

9-3-2015

The Big Data Debate Today

Bridget Fahey
Santa Clara University

Follow this and additional works at: http://scholarcommons.scu.edu/engl_176

 Part of the [American Popular Culture Commons](#), [English Language and Literature Commons](#), [Film and Media Studies Commons](#), and the [Nonfiction Commons](#)

Recommended Citation

Fahey, Bridget, "The Big Data Debate Today" (2015). *Advanced Writing: Pop Culture Intersections*. Paper 3.
http://scholarcommons.scu.edu/engl_176/3

This Article is brought to you for free and open access by the Student Scholarship at Scholar Commons. It has been accepted for inclusion in Advanced Writing: Pop Culture Intersections by an authorized administrator of Scholar Commons. For more information, please contact rscroggin@scu.edu.

Bridget Fahey

Professor Hendricks

English 176

September 2, 2015,

The Big Data Debate Today

“...Big data refers to things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value, in ways that change markets, organizations, the relationship between citizens and governments, and more.”¹

Today, technology is more a part of our lives than ever before. With more and more people all over the world gravitating towards social media and using sites such as Twitter and Facebook, more of our private lives is available to others than ever before. In his article “Big Data and Privacy” Tom Price explores the controversial role big data plays in our society today. While Price presents many of the benefits this data offers, he also notes its controversial nature. The main issue that big data collection and usage creates is that of consumer privacy. As Edward Snowden discussed in an interview with TV host John Oliver and throughout the documentary CitizenFour, organizations such as the NSA have access to unlimited amounts of data on the American public. Thus, while new technology has allowed for numerous advancements, it has also opened the doors for a debate on issues of privacy infringement. The world of big data has led society to a difficult crossroads, where people desire convenience and better, faster technology but

¹ Mayer-Schönberger, Viktor, and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. (Boston: Houghton Mifflin Harcourt, 2013)

simultaneously want to maintain their privacy. Although a huge issue, it seems many people do not recognize it as such, due to the fact that people are simply not aware of the existence, much less the use, of vast amounts of information, especially since big data is a relatively new concept. Additionally, it is often downplayed or presented in a very skewed manner. The government and other organizations that use this data for either surveillance or business practices do not want consumers to know how their private information may in fact be used. Thus when the risk of privacy infringement involves the government, it is usually justified for reasons of defense. Similarly, when companies are accused of misusing people's data, they often argue that its for the good of the consumer and allows them to better serve their customers.

Through analyzing Edward Snowden's data leak alongside the benefits of big data collection discussed by Tom Price and Viktor Mayer-Schonberger and Kenneth Cukier, it is evident that our society faces a difficult position concerning big data usage. Perhaps the most problematic part of this issue is the overwhelming lack of understanding the general public has about the matter. This lack of knowledge is largely due to the complexity of the subject matter as well as the media and government's approach to presenting the issue. In addition, there is the fact that the existence of such troves of data is so new. Big data, in the sense we now use the term, did not exist a decade ago. As we move forward, technological advancements will continue and our privacy will be further called into question. Therefore, it is important that the subject be explained in more relatable terms so that the public understands how their data is used and clearly sees both sides of this controversial big data debate.

Over the last few years big data has allowed for a great deal of innovation. In his article, "Big Data and Privacy," Tom Price introduces the topic of big data or "the collection and analysis of enormous amounts of information by supercomputers."² Although this data has allowed for numerous advancements it has also created concern amongst the public over issues of privacy. With more and more people using the Internet, entering information into social media sites such as Facebook and Twitter, the protection of the users privacy is called into question. In his article Tom Price presents both sides of the big data debate, providing his reader with an overview of both the pros and the cons.

Over the last few years, big data has allowed for numerous advancements and benefits. Besides the obvious technological innovations, this data has also paved the way for medical, business, security and even scholarly advancements. As Price mentions in his article, "Big data has led to cutting-edge medical discoveries and scientific breakthroughs that would have been impossible in the past: links between genetic traits and medical conditions; correlations among illnesses, their causes and potential cures; and the mapping of the human genome."³ Additionally, big data collection and analysis has allowed for scholars to gain new perspectives and gather information on historical figures. Now, scholars are able to study and access numerous books quickly and make conclusions about literary periods and genres that in the past would have taken years or perhaps may not have even been possible because they lacked access to the information.

² Price, Tom. "Big Data and Privacy." *CQ Researcher* 23, no. 38 (October 25, 2013): 909-32. <http://library.cqpre,ss.com.libproxy.scu.edu/cqresearcher/cqresrre2013102500>.

³ Price, "Big Data and Privacy."

Similarly, websites are able to provide their consumers with helpful information that is geared specifically toward the particular customer, based on the information that individual provides.

Perhaps the most significant of the advancements this data has allowed are related to public health and medical advancements. A positive example of this in real time use occurred during the H1N1 flu outbreak that took place back in 2009. As Viktor Mayer-Schonberger and Kenneth Cukier discussed in their book, “Big Data: A Revolution that Will Transform How We Live, Work, and Think,” Google’s search engine played a key role in helping the CDC determine where flu outbreaks were occurring. Google has about three billion searches per day and saves these searches for data analysis.⁴ This data allows us to track more information than ever before and most importantly utilize that information in a helpful manner. It is clear that big data has completely altered the way we approach all different aspects of our lives. The advancements discussed by both Price, Mayer-Schonberger and Cukier are just a few made possible by this data collection and in the upcoming years, more and more will surely become possible.

Although the collection of big data has allowed for numerous advancements, its usage has also led to issues of privacy infringement. As Price explains, “Internet platforms, such as Facebook or Google, act as big data ‘sensors,’ gathering information about people just as a thermometer gathers temperature information⁵. Back in 2010, Facebook admitted that many of its third party applications were tracking user data and passing it on for marketing and advertising purposes. Again, “in June of 2014, Facebook

⁴ Mayer-Schönberger, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*.

⁵ Price, "Big Data and Privacy."

‘disclosed that it had tested to see if emotions were contagious [by] deliberately manipulating the emotional content of the news feeds for 700,000 people.’ OKCupid, a dating website, published results of three experiments it ran on its users. The company’s president subsequently stated, ‘If you use the Internet, you’re the subject of hundreds of experiments at any given time, on every site... That’s how websites work.’⁶” With so much information being recorded, it is easy for people to feel as if they have no real privacy anymore. Therefore, although the collection big data has numerous benefits, it also has its share of drawbacks. Thus, moving forward it is important for consumers to carefully evaluate the role that this data plays and understand the way their information may be used.

Although there has been a debate around advertisers using big data, many consider big data's use in government surveillance to be a bigger threat to privacy. In an episode on government surveillance TV host John Oliver discussed the reauthorization of the patriot act and the controversial actions of Edward Snowden. His main focus during his first part of the show was on the controversial patriot act and specifically section 215. This section of the act allowed access to records and other items under the foreign intelligence act. The initial version of section 215 allowed for the National Security Council (NSA) to access "any tangible things for an investigation to protect against international terrorism..." And due to its overwhelming vagueness was what John Oliver called a essentially a "blank check."⁷ Today, people do not often realize the extent of the

⁶ Baranetsky, Victoria D.1. 2014. "Social Media and the Internet: A Story of Privatization." *Pace Law Review* 35, no. 1: 304-343. *OmniFile Full Text Mega (H.W. Wilson)*, EBSCOhost (accessed September 2, 2015).

⁷ Snowden, Edward, interview by John Oliver, *Government Surveillance*, YouTube. Last Week Tonight, 9 Apr. 2015

reach organizations like the NSA has when it comes to personal information. In a press conference President Obama assured the public that although the NSA has access to certain information they are "only looking at general info such as phone numbers and duration of calls and not names or content."⁸ However, in June 2013, a 29 year-old contractor for the NSA named Edward Snowden leaked numerous documents revealing just how much information the government has access to. His actions were incredibly controversial as he put American security at high risk yet simultaneously revealed some startling and important facts to the American public.

When it comes to big data debate is that people largely underestimate what kind of information the government and other organizations have access to. In his interview with John Oliver, Edward Snowden discussed the controversial access the NSA has to personal information shared via the phone and Internet. In terms of phone calls, he explained that the government has access to information regarding phone calls including the length of the call but also the content in some cases. Additionally, they can see the numbers you call and how often you call, so although the information they access may seem vague, anyone could easily look it up and infer a lot about you. In terms of data shared over the Internet including pictures, email messages etc. the NSA has the ability to pull this information if it leaves US borders. For instance, if you have a Gmail account and Google moves your information to another server for storage located beyond US borders, that information passes through the NSA database. Therefore, it is quite possible

⁸ Snowden, *Government Surveillance*

that a private message or phone call might end up in their system. CitizenFour,⁹ a documentary on Edward Snowden and his mass data leak, takes an even deeper look at the implications of this big data and the possible breaches in privacy that may occur. Snowden chose to release all of the data in stages and through the help of a select group of journalists. The second set of documents he released exposed just how much data the NSA actually has the potential to access and more importantly, how easy it is for them to justify investigations into people's private information. The documents basically proved that the NSA had the ability to reach into servers and extract a wide range of data including, audio, video, and email messages and did not require a permit to do so if the person was communicating with someone outside U.S. territory or anyone considered a foreigner. This of course would include a large number of people, especially since so many people have family outside of the U.S. whom they regularly communicate with. It also raised questions about the other information that could be potentially intercepted within U.S. borders between fellow Americans. While many people would likely support the NSA pulling these records under suspicion and therefore acting in the best interest of the American public, according to Snowden this was not always the case. Both CitizenFour and the Interview with John Oliver made one thing very evident: most people aren't even aware that so much information about them is captured and used; and they aren't informed about privacy risks. At the end of his interview with Edward Snowden, John Oliver pointed out that if the American public fully understood all that the NSA has access to, many more people would be upset. Before traveling to visit Snowden, Oliver interviewed some random people on the streets of New York and asked

⁹ Julia Lane, Citizenfour, Documentary, Laura Poitras (2014; City: The Weinstein Company, 2014), Film.

them about Edward Snowden and data privacy and these were just a few of their responses: "I have no idea who Edward Snowden is," "I've heard the name but can't picture who it is," "I think he sold some information to people," "He leaked documents about our operations in Iraq."¹⁰ It quickly became evident that many people are not informed about the privacy issues when it comes to their data. Additionally, most people he asked not only had no idea who Edward Snowden was but they also couldn't answer what information he had leaked, all they knew was that it had been bad for our country. It is likely that many individuals are not even aware that the government has access to their "private" information. Furthermore, for those who do feel unsure about their privacy it is unlikely that they know exactly what information is at risk. The big data debate begs the question of "do we have a right to know?" As Snowden himself explained, his reason for releasing the information he did was to educate people and ensure they were aware of how their information may be at risk. "I was motivated by a belief that citizens of the world derive to understand the system in which they live. My greatest fear was that no one would listen to my warning" (Snowden). While, yes it can be argued that the government accessed our personal information for our own protection, it can also be argued that we at least have a right to know about it.

There are a couple primary reasons that these big data issues are not so widely understood. For starters, the topic is rather complex and most explanations tend to be very technical, thus the average person is unlikely to fully understand the extent of the issue. As John Oliver pointed out during his interview with Edward, Snowden, although the data Snowden was presenting was incredibly important, most of it was being

¹⁰ Snowden, *Government Surveillance*

explained in a manner that escaped most people. Oliver suggested putting the information into terms that the general public could understand. Oliver of course found a comedic but very effective method of communicating the potential severity of the situation. Instead of presenting the NSA's capabilities in such technical terms as Snowden had, he chose to explain it in a manner that would grab everyone's attention: 'Dick Pics.' While it seems silly, his method proved to me very successful. When he first interviewed people on the street about privacy infringement they had originally responded in a very unconcerned manner and most explained that they did not think anyone had access to their private messages. However, when John Oliver asked them how they would feel about the government intercepting pictures of their 'junk,' everyone agreed they would be furious. Although this is obviously not what the NSA targets when they are looking for suspicious activity, if they wanted to they would have no problem looking at peoples private messages including any 'Dick Pics.' As was evident through the immediate change in reactions of the interviewees, if the issue at hand was always put in simpler, more relatable terms, more people would see its truly controversial nature.

Another way to put this data issue into relatable terms is explaining the situation in relation to people's social media use. Almost anyone with regular access to the Internet likely has some sort of social media account or at least uses major search engines such as Yahoo and Google. As Viktor Mayer-Schonberger and Kenneth Cukier make clear in their book, "Big Data: A Revolution that Will Transform How We Live, Work, and Think" is that although websites like Facebook and Google are "free" in a monetary sense, people unknowingly provide personal information about themselves to these companies. If you think about it, platforms such as Facebook are constantly collecting

data about their users to improve site content and in some cases providing this data to advertising agencies as was pointed out earlier in Tom Price's article. In their book, Mayer-Schonberger and Cukier make the argument that the reason those sites are free is because you are supplying them valuable information about yourself. So, although we do not usually recognize this exchange, it is in fact occurring¹¹. Facebook and others sell information about you to other companies who use it for targeted advertising. Facebook charges these entities rather than the user. But the user is not usually aware of what is being collected, whom it is shared with and how they use it. If you think about it, how often do advertisements for products or services that you recently searched for on the Internet show up while you are using sites such as Facebook? While in some ways this may be viewed as a convenient service, I think most people would agree it is also a bit creepy. When using the Internet, we like to think that the information we are searching for is private but seeing 'relevant' ads appear on the side of the screen proves otherwise. It also implies that these major sites and search engines are essentially connected in the sense that information from one may be transferred or perhaps accessed by another. If this is the case, it seems only fair that consumers should have knowledge of this and that more sites should make it ways to block such access if desired. In the past, sites like Facebook for instance have made their privacy settings fairly difficult to access and even more difficult to understand. While today it is definitely laid out in a simpler format, many people are still unaware of just how much information they are signing over to these sites by using them. Even if the 'Dick Pic' explanation did not resonate,

¹¹ Mayer-Schönberger, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*.

anyone with a Facebook or even just an email account can relate to the argument Mayer-Schonberger and Cukier make present.

The government and the media outlets largely control the angle at which information concerning big data is presented to the public, and thus control how much impact this information has. The people speaking out against privacy infringement caused by big data are often brushed aside or portrayed, just like Snowden as the ‘bad guy.’ When John Oliver asked people on the street who Edward Snowden was, many replied with “I have no idea.” Several others replied “the information he had leaked had been detrimental to the American military and had put our soldiers at risk.¹²” Not a single person replied that an answer that cast Snowden in an appealing or heroic light. In cases like this one, the media plays a huge role, releasing the stories and putting the person responsible for the leak in the spotlight. When this story first hit, almost every news channel was reporting it and even questioning the government’s role and the power of the NSA. However, despite the outrage this case initially caused, six years later it seems to no longer be a topic of concern. This is where the problem lies, as most people have forgotten all about Snowden and most importantly the information he shared. Moving forward it is crucial that the public better understands this problem at hand and takes precautionary steps to protect their privacy. Although Snowden’s information created a great deal of buzz when it was first released, because the media stopped reporting about it and the government portrayed him in such a negative light, people have returned to thinking that there is no issue. It seems that today, people are once again not aware of the data others have access to, and have thus lost control over their own personal data. Most

¹² Snowden, *Government Surveillance*

importantly, if they are not aware, then they have surely not given their permission for others to use their information.

People quickly classify whistleblowers, like Snowden as traitors because the government paints them as such. As Snowden proved, they are quickly shunned and painted in a very negative light. For instance, when John Oliver asked people on the street who Edward Snowden was, most of their answers were easily synonymous with the word traitor. Almost everyone agreed that what he had done had been bad for the United States and hardly any portrayed his actions as honorable. When the issue of data usage is linked government surveillance, it is likely that the issue quickly takes a political turn. "I don't think that Mr. Snowden was a patriot, I called for a thorough investigation of our surveillance operations before Snowden made these leaks. My preference and I think the American people's preference would have been for a lawful, orderly, examination of these laws a thoughtful, fact based debate that would then lead us to a better place" (Obama). When the President of the free world is calling someone like Snowden a traitor, the public surely begins to view him as such. In moments of controversy it is only natural that the American public keys into the media and the government's response to the events when forming their own opinion. It is particularly difficult for people to see the big picture and even know what to believe when the matter is so clearly skewed in favor of one side. When it comes to issues of privacy that involve government surveillance, it is not shocking that the government and eventually the media outlets turn the public against the person speaking out against the system. This is done in defense of the current system in place and ironically to protect the privacy of organizations such as the NSA. It is a very controversial matter, for on one hand, it is completely understandable, that the U.S.

government feels obligated to protect information regarding national security. However, at the same time the United States is a country that prides itself on free speech but the fact that these people like Snowden seem to be heavily persecuted for speaking up seems to suggest otherwise.

The angle that media outlets take on reports also directly influences the way the message is perceived to the public. Although when Snowden's story initially leaked, people across the world were enraged. However, today it seems that Snowden and what he released have been long forgotten. This was likely due to the fact that even news outlets backed away from his story or began persecuting him as an American traitor. In a private message to CitizenFour director, Laura Poitras, Snowden revealed the influence politics may have over media and news outlets. In his message, Snowden warned Poitras, "the FBI has authorization to work with the CIA and a number of unnamed foreign partners to team up in finding out my plans and the location of people in contact with me worldwide...They've been tasked to use 'all appropriate means'—including government pressure 'where appropriate'--to persuade media to refuse publication"¹³ According to Snowden's source, the news, which is supposed to expose the truth and share all sides of a given story, is clearly capable of being manipulated. As fewer and fewer news outlets continued to report this story or began to report Snowden as a traitor, the less influential the message he so desperately hoped to share became.

If people like Snowden were portrayed in a more heroic light or even just given a greater chance to speak their mind, perhaps their information would be better received.

¹³ Julia Lane, *Citizenfour*, Documentary, Laura Poitras (2014; City: The Weinstein Company, 2014), Film.

Perhaps the most important component of CitizenFour is the way Snowden is portrayed in a more heroic light and the fact that the director Laura Poitras, allows us to hear and understand his reasoning for his actions. The film itself could have easily been very dry because the topic of big data itself is very technical and not something that most people easily follow. Although the majority of the documentary takes place in a small Hong Kong hotel room, Poitras captures an exciting, emotional and eye-opening side of this story. While the news portrayed Snowden in a very negative light and eventually pushed him out of the picture, Poitras gives him a voice and a story. Most importantly, the film allows us to better understand *why* he made this choice. It becomes quite clear that he knowing sacrificed everything he had including his family, friends and life in the United States to release this information that he thought was crucial for the American public to know. If anything, he appears to be a patriot rather than a traitor or security threat. Perhaps if people like Snowden were depicted in this same heroic light more often, their message would have been better understood and more importantly would still be something widely talked about today. Topics concerning personal privacy should be something we can openly discuss and not be something forbidden or shunned in the manner it most frequently is.

It is one thing if people know exactly how their data is being used and then make an informed decision not to be concerned about the matter but it seems that many people today are not fully informed. This lack of awareness was by far the most startling point brought out by the Snowden interview. As John Oliver points out regardless of whether or not we feel that Snowden did the right thing or not, we cannot forget or ignore this new knowledge that we now have. When it comes to the protection of our country, we

all want to obviously feel safe and no one can deny that security measures taken to protect the American public are important. However, the manner in which data is currently collected may be overstepping some privacy boundaries. The most important part of this big data debate is educating people about the matter and making sure they are informed about the way their private information may be affected.

When individuals are informed about the way in which their personal data may be accessed, they take a firm stance on the issue. Perhaps the best example of this rise to action is German Chancellor, Angela Merkel's response to evidence of NSA surveillance on Germany and other European countries. Although she argued that Germany already had strict regulation concerning data and privacy, this knew knowledge allowed her to call for wider action across Europe. In response to this information, "Ms. Merkel said she now believed that only a broader pact could be effective. 'That has to be part of such a data privacy agreement because we have great regulation for Germany, but if Facebook is registered in Ireland, then it falls under Irish jurisdiction,' she said. 'Consequently we need a common European agreement.'¹⁴ When people are aware of how their data may be manipulated, then they can begin to take precautionary steps to protect it. Today big data and all other aspects of our lives are intertwined. And when people understand this they may then recognize how this "new data paradigm compromises our understanding of the social anniversary of the social and political implications of the compromises."¹⁵

¹⁴ Eddy, Melissa, and James Kanter, H, "German Chancellor Urges European Union to Tighten Rules for Internet Privacy.," *New York Times* July 2013: A6(L).

¹⁵ Elmer, Greg; Langlois, Ganaele; Redden, Joanna. *Compromised Data : From Social Media to Big Data*. New York: Bloomsbury Publishing, 2015. <http://santaclara.ebib.com/patron/FullRecord.aspx?p=2077089>

While it is very clear that big data has numerous benefits, it also has a fair share of risks. As Tom Price discussed in his article, big data provide our society numerous advancements including technological, medical and scholarly ones. However, as Edward Snowden revealed it has also opened the door for some controversial debates concerning personal privacy. As Ted Zarsky points out society is currently in a “privacy innovation conundrum”¹⁶. In other words we are stuck between wanting more freedom and more innovation and wanting to maintain our privacy. As a society, it is important to know exactly how our data is being used and what information may be susceptible. This is a fairly new issue and is not something taught in schools. It’s complex and new. And it involves technology and many people do not really understand digital, electronic technology. They use it, but they don’t fully understand its capabilities. Using big data often relies on sophisticated computer engineering and statistics and most people are not computer engineers nor sophisticated statisticians. Therefore it is critical that the matter be put into terms that everyone may understand. As we move forward, technological advancements will continue and it is crucial that the public understands how their data is used and clearly sees both sides of this controversial issue. It is not an easy matter to classify as good or bad and right now the most important thing is just that people understand exactly what is going on. Only once people are fully informed can they take

¹⁶ Ted Zarsky, Z.1. "The Privacy-Innovation Conundrum." *Lewis & Clark Law Review* 19.1 (2015): 115-168.

an educated stance on this issue, and decide whether to support government policies and business practices that give citizens and consumers control of their information.

Works Cited

1. Baranetsky, Victoria D.1. "Social Media and the Internet: A Story of Privatization." *Pace Law Review* 35, no. 1 (Fall2014 2014): 304-343. *OmniFile Full Text Mega* (H. W. Wilson), EBSCOhost (accessed September 2, 2015).
2. Elmer, Greg; Langlois, Ganaele; Redden, Joanna. *Compromised Data : From Social Media to Big Data*. New York: Bloomsbury Publishing, 2015.
<http://santaclara.ebib.com/patron/FullRecord.aspx?p=2077089>
3. Julia Lane, *Citizenfour*, Documentary, Laura Poitras (2014; City: The Weinstein Company, 2014), Film.
4. Kanter, Eddy, Melissa, and James, H, "German Chancellor Urges European Union to Tighten Rules for Internet Privacy.," *New York Times* July 2013: A6(L).
5. Mayer-Schönberger, Viktor, and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. (Boston: Houghton Mifflin Harcourt, 2013)
6. Price, Tom. "Big Data and Privacy." *CQ Researcher* 23, no. 38 (October 25, 2013): 909-32.<http://library.cqpress.com.libproxy.scu.edu/cqresearcher/cqresrrre2013102500>.
7. Snowden, Edward, interview by John Oliver, *Government Surveillance*, YouTube. Last Week Tonight, 9 Apr. 2015
8. Zarsky, Ted Z.1. "The Privacy-Innovation Conundrum." *Lewis & Clark Law Review* 19.1 (2015): 115-168.

