government, that want to use to the data to potentially incriminate people have a more difficult time with privacy laws because the user is not allowing the data that is collected on them be used for that purpose, unless it is stated otherwise in the agreement they accepted.

3. *Legal Reviews and Cases Analyzing the Issue of Big Data*

Some of the first examples of privacy related issues involving collected data date back to the 1970's. One example of this was the case of Whalen v. Roe, 429 U.S. 589 where a pharmacy kept records of a person that was a potential abuser of their prescriptions.[27] The pharmacy thought it would be a good idea to keep on record that the person may be abusing their prescription, the person did not like that and took them to court over it under their right to privacy. The court decided that keeping those records were violating the man's rights to privacy. While this is an issue with personal data, and the collection of it, there is much more happening now with the power of the internet.

As time went on, there were more privacy acts that were made to protect the people. For example 100 P.L. 618, 102 Stat. 3195, 100 P.L. 618, 102 Stat. 3195 was set into place in 1988.[28] It is called the Video and Library Privacy Act of 1988, and it was put in place to protect the personal privacy of people's rentals, purchases, or delivery of video tapes or similar audio visual materials. The courts believed that what a person was watching was private and should not be monitored. This statue was one of the beginning of the type of privacy issues that eventually evolved into what we call Big Data now.

Another example of medical privacy issues was the case of Sorrell v. IMS Health.[29]This case was about restricting the sale, disclosure, and use of pharmaceutical records that reveal the prescribing practices of other doctors. This is is to protect the doctor of judgment from other doctors on how they treat their patients and how they prescribe medications. It was thought that each doctor should be able to privately prescribe their patients medicine without being concerned of what other doctors thought, or that information being collected.

The collection of data can be looked at many different ways, two of which are small data surveillance and big data cybersurveillance. Data collection and mining is a whole new area that is still expanding rapidly, the law review Small Data Surveillance

[27] Whalen v. Roe, supra, note 6.

[28] 100 P.L. 618, 102 Stat. 3195, 100 P.L. 618, 102 Stat. 3195.

[29] Sorrell v. IMS Health.

v. Big Data Cybersurveillance, + Copyright (c) 2015 Margaret Hu., 42 Pepp. L. Rev. 773 goes in depth of comparing those two types of data collection and how different corporations use each. For example typically data collection companies use big data to collect massive amounts of information on many people to discover what they should focus on for things like advertising. While small data surveillance is used for things like determining if a person is involved in illegal activities by governmental agencies. However neither category is restricted to either group of data mining.[30]
The law review Consumer Protection in the Age of Big Data, 93 Wash. U. L. Rev. 859, holds an interesting perspective with regard only not to the aspects of users security, privacy and consent but to Big Data analytics potential effect on the insurance markets.[31] Specifically, the power insurance companies given the necessity of life and oligopolistic natures of the insurance market.[32] Big Data can be a major factor in the decision making process of who they choose to insure. Consumers can be placed at a disadvantage with the fact that insurance companies are collecting data from sources such as social media. Given the abundance of information posted on social media websites such as Facebook, Instagram and Snapchat, insurance companies have access to the behaviors potential policyholders engage in that could put them at risk for health issues and thus increasing costs for insurance companies.
The case Spokeo v. Robins highlights the recklessness that can result from the users of big data and the responsibility that collectors have not only to secure big data but to ensure its accuracy. In this case, Spokeo Inc., an online source that collects statistics about individuals, collected and displayed inaccurate information about Thomas Robbins, Robbins alleged. Robbin's believes that Spokeo's data about him was accessed by employers which raised Robbins' concern that the inaccurate information harmed his career. Convinced that it had, he took the defendant to court for the potential damages under that Federal Fair Credit Reporting Act (FCRA), arguing that Spokeo failed to make reasonable efforts the ensure that the information reported about him was true. Robbin's case was originally dismissed in the district court due to his inability to prove "injury in fact" under Article III of the constitution.[33] It was later reversed by the Ninth Circuit (appellate) court due to the individualistic nature of Robbins's injuries, The Supreme Court established that even though the appellate court established Robbins and an individual target, it did not fulfill sufficiency of Robbins' injury.[34] The judgement for this case was issued in 2016 finding that Robbins amended that Robbins did in fact prove injury under the FCRA. A positive highlight from 131 Harv. L. Rev. 894 states that relatives to data breaches is that the burden of proof for a plaintiff is lightened in that they only have to prove that their data was put in a position to be compromised not that it actually had been.[35]

[30] Margaret Hu., 42 Pepp. L. Rev. 773.
[31] Article: consumer protection in the age of big data, 93 Wash. U. L. Rev. 859.
[32] Id.
[33] 131 Harv. L. Rev. 894.
[34] Id.
[35] Id.

The history and future of big data

1. *When did Big Data Become a Public Concern?*

Issues with storing personal data has been around a very long time. Before computers were common household items, the most common area of using personal data when there was a grey area was in the medical field. For example there was a person who had been noted for potential substance abuse on their doctor's prescription.[36]The term Big Data came around in the late 1990's when computers could not store all the information that was needed on the local memory, that meant they had to acquire more resources.[37] The explosion of the amount of data that was available, that also could be useful, became known as Big Data. It really became a common name when technology advanced to the point of quick and easy data recording and storage. In the 1990's it was used for much smaller amounts of information due to the capacity they had at the time to store data. It was around the mid 2000's that computer storage technology became good enough to store large amounts of data, leading to a boom in the size of the internet as well as the data that was collected from it. The collection of personal data was not an important thing for the average person until the early 2010's. The first data collection agencies began appearing due to cheap data storage solutions for consumers and the idea of collecting people's behavior to analyze what products or services people want when accessing the internet.

2. *The Future of Big Data*

The future of Big Data is still unknown for the everyday consumer as the new area of technology continuously expands at an incredible rate. As time passes and more cases are brought to court, hopefully there will be an increase in the protection of people's privacy because at this point in time there is not much at all protecting people from the collection and mining of their personal data for the resale to other companies. Big Data and data analysis has become such a major part of life everyday life, and will only continue to increase in size and importance that some schools are beginning to offer programs to students. Krannert, in Purdue University, has begun to offer a new graduate program in business analytics and information management.[38] They have seen the need for more data scientists and want to prepare students for that opportunity. This field is growing so dramatically that it is estimated there will be an increase in demand for data scientists by 28 percent by the year 2020. That is an increase from a little over 364,000 job openings to a new total of over 2,700,000 job openings.[39]

In addition, many politicians are noticing the need for a more definite legal oversite in

[36] Whalen v. Roe, 429 U.S. 589, 97 S. Ct. 869, 51 L. Ed. 2d 64, 1977 U.S. LEXIS 42 (U.S. Feb. 22, 1977)

[37] Press, Gil. "A Very Short History Of Big Data." Forbes, Forbes Magazine, 21 Dec. 2013, www.forbes.com/ sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#7ffb63e265a1.

[38] "Data Masters." Data Masters: Krannert Introduces New Grad Program in Business Analytics and Information Management - Purdue Krannert Magazine,www.krannert.purdue.edu/konline/2016f/features/BigData_sidebar1.php.

[39] Singer, Natasha. "Acxiom Lets Consumers See Data It Collects." The New York Times, The New York Times, 4 Sept. 2013, www.nytimes.com/2013/09/05/technology/acxiom-lets-consumers-see-data-it-collects.html.Accessed 21 Feb. 2018.

the world of Big Data. Last year Senator Edward Markey even introduced a bill called the Data Broker Accountability and Transparency Act of 2015. He spoke about his bill and claimed that consumers should own full rights to their data, not corporations. The bill would allow individuals to review and to follow necessary steps to correct any illegitimate information. Senator Markey isn't the only governmental figure looking into more legal regulations. The Senate Subcommittee on Privacy, Technology, and the Law has held multiple hearings with representatives from large data brokerage companies. The government officials that spoke out demanded more transparency in data collection. In the future, these conversations will become more regular on the floors of the House and Senate in attempts to standardize mass data collection.[40]

Conclusion

Big data is a powerful in that it allows for users with ethical intentions to analyze data to reach their objectives however, the other side to that positive power is that there are users who seek big data to cause harm to others. Because Big Data is so powerful, it is obvious that the demand to acquire and analyze it will not subside from both those who plan to use it for good and evil. With issues that have emerged with regard to data insecurity and susceptibility to breaches, the commonality and frequency of such breaches poses great concerns and risks for people's information to be exposed. News stories such as Richard's experience, the Yahoo hacks, and IBM's new policy have demonstrated first hand the need for some regulation with Big Data. Governmental figures, such as Senator Markey have begun leading the crowd in demanding more regulations for Data Collection. In the future, it will be critical for laws and regulations to stay up to date with Big Data because as time goes on, the more data is collected and the more data available the bigger the risk.

[40] 37] Boutin, Paul. "There's Very Little Oversight in the Industry of Data Brokers." Newsweek, 16 June 2016, www.newsweek.com/secretive-world-selling-data-about-you-464789.