Data-Mining and Analytics: Rising Concerns over Privacy and People's Security

E.K. Jaisal *
Department of Political Science, University of Hyderabad

Preprint manuscript of the paper presented and submitted to the ***Indian Social Science Academy (ISSA)*** *in the Political Science Research Committee of the* ***XLII Indian Social Science Congress*** *with the focal theme* ***'Human Future in Digital Era'*** *organised by the ISSA under the auspices of KIIT Deemed to be University, Bhubaneswar, Odisha from* ***27th to 31st December 2018****.*

Abstract

An individual's privacy being sold to the corporates for customized advertisements and commercial micro-targeting is essentially a disturbing idea as one's personal information is being commodified and marketed. It becomes dangerous and poses a threat when analytics pushes the limit further and engages in predicting the real-life activities of the people such as electoral behaviour and employs tools of political micro-targeting to manipulate the electorate. Hence, the recent developments in data science such as data-mining and analytics have become the catchwords for those concerned with the question of privacy infringements and state surveillance. The paper attempts to evaluate the potential threats to individual privacy and security when data-mining and analytics are clubbed with artificial-intelligence-powered election campaigns, surveillance mechanisms, warfare strategies *etc.*

Keywords

Political Micro-Targeting, Big-Data, Data-Mining, Analytics, Aadhaar, Privacy, Security

*E-mail: ekjaisal@outlook.com — ORCID: https://orcid.org/0000-0003-3535-0273

Introduction

The discovery of fire and invention of the wheel and steam engine were watershed developments in the course of the human race's progress to become what it is today. The fire changed radically the form in which we produce, process and consume food, and the wheel redefined our modes of transport and nature of work. The steam engines in the near past revolutionized the methods in which we produce and distribute goods for consumption, by triggering large-scale industrialization. All of these are innovations that have tagged and marked eras to their names and similar is the case with the computer. This concept, which struck the mind of Charles Babbage in the early 19^{th} century is now revolutionizing the way people live, think, learn, bank, travel *etc.* The influence of computer and technology of the like has crept in to every nook and corner of our day to day life, to such an extent that our era is now being celebrated as the *digital era.* As the steam engines in the 18^{th} century had brought along with it the burden of industrial pollution, Information Technology (IT) too brings with it huge pit-holes and disasters, perhaps of even greater magnitude or gravity, out of which includes the breach of individual privacy and the manipulation of people's behaviour in the social and the political realm which will be discussed below in detail within the scope of the paper.

The famous statement *"Information is the oil of the 21^{st} century and analytics is the combustion engine"* (Sondergaard, 2011) had clearly hinted on where the world was heading to, in the years to come. Today, even in the countries belonging to the global south, data occupies the centre stage of everything ranging from politics to trade. The use or misuse of data is becoming a key determining factor in the day to day affairs of the world, where political parties employ it to consolidate power, huge multinational corporations use it for personalizing advertisements to sway the customers, and the elite-power-wielding classes, to manipulate and form the common-sense of the subalterns, in a Gramscian understanding, to reinforce their hegemony.

The term *data* refers to *"the collection of items of information such as the representations of facts, concepts, or instructions in a manner suitable for communication, interpretation, or processing by employees or by automatic means. Data can be in the form of files in a computer's volatile or non-volatile memory, in a data storage-device, or in the form of data in transit across a transmission medium".* (Gattiker, 2004) In the scope of this paper, it simply means, information that is created, stored, transmitted, disseminated and consumed in digital form *i.e.* at the most rudimentary level in binaries as *zeros* and *ones.*

Data-Mining and Analytics

Data mining is a process by which mathematical and computational algorithms are employed to structure and recognize patterns in the raw data that is available. It generates information and reveals trends that the data displays. Data mining, by nature, is either descriptive - explaining the trends that the data

exhibits or predictive *i.e.* to predict the trend which the data would exhibit in future. It provides a broad overview, chalking out the general nature of the *big data* aggregated at the data warehouse and does not answer to specific questions. It employs sophisticated and advanced tools of artificial intelligence and statistical analysis.

Analytics, whereas, is the art of exploring the facts and figures from the data with specific answers to specific questions. It draws conclusions from the information that is available so as to enable the employer of analytics to arrive at the right decisions. It analyses, for instance, the response of the customers towards a product or business model and may predict, on the basis of the available information, the success or amount of acceptance a similar case may receive in future or in other markets; thus enabling its employer to make wiser decisions. Such exercises could be conducted in the realm of domestic politics and international relations as well. Intensive human resource and computing processes are involved in the employment of analytics.

Big-Data

According to Gartner Inc. *"Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation"*. Dealing with big-data requires the processing of high volumes of low-density, unstructured data which can be of unknown value, such as *Facebook* feeds, search or view history on *YouTube* or *Google,* or the signals received from a device equipped with sensors. It might range from a few terabytes to several petabytes. Such data flows into the domain of analysis in high velocity and extreme variety such as in the form of audio, video, text, images *etc.* The data is considered to be of intrinsic worth but remains to be of no use unless the value of it is discovered by processing it. The big-data of which the value has been discovered and the veracity has been authenticated assumes the form of a capital which can yield profits upon its proper utilization by data processing firms. Thus, the keywords characterizing big data are Volume, Velocity, Variety, Value, and Veracity.

The Extraction and Production of Big-Data

The big-data is procured, through several methods such as customer feedbacks, responses, reviews *etc.* But the larger chunk of the data is produced and extracted by popular social networking services like *Facebook, Twitter, Instagram, LinkedIn* and other tech giants such as *Google* and *Amazon.* A majority of the social networking sites and search engines offer their services free *of cost* and do not demand a single penny from their end-users; instead, they extract profits by monetizing the identity of the customers. They log everything from status updates, list of *Facebook* pages liked, posts shared *etc.* and monitor the

entire online behaviour of their customers. Their business model is based on extracting data which is further processed to generate profiles of the customers by employing analytics and then connecting the prospective seller or advertiser to reach out to the potential customer.

Data of the customers is processed to such an extent that every end-user is provided with personalized advertisements. Personalized advertisements are those to which the person engaging with can easily relate to, and the possibilities of the person availing the service provided by the product advertised, are high. Hence, the advertisements are tailor-made in such a fashion that no amount of advertising space is spent on irrelevant content and thus maximizing the profit of the advertiser and the product seller.

To achieve accuracy in the data processed for personalization, the service providers scan through almost each and every bit of data transferred through their servers. This might account for text messages or chats, images, audio clips, and videos. Location tracking services are also employed to improve accuracy. Search engines which record the search history of the user, process it to identify their interests. Thus, the end-user is totally ripped-off of privacy by availing these services as his/her personal details are *commodified* and marketed in a form for profits.

Hence, the user pays for availing these services by monetizing his/her own interests, habits, preferences, relationships, emotional orientation, and vulnerabilities. These platforms partner with other agencies to determine who is more susceptible to advertising campaigns, by tracking their engagements with the advertisements and then flush the customer with more and more so as to efficiently influence the decision making of the customers towards purchasing the advertised products. The end-user is thus decimated to the level of a mere product which is *sold* over and again to marketing agencies. This is the most invasive and successful model of advertising ever created. The efficacy of this model has revolutionized the advertising industry and has been a breakthrough with conventional marketing strategies to *smart-marketing.* Albeit, it poses an ethical question of the commodification of individual interests, this advertising strategy apparently does not produce any physical harm of serious consequence on the end-user, other than its influence on the market itself; this model but assumes a highly influential role when it is taken further to the realm of politics. The application of big-data in politics will be discussed in greater detail later.

Data extracting exercises, as mentioned above, are conducted by almost all of the social-networking enterprises, search-engines and video streaming websites which render services free of cost. A quick look at two of the most widely used platforms *viz. Google* and *Facebook* would give further insights into the gravity and extent to which the user data is stored and analyzed.

By turning on location services on devices, *Google* stores the location of the user and can track where a user has been at any point in time. This can be restricted by turning the location services off but several mobile applications (especially those running on *Android* and *IOS*) require location services turned on to be able to function. Hence, the end-user is required to turn the services on at least once or several times in a day/week (depending on usage), thus enabling

Google to keep track of the location of a person over a long period of time.

The search history of a user is stored and maintained by *Google across* devices. This implies that the user cannot delete his/her search details from Google's servers by simply clearing the history on a device or a web browser. It involves a bit more complicated process to delete and clear the search history from what Google calls its *dashboard,* though these are being simplified recently upon the rising privacy concerns from the customers. By availing some of the several services that *Google* offers, such as *Gmail, Google Photos, Google Drive, Google Play Music, Google Play Store, YouTube etc.*, it collects information such as details of e-mails sent/received (including the content within), contacts, photos, files, calendar, music, books, *YouTube* watch histories and also passwords. It can identify the devices used by a customer and also record details about the applications installed on devices. Details such as phone numbers and the International Mobile Equipment Identity (IMEI) numbers and the like are also collected. *Google* allows the customer to download the entire information that it has about him/her and an attempt to do so would reveal how big a pile of information is being stored by *Google* about a customer.

Collecting information from customers to provide services is by every means fair, as no effective output can be generated without the required input. But the keyword here is *storage* and not *collection.* The data is stored even after the customers' intended purpose of furnishing the inputs is served. *Google* accumulates all the relevant data, process it and create an advertisement profile based on the information, which can include demographic details, physical and emotional vulnerabilities, political orientation, relationship status *etc.* It is based on these advertisement profiles that *Google* generates profit by allowing advertising agencies to feed their advertisements to the target audience.

Upon being questioned by the Judiciary Committee of the House of Representatives of the US Congress on the 11^{th} of December 2018, the Chief Executive Officer of *Google Inc.,* Sundar Pichai testified that the employees could look into the contents of individual users including those from *Google Docs,* but then such activities are governed by *comprehensive policies* to prevent misuse or leakage of information.

Facebook too, like *Google* stores tonnes of information about an individual customer. *Facebook* extracts content from each and every message sent through its servers, including images, audio files, and videos. *Facebook* being a social media website, has, by default, access and permission to store photos and videos and personal details such as date of birth, gender, interests, political opinions, movies, music, books of interest *etc.* It logs the details of every page or post liked and also analyses the status updates for generating a profile. *Facebook* has provisions to find out and store details about events attended by individual customers and the locations checked in by them. *Facebook,* like *Google,* has access to microphones and webcams as well. By virtue of the nature of its service, *Facebook,* as on date is the largest repository of personal data.

Upon being questioned by the Senate Commerce and Judiciary Committees on privacy, data mining, regulations and *Cambridge Analytica,* of the US Congress on 10^{th} of April 2018, Mark Zuckerberg, the Chief Executive Officer

of *Facebook Inc.* testified that they store user details and data, some of which with the permission of the user. He also stated that there are widespread misconceptions that *Facebook* sells data to advertisers. He reaffirmed that *Facebook* does not sell user data, but instead it allows advertisers to choose whom they want to reach out and then *Facebook* on its own does the placement.

The business models of *Google Inc.* and *Facebook Inc.* in relation to advertising and data collection provides insights into the much-a-lot similar models adopted by most of the other search engines, social-media websites and networks rendering services free of cost. Tracking of information is not restricted to these free services; trackers of such massive data extractors are present on several other personal and private websites as well.

Backdoors to Data Extraction

All of the methods to produce big-data that was discussed hitherto involved the use of free services by the end-user. But there are instances of data-extraction through the backdoors, with the user not even hinted on the fact that their information is being sent to external servers which may be located even on foreign territories. Such a method of data extraction is performed by applying *backdoors* on the firmware (software) of mobile devices, especially smartphones, which communicate with the servers concerned, at regular intervals.

Such an instance was reported by *Kryptowire,* a security firm, in 2016, where it found software which came pre-installed with the phone sending full contents of text messages, contact list, call logs, location information, and some other data to Chinese servers every 72 hours. The code was written by *Adups,* a technology firm based in Shanghai, China and the details about the surveillance activity were not disclosed to the users in the terms and conditions of service. Tom Karygiannis, vice-president of *Kryptowire* said to *The New York Times* that *"Even if you wanted to, you wouldn't have known about it".*

Such methods of data-extraction through the backdoor are not always done merely for advertising purposes. It could be a means of mass surveillance and espionage and is a matter of a serious threat to both personal and national security in an age when artificial-intelligence-based warfare is developing at a quick pace.

Big-Data Redefining Politics, Surveillance and the Future of Warfare

Access to big-data has now revolutionized the way election and mobilization campaigns are conducted. Earlier, politicians and parties knew very little about the electorate or the people whom they are dealing with. Campaigns used to be based on general trends and demands of the electorate as individual preferences and choices were difficult to isolate and identify.

With the advent of big-data, campaigners are much more familiar with the mood of the electorate and know the preferences and choices of each individual, gathered and profiled through data-mining and analytics. Hence, the amount and nature of campaigning that could work upon people is easily identified and campaigns focus on mobilizing those who tend to support them just by throwing up a few advertisements and posters and ignores that part of the electorate which does not tend to lend support to them. This has resulted in a much more efficient campaign which could yield better results. But such methods of smart-politics leads to an increase in the fissures existing in societies as the political leaders tend to totally ignore those who stand in their opposition, who are now much more easily identified.

According to Dr Eitan D. Hersh, Associate Professor of Political Science at Tufts University and the author of the book *Hacking the Electorate: How Campaigns Perceive Voters,* possession of properly profiled data can make the task of mobilizing and manipulating the voters easier for the campaigns. The electoral behaviour of the voters can be predicted beforehand and suitable measures can be employed to orient the campaigning to manipulate the electorate to generate favourable results. The conventional methods of relying on sample or survey data are replaced with population data which accounts for the information about the entire electorate. This enables the campaigns to engage in more individual level campaigning than the general speeches and promises. It enables them to identify the non-supporters or opponents with much more precision thereby making it easier to dismiss them off and focus on appeasing their supporters *i.e.* micro-targeting.

The alleged role of *Cambridge Analytica,* a British political consultancy, in the 2016 US Presidential Elections and the *Brexit,* brings to the limelight the immense potential of big-data to influence politics. The firm had procured data from *Facebook through* a workaround method, as *Facebook's* privacy policy does not apparently allow third-party platforms to access their data. Aleksandr Kogan, a Russian-American academic, built an app that encouraged *Facebook* users to take a quiz and Kogan had gained permissions from *Facebook* to collect the data of the users and their extended friend's circle by means of the app for *academic purposes. Cambridge Analytica* harvested this data and employed their algorithms to build a character profile of around 87 million *Facebook* users. This helped to predict voter behaviour, which was then allegedly manipulated to influence the 2016 US Presidential elections by Russian agencies which ran huge advertisement campaigns through stealth on *Facebook.*

The modus operandi of the firm that harvested the data of *Facebook* users without consent was revealed to *The Observer,* a subsidiary of *The Guardian* newspaper by an ex-employee turned whistleblower of the firm Christopher Wylie. *Facebook,* as they claim, was unaware of their data being harvested for political purposes and had to face staunch criticism from across the world for their alleged involvement in the *Cambridge Analytica Scandal.*

Upon being questioned by the US Senator Tammy Baldwin on whether Aleksandr Kogan had sold the data to anyone besides the *Cambridge Analytica,* Zuckerberg replied: *"Yes, he did".* Thus, the *Cambridge Analytica Scandal*

could be just the tip of an iceberg, with the entire gravity of data-harvesting operations still to be uncovered, with possible implications in similar or smaller scale across the globe.

Like marketing and politics, big-data is also creating a breakthrough in future techniques of warfare and State surveillance as well. *The British Broadcasting Corporation* (BBC) had reported in December 2017 about the world's biggest camera surveillance system that China has been building across its territory. John Sudworth a BBC correspondent was given special access for a demonstration to one of the hi-tech police control rooms in the city of Guiyang. Most of the cameras are equipped with face-recognition features and artificial intelligence, which could in real-time match the details captured by the cameras with the data of the population available in their digital catalogues and thus in a matter of minutes could identify and locate any person and track their movements and interactions with a high level of accuracy.

The US military's employment of drones to target militants in Afghanistan, Pakistan and Middle-Eastern Africa had garnered huge attention while stirring up several controversies as well. Drones make it easier to engage in war as it minimizes the need to deploy troops on the ground. Drone strikes are much more effective as it reduces the causalities on the side of its employer and also is relatively much cheaper to produce. With the future developments in technology, drones which are now being controlled by drone pilots remotely can acquire autonomy in locating and hitting the targets based on artificial intelligence and face-recognition. A drone could range from the size of a hummingbird to that of a fighter aircraft thus making them highly susceptible to be misused for espionage and targeted attacks on foreign lands, violating their sovereignty and tampering with their mechanisms of national security. Also, such capabilities falling in the hands of terrorist organizations could result in its indiscriminate employment in terrorism, making it even more difficult to contain and prevent.

Slaughterbots, an arms-control advocacy video directed by Stewart Sugg in 2017 provides insights to the future where the potential of Artificial Intelligence (AI), when employed in warfare, facilitates to identify individual targets with the help of collected data that are programmed into the drones, which then autonomously identify and attack the targets. Development of technology of the like in warfare could mean that the peace and order achieved in international relations through nuclear deterrence and the threat of Mutually Assured Destruction (MAD) could be superseded, as nuclear arsenals become redundant to employ, thereby opening new dimensions of war to engage in.

The case of China's surveillance camera network and the development in the US military's research towards enabling drones with more autonomy clearly indicates how significant and sensitive the data relating to biometrics such as face-recognition is. In an age when *Google* and *Facebook* use face-recognition techniques to auto-tag individuals in pictures, it becomes a matter of introspection on how much personal data is being opened up to such tech giants for free usage. The dependence of face-recognition techniques in developing the surveillance network and AI-based drones, on its own, testifies the extent to which such data can be misused if fallen into evil hands.

Some Reports of Data Mishandling in India

Considering the role of data in the present age in defining the market, politics, surveillance and even warfare, the security of data collected and stored by the government and its agencies too assumes high significance. Even minute susceptibilities to leakage could prove to be fatal as these databases contain structured information that could range from demographic, contact and banking details to even biometric data. The vulnerabilities, orientation, and behaviour of individuals or groups of people could be easily identified and misused by employing the required algorithms over such structured and veracious data, if breached. Reports on *Aadhaar* data breach that surface in the media are hence alarming.

The Tribune, in January 2018, conducted an investigative operation and reported on a group who were allegedly tapping the data from the Unique Identification Authority of India's (UIDAI) database and providing access to it upon charging a small amount of Rs. 500. The group which was running a racket communicated through *WhatsApp,* a popular instant messaging service, and upon receiving the money, created a gateway with *Login ID* and *Password.* This gave unrestricted access to the details of more than 1 billion *Aadhaar*numbers created in India till then and contained sensitive information such as name, photo, address, phone number, email, and postal code. Upon paying Rs. 300 more, the group provided software which could enable the printing of *Aadhaar* cards of any individual by simply entering 12 digits Unique Identification Number.

Further investigations by *The Tribune* revealed that the group had apparently provided illegal access to around 1 lakh users. The group had targeted the *Village Level Enterprise* (VLE) operators who were hired under the *Common Service Centre Scheme* (CSCS) by the Ministry of Electronics and Information Technology for the task of enrolling and generating*Aadhaar* numbers across India. In a later stage when this activity was limited to Post Offices and Banks, these operators lost a means of earning a substantial income and hence, providing them again with access was an easy catch for the racket.

The UIDAI responded to this by denying any case of a data breach. It held that *Aadhaar* data including the biometrics were secure and stays encrypted in its federated servers. The alleged breach was only a case of the misuse of a grievance redressal search facility which was provided to designated personnel and State Government officials. It reiterated that such activities are properly logged by the UIDAI and are traceable; hence action can be easily initiated against the fraudulent activity.

Another report released by the Bangalore based *Centre for Internet and Society* (CIS) indicated that the irresponsible handling of information has exposed almost 135 million *Aadhaar* numbers and personally identifiable information such as Name, Gender, Father's/ Husbands Name, Age, Bank or Post Office Account Number *etc.* The information was leaked out from the databases of two schemes under the Central Ministry of Rural Development – the *National Social Assistance Programme,* through its dashboard and the *National Rural Employment Guarantee Act* (NREGA), through its portal.

The Andhra Pradesh State Government's NREGA portal and the *Chan-*

dranna Bima scheme dashboard too provided the details of its beneficiaries publicly. This did not account for any breach from the side of the UIDAI but the Unique Identification Number issued by the UIDAI was the key factor that linked and glued the information from the different government departments and databases to a comprehensive whole, as the *Aadhaar* number was seeded with every facility. This provided scope for constructing a 360-degree picture of the people and of profiling them with attributes.

According to *The Huffington Post,* the dashboard on a website maintained by the Andhra Pradesh government allowed people to search and identify the homes of 51,66,698 families in 13 districts of Andhra Pradesh on the basis of religion, caste *etc.* The geo-location of the homes could be accessed with precise latitudinal and longitudinal coordinates. The CIS report says that though the details were masked from public view, a simple tweak in the URL parameter from *nologin* to *login* provided complete access to the information without a password.

Conclusion

It implies from the cases and conditions aforementioned that the question of privacy is being decimated to a mere joke in the digital era as details including personally identifiable information of a huge section of the population are already available on the public domain with or without the consent of the individuals concerned. Interestingly, a huge majority of the population has tacitly approved the infringements of privacy and has internalized the concept of one's privacy and personal preferences being monetized for availing services *free of cost.*

The major segment of the advertising industry with the potential to influence customer behaviour is now feeding itself on the profiled data to which the service providers allow access to engage with. Also, the formidable role of big-data, data-mining, and analytics in politics reveals the extent to which an individual's details could be analyzed to predict his/her behaviour in the real world and to expose their vulnerabilities which could jeopardize personal security as well. Thus, the growth of data-mining and analytics to the extent of assuming a key role in defining the market trends and the course of political developments have effectively ignored the people's privacy concerns and potential security threats as a trade-off for employing these strategies which draw the largest possible audience, that too with no apparent charges for the end-user and minimal costs for the employer.

A glimpse on the extent to which big-data and Artificial-Intelligence is defining the concepts of future warfare hints that the potential of technology to build upon the concept of data still remains unexplored to a large extent. Trading-off one's privacy and personal details for smaller gains in the present could prove to be acts of idiocy as it cannot be conveniently reversed upon sensing the arrival of something catastrophic to people's security; inconceivable and unthinkable for the genius of today, but perfectly appealing to the senses of a layman of

tomorrow. Such is the pace at which technology is advancing and gaining currency.

References

Gattiker, U. E. (2004). *The information security dictionary: Defining the terms that define security for E-business, Internet, information, and wireless technology.* Boston: Kluwer Academic.

Muoro Infotech. (2016, September 24). *Data analytics vs Data mining what's the difference?* Retrieved December 09, 2018, from https://yourstory.com/mystory/083dab83a0-data-analytics-vs-data-mining-what-s-the-difference-

Big Data. (2016, December 19). Retrieved December 09, 2018, from https://www.gartner.com/it-glossary/big-data/

Oracle Corporation. (n.d.). *What Is Big Data?* Retrieved December 09, 2018, from

https://www.oracle.com/big-data/guide/what-is-big-data.html

Montcheuil, T. Y. (2015, August 07). *Facebook: A Decade of Big Data.* Retrieved December 12, 2018, from

https://www.wired.com/insights/2014/03/facebook-decade-big-data/

Watson, C. (2018, April 11). *The key moments from Mark Zuckerberg's testimony to Congress.* Retrieved December 12, 2018, from https://www.theguardian.com/technology/2018/apr/11/mark-zuckerbergs-testimony-to-congress-the-key-moments

Curran, D. (2018, March 30). *Are you ready? This is all the data Facebook and Google have on you.* Retrieved December 12, 2018, from https://www.theguardian.com/commentisfree/2018/mar/28/all-the-data-facebook-google-has-on-you-privacy

Ganjoo, S. (2018, December 13). *Google CEO Sundar Pichai testifies before US Congress, here are 5 important things he said.* Retrieved December 13, 2018, from

https://www.indiatoday.in/technology/news/story/google-ceo-sundar-pichai-testifies-before-us-congress-1407703-2018-12-12

Apuzzo, M., & Schmidt, M. S. (2016, November 15). *Secret Back Door in Some U.S. Phones Sent Data to China, Analysts Say.* Retrieved December 13, 2018, from https://www.nytimes.com/2016/11/16/us/politics/china-phones-software-security.html

Illing, S. (2017, March 16). *A political scientist explains how big data is transforming politics.* Retrieved December 13, 2018, from https://www.vox.com/conversations/2017/3/16/14935336/big-data-politics-donald-trump-2016-elections-polarization

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). *How Cambridge Analytica turned Facebook 'likes' into a lucrative political tool.* Retrieved December 13, 2018, from https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm

Rosenberg, M., Confessore, N., & Cadwalladr, C. (2018, March 17).*How Trump Consultants Exploited the Facebook Data of Millions.* Retrieved December 13, 2018, from https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html

Cadwalladr, C., & Glendinning, L. (2018, September 29). *Exposing Cambridge Analytica: 'It's been exhausting, exhilarating, and slightly terrifying'.* Retrieved December 13, 2018, from https://www.theguardian.com/membership/2018/sep/29/cambridge-analytica-cadwalladr-observer-facebook-zuckerberg-wylie

In Your Face: China's all-seeing state. (2017, December 10). Retrieved December 13, 2018, from https://www.bbc.com/news/av/world-asia-china-42248056/in-your-face-china-s-all-seeing-state

Tarnoff, B. (2018, October 11). *Weaponised AI is coming. Are algorithmic forever wars our future?* Retrieved December 13, 2018, from https://www.theguardian.com/commentisfree/2018/oct/11/war-jedi-algorithmic-warfare-us-military

Council on Foreign Relations. (2010, June 02). *Raising the Curtain on U.S. Drone Strikes.* Retrieved December 13, 2018, from https://www.cfr.org/interview/raising-curtain-us-drone-strikes

Sugg, S. (2017, November 16). *Keep Killer Robots Science Fiction.* Retrieved December 13, 2018, from

https://autonomousweapons.org/slaughterbots/

Khaira, R. (2018, January 04). *Rs 500, 10 minutes, and you have access to billion Aadhaar details.* Retrieved December 13, 2018, from https://www.tribuneindia.com/news/nation/rs-500-10-minutes-and-you-have-access-to-billion-aadhaar-details/523361.html

The Tribune. (2018, January 04). *UIDAI denies any breach of Aadhaar data.* Retrieved December 13, 2018, from https://www.tribuneindia.com/news/nation/uidai-denies-any-breach-of-aadhaar-data/523469.html

The Wire. (2017, May 01). *130 Million Aadhaar Numbers Were Made Public, Says New Report.* Retrieved December 13, 2018, from https://thewire.in/tech/aadhaar-card-details-leaked

Sethi, A. (2018, September 11). *Aadhaar Seeding Fiasco: How To Geo-Locate By Caste and Religion In Andhra Pradesh With One Click.* Retrieved December 13, 2018, from https://www.huffingtonpost.in/2018/04/25/aadhaar-seeding-fiasco-how-to-geo-locate-every-minority-family-in-ap-with-one-click_a_23419643/