

Ethics in Analytics and Social Media

Ed Lindoo, Ph.D
Associate Professor
Regis University
3333 Regis Blvd.
Denver, CO 80221-1099
303-964-6358
elindoo@regis.edu

Abstract. In Data Sciences and Computing, bias, conscious or subconscious, can be introduced in a variety of ways. The need for ethics is essential as bias can be found in the data collection phase of big data, analysis phase, and insight phase. Bias can also be found in the outcome phase, assessment phase and improvement phase., Analysts must be aware of possible bias to prevent it in their analysis. With the introduction of social computing, bias, today considered “fake news”, must also be considered. Bias can be a big issue with sites like Facebook, Twitter and even the news media. It has been found that our young, particularly college students, can easily be swayed by social media information, including information that is ultimately proven incorrect. This can be dangerous to society. We have seen countries, for example, China, Saudi Arabia among others blocking social media sites.

Keywords—Ethics; Big Data; Bias; Analytics; Fake News; Social Computing;

1 Introduction

1.1 The Introduction of Bias in Data

In Data Sciences and Computing in general, bias can be introduced in a variety of ways. Conscious or subconscious prejudice, bias can be found in the data collection phase, analysis phase, and insight phase. Bias can also be found in the outcome phase, assessment phase and improvement phase. Analysts must be aware of bias and cautious to prevent it in all phases of their analysis. Analytics and ethics are not necessarily enemies but can be if not properly treated. For example, releasing data gained through analytics, whether it is technically correct or not, could result in a company’s loss of their reputation, create competitive weakness, or possibly result in legal sanctions [1]. Big data and analytics extend well beyond organizational benefits, Gartner [1] predicts that by 2019, big data analytics will cause half of all business violation ethics.

Following is an example of a recent case involving Target Corporation. The retail giant regularly collected data on their customers, assigning a unique customer number in the process. They also collected demographic and geographic data on each of the customers. Realizing they had a data goldmine, Target hired economists and statisticians to identify patterns in data that were predictive, that would provide Target with marketing purchase opportunities. For example, if they found people buying baby

aspirin, they could theorize that these people are either in the market for baby toys and clothes or soon would be [2]. The analysts hired were placed into a new department called “Target’s Guest Marketing Analytics department” and were quickly put to work to develop new algorithms.

Develop they did, producing satisfying marketing data. One of the areas that Target wanted to delve into was the baby market, specifically those women who are pregnant that would need clothes for themselves, baby clothes, cribs, toys etc., in the future. Birth records and public records can be easily obtained but determining pregnancy for marketing is nearly impossible. Target saw this as an opportunity to maximize their profits by creating an “edge” over their competition [2]. It didn’t take long for the Target Analytic team to create a prediction algorithm that could figure out who was pregnant (based on spending habits), and then market these folks with coupons and information for baby items. The strategy backfired however when a father walked into a local Target store and asked to speak to the store manager. He had numerous coupons targeting his 15-year-old daughter whom he did not know was pregnant, in fact, no-one in the family knew!

Target’s management faced an internal and public crisis. The issue at hand is where should the line be drawn? That is, did Target have the right to collect information and use predictive analytics to target shoppers? Where do the ethical boundaries of privacy stop? Should Target have tried to ensure that they were not soliciting minors? Could they, without reasonable doubt?

2 Ethics

2.1 Data Collection Ethics

Target is not the only one involved in predictive analytics. In fact, many retailers, - investment firms, grocery chains and even the U.S. Postal service have a department that’s devoted to learning consumer habits, be they personal habits or shopping habits. The science of habit formation has become a major field of research for major medical centers, universities and corporate labs. “It’s like an arms race to hire statisticians nowadays,” said Andreas Weigend, the former chief scientist at Amazon.com [3]. With the ability to analyze data increasing exponentially, the push to understand consumer habits and decisions is at the forefront of analytic departments around the world. A study from Duke University estimated that habits shape 45 percent of the daily choices we make. Therefore, if a company can harness one’s habits, they have a good chance of predicting what one will do and/or purchase. This research is transforming our knowledge of how habits function across organizations and societies. An example of this is Tony Dungy, head coach for the Indianapolis Colts. At the time he was hired, the Colts were one of the worst teams in the NFL. Coach Dungy transformed the team to Super Bowl stardom by focusing on how his players habitually reacted to on-field cues. Likewise, Paul O’Neill, prior to becoming Treasury Secretary, overhauled a poorly performing company, Alcoa, turning it into a top performer on the Dow Jones. He did this by simply attacking the one habit that was greatly hurting the company.... lack of good worker safety. O’Neill’s pursuit of changing a bad habit turned the company around. However, companies need to be careful of the type of data collected and how it is analyzed. Asking a simple question, like “what’s your zip code” can cost the company millions of dollars. It occurred at Urban Outfitters, who settled a class action

lawsuit for \$731,180, plus attorney and legal fees, amounting to over one million dollars! [4] After making a credit card purchase at Urban Outfitters stores, consumers were asked for their zip code, assuming it was necessary for the completion of their credit transaction, however, Urban Outfitters was using this information to market the consumer. Apparently, not a legal practice in Washington, D.C., where the lawsuit was filed. It doesn't stop with Urban Outfitters. Retail giant Michaels agreed to pay \$875,000 in a Massachusetts court, and OfficeMax, \$600,000 in a California settlement, all for requesting zip code information.

Nordstrom has also found themselves in legal battles. In 2012, Nordstrom started testing a new technology that tracks customer movement by following their cell phone Wi-Fi signals. Using video surveillance combined with Wi-Fi signals and their proprietary app, Nordstrom learned a lot about their customers, including sex, amount of time certain merchandise was looked at prior to purchase, even the customer's mood. A London based company analyzes customer facial cues, "happiness level," while shopping and reaction at the register. Another company based in St. Petersburg, Russia, specializes in software for checkout devices that produce marketing messages to customers based on gender, age, mood, all measured by facial recognition and/or combined with Wi-Fi tracking. Swipe your credit card and retailers have a bank of information in addition to name, address, age, sex, cell phone info, if they tracked you via Wi-Fi, previous shopping habits (which can infer many things) and a host of other potential personal data [5].

How is this possible? Your smartphone has a unique identifier code--a 12-character string of letters and numbers called a MAC address, used for Wi-Fi and Bluetooth. It's like a vehicle identification number which is unique to each vehicle, traceable to its current registered owner. Likewise, your phone could be linked to your name and address if the cell companies provide that, supposedly they do not. Laws are currently being implemented to ensure they do not. When your phone is on, it is broadcasting the MAC address which can easily be picked up by any store's Wi-Fi and stored in their database [6]. Since you can be tracked throughout the store, and if you check out with a credit card, your information and habits can be stored and shared among stores to create a "super" database of customer information. Nordstrom started this in 2012, posting signs at their doors and advising customers that they could opt-out by turning off their cell phone Wi-Fi. Although there was backlash, today many retailers such as Family Dollar, Cabela's, Mothercare (a British company), even Nomi, of New York, use Wi-Fi in their stores to track customer's habits. While it is unknown as to which chains are using the technology, assume they are, and turn Wi-Fi off, not to share the information as even malls track consumers to interpret which stores consumers prefer to frequent.

This is where ethical analytics is essential. If you are an analyst at one of these stores, you have accrued many years' worth of Wi-Fi tracking, credit card matching, and the ability to link to other databases such as the DMV (Department of Motor Vehicles), the US Department of Justice Sex offender database at <https://www.nsopw.gov>, and the ability to scan social media sites that have the ability to acquire customers based on this information. Because the store you work for sells dietary products in one of their departments, your boss asks you to link to the DMV and use driver's license information to create a report of customers who are overweight. This information will then be used to target overweight customers through marketing. While there are no current legality issues, is it ethical? Is it ethical to track sex offenders

in stores? The data is available through the nsopw website, and if the information has been tied to their cell phone, it is known upon their entry into the store, upon. Also, is it ethical to track and analyze buying habits? Because we have, or can get the data, is it ok to access the data?? What would be the ramifications if you tell your boss, “NO!”

In the late 1990’s when newspaper companies were at the peak of their revenue, a technology vendor created a system for the classified department that at the time was quite unique. It was basically a voice mail system with several added features. For example, if you were listing a car for sale in the newspaper, for an upcharge you could put a phone number in your advertisement that would ring into the system. You would have access to the system to create a greeting, leave audio details about the car, and ask the caller to leave a message if they had questions, etc. The program also had the ability to forward the call to your home or cell phone, and the capability to set the desired time for a call back. It seemed like a harmless system, yet it collected data on everyone who used it, which today is seen as alarming. At that time newspaper companies ran classified headings such as men looking for men, women for women, swingers, etc., and anyone using this add-on phone system was unknowingly victim to their information being stored in a database.

After just a few months of gathering data, the newspaper companies had thousands of customers on file of which data could be extracted on religious affiliations, sexual preferences, car shoppers (they, too, seek apartments, houses, furniture, etc.) Truly an endless amount of customer information of possible value to a number of sources was created. So much so, that some newspaper companies started selling this information to car dealers, realtors, furniture stores, even gay and swinger magazines. This begs the question, how ethical was it for newspaper companies to sell customer information when these customers did not consent for their information to be utilized for other purposes?

2.2 HealthCare Analytics and Ethics

In the HealthCare arena, predictive analytic models designed to improve overall health outcomes can now make treatment recommendations. However, these recommendations have the potential to conflict with physicians’ ethical values when acting in the best interest of their patients [7]. For example, a predictive model may recommend withholding a potential treatment from a patient or patients with a given condition because there is a lower probability of a positive outcome compared to another patient or patients with a higher probability of success. Likewise, the use of predictive analytics could raise concerns that those with illness, lack of access to health care, or poverty, will be left out. While predictive analytics promise to make dramatic changes in the way HealthCare is practiced, it will likely run into legal, policy, and ethical challenges. Developers of predictive analytic models will need access to vast amounts of big data HealthCare as acquiring that data will certainly raise questions about consent, privacy, and fairness. Predictive analytics will certainly help build a stronger and more dynamic HealthCare system in the next few years and for decades to come. This will present an enormous impact on the roles that technology will play in reforming the HealthCare industry.

2.3 The UPB Phenomena

At the beginning of the 21st century, accounting scandals among other events, led to the collapse of billion-dollar companies. These scandals only confirmed that unethical acts were being conducted, often flourishing within organizations. Research into unethical behavior within organizations has found assorted reasons as to why employees might engage in unethical acts, for example, personal gain, the gain of the organization, or even to cause harm against the organization [8]. These acts fall under the title of unethical pro-organizational behavior (UPB) [8]. UPB includes acts of commission (e.g., “cooking” numbers to boost analyst projections and stock values) and omission (e.g., withholding information about the hazards of a pharmaceutical product) that are typically considered unethical.

UPB is predominantly considered pro-organizational behavior which is neither specified in formal job descriptions nor ordered by superiors, yet, is carried out to benefit or help the organization. Many theorists believe that the more an individual identifies with their organization, the more likely they are to disregard their personal moral standards and engage in acts that favor the organization. Further research has found that the stronger the employees’ organizational identification, the more influence it will have on productive work behavior, such as increased extra-role behaviors, and job performance. Umphress, Bingham and Mitchell [8] believe that individuals who identify strongly with their organization may be more likely to engage in UPB.

As a public example, we have recently noted journalists who have been found to be lying or “bending” the truth. In 2015, NBC Nightly News anchor, Brian Williams, was caught lying about riding in a helicopter that was shot down in Iraq in 2003. NBC suspended him for six months without pay and the Peabody Award-winning newsman, became the top story himself. In 2015, Fox News Channel star, Bill O’Reilly, was caught lying about experiencing war involving Argentina and England in 1982. In 2011, MSNBC’s Rachel Maddow, was caught lying when she accused Rush Limbaugh of racism. She was later forced to apologize to Limbaugh on the air. Be it public figures, or an analyst within an organization, people tend to get caught up in the moment, the organization, personal gain, or a number of reasons that might fall under UPB. Imagine then, how an analyst could add an extra day of revenue to the monthly or quarterly reports for a company. Depending on their position within the company, unethical behavior may be committed.

2.4 Presenting Facts in a not so Factual way

Although manipulating numbers in analytics can easily be done, presenting true numbers can also be misleading. For example, in his April 14, 2017, article on the lack of African Americans in Major League Baseball (MLB), Ken Rosenthal came across as though there is a huge problem in MLB because African Americans represent only 7.73% of the players, but according to the U.S. Census Bureau, African Americans represent 13.3% of the U.S. population. On the surface, this could also appear to be a problem, 13.3% versus 7.73%. Rosenthal interviewed MLB commissioner Rob Manfred, who said, “The numbers are low, but it is not a phenomenon that is shocking to me.” Rosenthal went on to point out all the programs in place to attract more African American players. He found that in 2016, baseball spent approximately \$30 million on youth development for underserved communities, but still, the numbers show that

baseball's initiatives are years away from making a significant impact at the major-league level.

If MLB numbers are compared to the population, i.e. the race population in the United States, it is necessary to establish what exactly is being presented. For example, Rosenthal stated that the MLB is made up of only 7.73% African American players, however, the US population of African Americans is 13.3%. We need to ask the question, "who plays in MLB baseball?" Men. With that, we have just cut the population in half, specifically 48% black males versus 52% women. In other words, we are now down to 6.83% of the population which is a reasonable amount to see in the MLB. However, as analysts, we need to dig deeper. Of the 6.83% African American male population in the U.S., what population might play major league baseball? The answer is typically 18-35-year old's, and this group makes up 48% of the 6.83% U.S. population of African American males [11]. 48% of 6.83% is 3.28%, thus we can ascertain that 3.28% is an acceptable number for African American players in the MLB.

Figure 1 is an excellent representation of what has happened ethnically in major league baseball since 1946. We can see that in 1946 MLB was made up of 100% white. In 1947 Jackie Robinson came on board, and other African Americans slowly followed, reaching a peak from 1976-1986. We can also see that at the same time African Americans started moving into baseball, Latino's did as well. However, the Latino expansion continued to grow to about 27% by 2016.

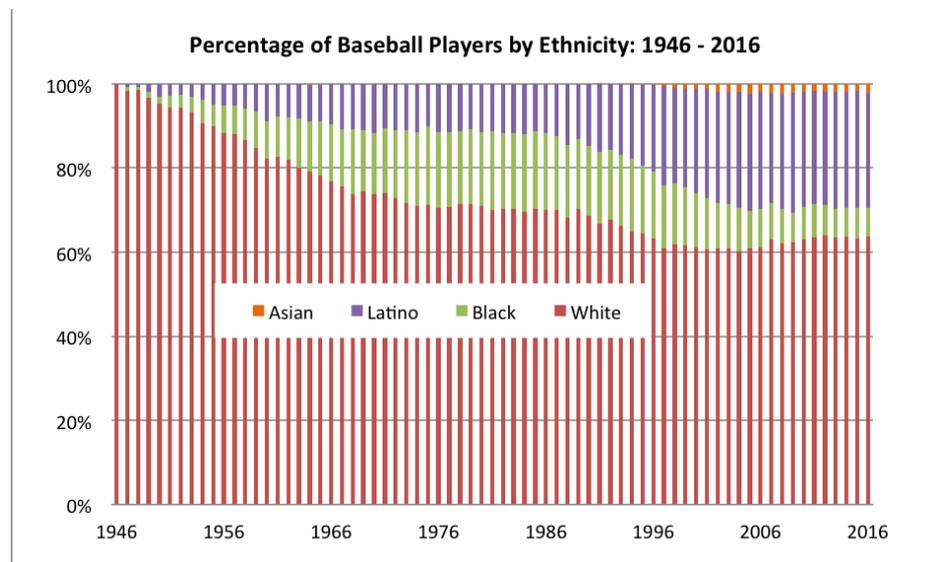


Figure 1 Baseball Demographics, 1947-2016. Society for American Baseball Research (SABR)

Is it ethical to report on numbers like this? Yes, because these are simply stats and do not single out any single person or traceable group. However, as an analyst it is important to know that the stats may include sensitive data. For example, if you have been given a database of female statistics by state and are then asked to report on the number of pregnancies by state in a given database.

State	
ARKANSAS	97
IOWA	97
MINNESOTA	97
MONTANA	97
COLORADO	96
MARYLAND	96
DELAWARE	95
KANSAS	95
TEXAS	95
SOUTH CAROLINA	94
VERMONT	93
MAINE	92
GEORGIA	91
NEW HAMPSHIRE	91

State	Disease		N
	AIDS	canc..	
ARKANSAS	1	6	90
IOWA	3	11	83
MINNESOTA	4	3	90
MONTANA	1	8	88
COLORADO	1	3	92
MARYLAND	2	4	90
DELAWARE	4	3	88
KANSAS	3		92
TEXAS	3	1	91
SOUTH CAROLINA	5	4	85
VERMONT	4	4	85
MAINE	2	5	85
GEORGIA	3		88
NEW HAMPSHIRE	3	3	85

Figure 2. Left data shows pregnancies by state, right figure shows pregnancies with AIDS and/or Cancer

This is a straight forward task rendering the top four states (Figure 2 left table) with 97 pregnancies. The data also contains information on diseases such as AIDS and/or cancer, thus we can add in a filter for disease and see which state has the highest rate (Figure 2 right table). In this analysis, Iowa shows 11 pregnant women with cancer, and five pregnant women in South Carolina with AIDS. The table shown in figure 3 represents further data acquired of various ethnic groups that have AIDS and/or cancer.

State	AIDS					Disease / Race cancer							
	Ame..	Ame..	Asian	Black	Hisp..	White	Ame..	Ame..	Ame..	Asian	Black	Hisp..	White
ARKANSAS				1					1		1		4
IOWA			1		1	1					3	2	6
MINNESOTA				1		3							3
MONTANA						1						2	6
COLORADO		1									2		1
MARYLAND				1		1					1		3
DELAWARE				2		2							3
KANSAS				1		2							
TEXAS					2	1					1		
SOUTH CAR..			2			1	2		1		2		1
VERMONT				1	1	2					1	1	2
MAINE				1	1						2	2	1
GEORGIA				1		2							
NEW HAMP..				1		2					2		1

Figure 3 showing states of state pregnancies with various diseases based on race

So far, basic data has been discussed, nothing that singles anyone out or, is too sensitive to publish. After providing a benign analysis to an employer, you are asked to elicit the names of women in the states of Iowa and Arkansas that are African American, pregnant and have cancer. Not considering the impact of this request, queries are run

State	Last	First	Disea.. canc.. Black
IOWA	CUEVAS	PILAR	1
	SCHULTE	DANNETTE	1
	WILLIAM	MINERVA	1
ARKANSAS	MADDOX	DONA	1

Figure 4, Data has been narrowed down, singling out people by name

and the requested information is produced. As observed in Figure 4, there are three women in Iowa and one in Arkansas. This data creates an issue, however, as specific individuals have been singled out. If this information is not handled sensitively, the individual running the report, the person requesting report and the employer, can be subject to a potential lawsuit (Figure 4).

3 Four Aspects of Big Data Ethics

3.1 Big Data Ethics

Today companies need to be aware of the data they possess and the ramifications they may endure if they use this data unethically. Davis [12] proposed four aspects of big data (Figure 5) in ethics in which businesses should balance: identity, privacy, ownership and reputation.

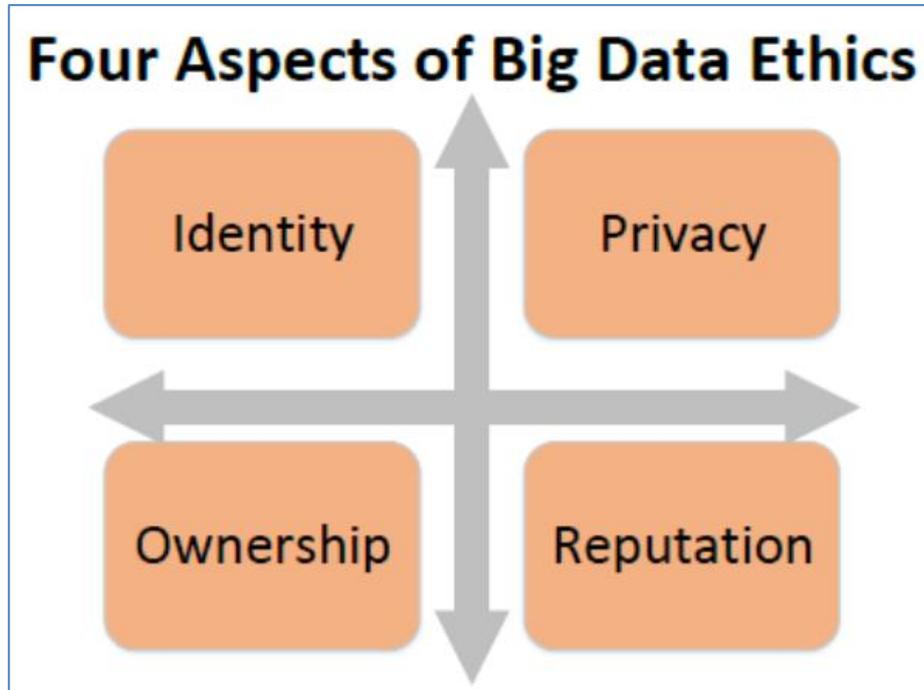


Figure 4, *Balancing Ethics as presented by Davis [12]*

One thing to understand about these four aspects is that they are all equally important as they all contain important ethical characteristics. For example, the identity questions focus on the question, “is our offline existence identical to our online existence?” Davis [12] pointed out that some people say the obvious answer is that both should be the same, while others disagree. Looking at the aspect of privacy, the question is “who should control the access of data on an individual?” It’s not just control of access, but the fiduciary responsibilities that companies have in protecting data. For example, the Equifax breach of 2017, exposed the data of over 143 million people! It’s hard to imagine that all this data was stored in plain text files instead of being encrypted. While the company will undoubtedly be sued for years to come, one would think that there would, or should be, consequences for the top officers of the company. However, the CEO and CIO retired just days after the breach announcement and the CEO, Richard Smith, walked away with over 100 million dollars in severance and retirement compensation.

Regarding the ownership aspect, in a non-digital world it probably doesn’t matter much to anyone that someone is 5’11, a white male, over 60 and overweight. These are facts that are seldom looked at in a non-digital world. However, in a digital environment these facts can be stored and possibly used in ways that might not be appreciated by the person. For example, if a marketing company has this information, they might market the person for weight loss, or, send senior citizens coupons. Looking at the reputation aspect, we again have the non-digital vs digital world. In the non-digital world, people close to a person, know the person, e.g., their religious beliefs, political preferences, what they believe in, stand for, etc. However, in the digital world,

someone across the country can find online information regarding the individual and form an opinion without ever meeting them. A good example of this is a prospective employer looking at Facebook or other social media accounts.

4 Social Computing

4.1 Previously Ethical News

Prior to the 1900's, local, national and world news came to us primarily by way of newspapers. In the early 1900's, radio introduced a new medium by which people obtain news. These mediums had editorial boards that scrutinized articles written by trained professionals. These professionals often went through years of college preparatory courses and internships prior to attaining a job in the world of news. Television appeared in the 1950's, and the footsteps of radio were followed when it came to news integrity. When the Internet arrived in the late 90's, so did Blogs, Facebook, Twitter, YouTube, and a host of other social media sites. With social media, came reporter "wanna-bees", with little or no training in the world of news reporting, and frequently, with little, if any, reporting ethics, presenting information that is often hard to substantiate.

4.2 Fake News

It's not known which came first, poor reporting on the Internet, or, poor reporting on the major cable networks such as CNN, MSNBC, FOX, BBC. As witnessed prior to the 2016 presidential election, everyone appears to be scrambling to be the first to report news, be it accurate or not. As a result, we regularly see, companies such as The New York Times and CNN retracting stories upon realizing the stories are incorrect. How much damage is done when a wrong (Fake News) story goes out? Researchers at MIT researched this question, and recently concluded a five-year study on "Fake News on Twitter" [13].

The study found that a false story reaches 1,500 people six times quicker, on average than a true story does. They also found that false stories outperform the truth on every subject including business, terrorism and war, science and technology, and entertainment, determining that fake news on politics routinely outperform the other categories.

Ultimately the researchers found that between 2006 and 2016, about 126,000 tweets had been retweeted more than 4.5 million times. Some of these were linked to "fake" stories hosted on other websites while others started rumors themselves, either in the text of a tweet or in an attached image. For example, a rumor was tweeted in February of 2016, reportedly by a recently deceased elder cousin of Donald Trump, and opposing the presidential bid, stated in his obituary, "As a proud bearer of the Trump name, I implore you all, please don't let that walking mucus bag become president". However, Snopes could not find evidence of the story and rejected it as false. Nonetheless, roughly 38,000 Twitter users re-tweeted the story. It's not known how many might have shared this in the form of an email or other medium. [13]

In a 2017 study of Higher Education Students, Chandra, Surjandy and Ernawaty [14] found that higher education students can become emotional and be misguided by fake news. Alarmingly they can disseminate this fake news faster because of social media and instant messaging applications. Because social media and instant messaging are used by almost everyone who has a smartphone, to interact and communicate with just about anyone, they can quickly disseminate information, motivations, and promotions, good or bad. Through their research, it was determined that a large proportion of their sample population of Higher Education students actually modify the news and made it look legitimate by including a reliable source! [14]

4.3 Misinformation and Disinformation

False information comes in two forms: misinformation and disinformation. Disinformation is false information that is purposely spread to deceive. Misinformation is simply incorrect information, for example, “I was misinformed about when to meet you for lunch, however I know it was not deliberate.” One study, as reported by Kshetri [15] found that roughly 62% of US adults get their news from social media sites, and of those 40%, from Facebook. Kshetri [15] also determined that in the final three months of the 2016 US presidential campaign, the top performing fake election news stories on Facebook attracted more views than the top stories from major news outlets such as the New York Times, Washington Post, Huffington Post, or NBC News. During that time 8.7 million shares, reactions, and comments were generated on Facebook as compared to just over 7.3 million from 19 major news websites. As we all know now, Russia was involved.

Russia, however, was not the only country spreading fake news. We have recently learned that other fake news creators were known to be operating from countries such as the Republic of Georgia and Macedonia. For example, during the one-year period before the 2016 US presidential election, residents of the Macedonian town of Veles (population 45,000) launched more than 140 US political websites. Most of the domain names looked American, such as WorldPoliticus.com, TrumpVision365.com, USConservativeToday.com, DonaldTrumpNews.co, and USA DailyPolitics.com. [15]

4.4 The China Syndrome

Realizing that social media can sway stock markets and has the potential to overthrow governments, the Chinese central government has placed bans on many social media URLs. It is common knowledge that Facebook, Google, YouTube and Twitter are banned, in fact, over 8,000 sites are banned in China [16]. Studies [17] suggest that all media content harmful to the legitimate ruling of the communist state is censored. To protect its rule, the Chinese Communist Party (CCP) has identified a few priorities including the sustaining of economic growth, nationalism, social stability, and rational legal authority and electoral legitimacy. This is also what sustains the current political regime. Therefore, the news most likely to be censored is news that the propaganda authorities believe will have a negative effect on the legitimate rule of the Communist Party state. Further, Chinese law mandates that all Chinese news media must contribute to the enhancement of party ruling [17].

China, however, is not the only country to ban social media sites. Figure 6 is a good representation of the major sites banned by numerous countries around the world [18]. North Korea is not shown as everything is blocked. Until recently, Cuba was very much like North Korea, although presently more and more sites are being allowed by the Cuban government. However, all the Internet access is controlled by the Cuban government and opposition sites are blocked [19].

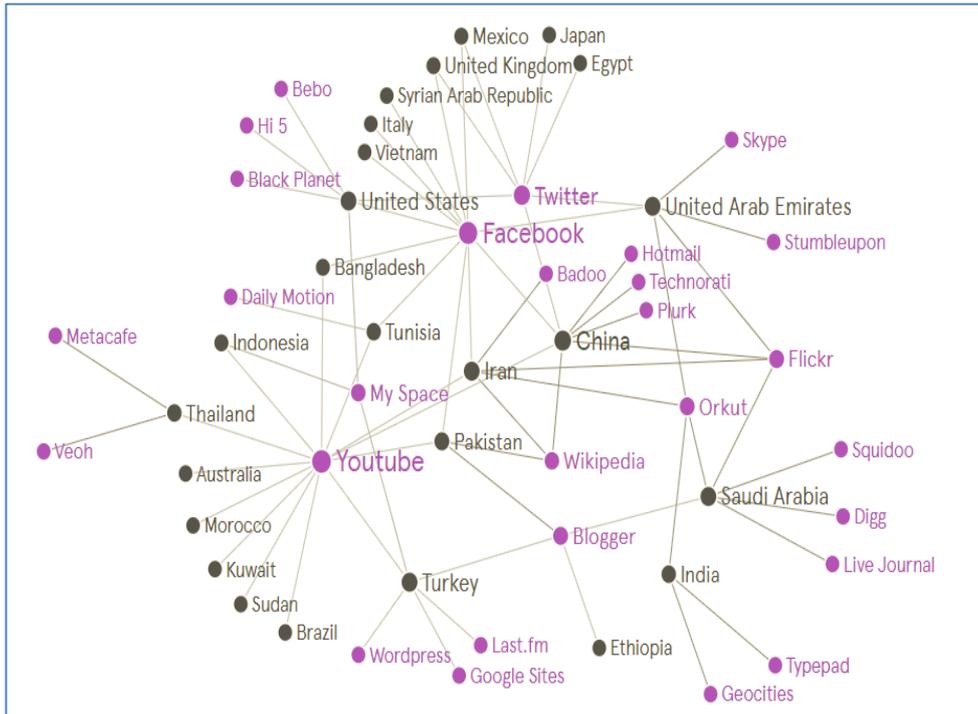


Figure 5 [18], a high-level view of countries that block various websites

5 Conclusion

5.1 The Future of Information and Communication

As demonstrated in this paper, data analytics can and is often swayed by bias. Managers of data and analytics need to be aware of this and ensure that the findings they present are accurate and without bias. It remains to be seen if the UPB phenomenon will continue to grow, or if those in charge of data will come to realize the ethical and moral values of accurate reporting. It also remains to be seen if college students, news organizations and others will continue to spin “fake news” across social media sites, or if they too will come to realize the ultimate ramifications of putting out un-validated stories.

This paper presented companies such as Target, Michaels, Urban Outfitters and Nordstrom and their use and misuse of customer data. It is apparent that companies

need to be very aware of negative possibilities and be protective of customer data. Also, introduced was the UPB phenomenon, pointing out that it is quite easy to fabricate data, be it for personal reasons, or to benefit a company. Data can be analyzed and presented in a way that may be incorrect and more importantly, analysts to be aware of sensitive data, such as the data looked at regarding the names of pregnant women with diseases. Sensitive data like this must be well secured, well managed, and not abused. Finally, it was determined that social media and “Fake News” are presenting and will continue to present dangers to our society, especially through our young, college-level students.

References

- [1] Noyes, K. Mind your analytics, or you could soon be in ethical hot water. IDG News Service, October 15, 2015.
- [2] Cox, S. Target: The challenge of data mining. *Journal of Critical Incidents*, volume 6, 2013.
- [3] Duhigg, C. How companies learn your secrets. *The New York Times Magazine*, February 14, 2012.
- [4] Groden, C. Urban Outfitters, Anthropologie face D.C. class action suit. *Time*, August 9, 2013.
- [5] Clifford, S. and Hardy, Q. Attention, shoppers: Store is tracking your cell. *The New York Times*, July 14, 2013.
- [6] Kerr, J. Stores can see where you go by tracking phones, *USA Today*, February 22, 2104.
- [7] Cohen, G., Amarasingham, R., Shah, A., Xie, G., and Lo, B. (2014). The legal and ethical concerns that arise from using complex predictive analytics in health care. *Health Affairs*, <http://content.healthaffairs.org/content/33/7/1139.full>, 2014.
- [8] Umphress, E., Bingham, J., and Mitchell, M. Unethical behavior in the name of the company: The moderating effect of organizational identification and positive reciprocity beliefs on unethical pro-organizational behavior. *Journal of Applied Psychology*. Vol. 95, No. 4, 769 –780, 2010.
- [9] Rosenthal, K. The African-American talent pool remains a troubling problem for baseball. *Fox Sports*, April 14, 2017
- [10] Armour, M. and Levitt, D. *Baseball Demographics, 1947-2016*. Society for American Baseball Research (SABR), <https://sabr.org/bioproj/topic/baseball-demographics-1947-2012>. 2017.
- [11] *Black Demographics* (2017), <http://blackdemographics.com/population/black-male-statistics/>
- [12] Davis, K. *Ethics of big data: Balancing risk and innovation*. O’reilly, Boston, MA., 2012.
- [13] Meyer, R. The Grim Conclusions of the Largest-Ever Study of Fake News. *The Atlantic*, <https://www.theatlantic.com/technology/archive/2018/03/largest-study-ever-fake-news-mit-twitter/555104/>, March 8, 2018.

- [14] Yakob, C., Surjandy, Ernawaty. Higher Education Student Behaviors in Spreading Fake News on Social Media: A Case of LINE Group. International Conference on Information Management and Technology (ICIMTech), 2017.
- [15] Kshetri, N. and Voas, J. The Economics of “Fake News” IT Pro-IEEE. <https://ieeexplore-ieee-org.ezproxylocal.library.nova.edu/stamp/stamp.jsp?tp=&arnumber=8123490>, 2017.
- [16] GreatFire.org. Bringing Transparency to the Great Firewall of China. <https://en.greatfire.org>, 2018.
- [17] Kuang, X. Central State vs. Local Levels of Government: Understanding News Media Censorship in China. Fudan University and Springer Nature Singapore Pte Ltd., 2018.
- [18] Yasaklanan, D. and Ağlar, S. Mapping the banned social media by country. <https://graphcommons.com/graphs/52eadc97-6bc3-48a2-8526-741f358c9146>, 2018.
- [19] San Pedro, E. Cuba internet access still severely restricted. BBC <https://www.bbc.com/news/world-latin-america-35865283>, 2016.