Makale gönderilme tarihi: 01.09.2019
Makale kabul tarihi: 05.12.2019

Big Data, Artificial Intelligence and Machine Learning Algorithms: A Descriptive Analysis of Digital Threats in the Post-truth Era

Tirşe Erbaysal Filibeli

Ph.D.
Bahçeşehir University, New Media Department
tirse.erbaysalfilibeli@comm.bau.edu.tr
Orcid:0000-0003-4642-2279

Abstract

Did the utilization of big data change everything about how information circulates? Our digital data have been kept in big warehouses that we name 'big data'. All the things that we do in virtual life leave a digital footprint and thanks to machine learning algorithms; in our newsfeed, we mostly face content, which is similar to the subjects that we looked for before. Big data are being used for manipulating people to buy new products, to travel to new places, to read new books, etc. However, as it emerged in 2016 with Cambridge Analytica Scandal of Facebook, sometimes those technologies construct a threat for democracy. The underlying reason is that in our days, big data and AI algorithms have been used by political campaign managers to manipulate and/or persuade people through diffusion of promoted 'false' content. The aim of this study, by doing a descriptive analysis of very recent historical events like the failure of Microsoft's AI Tay and Youtube's effects on the presidential election in Brazil, is to define very current digital threats against democracy. Additionally, to better describe and discuss these digital threats we conducted semi-structured interviews with four experts who work on AI algorithms, big data, and social engineering. Our analyses and findings that we gathered from semi-structured interviews showed that there are several digital threats in the post-truth era that we live in like digital manipulation, violation of data privacy and misuse of big data, personalized search engine algorithms that create filter bubbles, the ease of production and diffusion of fake content.

Keywords: *Digital threats, big data, artificial intelligence, machine learning algorithms, disinformation*

DOI:10.16878/gsuilet.626260

Big data, intelligence artificielle et algorithmes d'apprentissage automatique: une analyse descriptive des menaces numériques à l'ère post-vérité

Résumé

L'utilisation du big data at-il tout changé concernant la circulation de l'information? Nos données numériques sont conservées dans un grand entrepôt que nous avons nommé 'big data'. Tout ce que nous faisons dans la vie virtuelle laisse une empreinte numérique et à cause des algorithmes d'apprentissage automatique; dans notre fils d'actualités, nous sommes principalement confrontés à des contenus similaires aux sujets que nous avions précédemment recherchés. Big data est essentiellement utilisé dans le but de manipuler les internautes afin d'acheter de nouveaux produits, visiter de nouveaux endroits, lire de nouveaux livres, etc. Cependant, comme il s'est avéré en 2016, avec Le Scandale Cambridge Analytica de Facebook, ces technologies constituent parfois une menace pour la démocratie. La raison sous-jacente est qu'aujourd'hui, les responsables des campagnes politiques utilisaient des big data et des algorithmes d'IA pour manipuler et/ou persuader les gens en diffusant du 'contenus faux'. Le but de cette étude est de définir les menaces numériques qu'affronte la démocratie et cela en effectuant une analyse descriptive d'événements historiques très récents, tels que l'échec de Microsoft AI Tay et les effets de Youtube sur l'élection présidentielle au Brésil. De plus, pour mieux décrire et discuter ces menaces numériques, nous avons réalisé des entretiens semi-structurés avec quatre experts travaillant sur les algorithmes d'IA, le big data et l'ingénierie sociale. Nos analyses et résultats tirés d'entretiens semi-structurés ont montré qu'il existe plusieurs menaces numériques dans l'ère post-vérité, telles que l'ingénierie sociale, la violation de la confidentialité des données et l'utilisation abusive de big data, des algorithmes de moteur de recherche personnalisés qui créent des bulles de filtre, et la facilité de production et de diffusion de faux contenu.

Mots-clés: *Menaces numériques, big data, intelligence artificielle, algorithmes d'apprentissage automatique, désinformation*

Büyük Veri, Yapay Zeka ve Makine Öğrenimi Algoritmaları : Hakikat Ötesi Çağda Dijital Tehditlerin Betimleyici Bir Analizi

Öz

Büyük verinin kullanımı, bilginin dolaşımına dair her şeyi değiştirdi mi? Dijital verilerimiz "büyük veri" adını verdiğimiz büyük depolarda saklanıyor. Sanal hayatta yaptığımız her şey dijital bir ayak izi bırakıyor ve makine öğrenimi algoritmalar sayesinde haber akışımızda çoğunlukla daha önce aradığımız konulara benzer içeriklerle karşılaşıyoruz. Temel olarak büyük veriler, insanları yeni ürünler almaya, yeni yerlere seyahat etmeye, yeni kitaplar okumaya vb. yönlendirmek için kullanılıyor. Bununla birlikte, 2016'da ilk kez Facebook'un Cambridge Analytica Skandalı ile ortaya çıktığı üzere, bazen bu teknolojiler demokrasi için bir tehdit oluşturmaktadır. Bunun nedeni günümüzde, büyük veri ve yapay zeka algoritmaları politik kampanya yöneticileri tarafından promosyonlu yanlış içeriklerin dolaşıma sokulması yoluyla insanları manipüle etmek ve/veya ikna etmek amacıyla kullanılmaktadır. Bu çalışmada, Microsoft'un yapay zekası Tay'in başarısızlığı ve Youtube'un Brezilya'daki cumhurbaşkanlığı seçimleri üzerindeki etkisi gibi son zamanlardaki tarihi olayların tanımlayıcı bir analizini yaparak, demokrasiye karşı ortaya çıkan güncel dijital tehditleri tanımladık. Bunun yanı sıra, bu dijital tehditleri daha iyi tanımlamak ve tartışmak için yapay zeka temelli algoritmalar, büyük veri ve sosyal mühendislik üzerine çalışan dört uzmanla yarı-yapılandırılmış görüşmeler gerçekleştirildi. Yapmış olduğumuz analizler ve yarı yapılandırılmış görüşmelerden elde etmiş olduğumuz bulgular, içinde yaşadığımız hakikat ötesi çağda sosyal mühendislik, veri gizliliğinin ihlali, filtre baloncukları yaratan kişiselleştirilmiş arama motoru algoritmaları, doğru olmayan içeriklerin üretiminin ve dolaşıma sokulmasının kolaylaşması gibi birçok dijital tehdit bulunduğunu göstermiştir.

Anahtar kelimeler: *Dijital tehditler, büyük veri, yapay zeka, makina öğrenimi algoritmalar, dezenformasyon*

Introduction

On March 23, 2016, Microsoft's AI (artificial intelligence) chatbot Tay was released via Twitter. Microsoft declared "Tay is designed to engage and entertain people where they connect with each other online through casual and playful conversation." The idea was that "the more you chat with Tay, the smarter she gets." However, just in hours, she turned into a Hitler loving sex robot who also supports Donald Trump. Tay was simply based upon machine learning algorithms fed from human bias. So, thanks to algorithms, it had been affected by malicious conversations and infected by racist data created by humans. In the end, Microsoft shut down Tay only 16 hours after its launch (Hunt, 2016; Molly, 2016; Wakefield, 2016). If we think in adverse way, it is possible to say that humans' bias might also be shaped with the fake content created and circulated with the use of digital technologies.

Within this context, in this study, by accepting the hypothesis, "in the post-truth era that we live in, there are some digital threats to democracy," we are seeking answers to two questions: "How are those technological developments constructing digital threats to democracy? Might humans' ideas be manipulated with 'fake content' by making use of big data, artificial intelligence, and machine learning algorithms?" To better understand the advantages and limitations of information technologies, besides defining digital threats in the post-truth era, four semi-structured interviews had been conducted with well-known academics from engineering departments, who are experts on computer science and information technologies. Within the interviews, we aim to have and reveal the knowledge about how digital technologies have been used to influence people for both producing and circulating information, and we explore answers to questions such as: "Is it possible to manipulate people by the use of technological assets? Might algorithms cause a loss of media pluralism? Might computer technologies be used for fighting with digital threats to develop a more pluralistic and democratic environment?"

Misuse of big data and digital manipulation[1] in the post-truth era

In 2016 the term "post-truth" was named the word of the year by Oxford Dictionaries. This term was firstly used in 1992 by the Serbian-American playwriter Steve Tesich. In 2004 for the first time, an academic Ralph Keyes used the term in the title of his book on post-truth politics and lies. However, the term became very popular with the effects of post-truth politics in the digital era, especially in 2016. Since there were Brexit Referendum in United Kingdom (UK) to leave the European Union (EU) and the presidential election in the United States (US), this notion drew people's attention all around the world (Flood, 2016; Keyes, 2004).

1 Digital manipulation is a term mostly used to define 'digital photo manipulation.' In this study this term has been chosen to use instead of the term 'digital propaganda' since it better describes the current phenomenon, which includes both visual and verbal manipulation besides algorithmic manipulation.

Keyes (2004) says that we live in a post-truth era where lies are not named as lies anymore. If lies are believable, they might be considered as truth. So, you do not have to lie anymore. Despite lying, you might 'exaggerate' or maybe 'misjudge' the situation. The term is defined by Oxford Dictionaries, as "objective facts are less influential in shaping public opinion than appeals to emotion and personal belief." In this era, it is easy to make people believe in something by using digital technologies. The most referenced and visible example of this situation may be the Cambridge Analytica Scandal and the role of Facebook to diffuse disinformation.

The scandal showed everybody how tech guys help campaign managers run computational propaganda by breaking data privacy, and with the use of algorithmic manipulation. The propaganda campaign of Donald Trump, named "Project Alamo" that was originally adopted from Cambridge Analytica since they came up with the Alamo data set. The firm harvested millions of people's Facebook profiles through an app named 'thisisyourdigitallife' which was developed by Aleksandr Kogan who had been working at Cambridge University as a data scientist at this time. Hundreds of thousands of people took the personality test and agreed to have their data collected for academic purposes by using this application. However, test takers unknowingly shared their Facebook friends' data, as well. With this app, they gathered data on the digital footprints of millions of people and they created psychographics to describe people's attitudes. In this way, a system had been built to profile voters. Based upon those psychographics, they produced different kinds of news that hold different points of views on the same topic. For example, about the migration policy of Donald Trump, they produced several news stories. By using psychographics, they targeted the potential voters. Within the Project Alamo, they have called these potential voters as 'persuadable voters' and by using micro-targeting technics and algorithms they made visible the most suitable content to influence those voters support Donald Trump. In short, they personalized political advertisements and they had tried to affect political choices of the crowd by using AI-based algorithms to promote ideas and promote politicians, namely Donald Trump (Bartlett, 2018; Amer & Noujaim, 2019; Cadwalladr & Graham Harrison, 2018; Rampling, 2017).

According to Chris Wylie, the whistleblower of the scandal, the company worked as a full-service propaganda machine (Sich, Bullock & Roberts, 2018). Theresa Hong, who ran the digital campaign for Donald Trump explained how they determined persuadable voters to change the election results and said that without Facebook they wouldn't have won the election (Rampling, 2017). According to Jamie Bartlett (2018, p. 69) the 2016 Presidential Election showed how big data and micro-targeting could win votes. Donald Trump's campaign was not the only example of digital manipulation, but it was the publicly known one. Like Trump Campaign, Dominic Cumming who is the campaign manager of the 'Vote Leave Campaign' utilized billions of targeted adverts during the Brexit Referendum Campaign (negotiations). However, even though Facebook agrees

to pay fine to UK for Cambridge Analytica, campaigners have never accepted that they used those kinds of adverts (Bartlett, 2018; Amer & Noujaim, 2019; Cadwalladr, 2017; BBC, 2019).

The emergence of Cambridge Analytica Scandal has made people think about from which countries the company harvested data. Facebook notified almost 87 million people all around the world that their information had been collected by Cambridge Analytica (Hern, 2018). Even in Turkey, 223 people had used the application and with their friends in total 234.584 people's information shared with Cambridge Analytica (T24, 2018). Data scandal of Facebook and misuse of information about people's digital footprints directed academics and digital activists to put new questions on the circulation of information, namely 'false information' and how the whole process has affected the democracy.

Truth matters for democracy: propaganda is not a new phenomenon

Data based information war plays a significant role in this era. Jamie Bartlett (2018) said that whoever owns the data owns the future. Is it right or wrong? Jennifer Pybus (2019) claimed that Donald Trump became the first Facebook President. When we just look at the results and think about their marketing technics, it makes sense. On the other hand, as she outlined, big data politics have many different angles that we should consider while discussing the role of behavior analysis that made it possible to influence people. Pybus assumed that there is a capital market behind data. Not only, Donald Trump, but many politicians also spent money to promote their campaigns. For this reason, sponsored content showed up on people's newsfeed. Most probably, these kinds of manipulation technics have not been used only in the USA. Just because of Facebook's disinformation scandal of Cambridge Analytica having happened there and the success of Trump Campaign, it is visible. Ted Cruz also utilized the same tools to be nominated for the Republicans but he failed. In ex-prime minister of United Kingdom (UK) Theresa May's Campaign, campaign managers also utilized big data and promoted content for her campaign, but in the 2017 election she lost the overall majority in the UK Parliament (Pybus, 2019). So from how much money you paid for sponsored content to how you determine your target audience; there is a hall process that is very similar to digital marketing professionals of trade-based companies utilized for promoting their products.

Propaganda, especially political propaganda is not a new subject to discuss for social science scholars. Media have always been used to manipulate people. Herbert Schiller explained the role of 'mind managers' in 1973. Chomsky and Herman defined traditional media's propaganda filters considering the role of the political economy of media in their book "Manufacturing Consent" in 1984. From Marx to Gramsci, Gramsci to Althusser, Althusser to Foucault, many different thinkers discussed the role of hegemony on the ideology during the 19th and 20th centuries. For this reason, saying propaganda is a new thing; or saying 'the term

"fake news" is being utilized to define a new phenomenon' is going to be an illusion. In this era the major discussion is based upon on 'the digital manipulation of masses'. These new digital technologies give people some digital opportunities to create different type of content and also to diffuse this misleading content to more people than they imagine. In this way, those technologies make easier to manipulate the crowds. Digital manipulation is a process which starts with determining the target audience, continuing with the production of the digital content, diffusing the content via digital wires, and making it visible with the use of social media platforms' algorithms. So, this process consists of digital content manipulation and algorithmic manipulation. Hal Berghel (2018) claims that this digital manipulation is a form of abuse, and all forms of abuse (physical abuse, mental abuse, verbal abuse, digital abuse, etc.) have similarity with Machiavellian roots: 'the desire to impose one's will or belief set on others.' So, here it might be said that it is digital abuse. Somehow, digital media users just as being connected to the networks give some advantages to people who use media to run political campaigns or to promote some ideologies.

The most important thing for political advertising is being good enough to persuade people (Pybus, 2019). Donald Trump professionally burst the mainstream media bubbles, and weaponized the reality to win the election. So after the election, discussions going on 'made up promoted content to persuade people' left its place to discussions on how mainstream media diffusing lies. Mark Zuckerberg didn't get the responsibility since Facebook is not a mainstream media company. It has always been based upon user-generated content, for this reason, Zuckerberg defended Facebook's policies and claimed that the social media giant Facebook might not be responsible for disinformation (Happer et al., 2019). However, there is a reality that everybody faced in 2016; people became victims by being exposed to 'promoted fake content'. Maybe the most visible digital threat for democracy is the circulation of 'fake news' on a purpose. On the other hand, some other digital threats make people unwilling victims of this digital manipulation.

(Un)willingly being the victim of digital manipulation

Not only in the USA but also in many other countries, digital technologies have been used for ideological purposes. The very recent example of this use might be the political radicalization of Brazil. In October 2018 far-right politician Jair Bolsonaro was elected as the president of Brazil. About the election of Bolsonaro, some discussions on 'how YouTube has played a role to radicalize Brazil' appeared. YouTube's powerful artificial intelligence system that learns from users' behavior recommends videos to people. However, in Brazil, most of the videos that were recommended were conspiracy videos, which were produced by far-right figures. You might search for different topics than politics but in the end, you'll find yourself watching those kinds of videos (Fisher & Taub, 2019).

Our willingly made choices have made us (un)willing victims of digital manipulation, since social media platforms work with machine learning algorithms. As in the YouTube example, its recommendation system is engineered to maximize watch-time. Before the Presidential Election in Brazil, Google's video site YouTube has already been started being criticized for promoting misleading videos and isolating people in filter bubbles (Nicas, 2018; Silva, 2019). Zeynep Tufekci (2018) criticizes the video recommendation system of YouTube that promotes conspiracy videos to increase the time people spent on the site, and she defines Youtube as the great radicalizer. Guillaume Chaslot, an AI specialist who worked on the recommendation engine of YouTube says that the algorithms aren't there to optimize what is truthful or honest but to optimize watch-time (Silva, 2019; Bartlett, 2018). It's because if users spend more time on their platform, they will earn much more money. So, machine-learning algorithms work for keeping users in the system, for this reason, they are recommending people videos that they might like. However, filters that we create with our choices and misuse of these filters and algorithms of YouTube for ideological manipulation might radicalize people's ideas. Those kinds of algorithms are not only utilized by YouTube, from Facebook to Twitter, each social media platform aims to develop algorithms that keep users on the platform. So, capitalist social platforms' machine learning algorithms, which make people spend more time in the system, is constructing a threat for easy circulation of false content. According to Bilge Narin (2018), because of the automatic filtering system, users have never encountered the information that could lead to overcoming prejudices, since they are only exposing to voices that are close to their voice, so algorithms reinforce their prejudices.

Herein, it is necessary to think about big data and theories like echo-chamber and filter bubbles. To start to explain what is filter bubble and how it damages the informational sphere, we need to understand the function of 'big data'. Today the volume of information has become so large. For this reason, both to store and to analyze this amount of data engineers need to develop new tools. So, they created huge information warehouses where all the people's digital footprints are being kept. In our days, with the data kept in these warehouses, it is easy to do qualitative analyses (Mayer-Schönberger & Cukier, 2013). Facebook is the most popular social media network in the world with more than 2.41 billion monthly active users.2 YouTube is following Facebook with almost 2 billion active users, and Instagram is following YouTube with 1 billion users.3 That means, every day billions of people share something or like/dislike those sharings. As of May 2019 every minute, more than 500 hours of video have been uploaded to YouTube.4 It can be assumed

2 Number of active Facebook users worldwide as of 2nd quarter 2019 (in millions) https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

3 Most popular social networks worldwide as of July 2019, ranked by number of active users (in millions) https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

4 Hours of video uploaded to YouTube every minute as of May 2019, https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/

from the numbers, the size of data is unimaginable. Thus, in today's world, data is the most important value for marketing products, and running digital campaigns and/or maybe manipulating ideas.

On December 4, 2009, Google introduced the "personalized search engine". Eli Pariser (2011, p. 6) says that with this technology, Google would use everything from 'where we were logging in' to 'which browser we were using' and to 'what we had searched for before' to guess about who we were and what kinds of sites we would like. He indicates that although most of us assume that when we search a term on Google we see the same results, Google is not working like what it was before since with the era of personalization began. Today algorithms are observing what we click, and our searches have been reflecting our own interests. Those new generation Internet filters look at the things we like and work like 'prediction engine', at the end, filter bubbles fundamentally alter the way we get information as Pariser said (2011).

As Cass Sunstein (2009) said in a democracy, people do not live in echo chambers or information cocoons. What does that mean? People might not want to see or hear topics and ideas that they don't like, however as a need for having a democratic point of view, we need to create a more pluralistic environment. In this era, a certain use of new technologies raises new questions on democracy. It is because, these technologies give us to right to choose what to watch and what to not watch, with who befriend and with who be unfriend, from whose sharing we like from whose sharing we dislike, etc. So, as Sunstein (2009) claimed, technology has greatly increased people's ability to 'filter' what they want to see, read and hear. Basically, we create our media according to our personal choices. In this sense, we are creating our echo chambers by being friend and/or following people who think like us. Besides, we create our information cocoons, which are feeding us with the same types of information. With the likes and dislikes, we are creating our filter bubbles (Burns, 2017). Herein, as Sunstein (2009) explained, a well-functioning democracy does not benefit from echo chambers or information cocoons. For this reason 'personalize search engines' can be defined as a digital threat for pluralism and democracy.

Since the personalization engine algorithms utilized by social media platforms don't display different ideas, users mostly see similar perspectives. This digital evaluation carries the risk of the extinction of diversity. As digital media users, when we do a search or when we use Facebook, YouTube, Twitter and/or other social media platforms, we get biased information according to our priorities. So users who have far-right ideology will never see news on left-wing politics or vice-versa. It harms not only pluralism but also democracy. Finding new ideas (or for the digital market finding products accept algorithms recommended us) will be hard unless we didn't escape from the filters that we created with our choices (Pariser, 2011; Bauman & Lyon, 2013; Haim et

al., 2018; Bartlett, 2018). In a way, it is auto-propaganda like Bauman and Lyon (2013) said, besides it is the violation of people's right to know (Pariser, 2011; Bozdag & van den Hoven, 2015).

As digital media users; with our shares, our likes, and dislikes, we create filter bubbles but also we become a part of content production. Today not only in Turkey in many other countries people get news on daily issues from social media and/or user-generated content based apps like WhatsApp (Newman et al., 2019). Digital media makes it easy to share any type of content but not only true one, but also false one. Fast spread of information might be a good thing for creating the digital public sphere if we could use social platforms as they meant to be. However the fast spread of wrong information creates another digital threat in this era (Tandoc et al., 2018). Since, once you fall in the loop of false information, it will be hard to get out of this loop.

Especially, if people are getting those wrong information from the sources that they trust namely from their echo chambers, they trust the information that they get without a doubt. Within this, if a post is popular, more people will see it. For this reason, using software robot accounts (bots) to get more likes, shares and comments is very common in this era. Social bots have been created with codes that learn from human interactions and simulate human behaviors (Tandoc et al., 2018; Erbaysal Filibeli & Şener, 2019, upcoming; CITS, n. d.). If we have biases, bots will have biases, if we share fake content most probably bots will rewrite and share this fake content. So, we can say that AI bots can be defined an another digital threat to democracy in the post-truth era, and all those digital threats make people (un)willing victims of digital manipulation.

Approaches and consequences: digital threats in the post-truth era

For the people who have no knowledge on psychographic and how algorithms work, digital manipulation seems like an illusion or a conspiracy theory. Accordingly, for understanding how technological developments have been used for manipulation and how it creates threats against democracy, it is necessary to comprehend dimensions of technological developments; namely big data, data mining, artificial intelligence, machine learning algorithms, etc. for this reason, within this perspective, it is a need to define digital threats in a more comprehensive way. In this context, we conducted semi-structures interviews with four professors of engineering to discuss digital threats against democracy. At first, we ask them to explain how technological tools turn into propaganda tools. Then to better describe digital threats we exemplify very current affairs like fake academic papers written with AI text generators, politicians' deepfake videos, etc. and ask them to define digital threats. According to their answers, we categorized digital threats that we face in the post-truth era and enumerated them.

'Social Engineering'[5] is for real

'Utilization technological tools to hack people to collect digital data is called social engineering. It is a serious threat in virtual communities, since there are many ways to gather personal data of people (Krombholz et al., 2015). In our days, if you combine social engineering with digital psychological manipulation technics to convince people to do something or believe something, it might be defined as 'digital manipulation'.

Serkan Ayvaz6 (personal communication, 2019) said that by using technological tools, doing mind management is technically possible and from marketing to politics, from politics to media sector AI algorithms have already been widely used in our days. Ayvaz indicates that through the big data and data mining, algorithms started to work better than before, since when the amount of the data rises, especially when the diversity of data increases, these algorithms have started to function better and better.

Sait Ölmez7 (personal communication, 2019) says that technological devices are just tools, these kinds of manipulation might be done only with the analysis of users' data. In this manner, according to Ölmez it is sure that digital data of users and AI algorithms have been used for commercial purposes. Ölmez indicates that algorithms look at what kinds of news we are reading, what kinds of daily practices we have, what kinds of videos we watch and what we like or dislike. In that way, algorithms show us only the subjects close to us. If we continue to read, watch, like and dislike similar news/sharings that algorithms recommend, algorithms will develop themselves according to users' preferences. So, we strengthen algorithms and in this sense, it is possible to manipulate people by using data on what we see, what we read, what we look, where we go, etc.

İlker Birbil8 (personal communication) affirms that it is not hard to manipulate people if we have the right tools. Birbil determines that in the history many kinds of manipulation technics were being used to manufacture consent, howev-

5 'In social sciences, 'social engineering' means engineering the public to influence social behaviors of masses. In the context of information security it means, hacking confidential information of people or companies to use for malicious purposes. Here it refers both definitions. Tüfekçi (2014) calls this phenomenon as 'networked based social engineering'.

6 Assistant Professor Serkan Ayvaz is the faculty member at the Software Engineering Department of Bahçeşehir University. He is an expert on big data and the coordinator of the Big Data Analytics and Management Master's Program at Bahçeşehir University.

7 Prof. Dr. Sait Ölmez is the faculty member of Sabancı University. His research areas are numerical techniques, data communications and security, Data Analytics and applications of Big Data. He is the director of Professional Master's Program in Data Analytics at Sabancı University.

8 Prof. Dr. İlker Birbil worked as a faculty member at Sabancı University, Industrial Engineering Program for 14 years. He is a faculty member at Erasmus University Rotterdam at the Econometric Institute. His research interests are data science, optimization in machine learning, algorithms for large-scale optimization and data privacy in decision-making. He worked as a columnist for Radikal Newspaper between 2012-2014 and for BirGün Pazar between 2016-2018.

er in our days if we want to manipulate masses, we need to use big data, since the tools that utilize to inform ourselves have been changed. According to Birbil it is possible to show people news that might affect them in the way that we want, so if we feed people's newsfeed with fake news, we can do much more effective manipulation. Thus, in this case, the reality loses its meaning and a new kind of reality have been created.

Cem Say9 (personal communication, 2019) claims that results of the utilization of psychological categorization of people are scientifically surprising; because via psychographics, it turns out that responses of people to the different kinds of news become understandable. In that case, it might be said that using psychographics is very functional to promote products or ideas. As remarked by Say, if you say "I want to show this content to these types of people" and if you pay enough money to Facebook for promoted content, it gives you that service. So, once you manage to classify users according to their sharings, likes and/or dislikes the circle is complete to give this service.

Digital manipulation of people looks like a conspiracy theory but when engineers namely data miners analyze data, they come up with very determined findings on users' behaviors. As mentioned by Ayvaz, today even our simplest mobile devices instantly collect data about all our movements and behaviors; moreover, cameras constantly collecting data and when we did a search on Google, our search leaves a digital mark. Ayvaz says that algorithms are mathematical models, which might see templates that we cannot see since as humans we constantly share data about our lives; for this reason, technically algorithms can clearly predict behaviors or feelings of societies.

Violation of data privacy and misuse of big data

According to Zeynep Tufekci (2014) without behavioral science models of how to persuade, influence and move people to particular actions, predictive analytics of big data wouldn't be as valuable as it is. Sait Ölmez (personal communication, 2019) says that commercial digital manipulation has been done during a long time that is why when we look at something in Amazon, it recommends us another product by saying "people who looked at this product, also looked at this one." With a broad definition of the term, this is also a manipulation but Ölmez says that this service doesn't really bother most of the users. In this manner, he underlines the importance of data privacy and data policies of social platforms, since the consumers need to know what these platforms learn from people to give such a service. However, most people don't read the policy of social platforms, even if they read policies it is hard for most people to understand such a long text. So, without what we share, we accept policies (personal communication Ayvaz 2019; Ölmez 2019; Birbil 2019; Say, 2019).

9 Prof Dr. Cem Say is a faculty member at Boğaziçi University, Department of Computer Engineering. His research interests are theoretical computer science, artificial intelligence, quantum computing and natural language understanding.

All of the four scholars underlined the importance of data privacy. So, "who holds data and who uses data for which purposes?" is important. To protect data privacy we need to develop better policies that might be understood by all the people. Since nobody read privacy policies written with small letters and pages long. For this reason, critical discussion of data politics might be the first and the most important step. All scholars mentioned the importance of General Data Protection Regulation (GDPR) and Personal Data Protection Law (KVKK in Turkey). According to scholars, this is a new area and we need to develop data policies. Besides, İlker Birbil (personal communication, 2019) says that some scholars work to develop some algorithms to protect data privacy. For this reason, they say that tech scholars and social scientists need to work together to develop up to date policies and strategies.

Disinformation 2.0: deepfake and AI text generators

In 2017 Mutlu Binark underlined that in the near future, algorithmic propaganda will not be used only in our social media news flow, but it will also be integrated into virtual reality (VR) and augmented reality (AR) applications. The application of VR and AR is still limited but another technological development that affects our understanding of reality has been utilized much more than these technologies: deepfake. According to Cem Say (personal communication, 2019), the 'generative adversarial network' might be the worse thing in the world since people started to race each other to generate the worse one. Generative adversarial network (GAN) is the tech behind deepfake. So, with GANs it is easy to produce real people's fake videos. Mona Lisa's or Salvador Dali's deepfake videos might be funny to watch since Mona Lisa is not a real person, and Salvador Dali passed away thirty years ago; however real people's deepfake videos, especially politicians might cause major disinformation. The most popular example of deepfake videos is Obama's video that is generated by comedian Jordan Peele. In the video, fake Obama was swearing to Donald Trump (Chivers, 2019; Parkin, 2019).

In this era, another problem is the ease of creating fake written content via AI fake text generators. There are many examples of these generators. Four scholars gave the same example on this issue. In 2005, three MIT students developed a computer program named SciGen to generate research papers and they got acceptance from many conferences (Sample, 2014). In our days, text generators are functioning more effective and more importantly, they are easily accessible on the web. For this reason, they create another threat. However, according to İlker Birbil (personal communication, 2019) they don't generate something original, for this reason, texts created by text generators are far away from being persuasive. Yet, one day in the not so distant future, these text generators might function better. For this reason, as Cem Say (personal communication, 2019) said, despite tackling fake content by developing technological tools, we also need to build critical thinking in the society.

'Filter bubbles' and capitalist digital media system

In addition to digital manipulation, violation of data privacy and disinformation, there is another threat in this post-truth era: filter bubbles. Even so, some researches show that negative effects of filter bubbles' aren't visible as much as we think, (Haim et al., 2018; Borgesius et al., 2016) we shouldn't ignore their presence. These digital information bubbles might be defined as a threat to media pluralism and diversity. According to the scholar that we talked to, algorithms show us mostly popular content, which get much more clicks than others or which make people spend more time in the system. In that way, the platform might earn more money. With machine learning algorithms, tech firms understand our behaviors and they might propose us to show specific content, which might willingly keep us in the system (personal communication Ayvaz, 2019; Ölmez, 2019; Birbil, 2019; Say, 2019). For this reason who we are friends with and from which sharing we like is important. We mostly give these data to social platforms by our hands. Our social media friends create our echo chamber and our dislikes/likes construct filter bubbles. These bubbles are great tools to understand our behaviors. In that manner, we mostly get the same news stories from similar kinds of people (personal communication Ayvaz, 2019; Ölmez, 2019; Birbil, 2019; Say, 2019). Thus, it is not wrong to say that filter bubbles create obstacles against the diversity of content and this one way, one type of communication cause the dysfunction and/or breaking down of democracy.

It seems like, to guarantee the diversity, the design of diversity-sensitive algorithms might be a solution (Bozdag & van de Hoven 2015). Sait Ölmez (personal communication, 2019) says that algorithms look at with who we are friends, what kind of person we are, with who we share similar political view or we have similar values etc. and based on these assumptions algorithms propose us places that we might like to go, books that we might like to read, news that we might like to watch or read. He indicates that algorithms might be designed to blow up those filter bubbles, but here is the question: do users like it? More importantly, do social media companies want to develop those kinds of algorithms that might annoy their users? The answer to these questions is "no". Ölmez also says that to provide pluralism, we don't need any algorithm, because if we randomly select news stories and give them to people, we'll provide media pluralism and diversity. But here again, we need to have a proper sample, which is already diverse. At this point, we need to sustain media pluralism and diversity at all.

So, scholars that we talked to agree about one thing; it is easy to create a pluralistic environment with codes (or without codes) but also they agree about one another thing neither tech companies nor people who like personal recognizer's services would want to ignore the personal recommendation. Cem Say says that with AB testing, YouTube found a way to keep people on the system in that way they have earned more money. So this capitalist digital media system is an obstacle to start a fight against filter bubbles.

Filter bubbles and machine learning algorithms fed from people's bias

İlker Birbil (personal communication, 2019) says that social media algorithms are man-made, so if we want, we can develop a recommendation engine based upon diversification. However, people have biases, for this reason, it doesn't mean everybody reads every content on their newsfeed. According to İlker Birbil, to create a more pluralistic virtual environment at first we need to have suitable conditions. In other words, we have to have a well-functioning democracy and a more pluralistic society, since algorithms learn from people's behavior.

Serkan Ayvaz (personal communication, 2019) also says that algorithms haven't recommended more pluralistic content since machine learning algorithms learn from real people's bias. He states; some studies showed that machine learning algorithms have biases like humans, for this reason to better understand algorithms we need to understand input data. Ayvaz says that after all, machines learn from people and human behaviors, then it makes itself more optimized; so prejudices, evil and goodness in humans affect algorithms biases.

Sait Ölmez (personal communication, 2019) gives an example about how people's bias affect machine learning algorithms. In 2016, in China two academics named Xiaolin Wu and Xi Zhang published their research called "Automated Inference on Criminality using Face Images". They utilized 1,856 criminals' pictures to develop an algorithm to determine which one is criminal and which one is not. Ölmez says that this study has been criticized a lot since there were human judges who decided these people are criminals or not. Additionally, humans who might have prejudices wrote algorithms which learn from data coming from people who might have similar prejudices: So, at the end, people thought that algorithms might replicate the decision of others.

Cem Say denotes that it is like a series of mirrors and it is what happened to Microsoft AI Tay or the real people. Machine learning algorithms learn from us: human beings. If we feed AI accounts with fake content or false biases or malicious ideas, they become like Tay. It is because AI accounts reflect what they learn from people. It is working in an adverse way; at first, we feed algorithms and then algorithms feed us. On the other hand, if we use psychographics, do micro-targeting and by using recommendation engines feed people with fake content to manipulate them, we can do it as well. As a consequence, like Birbil (personal communication, 2019) said, propaganda is not a new thing but the shape of our tools had changed and for this reason, the manipulations technics changed.

Discussion and conclusion: facing digital threats

"We shape our tools and, thereafter, our tools shape us."

This very famous quote is said by John Culkin (1967) to explain Marshall MacLuhan's ideas on how media have changed the human's environment. In

this era, we have so many digital media platforms to communicate with each other, and thanks to digital tools it is very easy to produce any type of content just in seconds. Moreover, there are many ways to diffuse or promote content that we want to show masses. It makes us think about digital manipulation because in this era new media have a great effect on human behavior. It is not possible to measure the real effect of digital manipulation, since it is impossible to test its effects with the same samples, because there is no possibility to do reelections under the same conditions and run the same digital campaigns. Yet it is possible to understand how companies like Cambridge Analytica work. So, one more time we need to ask "what happened in 2016?" Via a Facebook engaged app, Cambridge Analytica hacked people's data and violated millions of people's data privacy. With the data they hold, they did micro-targeting. Campaign managers determined their target audience and they produced information according to the ideas of their target audience by using psychographics and with promoted political adverts, they shared those fake news. When someone liked or shared promoted/fake content, this content became more visible. In the end, thanks to algorithms, they appeared in the news feed of masses. In that way, they had diffused not only true information but also false ones. It caused major disinformation. In short, it is a process starting with data mining, ending with the fast circulation of disinformation.

It is the most visible example of digital threats against democracy that we faced in this post-truth era. However, there are least visible examples, since Cambridge Analytica is not the only company that had run those kinds of digital campaigns and Cambridge Analytica was also active in countries such as India, Kenya, and Mexico, etc. It is hard to see how digital manipulation affected those countries. So, as Jamie Bartlett (2018) said smart machines are transforming humans' decisions and invisible algorithms create new hard-to-see sources of power and injustice. For this reason, we need to define threats and discuss the ways of struggling with those threats as soon as possible.

According to Serkan Ayvaz (personal communication, 2019), artificial intelligence is not capable of taking over the world now, and it won't happen in the near future, but that shouldn't prevent us from seeing a potential problem. Ayvaz indicates that the manipulation of elections and the circulation of fake news are actual problems. For this reason, we have to find a way to start to fight against those real problems. İlker Birbil (personal communication, 2019) claims that somehow they manipulate people and it hurts democracy. He also says that in the end, engineers invented all of those technological tools that cause some future threats, and also people who use these technologies founded Cambridge Analytica to manipulate people. All of them are human-made. By taking all these considerations, Birbil asks; "as humans why shouldn't we fight against those threats with our human made technological tools?"

As a conclusion, our descriptive analysis and findings showed that there

are digital threats like digital manipulation, violation of data privacy and misuse of big data, personalized search engine algorithms that create filter bubbles, the ease of production and diffusion of fake news. In the near future these threats might hurt (or already hurt) democracy; but more importantly, in the distant future it might kill the democracy. In this context, our interviewees said that unless if engineers and social scientists work together, we can struggle with the current and also future digital problems. For this reason, we need to come together and discuss how to diminish the negative effects of those digital threats.

References

Amer, K. & Noujaim, J. (2019). The Great Hack. [Documentary Movie]. United States: Netflix.

Bartlett, J. (2018). The People vs. Tech. how the internet is killing the democracy (and how we save it). London: Penguin.

Bauman, Z., & Lyon, D. (2013). Liquid Surveillance. Cambridge, UK: Polity Press.

BBC (2019, October 30). Facebook agrees to pay Cambridge Analytica fine to UK. BBC News. Retrieved from https://www.bbc.com/news/technology-50234141

Binark, M. (2017). Algoritmaların Yarattığı Yankı Odaları ve Siyasal Katılım Olanağı veya Olanaksızlığı. Varlık Dergisi, 1317, 19–23.

Berghel, H. (2018). Malice Domestic: The Cambridge Analytica Dystopia. Computer, 51(5), 84–89. doi: 10.1109/MC.2018.2381135

Bozdag, E., & van den Hoven, J. (2015). Breaking the filter bubble: democracy and design. Ethics and Information Technology, 17(249). https://doi.org/10.1007/s10676-015-9380-y

Burns, A. (2017). Echo chamber? What echo chamber? Reviewing the evidence. 6th Biennial Future of Journalism Conference (FOJ17). Cardiff, UK. Retrieved from https://eprints.qut.edu.au/113937/

Cadwalladr, C., & Graham-Harrison, E. (2018). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian. Retrieved from https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

Cadwalladr, C. (2017). The great British Brexit robbery: how our democracy was hijacked. The Guardian. Retrieved from https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy

Center for Information Technology & Society (CITS) (n.d.). How is fake news spread? Bots, people like you, trolls, and microtargeting. Retrieved from https://www.cits.ucsb.edu/fake-news/spread

Chivers, T. (2019). What do we do about deepfake video?. The Guardian. Retrieved

from https://www.theguardian.com/technology/2019/jun/23/what-do-we-do-about-deepfake-video-ai-facebook

Culkin, J. M. (1967, March 18). A schoolman's guide to Marshall McLuhan. The Saturday Review, 51–53. Retrieved from http://www.unz.org/Pub/Saturday-Rev-1967mar18-00051.

Tandoc, E., Lim, Z. W. & Ling, R. (2018). Defining "Fake News". Digital Journalism, 6(2), 137–153. doi:10.1080/21670811.2017.1360143

Erbaysal Filibeli, T. & Şener, O. (2019, upcoming). Manipüle Edilmiş Bir Enformasyonel Vitrin ve Popülist bir Enformasyon Alanı olarak Twitter, Moment Dergi.

T24 (2018, April 6). Facebook'taki veri skandalı, Türkiye'de 234 bin kişiyi etkiledi. T24. Retrieved from https://t24.com.tr/haber/facebooktaki-veri-skandali-turkiye-de-234-bin-kisiyi-etkiledi,599408.

Fisher, M. & Taub, A. (2019, August 11). How YouTube Radicalized Brazil. The New York Times. Retrieved from https://www.nytimes.com/2019/08/11/world/americas/youtube-brazil.html

Flood, A. (2006, November 15). 'Post-truth' named word of the year by Oxford Dictionaries. The Guardian. Retrieved from https://www.theguardian.com/books/2016/nov/15/post-truth-named-word-of-the-year-by-oxford-dictionaries.

Haim, M., Graefe, A., & Brosius H. B. (2018) Burst of the Filter Bubble? Effects of personalization on the diversity of Google News. Digital Journalism, 6(3), 330–343. doi: 10.1080/21670811.2017.1338145.

Hern, A. (2018, April 10). How to check whether Facebook shared your data with Cambridge Analytica. The Guardian. Retrieved from https://www.theguardian.com/technology/2018/apr/10/facebook-notify-users-data-harvested-cambridge-analytica#img-1

Happer, C., Hoskins, A., & Merrin, W. (2019). Weaponizing reality: an introduction to trump's war on the media. In Happer, C., Hoskins, A. & Merrin, W. (Eds.), (2019). Trump's Media War (pp. 3–22). London: Palgrave Macmillan.

Hunt, E. (2016, March 24). Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter. The Guardian. Retrieved from https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter.

Keyes, R. (2004). The post-truth era. New York: St. Martin's Press.

Krombholz, K., Hobel, H., Huber, M., & Weippl, E.R. (2015). Advanced social engineering attacks. J. Inf. Sec. Appl., 22, 113-122.

Mayer-Schönberger, V. & Cukier K. (2013) Big Data: A Revolution That Will Transform How We Live, Work, and Think. Boston, New York: Houghton Mifflin Harcourt.

Narin, B. (2018). Kişiselleştirilmiş Çevrimiçi Haber Akışının Yankı Odası Etkisi, Filtre Balonu ve Siberbalkanizasyon Kavramları Çerçevesinde İncelenmesi. Selçuk Üniversitesi İletişim Fakültesi Akademik Dergisi, 11(2), 232–251. doi:10.18094/josc.340471

Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2019) Reuters Institute Digital News Report. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-06/DNR_2019_FINAL_0.pdf

Nicas, J. (2018, February 7). How YouTube drives people to the Internet's darkest corners. The Wall Street Journal. Retrieved from https://www.wsj.com/articles/how-youtube-drives-viewers-to-the-internets-darkest-corners-1518020478.

Metz, C., & Blumenthal, S. (2019, June 7). How A.I. could be weaponized to spread disinformation. The New York Times. Retrieved from https://www.nytimes.com/interactive/2019/06/07/technology/ai-text-disinformation.html

Molly, M. (2016, March 24). Microsoft 'deeply sorry' after AI becomes 'Hitler-loving sex robot'. The Telegraph. https://www.telegraph.co.uk/technology/2016/03/26/microsoft-deeply-sorry-after-ai-becomes-hitler-loving-sex-robot/

Pariser, E. (2011). The filter bubble: What the internet is hiding from you? New York: Penguin Press.

Parkin, S. (2019, June 22). The rise of the deepfake and the threat to democracy. The Guardian. Retrieved from https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy

Pybus, J. (2019). Trump, the first Facebook president: why politicians need data too. In Happer, C., Hoskins, A. & Merrin, W. (Eds), (2019). Trump's Media War (pp.227-240). Switzerland: Palgrave Macmillan.

Rampling, J. (2017). Secrets of Silicon Valley: The Persuasion Machine. [Documentary Movie]. UK: BBC.

Sample, I. (2014, February 26). How computer generated fake papers flooding academia. The Guardian. Retrieved from https://www.theguardian.com/technology/shortcuts/2014/feb/26/how-computer-generated-fake-papers-flooding-academia

Schwartz, O. (2019, July 4). Could 'fake text' be the next global political threat? The Guardian. Retrieved from https://www.theguardian.com/technology/2019/jul/04/ai-fake-text-gpt-2-concerns-false-information.

Sich, A., Bullock, J. & Roberts, S. (2018, March 19). What is the Cambridge Analytica Scandal? The Guardian. Retrieved from https://www.theguardian.com/news/video/2018/mar/19/everything-you-need-to-know-about-the-cambridge-analytica-expose-video-explainer

Silva, M. (Presenter) (2019, May 25). How YouTube decides what you should watch. Retrieved from https://www.bbc.co.uk/programmes/w3csyvmt.

Sunstein, C. R. (2009). Republic.com 2.0. Princeton, N.J.: Princeton University Press.

Thurman, N. (2011). Making 'the Daily Me': Technology, Economics and Habit in the Main- stream Assimilation of Personalized News. Journalism, 12 (4), 395–415. doi:10.1177/1464884910388228.

Tufekci, Z. (2014). Engineering the public: Big data, surveillance and computational politics. First Monday, 19(7). https://doi.org/10.5210/fm.v19i7.4901

Tufekci, Z. (2018, March 10). YouTube, The Great Radicalizer. The New York Times. Retrieved from https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html.

Borgesius, Z., Trilling, D., Moeller, J., Bodó, B., de Vreese, C. H. & Helberger, N. (2016). Should We Worry About Filter Bubbles? Internet Policy Review. Journal on Internet Regulation. 5(1). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2758126

Wakefield, J. (2016, March 24). Microsoft chatbot is taught to swear on Twitter. BBC News. Retrieved from https://www.bbc.com/news/technology-35890188