A REVIEW ON THE RELATIONSHIP OF BIG DATA AND CAMBRIDGE ANALYTICA

Mustafa Eren AKPINAR
İstanbul Aydın University, Turkey
merenakpinar@stu.aydin.edu.tr
https://orcid.org/0000-0002-3917-3203

ABSTRACT

It has become a necessity to keep up with the digitalized world in the context of today's conditions. Because the world is changing and transforming. In the context of this change and transformation, it has been revealed in many academic studies that our lives have turned into data. In the light of these academic studies, a concept called big data has entered our lives. The importance of this concept, which is called and expressed as big data, is increasing day by day with the developing technologies. Especially after the Cambridge Analytica scandal, the meaning world of this concept, which is expressed as big data, is expanding, growing and creating a new scope for itself. In addition to this, it is observed how big data, whose importance is already understood, points to a situation with a wide range of influence. For this reason, within the scope of this study, it is focused on what big data is within the framework of the Cambridge Analytica scandal, and concepts such as privacy and participatory culture are examined. Then, within the scope of the current scandal, how big data is used and what it causes, a literature review was made with the selection of keywords for these concepts and the information obtained was compiled in the study.

Keywords: Big Data, Cambridge Analytica, Critical Approaches, Violation of Privacy.

BÜYÜK VERİ VE CAMBRIDGE ANALYTICA İLİŞKİSİ ÜSTÜNE BİR İNCELEME

ÖZ

Günümüz şartları bağlamında dijitalleşen dünyaya ayak uydurmak bir zorunluluk halini almıştır. Çünkü, dünya değişmekte ve dönüşmektedir. Bu değişim ve dönüşüm bağlamında ise hayatlarımızın birer veriye dönüştüğü birçok akademik çalışmada ortaya konmuştur. Bu akademik çalışmaların ışığında ise hayatlarımıza büyük veri olarak adlandırılan bir kavram girmiştir. Büyük veri olarak isimlendirilen ve ifade edilen bu kavramın gelişen teknolojiler ile birlikte önemide her geçen gün artmaktadır. Özellikle Cambridge Analytica skandalının ardından büyük veri olarak ifade edilen bu kavramın anlam dünyası genişlemekte, büyümekte ve kendine yeni bir kapsam yaratmaktadır. Bunun yanı sıra ise önemi zaten anlaşılmakta olan büyük verinin ne kadar etki alanı geniş bir duruma işaret ettiği gözlemlenmektedir. Bu nedenle de bu çalışma kapsamında Cambridge Analytica skandalı çerçevesinde büyük verinin ne olduğu üzerinde durulmakta, mahremiyet, katılımcı kültür gibi kavramlar incelenmektedir. Daha sonrasında ise mevcut skandal kapsamında büyük verinin nasıl kullanıldığı ve nelere yol açtığı üzerinde durularak bu kavramların anahtar kelime seçimi ile literatür taraması yapılmış ve ulaşılan bilgiler çalışmada derlenmiştir.

Anahtar Kelimeler: Büyük Veri, Cambridge Analytica, Eleştirel Yaklaşımlar, Mahremiyetin İhlali.

INTRODUCTION

With the digitalizing world within the framework of today's changing structure, the concept of big data is also becoming integrated into the lives of individuals and gaining a place in today's life. In this context, what big data is gains more importance. Because big data, which can be considered as a relatively new concept, is becoming a matter of curiosity both in academic studies and in daily life. Therefore, there is a need for an in-depth investigation.

In the light of in-depth examinations, internet technologies that have become important in today's world are mentioned. Along with these references, what big data is, what the Cambridge Analytica scandal is and what it evokes, and how the emergence adventure took place are examined. For this reason, in this study, the details of the big data concept are analyzed and revealed, and its practical process is expressed as well as its conceptual dimension. In the study, what big data is and what the concept of big data means are detailed. In this way, the literature support of the study is tried to be strengthened and the threats and advantages of big data in today's internet world are examined. One of these threats, the data gap, will be exemplified by the Cambridge Analytica scandal and the relationship between big data and big data will be examined.

In the review, it is also mentioned what the Cambridge Analytica scandal is, and within this scope, it is tried to be detailed by examining how it uses big data. At the same time, it is aimed to take the big data and Cambridge Analytica scandals, which are the two main subjects of the study, into a framework. Because it is thought that the relationship between big data and Cambridge Analytica has a deep structure and has a scope that needs to be examined. For this reason, the Cambridge Analytica sample was used as the main sample of the study and the study was tried to be detailed by emphasizing the concepts of digital labor exploitation and privacy.

At the end of the study, the study was made more detailed with all the details mentioned as big data, Cambridge Analytica, digital labor exploitation, privacy and participatory culture and so on. In addition, in the conclusion and recommendations section, it is mentioned whether it is possible to avoid such scandals. On the other hand, in the suggestions part, it is focused on how to bring solutions to the points mentioned in this study on an individual basis.

An Overview of the Cambridge Analytica Scandal

Cambridge Analytica is a data policy consultancy company founded by Alexander Nix in London in 2013 and whose mission is to understand what motivates individuals and what motivates them, with "data-based behavioral analysis" (Act. Güden, 2019: 217) and especially in the United States in 2016. It is a company that comes to the agenda during the United States (USA) Presidential elections.

Cambridge Analytica has a particularly massive impact with the data analysis methods it followed during the US presidential elections and the techniques of using big data, and allegedly, its effects on Donald Trump's victory in the 2016 US presidential election are also quite large.

In summary, the Cambridge Analytica scandal has resulted in the collection of the personal data of the users in exchange for the protection commitment of the Facebook company in order to expand their personal social networks, by collecting them through the research firm Cambridge Analytica (CA) without the knowledge and consent of the users and presenting them to various politicians during the election periods, and in the light of this data, the politicians. It is an event that includes the publication of regional and even city-based internet advertisements and the opportunity to develop political discourse. This process, which seems very innocent at first, is a privacy-oriented scandal in terms of the unauthorized collection of many data from users' social media shares to their likes, profile information and private correspondence, as well as political and political science and democracy theory when considered in terms of concepts such as voter behavior, democratic conventions and political populism. It is evaluated as an event that can bring serious discussions on the agenda" (Aksoy Abdulkadir & Türkölmez, 2020).

The Emergence of the Cambridge Analytica Scandal

The Facebook-Cambridge Analytica Scandal, which had a huge impact around the world, basically emerged when Facebook shared user information with other applications. Thanks to this feature, it can be expressed not only as a political event, but also as a technological and sociological event (Fidan, 2019: 96). This actually created one of the starting points of the scandal.

After Facebook shared the data with other third parties, Cambridge Analytica processed the information about individuals with the help of big data thanks to personalized content and, according to Fidan's (2019: 96), "data campaign that turned the world upside down" into action. Thus, the first seeds of scandal were planted. In this way, he succeeded in displaying a manipulative attitude on the target audience. Along with this attitude, it aroused repercussions in the eyes of the whole world public opinion, and according to many opinions, it played a major role in the 2016 elections of the USA.

Cambridge Analytica is also a scandal that is thought to be effective in the Brexit event, which is the event of Britain leaving the European Union. The Analytica scandal continued to arouse a great deal of repercussions afterward. It is also stated that there are days when 35 thousand news reports about the scandal are published in the USA (Amer & Noujaim, 2019). Another event where Cambridge Analytica is mentioned is the role it played in the presidential elections in the Central American country Trinidad and Tobago. The starting point of this event is that *"The process that started with Michal Kosinski and David Stillwell's doctoral studies at Cambridge University trying the behavior theory known as the "big five" on Facebook users, brought many names from the world of politics and technology to the agenda, and how important an instrument data is. underlined"* (Fidan, 2019: 97).

Although Cambridge Analytica was thought to be trying to create a new social structure in the main context, all the company actually did was pull individuals' information from the big data pool. However, the primary purpose of the company is considered to be a social intervention, as they pull this data extraction from illegal contexts and aim to win undecided voters. However, thanks to the ways the company followed, the election in the USA resulted in a 6 percent shift in the equal votes in Trinidad and Tobago, where the two candidates competed. This shows that the company is not recreating society or politics, but merely directing it. However, the use of data without the consent of individuals caused the company to be closed and the company name to be referred to as a scandal.

Cambridge Analytica Scandal in the Framework of the Culture Industry Approach

Before transferring the culture industry, it is necessary to look at what the culture and industry is. Because in order to form a unity, first of all, the concepts in their most basic form must be understood. Culture, the concept of culture (Ünür & Apak, 2021: 190), which was generally accepted as synonymous with the word civilization in the 16th and 19th centuries, over time "all the material and spiritual values created in the historical and social development process and in creating them, helping the next generations. The whole of the tools used in communication, showing the extent of human dominance over his natural and social environment, is defined as "hars, ekin" (Türk Dil Kurumu, 2011: 2558). In the light of this definition, it is possible to express the concept of culture as the social conditions of a society. The concept of industry, on the other hand, can be expressed as "industry" (Türk Dil Kurumu, 2011: 799) with the simplest definition. In other words, in the light of these definitions, it is not wrong to convey the culture industry as the industrialization of culture.

The Frankfurt School, in other words, the Critical Theory, which takes its place in the world of thought, when evaluated within the framework of the totalitarianism claims directed towards the enlightenment tradition, which basically establishes superiority over nature, distances people from legends and myths, and in this context, exposes the person to the law of non-human determined relationships (Larrain, 1995: 80) can be considered as a formation. In fact, the starting point of the culture industry is based on the Frankfurt School.

Although the Frankfurt School is a school that works in the field of social sciences (Kara, 2014: 52), which is generally based on Marxist orientations, it is where the understanding of critical theory emerges. The concept of culture industry, which is thought with critical theory, is mentioned for the first

time in the study "Dialectic of Enlightenment" written together by Thedor Adorno and Marx Horkheimer (Adorno, 2003). Along with this study, Adorno and Horkheimer, other Frankfurt School thinkers, made various studies on the cultural dimension of domination in the 20th century, the aestheticization of daily life, the commercialization of culture, standardization and social alienation (Kara, 2014: 52). In short, Adorno and Horkheimer see the culture industry and the media as an industry in which power relations are created, protected, multiplied and distributed, that is, actually affirmed by and nourished by power (Adorno, 2007: 74).

In this context, it is possible to say that there is a deep relationship between the culture industry and Cambridge Analytica. Because, when the Cambridge Analytica scandal is examined, the effect of the culture industry is observed in all the election campaign processes carried out by using power, media and big data, together with the effective way it played in the United States presidential elections in 2016. The clearest indicator of this is the acquisition of Facebook's data and the creation of personalized content during the election campaign process and the use of big data while doing this. For this reason, the deep, manipulative relationship between the culture industry and Cambridge Analytica is actually an indication that the privacy rights of individuals are also violated through big data and the internet. As a result, there is a relationship between big data, Cambridge Analytica and privacy.

Right to Privacy Violated by Big Data
With the rise of big data, it is possible to come across examples where individual privacy rights are violated. However, before addressing the breach of privacy with big data, it is first necessary to explain what data is. Data can be defined as a set of values (Şeker, 2013: 22) that can be obtained through observation or measurement methods without being subjected to any processing. According to another expression, data can be transferred as a preferred general expression for the information available in databases (Prytherch, 2005: 195). Within the framework of these transfers, it would not be wrong to say that data is actually information obtained through various methods.

Big data, on the other hand, can be expressed as a pool where the data obtained by various methods are integrated. However, it should be underlined that there is no complete definition of big data. According to another definition, big data is expressed as a phenomenon that covers various aspects of a phenomenon (from technological structure to economic models) (Altunışık, 2015: 48). In this context, it is possible to define big data as very large, very raw and exponentially growing data sets (LCIA, 2011), which are generally unstructured, continue to accumulate endlessly, are too unstructured to be analyzed with the help of traditional relationship-based database techniques (LCIA, 2011).

For this reason, it is correct to express big data as a set of data that develops and accumulates independently of individuals. From this point of view, it should be said that there is a relationship between privacy violation and big data. When the concept of privacy is briefly examined, it is observed that it can be expressed as the rights that individuals and individuals prefer to hide or do not prefer to say about their private lives. For example, sneaking into the houses of individuals, messing with their belongings or listening to their conversations (Yüksel, 2003: 184) can be a clear violation of privacy and such situations should not be experienced. However, the exemplary description conveys what privacy is in general, and therefore it is insufficient to explain privacy in the internet world or in the context of big data.

The concept of privacy is also shown as "an area where people can be alone and decide for themselves the conditions under which they will interact with others" (Çelikoğlu, 2008: 12). In this context, in fact, privacy refers to a structure that protects and adopts the personal information and opinions of individuals within the framework of virtual worlds. However, the fact that Facebook and many similar platforms store our personal data in today's structure and share it with Cambridge Analytica is a clear violation of privacy. The main reason for this scandal to occur is undoubtedly the sharing of personal data with third parties without the consent of the individual. In this case, it can be an indication that every individual, which is defined as big data and frequently encountered today, is data consisting of 0 and 1.

The Cambridge Analytica Scandal in the Context of Digital Labor Exploitation and Participatory Culture

The concepts of digital labor and participatory culture are very important concepts in today's world. Because, from the perspective of the changing world, the importance of digital labor and participatory culture is understood more. However, in order for these concepts to have a more solid foundation, it is necessary to mention the meaning of each word in the concepts and how they take on a structure when they are a whole.

In this context, it is necessary to look at the meanings of the words before defining the concepts of digital labor exploitation and participatory culture. Digital is defined as the electronic display of data on a screen, digital (Türk Dil Kurumu, 2011: 657) and actually represents a process that proceeds through screens. According to another definition, it is generally defined as "continuity and theoretically divisible structures of analog indicators and codes" (Chandler & Munday, 2018: 92). Labor, on the other hand, can be expressed as a work process, overtime, toil (Türk Dil Kurumu, 2011: 793), which changes both the natural and social framework and the person himself/herself consciously engage in to achieve a certain goal. This actually refers to the process carried out to achieve a goal. When the word exploitation is examined, it is encountered that it is defined as the work of exploitation (Türk Dil Kurumu, 2011: 2150), and it can be basically expressed as putting on a work.

In this context, it can be thought that digital labor exploitation can be expressed as the exploitation of digital labor by someone else. In addition, when the formation processes of digital labor are examined, it can be said that it can be changed within the framework of today's capital flow model. According to these models, it can be said that the unpaid labor of the users has become the means of profit generation (Fuchs, 2010) of the internet economy in today's economy, with productions within the framework of many different platforms such as microblogs, wikis, social networks, and content sharing sites. For this reason, it should be stated that the Cambridge Analytica scandal is an example of digital labor exploitation. Because the use of digitally produced labor by different institutions or organizations for their own purposes and doing this over digital or internet is one of the clearest indicators of the exploitation of digital labor.

It is possible to examine digital labor in three different categories (Uzunoğlu, 2015: 187). According to Fuchs and Sevignani, digital labor can be grouped as cognitive digital work, communicative digital work, and cooperative digital work (Act. Uzunoğlu, 2015: 187). In the light of these groupings, the types of digital labor are actually illuminated, and the diversity of digital labor is revealed.

Participatory culture, on the other hand, is a concept that does not show any similarity and sheds light on different issues, although it has similar points with digital labor. However, before mentioning participatory culture, it is necessary to look at the meanings of the concepts of participatory and culture. A participant can be defined as anyone who interacts with any meeting or event (Türk Dil Kurumu, 2011: 1350) or anyone involved in any communication process (Chandler & Munday, 2018: 231). It refers to active users who can interact with 2.0. Culture, on the other hand, is the sum of all the material and spiritual values created in the historical and social development process and the tools used in creating and transmitting them to the next generations, showing the extent of man's dominance over his natural and social environment, hars, ekin (Türk Dil Kurumu, 2011: 1558). is expressed. At the same time, it can be defined as "the sum of the features that give human communities their identity and distinguish them from each other" (Mutlu, 2017: 219). In addition to these, and in the context of etymology of the notion of culture, it can be stated that it actually originated from the Latin word "culture" and meant agriculture when it was used (Çeçen, 1996: 11). From this etymological point of view, it would not be wrong to say that culture is a production process (Akpınar, 2021: 365). Likewise, in this context, it is possible to express culture as a production process created through digital or traditional environments or as a structure created by participants.

Participatory culture has eliminated the necessity of individuals to come together in the physical environment, which are the main features of new media channels such as interaction, de-massification and simultaneity, and every individual who is a content producer on the internet has been the most

Review Article - This article was checked by iThenticate

important factor in the change of individual and mass communication (Act. Çetin, 2019: 157). Today, it is necessary to say that there is a process in which power passes from media tools to the user (Çetin, 2019: 157). Henry Jenkins evaluates this change as the new aspect of changing communication and calls active user performances as "participatory culture". However, in his studies on these changes, he deals with the convergence framework. Convergence; It is considered as "a cultural change in which consumers are directed to seek new information sources and are encouraged to make connections between dispersed media contents, rather than merely a technological process that expresses the collection of different media functions in a single device" (Jenkins, 2006: 15).

In this context, it would not be wrong to state that participatory culture is a conceptual framework combined with changing technologies between media and active users. In this case, it is an indication that the media and culture are intertwined and that cultural forms are transformed and redefined as Manovichn said (Manovich, 2001). In this context, when we look at the Cambridge Analytica scandal from the perspective of participatory culture, it can be said that it tries to manipulate social perception through the transformation of its cultural forms and the use of data. This manipulation process, combined with the exploitation of digital labor and participatory culture, paved the way for a more effective environment. With this newly formed ground, Cambridge Analytica has achieved great success. However, after other developments, the background of this success emerged and the Cambridge Analytica scandal brought up an international data gap problem. In this case, it has been realized that big data also creates a data gap.

METHODOLOGY
This study examined Turkish and English studies through a literature review, using keywords specifically determined within the scope of the study, such as big data, participatory culture, privacy violation, Camrbidge Analytica, and brought them together through conceptual definitions and case studies from these sources.

In this respect, the method of this study is a compilation study that will help future studies at the point of literature review, which consists of scanning the specially selected keywords in the context of the subject from the literature with appropriate scanning methods and examining how these definitions are made in both English and Turkish sources.

CONCLUSION
It is a fact that today's world has evolved with digital, and when the study data are examined, it has been determined that this assumption is correct. At the same time, the development of internet technologies and the widespread use of mobility cause the masses to leave a greater trace in the data sea. As these traces increase, data users such as Cambridge Analytica get more data and more manipulative movements can be exhibited. It is determined that this situation causes an increase in privacy violations.

It has been realized that the increasing privacy violations due to companies such as Cambridge Analytica that store data and sell or use this data cause individuals to have no private life and to obtain private information of individuals by certain companies or organizations. It is seen that this confidential information obtained is used by companies such as Cambridge Analytica to guide the society in the Brexit or democratic election processes. In this context, it is observed that individuals' decision mechanisms are interfered with by transforming various information into personalized content.

Likewise, when the study is examined in the context of digital labor exploitation and participatory culture, it is determined that the Cambridge Analytica scandal exploits digital labor and creates cultural changes in the perspective of participatory culture. In addition, it has been determined that people direct the society as they wish with the content they produce within the scope of their data, and fuel polarization among people in the Trump elections, Brexit event and many other big data-based events. Because, it has been seen that the information, data and personal content produced create an information gap between people, and in fact, algorithmic structures are changed with guiding information rather than correct information, and the homepage is shaped according to the personal preferences and ideology of the person.

Within the scope of all these cases, it is seen that Cambridge Analytica achieved success thanks to the policies it followed and implemented, but the company was terminated due to the fact that it committed a data crime as a result of following an unethical path. It is seen that this situation actually causes each individual to have data consisting of 0 and 1 in the digital environment. In order to prevent this, it is recommended that individuals be more careful in the digital and internet media, not to circulate their data freely, and to circulate as little personal information as possible in the digital world.

In addition, while individuals are present in the digital world, it is recommended that applications be given the least possible permission, some permissions are not given at all, and some information is only allowed while using the application. In this way, less personal data can be circulated by individuals in the world of data, and they can be less affected by the manipulative power of big data. For this reason, it is foreseen that paying attention to these points will create a more effective process.

REFERENCES

Adorno, T. W. (2003), "Kültür Endüstrisini Yeniden Düşünürken", Cogito Dergisi, Çev: Bülent O. Doğan, Yaz, (36).

Adorno, T. W. (2007), Kültür Endüstrisi, İletişim Yayınları, İstanbul.

Akpınar, M. E. (2021). TikTok Akımlarının Sosyal Medyadaki Eleştirisi: Twitter Üzerinden Yapılan Eleştirilerin İncelenmesi. 19 Mayıs Sosyal Bilimler Dergisi, 2 (2), 363- 381.

Aksoy A. & Türkölmez, O. (2020). Dijital Çağa Demokrasiyi Çağırmak: Cambridge Analytica Skandalı. Journal of Political Administrative and Local Studies (JPAL), 3(1), 41- 59.

Altunışık, R. (2015). Büyük Veri: Fırsatlar Kaynağı mı Yoksa Yeni Sorunlar Yumağı mı? Yıldız Social Science Review, 1(1), 45- 76.

Amer, K. & Noujaim, J. (Yönetenler). (2019). The Great Hack [Sinema Filmi].

Chandler, D. & Munday, R. (2018). Medya ve İletişim Sözlüğü. (B. Taşdemir, Çev.) İstanbul: İletişim Yayınları.

Çeçen, A. (1996). Kültür ve Politika (2. b.). İstanbul: Gündoğan Yayınları.

Çelikoğlu, N. (2008), 'Mahremiyet' Kişiye Ait Özel Alanlar Tartışması, İskenderiye Yayınları, İstanbul.

Çetin, İ. (2019). Yeni Medyada Yükselen Ses "Katılımcı Kültür": Survivor Örneği. Dördüncü Kuvvet Uluslararası Hakemli Dergi, 2 (2), 150-173.

Fidan, M. (2019). İktidarın Panoptik Gözü Olarak Sosyal Medya: Facebook-Cambridge Analytica Örneği. Yüksek Lisans Tezi, İstanbul Üniversitesi, Gazetecilik, İstanbul.

Fuchs, C. 2010. Labour in Informational Capitalism. The Information Society 26 (3): 176-196.

Güden, O. (2019). Küresel Kuruluşların Kriz Yönetimi ve İletişimi Süreçlerine Yönelik Bir Değerlendirme: Facebook & Cambridge Analytica Veri Skandalının Retoriksel Arena Kuramı Bağlamında İncelenmesi. Galatasaray Üniversitesi İletişim Dergisi (31), 209- 231.

Jenkins, H. (2006). Convergence Culture. New York, NYU Press, 2006.

Kara, T. (2014). Kültür Endüstrisi Kavramı Çerçevesinde Medya Ürünleri: Eleştirel Yaklaşım. The Turkish Online Journal of Design Art and Communication, 4(1), 51- 60.

Larrain, J. (1995). İdeoloji ve Kültürel Kimlik, (Çev. N. Nur Domaniç)., İstanbul: Sarmal.

LCIA (2011): "Big Data: Big Opportunities to Create Business Value", http://poland.emc.com/microsites/cio/articles/big-data-big- opportunities/LCIA-BigData-Opportunities-Value.pdf.

Manovich, L. (2001). The Language of New Media. Cambridge: MIT.

Mutlu, E. (2017). İletişim Sözlüğü. Ankara: Ütopya Yayınevi.

Prytherch, R. (2005). Harrod's Librarians' Glossary and Reference Book A Directory of Over 10.200 Terms, Organizations, Projects and Acronyms in the Areas of Information Management Library Science, Publishing and Archive Management. Londra: Routledge.

Şeker, Ş. E. (2013). İş Zekâsı ve Veri Madenciliği. İstanbul: Cinius Yayınları.

Türk Dil Kurumu. (2011). Türkçe Sözlük (5 b.). Ankara: Türk Dil Kurumu Yayınları.

Uzunoğlu, S. (2015). Yeni Medyada Dijital Emek Sömürüsü: Tüketiciden Üreticiye Yeni Medya, Yeni Sömürü Pratikleri. Intermedia International E-journal, 2 (1), 181-194.

Ünür, E. & Apak, D. (2021). Maddelendirilmiş Anlatımla İletişim ve Medya Kavramları Sözlüğü. Ankara: Gazi Kitabevi.

Yüksel, M. (2003). Mahremiyet Hakkı ve Sosyo - Tarihsel Gelişimi. Ankara Üniversitesi SBF Dergisi, 58(1), 181- 213.