Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi
Hacettepe University Journal of Faculty of Letters
Aralık/December 2023 – *40*(2), 518-530
doi:10.32600/huefd.1229239

Araştırma Makalesi – Research Article
Geliş Tarihi / Received:04.01.2023 Kabul Tarihi / Accepted: 03.04.2023

Big Data, Digitization and New Surveillance Practices: New Decision-Making Culture and Panoptic Classification

Büyük Veri, Sayısallaştırma ve Yeni Gözetim Pratikleri: Yeni Karar Verme Kültürü ve Panoptik Sınıflandırma

Elif ÖZUZ DAĞDELEN[*], Tuğça POYRAZ[**]

Abstract

Big data brings along a new decision-making culture through data mining aided decision-making mechanisms. One of the biggest sources of this mechanism is meta-data and that creates a suitable condition for mass-monitoring surveillance. The belief in these systems creates a discussion about usage of big data as a social classification tool, big data hubris, and big data divide. The aim of this study is to discuss the possible destructive results of big data especially as a social classification tool and new surveillance practices. In that point, this study focuses on the idea that big data paves the way to the new surveillance or dragnet surveillance practices and the possibility of social classification, and these might construct a world of increasing disadvantage. Through this panoptic sort, it comes with the problem of a serious power imbalance, simplification and decontextualization. It might result in the danger of loss of privacy and erosion of long-term privacy norms, a routine classification, manipulation of the future behaviors, antidemocratic control system, panoptic missort, data-antisubordination, abuse of civil rights and security problems.
Keywords: Big data divide, surveillance, panoptic sort, social classification.

Öz

Büyük veri, veri madenciliği destekli karar verme mekanizmaları aracılığıyla yeni karar verme kültürünü beraberinde getirmektedir. Bu mekanizmanın en büyük kaynaklarından biri meta-veridir ve kitlesel gözetim için uygun koşulları oluşturmaktadır. Bu sistemlere olan inanç, büyük verinin bir sosyal sınıflandırma aracı olarak kullanımı, büyük veri kibri ve büyük veri ayrımı hakkında bir tartışma yaratmaktadır. Bu çalışmanın amacı, özellikle bir sosyal sınıflandırma aracı olarak büyük verinin olası yıkıcı sonuçlarını ve yeni gözetim pratiklerini tartışmaktır. Bu noktada, bu çalışma, büyük verinin yeni gözetleme veya gizli ağ gözetleme pratiklerine ve sosyal sınıflandırma olasılığına yol açtığı ve bunların giderek dezavantajlı bir dünya kurabileceği fikrine odaklanmaktadır. Bu panoptik sınıflandırma sebebiyle, ciddi bir güç dengesizliği, basitleştirme ve bağlamsızlaştırma sorunu ortaya çıkmaktadır. Mahremiyet kaybı tehlikesi ve uzun vadeli mahremiyet normlarının aşınması, rutin sınıflandırma, gelecekteki davranışların manipüle edilmesi, antidemokratik kontrol sistemi, panoptik yanlış sınıflandırma, veriyi ikinci konuma tabii tutma, medeni hakların kötüye kullanılması ve güvenlik sorunları ile sonuçlanabilmektedir.
Anahtar sözcükler: Büyük veri ayrımı, gözetim, panoptik sınıflandırma, sosyal sınıflandırma.

[*] Res. Asst., Başkent University, Faculty of Science and Letters, Department of Sociology. E-mail: elifozuz4@gmail.com, ORCID: 0000-0002-6567-268X
[**] Prof. Dr., Hacettepe University, Faculty of Letters, Department of Sociology. E-mail: tpoyraz@hacettepe.edu.tr, ORCID: 0000-0003-3717-4728

Introduction

Data mining assisted decision-making mechanisms are the mechanisms that emerge because of the process of collecting, storing and then evaluating the surveillance resources required for big data surveillance with other mechanisms to obtain big data, and then evaluating them with the necessary analysis tools. These mechanisms are used in many areas based on data surveillance. These mechanisms are used in different areas like “tracking and evaluating disease distributions, tracking business trends, mapping crime patterns, analyzing web traffic, predicting everything from the weather to the behavior of individuals in economic markets” (Andrejevic and Gates, 2014, p. 185). One of the biggest sources of this mechanism is meta-data. Meta-data is information that identifies, describes, locates, or facilitates access, use and management of a source of information in the digital ecosystem, such as e-mail (Lokke, 2020, p. 40). Meta-data is crucial because of its capacity to turn into content that can be used in decision making through cross-linking. Lokke sees the rise in the capability to process meta-data as the driving force behind mass monitoring by explaining that the storage of every e-mail, SMS and phone call resulted in mass surveillance in contrast to surveillance of small groups or individuals. He explains the transformation of meta-data to content with an example while showing that the content is a combination of metadata. If a person saw a doctor, then researched HIV information, and then emailed a pharmacy, cross-linking shows that a lot of information can be obtained about this person (Lokke, 2020, pp. 41-42). In this process, individual is only looking for information through digital media channels but there are no identified restrictions about personal privacy and consent. However, because of the transformation of meta-data to content through cross linking and other analysis systems of big data, any information about an individual become a number under mass surveillance systems. It is called digitization in digital sociology, it is a compilation of information gathered about people, namely big data, through algorithms, and the pouring of human movement and even emotion into numerical data. As a result of this situation, the process created by all social, cultural, political and economic conditions in society is called 'digitalization' (Özuz, 2018, p. 55).

Dataization is a current phenomenon that explains the digitization of human life through digital information. It operates with two key elements of dataization: the external collection, distribution and storage of data, and value creation processes that include monetization as well as state control and cultural production. Dataization consists of “two processes: the transformation of human life into data through processes of quantification, and the generation of different kinds of value from data” (Mejias and Couldry, 2019). This digitization proceeds not only in the form of systems evaluating the individual, but also in the form of taking this digitization into consideration while self evaluating. The development of digital technologies facilitates both data collection and access to digitization. Social media platforms and self-monitoring tools that contribute to production, value production and the process of self-evaluation of the individual in this way are explained by the concept of "commensuration" in digital sociology. The ability to produce goals and make predictions about individuals through these means is conceptualized as "algorithmic manipulation". Lupton takes sexual activity and self-monitoring practices as examples. She states that even a physical and privacy-related subject such as sexual experience has become a situation that can be compared with self-monitoring practices and transformed into digital data. In this way, not only digitization and dataization occur but at the same time it is possible to form certain types and patterns in sexuality (2016). This entire process is conceptualized as the ideology of dataism. In the words of Jose Van Dijck, the ideology of dataism is the increase in belief in its objectivity, the belief that “metadata is a raw material of social life, and the disappearance of the opacity of data science” (Andrejevic & Gates, 2014).

In such circumstances, mass surveillance system creates a situation in which it is unknown who watches whom and for what purpose. Furthermore, regarding the privacy issue in big data analysis, the consent of individuals or the awareness of the masses about the ways in which their data is obtained constitute the main critical points to be discussed. Lyon states that questions about privacy are much more complex than they used to be. Referring to the concept of secrecy in Simmel's sociological work, he reminds us that keeping secrets is especially important in terms of social interaction and that what we know about people gives information about how we are related to them (Bauman and Lyon, 2020, p. 39). That is why big data makes this issue more complex because of the owner of this data lost its control and the usage of

this data become blurred. Lokke says that human biology is social and humans intent to share their personal information with others, so privacy is not the opposite of the public, but the issue is here that whether the owner decides to share this information and with whom (Lokke, 2020, p. 21). Because of the development and usage of big data, algorithmic manipulation becomes much more easier and big data and new surveillance practices create a significant power inequality especially in accordance with the aim of the utilization because of new decision-making culture they create. To discuss the power of big data through digitization and new surveillance practices as a social classification tool, this article firstly looks at the emergence and development of big data and sociological background of it. Then, panoptic to post-panoptic, theoretical perspectives of today's surveillance practices as a tool for social classification, are explained briefly. Then an attempt is made to establish relationship between big data divide and panoptic sort. Lastly, in the discussion and conclusion section, the possible results of these processes are discussed with both technical and sociological aspects.

Utilization of Big Data

The most common use of big data and the mechanisms in which it is used are surveillance practices, as this study will be based it on. But the use of these surveillance practices is not "traditional surveillance practices." It involves "new surveillance practices". While the basic principle in traditional surveillance is the surveillance of people who are thought to pose a kind of risk, the new surveillance involves the collection of data about everyone and is called "dragnet" surveillance practices. The biggest development that provides surveillance practices in the form of dredging network is an automated, continuous and even routinized data collection and processing process, which is easily performed with big data from a distance, has low visibility or is completely invisible. The necessity of evaluating this issue and discussing its use in society is to ensure the adequacy of the practices that individuals must participate in today for their inclusion in this system. All messaging used for communication, phone calls, e-mails, travels, hospital records, use of social media accounts, use of internet banking, expenditures made with credit cards, all purchases made from shopping sites, etc. are part of big data stored and algorithms.

Stating that big data surveillance will have social consequences in terms of law and social inequality, Brayne summarizes these as the increase in the use of big data as risk assessments through risk, the use of data for predictive purposes, mass surveillance, and data surveillance covering everyone, including information from more different institutions of data systems (Brayne, 2017, p. 82). It is known that big data is not used only to spy on people or to ensure public safety through methods such as visual surveillance. One of the most up-to-date and frequently discussed examples of big data surveillance in the literature, and even the most used one, stands out as Google's Influenza Trends Project. The fact that in 2008, Google created algorithms based on IP addresses to detect where diseases such as influenza spread, the regions with a high frequency and to take precautions, which is just one example of the use of big data surveillance in the field of health.

Technologically supported surveillance practices have become a general management tool in private and public spaces. Stating that there are two theoretical approaches to the question of why big data is especially preferred as an institutional practice, Brayne classifies them as technical/rational approach and institutional approach. In both approaches, it is stated that institutions are based on their own interests, but the reactions of actors to big data analysis are different. From a technical perspective, big data is seen as a tool for corporate actors to improve forecasts, fill analytical gaps and use scarce resources more effectively. In the institutional approach, the role of culture is emphasized, big data analytics can be used to provide legitimacy, not to increase efficiency. Big data replaces human decision-making mechanisms with automatic decisions (2017, p. 980). For example, in the traditional surveillance practices, the aim was to surveil the criminals and provide safety for other people in the society like it is mentioned in the technical perspective but in the new surveillance practices everybody is included, and aim might turn into a different agenda like it is mentioned in the institutional approach. According to institutional theory, new undesirable consequences occur when a recent technology is built on top of an old institutional structure. In the sense of big data, the transformation of individual actions into "objective" data may lead to undesirable results

such as simplification, decontextualization and privileged measurable complex social phenomena in the big data environment (ibid. p. 1004). The use of big data is also related to space, resources and storage capacity, and who has access to cloud data and provides the material means is also effective.

In that point, the principle of functional efficacy comes around. It represents the benefit that people or institutions try to achieve in various goals, especially by using algorithms, social recommendation system and analysis of the click sequences. In a sense, the functional benefit might be about achieving for various goals like predicting or manipulating by investigating any information obtained through digital surveillance like a person's or masses preferences about buying books, clicking on some links, or watching movies. The question at that point is what the real benefit or real function is, and whom these benefits and or functions are for? It is seen that it is not that simple as basic targeting of consumption patterns. It is about the power of the processing of big data. Bridle explains it in a very striking way:

> Processing first maps and models the culture, and then takes over its control. Google first set out to catalog everything that humanity knew, and then became the source, the determinant of this information: it became what people thought. Facebook first attempted to map the connections between people—the social graph—and then became the platform on which these connections were made, and irreversibly reshaped social relations. Just like an air control system that mistook flocks of migratory birds as bomber squadrons, software cannot distinguish between its model of the world and reality, and once we are conditioned, the same goes for us. (Bridle, 2020, p. 49)

In addition to these situations, the new surveillance or dragnet surveillance practices create a problem of data hubris that is a critique to belief in the possibility of predicting everything through big data and datafication. The goal of correlation replaces explanation and causation to reach the goal of prediction in data mining-supported decision-making mechanisms based on big data analysis. The fact that correlation can be reached between the two variables is not sufficient to support this assessment as correct. "Algorithmic correlations and predictions: they do not provide us with underlying common-sense explanations, but in some cases, they enable us to understand the incomprehensible findings that make them more complex" (Andrejevic and Gates, 2014, p. 185). Bridle comments on that situation as the magic of big data in an ironic way because he sees big data fallacy as the logical consequence of scientific reductionism. He gives an example of "GPS Deaths." In an arid region where there are many roads with impassable warnings and the temperature reaches fifty degrees during the day, you will die if you get lost. In the examples mentioned above, the GPS signal was neither obscured nor deflected. "A simple question was asked to the computer, and it answered – and people believed that answer to death" (Bridle, 2020, p. 52).

Another example of utilization of big data might be the application of big data by states. Big data and its supportive technologies, especially with the support of social media tools and digital media technologies, enable states to evaluate, control and in some cases even surveil their citizens. One of the most recent examples of this is the Chinese State Council's redefining the concept of social trust in 2014 in line with the "Draft Program for the Construction of the Social Trust System" and creating a new system called the "Social Credit System" or "Artificial Intelligence Supported Social Scoring System", by taking advantage of all the innovations provided by digital. With this system, people's attitudes and behaviors in the places they are physically present and every movement in the social media areas, tools, or platforms that have now acquired a real "space" meaning are collected. The collected information is transformed into data by algorithms. While determining the scores of the individuals, many factors such as their social habitus, their use of social media, their participation in social responsibility projects, whether they comply with the norms, values and laws deemed appropriate in the society are brought together and scoring is conducted in this way. The thing that allows a system to be established in this way is digital surveillance systems (big data as the main source), which is one of the most fundamental subjects of digital sociology. Digital surveillance is not just using physical surveillance technology tools such as cameras in the application area to establish a social credit system as China did; it also tries to explain the thought systems operating in the background in social and individual sense based on surveillance. In this sense, this system causes the concept of social trust

to be questioned again in societies with the use of face recognition technologies, social media, artificial intelligence technologies and robot birds. Although China is shown as an extreme example, in digital societies, individuals' data in digital platforms, social media and e-mail accounts, which websites they browse, what kind of internet searches they make, phone conversations, conversations in messaging applications are under investigation all the time by both international companies and governments. It is seen that these data can be collected and used by algorithms. NetzDG law has been enacted just for this purpose in Germany, the establishment of the e-Security Commissioner in Australia, a special law for digital media for hate speech in France, the obligation to keep the data of its own citizens on servers within the borders of the country in Russia, the prohibition of Google, Twitter and Whatsapp and the obligatory use of applications such as "national" Weibo, Baidu, Wechat in China are also indicators of how current and important the issue is. In this sense, the data of the individual is faced with "datafication" and the "dataveillance", which means providing social control and social classification with data and this is done with surveillance technologies. It results in the gamification of trust.

While the concept of social trust is shaped within the digital culture created by digitalization in society, it is thought that in a sense, "social engineering" may appear with completely new sociological phenomena such as the use of power and subjectification technologies, reputation scoring, and profiling. This process points to the emergence of data-driven management and regulatory systems based on traceability or computability. As was the subject of the "Nosedive" episode of the Black Mirror series, which has many episodes with its scenarios that can be considered as dystopias in the field of digital sociology, as it was discussed years ago, the fact that mobile devices gradually become a part of the individual is followed by the process of the formation of "technological habitus", the quantification of the body and identity and "quantification" of individuals. The analysis of social parameters by distinguishing social habitus with terms such as sphere of influence and inner circle, and even the emergence of new professions with the responsibility of increasing the scores of individuals with low social parameters reveals that none of these are dystopias. The point here is, there comes the issue that big data and its supportive systems are doing “social sorting” by using mass surveillance products. This sorting might be discussed under the sociological issue of big data divide. Two of the biggest problems in big data divide might be thought as big data hubris and data antisubordination.

Emergence and Development of Big Data: Sociological Background

Big Data has emerged and has been developing since data became a new raw material in the social and the business world and a product almost equal to capital and labor. With the development of information technologies, the collection of individual and mass data actively by bringing together various systems, and the necessity of processing, storing, and analyzing these data, all processes have supported each other. Chen, Mao and Liu show that the amount of data obtained until 2003 could be produced in 2 days in 2011; 750 million photos uploaded to Facebook, storage capacity of the American manufacturing industry reaching 966 petabytes in 2009; RFID (Radio Frequency Identification Technology) tags, which are used to identify objects with electronic codes, reached 209 billion in 2021, from 12 million in 2011; the amount of data produced by smart city projects in China reaches 200 petabytes; personal location data reaching $800 billion in 10 years (2014, pp. 172-173).

By scanning the cholera epidemic in 1854 and revealing that the water pump on a street caused this disease, it is said that Dr. John Snow laid the foundations of advanced data processing and indeed Big Data in its modern form (Lokke, 2020, p. 59). Its historical development is as follows: The emergence of “database machines” in 1970, the establishment of the parallel database named “Share Nothing” in 1980 to meet the increasing data volume, the establishment of the Teradata System in 1986 as the first successful commercial parallel database system, and the establishment of the Teradata System in 1990. It is seen that there are more technical processes like the spread of parallel database systems (Chen, Mao, and Liu, 2014). In fact, because of the Vs (Volume, Variety, Velocity and Veracity) of the data, the processes such as storage, analysis and processing become more difficult, the usage area of big data has developed with the creation of statistical and modeling programs by large companies. One of them is Google's GFS and

MapReduce Programming models. In 2005, Roger Mogoulas of O'Reilly Media first introduced the concept of "Big Data" to describe the large amount of data that current data processing approaches cannot manage due to complexity and size. However, the development of Big Data as a research and scientific subject has been going on since the 1970s (Subudhi, Rout and Ghosh, 2019, p. 26131).

In addition to these technical processes, Tüfekçi evaluates the rise of big data on a sociological basis with five intertwined dynamics. These dynamics are: "Moving from demographics to individualized targeting, the power and obscurity of computational modeling, use of persuasive behavioral science, digital media's provision of dynamic real-time experimentation and the formation of new political brokers who have data or social media circles" (Tufekci, 2014, p. 1).

With all these processes, big data has become a system that has the power to collect data about "everyone" and can be validated in the public domain. The biggest controversial point of big data is that it allows mass surveillance and data collection as well as targeted groups. The fact that the collected data is "included with everyone concerned" is not an equality. The fact that it is based on the idea of spying on everyone and treating everyone as a "target" or "suspect" is one of the most important points in big data becoming a sociological issue. By whom the compiled data is collected, how it is gathered, whether individuals or societies participate in the approval mechanisms in this process or have a choice, the measures taken by the states in this regard or the level of awareness, how international relations are affected by these processes, and especially those who have power in terms of computational modeling these are the main questions. Questioning the power gained by the state or third-party organizations has also revealed the necessity of evaluating and questioning big data as big brother. The use of the obtained data for real persuasion or manipulation has manifested itself in many international events (such as Facebook and the American elections), the dynamic real-time experimentation opportunity of digital media has changed the balance of power on all social, economic, cultural, and political grounds.

What is tried to be emphasized in this study is that the foundations of mass surveillance and surveillance societies come together with historical social events as well as the initial technical processes and gain a different dimension. When the historical development of the functioning of big data and basically the surveillance process is investigated sociologically, the processes that contribute to these processes are expressed as follows: "The emergence of nation-states in the 19th century, the transatlantic slave trade, bureaucratization, rationalization, and modern forms of government in the 19th and 20th centuries were influential in the functioning of the surveillance process. In addition, risk management practices in the 20th century have also contributed to these processes" (Brayne, 2017, p. 978). To put it simply, all these historical developments have led the development and transformation of big data, both directly and indirectly, to the importance of surveillance and accelerated it.

The basis of big data and algorithmic management discussions is about the participatory culture approach, user-oriented content production and production processes that contribute to the changes and transformations that have occurred after Web 2.0. The data collection, storage, and processing power of social media platforms or third parties in the legal sense have also taken place as the most effective discussions on algorithmic management and data privacy issues in the transformation of big data into a sociological phenomenon. The UGC (User Oriented Content) has not only formed the basis of a participatory culture but also a new decision-making culture. Big data is creating a new culture of decision-making and in that when tough questions are asked such as what the data says, where the data came from, what kind of analysis it went through, how confident it is about the results, data now offers the best estimates (Mcafee and Brynjolfsson, 2012, pp. 4-9). However, evaluating that as best estimate might be rethought with big data hubris. While big data is thought to be an effective and powerful tool in all the aforementioned areas, it is also frequently stated in the literature that it has become a trendy concept recently and might be the next new disruptive technology and due to the big data hubris, researchers might completely abandon traditional methods in data analysis and reach wrong conclusions by focusing entirely on big data without considering validity and reliability or without considering the content of the data. To avoid this, an interdisciplinary approach should be taken, which also combines a technological, legal, and economic perspective (Scherman, Hemsen, Buchmüller et al., 2014, pp. 261-263). In a sociological point of view, like Bauman and Lyon put it in this way: "The statistics and software logic that drives today's surveillance yield

results with uncanny consistency. Not only are 'Arabs' and 'Muslims' subjected to more 'random' scrutiny at airports than other people; at the same time, as Oscar Gandy points out, the social classification created by contemporary consumer surveillance constructs a world of 'increasing disadvantage'" (Bauman and Lyon, 2020, p. 25).

Theoretical Perspectives of Today's Surveillance Practices as a Tool for Social Classification: Panoptic to Post-Panoptic

The word 'Surveillance' was formed by the combination of word roots meaning "sur" (from above), "veillance" (watching), watching from above, and gained meanings such as monitoring, especially with CCTV cameras for a long time. Panoptic and post-panoptic surveillance studies are based on French post-structuralist and postmodern philosophies. Poster used the concept of superpanopticon in 1990, Haggerty and Ericson developed the concept of "surveillance assemblage" in 2000 based on the work of Deleuze and Guattari in 1987, and Mann developed the concept of "sousveillance" (Bogard, 2012). But all these studies are based on the panopticon. The known Panopticon, as designed by Bentham, and used in Foucault's theories, is a scheme that allows guards to observe prisoners while they are invisible from a central surveillance or inspection tower (Bentham and Božovič, 1995, p. 6).

Looking at the foundations of surveillance theory, Bentham's liberal project shaping the architectural design of a prison and other buildings and Foucault's use of it with a disciplinary analysis to talk about institutions and society stand out. Post-panoptic surveillance theories, which move away from the panopticon and develop alternative theoretical frameworks, stand out with Deleuze and Guattari's bureaucratic control societies and the east of a computerized, network-based society, Haggerty, and Ericson's surveillance community concept and Zuboff's theories of surveillance capitalism. Contemporary surveillance theories are characterized by improvements and additions to the main conceptual framework, while new types of panopticons include concepts such as digital surveillance, more user-centered participation and resistance perspectives (Galic, Timan, & Koops, 2016). In the first, there is a type of surveillance in the form of the architecture of power and the self-discipline of power, not just direct; in the second, there is a dispersed form of monitoring, dealing with "data doubles" rather than physically individuals; in the third, it is a distinction that includes concepts such as data surveillance, access, social classification, and dataization. In that study, due to the aim of seeing the big data's role in shaping social classification actions, like in the third type of surveillance studies, Oscar Gandy's surveillance as panoptic classification and Bogard's simulated surveillance are emphasized by relating with other sociological theories of surveillance especially with digital surveillance and new surveillance practices.

Surveillance as panoptic classification[1], it is a type of surveillance defined by H. Oscar Gandy Jr. based on the theories of Michel Foucault, Karl Marx, Max Weber, and Jacques Ellul. He takes perspectives from Max Weber - the role of knowledge in legitimizing the efforts of institutions to maintain predictability in their operations, from Ellul - the definition of technology that will include the philosophy and ideology of efficiency, as well as the hardware, software, and expertise necessary to achieve the goals of advancing panoptic sorting and reducing uncertainty, from Foucault - an appropriate description of the role of surveillance in the technology of power. Though the idea of adopting panopticism as its recipe and the disciplinary logic of surveillance, Giddens' theory of structuration, the idea of routine patterns of interaction between individuals and institutions, the action of knowledgeable agents from Marx and Giddens, produces

[1] First edition of Gandy's book on Panoptic Classification, "The Panoptic Sort: A Political Economy of Personal Information" 1993; the second edition is in 2021. Gandy also says that he added many innovations to the second edition, as there were so many changes in communication technologies and digitalization processes during this process. The most important of these is the emphasis on enormous amounts of data, as it is within the scope of this thesis. "The thing that has changed the most since the first edition has been the development of internet-connected digital technologies developed to differentiate individuals and various groups, with data obtained by processing large amounts of data obtained algorithmically" (Penn Today, 2021). For more details, see Gandy Jr., 2021.

and reproduces the world that must be continually challenged, this world will never be faithful to our design (Gandy Jr., 1996).

Oscar Gandy considers Giddens' theory of structuration important in finding a middle ground between hegemonic theories of domination, George Gerbner's theory of cultivation, and the near-absolute denial of external influence by some of the extreme culture theorists. Haggerty and Ericson introduced the Giddens's concept of "disturbance" as part of the "surveillant assemblage" that facilitated the creation of data doubles, "a replica of persons and lives," without being limited to a control tower, where power could be used remotely. He considers it is particularly important to understand its functioning (Gandy Jr, 2021, p. 2). It is one of the most important characteristics of new surveillance practices, it is based on the logic that power could be used remotely when people who thought they are under a surveillance, especially with new digital media technologies and because of data doubles.

Surveillance assemblage today includes many operations such as sorting, processing, recalling information. When Deleuze and Guattari's concept of assemblage is examined, it is seen that there is a distinction in horizontal and vertical axes. While the horizontal axis consists of two parts, content and expression, there are borders and unrestricted areas on the vertical axis. Although it is a heavy concept conceptually, Bogard explains the surveillance assemblage in two ways as "machinic" and "enunciative" using Foucault examples, based on the definition of Deleuze and Guattari. Divergent material flows are compounded or demarcated by the machinic assembly of surveillance. The function of prison is splitting and surrounding space, discerning bodies, or attaching them together into larger operational unity, bringing their collective flows all together as a machine and that typifies its disciplinary function like Foucault asserts that. Besides, the enunciative assemblage, by allocating non-corporeal conversions or events to bodies such as embodied actions and passions, is identified as a collective attribute machine. The most obvious function of this is the "correct" and "normal" references over this example (ibid., 2006, pp. 103-104). With this theoretical background, the subject that Bogard tries to interpret is that disciplinary surveillance can be thought of as a machinic assemblage in the sense that it sets boundaries, adjusts time and space, uses surveillance and classification mechanisms as a control strategy, but in the new society, control changes with new modes of information and simulation, and control is shaped by them (Haggerty and Ericson, 2019).

With these new modes, Bogard considers simulation as a post-panoptic control strategy. Referring to Baudrillard's concept of simulation, he states that simulation is a reproduction that refers to the real model but becomes "more real than real" or "hyperreality" in the process. At some point, simulation is also defined as something that masks the absence of reality. When we look at post-panoptic societies with Baudrillard's logic, it is seen that data mining and information clouds replace the visible areas organized by the panopticon, and doors, locks, keys, and entry are replaced by passwords, pin codes and encryption tools. Bogard conceptualizes this type of postmodern control as 'simulated surveillance'. As Baudrillard states, the collapse of the reality principle reverses causality in panoptic control. In disciplinary machines, justification comes before judgment. While panoptic surveillance aims to produce automatic obedience, it still reacts to and categorizes the events it describes before transmitting this information to the authorities determining its ultimate importance. In control societies, reasoning is much more initiative-taking, the simulation model configures the production and meaning of events and decides in advance (Bogard, 2012, pp. 32-35). In terms of the future of simulation surveillance, it is highlighted that haptic technologies are emphasized, and it is emphasized that the dimension of surveillance is completely different from the panoptic order and reaches much different technical and social dimensions instead of confinement. One of the striking points in Bogard's discussion stands out as the fact that social control does not remain within certain limits thanks to new simulation technologies and new control mechanisms positioning themselves in global capitalism. With this process, control becomes more inclusive, discipline moves into cyberspace, turns into simulation, and becomes hyperreality, reconstructed. In this sense, surveillance takes place not with centralized hierarchical surveillance or centralized powers, but with simulated softer forms (Haggerty and Ericson, 2019).

These simulated softer forms of "new surveillance practices" take place in Bauman and Lyon's work of liquid surveillance. Liquid surveillance is Bauman's work in which the framework of today's surveillance practices is considered with the liquid modernity of his previous work. This study consists of Lyon and Bauman's discussions by e-mail on the subject. The liquidity of today's surveillance can be seen in ever-

changing social media platforms for its users to collaborate with providers to develop new forms of communication and identity creation. “As information about our daily lives becomes transparent to the agencies that spy on us, it becomes more difficult to understand their own activities. As power moves with the speed of electronic signals through the volatility of fluid modernity, transparency increases for some and decreases at the same rate for others” (Bauman and Lyon, 2020, p. 24). In particular, the concepts of DIY-do it yourself and adiaphorization stand out. While do-it-yourself surveillance represents the process that occurs with the transition of surveillance from the ruler to the ruled, adiaphorization is defined as the process of setting aside moral and ethical questions for action, it is stated that this action can be to take something or pull the trigger (Finn, 2014). The second dimension of adiaphorization is data that can be processed, analyzed, combined with other data, and then reintroduced as data copies, either bodily (such as biometrics and DNA) or body-induced (such as being online, using access cards, and showing identification) (Bauman and Lyon, 2020, p. 19).

In do-it-yourself surveillance, it is emphasized that consumers (and at that point everybody who are under surveillance) are commodified and recommodified, individuals transforming themselves into commodities and a salable commodity (Bauman and Lyon, 2020, pp. 46-47). Inspired by Ernst Gellner's thought, pre-modernity is compared as a wild culture that does not need care, and modern as a garden culture that behooves perpetual care and control for weeds. Surveillance and education are also responsible for the demolishment of popular culture and the cultivation of compatible plants in this process (Lyon, 2010). Lyon defines liquid surveillance as “something that defines today's regimes of visibility well, characterized by data flows, mutating surveillance agencies, and the targeting and categorization of everyone.” With the concept of Liquid Surveillance, the reduction of the body to data creates a situation where data doubles, on which life opportunities and choices depend, become more important instead of our real lives and the stories we tell about them (Lyon, 2010, p. 325). Bauman, who also published an article titled "Never Be Alone Again", evaluates the death of anonymity due to the internet as our killing of our right to privacy with our consent, and that new and improved UAVs (unmanned aerial vehicles) will be programmed to fly on their own, and will be updated with the aim of inconspicuous design studies. saying, “No one knows whether or when a bee-bird will land on its sill” (Bauman and Lyon, 2020, pp. 32-35).

Big Data Divide: Big Data as A Social Classification Tool

As the most important aspects of big data in terms of social sciences stand out when a literature review is made, the social subject gives a kind of "approval" to sharing his/her data with a sense of acceptance in order not to stay away from direct systems, or without his consent or lack of sufficient knowledge and consciousness, and this perception of consent is constantly questioned. It is mostly seen dangerous because machines create the perception that individuals know them better than themselves in all systems that support it, and its future is not known for what purposes it is used. The formation of personal information networks occurs not only with the movements of individuals in an area, but also with all personal devices, social networks, all other applications and devices. New surveillance practices come with much more serious big data divide problem when they are compared with the traditional ones. Big Data Divide refers to the asymmetrical relationship between the people who collect, store and mine the data and the target group whose data is collected (Andrejevic, 2014, p. 1674). What is important at this point is the power imbalances in the digital age and the public attitude towards the collection and use of personal data. One of the main reasons for the emergence of the big data divide is related to qualifications. Andrejevic conceptualizes the two groups that emerge depending on these competencies as “Big Data Rich” and “Big Data Poor” (2014, p. 1675). Digital surveillance practices widen the gap between people who have power in that sense or not. The reason for such a separation in terms of big data existence is based on the expensive technological infrastructures, expensive data sets, software, processing power and expertise needed to access and control big data analysis (ibid. p. 1675).

Today, it is striking that the collection of big data based on will and consent is done under the name of "Informed Consent Form", which is a process that supports big data divide. The weakness here is that the ownership and control over information and communication technologies is based only on the Big Data

Rich. It is emphasized that the group conceptualized as the Big Data Poor is also weak in terms of knowledge-based decision making and that the consent given is not actually a kind of consent or at least it is a subject that needs to be discussed. Systems that support the sense of powerlessness are based on the proliferation of "monitoring technologies" that allow data to be collected online and offline.

Big Data Divide is related with the Gandy's concept of "Panoptic Sort" in the development of discriminatory technologies and making future predictions. The point that Gandy bases this on is the idea that calculation systems and social classification began to occur in the era of modern bureaucratic rationality, and that it is a continuation of the bank, home and insurance industries from the Taylorist "Scientific Management" systems in the mid-20th century (Andrejevic, 2014, p. 1677). The fact that the data speaks for itself with the "End of Theory" discussion, which is the suggestion of Chris Anderson, who is known and widely included in the literature. Anderson calls this the "Petabyte Age". This divide and social classification through new surveillance practices should be thought again with big data hubris like it is explained above. The concept of "Surveillance Simulation," mentioned by William Bogard, helps to discuss how the system and indeed the Big Data Rich and Big Data Poor can manipulate the data of the Big Data Poor to use it for their own benefit economically, socially, culturally or politically. An example would be that health insurers use customer data to reduce coverage to avoid incurring large just-in-time healthcare costs. In terms of the sociology of digital surveillance, David Lyon and Bauman's conceptualization "liquid surveillance" that brings all these concepts together and gives meaning and definition of "Surveillance as Social Classification" examines in detail in terms of defining how new surveillance practices because of digitization and big data identify identities and attribute values.

Discussion and Conclusion

The panoptic classification, created with this theoretical background, is a complex discriminatory technology. Panoptic is called "all-encompassing" because individuals consider all information about individual status behavior potentially useful in producing intelligence about a person's economic worth. Based on these estimates, people are divided into categories (Gandy Jr., 1996, p. 133). In panoptic classification, personal information has three distinct functions. These are: identification, classification and assessment. After the identification phase is conducted using the information of the persons, certain types are created based on these data, and here, with the concepts of Foucault, the exercise of power and disciplinary isolation occur. In the classification phase, a differentiation process is aimed by minimizing the uncertainty about the future behavior of individuals. In the evaluation part, the classification is completed with preprocessing and estimating software (ibid. p. 136).

This classification highlights the data collection cooperation between companies and government and the process that this cooperation poses a threat to individual privacy. As in the example of Snowden, which is often mentioned in other surveillance discussions, the National Security Agency in the United States collects data about individuals and monitors all kinds of movements of its citizens with a program called PRISM, and the emergence of big data and control power reveals the danger of loss of privacy and social classification by group of big data rich. In Sullivan's words, Gandy said that at a time when Google was not yet built and computing power was extremely low, organizations such as Equifax, TRW, and the Direct Marketing Association were accumulating huge repositories of consumer data using individuals' data, combined with complex matching algorithms with government databases such as censuses. He warns that it is monitored, including sensitive privileged information about individuals. The crucial point here is that the routine classification of personal data has turned into a powerful form of institutional power (2014). Companies contract with independent agents and consultants to provide strategic information subsidies, and while these relationships are sometimes kept private and confidential, relationships are sometimes announced for their benefits to the public (Gandy, 2003, p. 289). In general, it is thought that such a panoptic classification, whether announced or not, will lead to an anti-democratic control system. The panoptic classification system is becoming a more effective system with developments such as social media platforms, which were developed in addition to the way it was designed at that time and were even more

convenient for collecting data. According to Sullivan, who states the four main reasons why the data panopticon has become more effective, these are listed as follows:

1. Software has a main place in social, commercial, and political systems and they all support each other
2.Increasing importance of metadata in routing, storing and classifying information
3.Global data collection and retrieval business
4.Blurring the lines between corporate and government data mining. (Sullivan, 2014)

Campbell and Carlson highlight the great inequality in the relationship between the providers and consumers of products and services in the market, as one of the most important points that Gandy makes with the panoptic classification. What makes this relationship so unequal is that individuals have difficulty in sharing their personal information (Campbell & Carlson, 2002, p. 591). In addition to this inequality, Blevins, who tries to examine Gandy's work with the current digital media technologies and especially the concept of big data and based on the traditional Marxist critique in Gandy's view, panoptic classification is a method that defines, classifies and identifies consumers by stealing a kind of additional value that occurs with their individual information. Stating that it is a power technology that evaluates power, he says that panoptic classification produces inequality not only based on economic value, but also based on race, gender, age, class and culture. In today's online data collection and panoptic machine logic, it is concluded that the personal information of individuals is misrepresented. In this sense, the panoptic classification becomes an incorrect classification (the concept of panoptic sort is considered as a panoptic missort at this point). While the data is processed, reproduced and standardized in terms of meaning, wrong measurement and miscalculation occur (Blevins, 2016, p. 29). As a result, according to Gandy's own observations, surveillance builds a cumulative disadvantaged world (Pero, 2015, p. 478).

The most discussed issue with the concept of big data in the literature is inclusion, emphasizing privacy and civil rights. However, it is also important to emphasize the issue of exclusion. Exclusion also needs to be considered based on the threat of equality and privacy brought about by big data. It is named as data antisubordination. It is stated that there should be a "*Doctrine of Not Subjecting Data to a Secondary Position*" (Lerman, 2013, p. 55). In the issue that is wanted to be discussed with the concept of data antisubordination, all people who are not included in big data and live outside these lines are meant. There is a possibility of isolation from the general population and even marginalization of those, whose poverty, geography, way of life or lives are somehow less given. It is thought that this marginalization may lead to the fact that voices of these people are not heard. It is among the strong predictions that economic opportunities, social mobility, and even democratic participation can develop in line with big data and recent technologies. In this sense, the Doctrine of Not Subordinating Data to a Subordinate Position should ensure that states benefit equally from public goods and services, especially so that big data does not lead to new social stratification, equality should be ensured in legal proceedings and political processes, and even measures should be taken for this in the private sector, rather than taking a state action (ibid., pp. 60-63). Increasing datafication enables situations such as predicting people's experiences and future attitudes and making various applications for this. But datafication is not the only dark side of big data. Big data causes, as suggested by Lerman, "abuse of civil rights by the state, erosion of long-term privacy norms and damage to the environment by server farms which are responsible for the great energy spent in processing big data" (2013, p. 56).

Increasing datafication and big data divide, result in "big data hubris," and the debate of "End of Theory", it is about where sufficient data speaks for itself, and the theory is not needed. While making this discussion, big data analysis is the necessity of analyzing all the available data, the emphasis here being on necessary and unnecessary, relevant and irrelevant collected data ("found data"). The data to be used in the analysis is mixed and dispersed data that is constantly updated and grows simultaneously, searches made from search engines, location information recieved from all mobile devices, accessible messaging, user-generated content produced from social media platforms (UGCs- User Generated Content). Big data, as Harford points out, has growing data problems that result from too much going on in its analysis. The fact

that the focus is not on causality but only on numerical correlations, as in the example of detecting Flu Trends that Google is trying to execute, removes logical suggestions and leads to an analysis that is far from theory. Statistically, such analyzes lead to sampling errors and sampling bias. While the first is the case where the sample to be inferred does not reflect the basic population, in the second case the sample is not even chosen randomly (2014, pp. 14-17). One of the biggest problems encountered in big data analysis, in this sense, arises from focusing directly on big data mining without question, instead of turning to data mining of the truth. Correlation and prediction purpose, which replace explanation and causation, stand out as the factors that most affect big data analysis and use of big data. In addition to that, big data has security problems too, especially the processing and collection of sensitive data in the collection and processing process is one of the examples thereof. It is necessary to ensure the control and protection of data and to develop the data location laws necessary for this. In addition to laws, encryption, tokenization, and data masking are required to prevent sensitive data from being read (Tankard, n.d., p. 5). As this article tried to discuss, digitalization and digitization have become the reality of society. To follow and adapt to the social, cultural, economic, legal, and political changes that these processes, big data, and new surveillance practices resulted in and to prevent the problems that might induce, both theoretical and applied research come into prominence.

Authorship Contribution:	Conceptualization: EÖD, TP; Data curation: EÖD, TP; Formal analysis: EÖD, TP; Investigation: EÖD, TP; Methodology: EÖD, TP; Resources: EÖD, TP; Supervision: EÖD, TP; Validation: EÖD, TP; Writing-original draft: EÖD, TP; Writing-review and editing: EÖD, TP.
Conflict of interest:	The authors declare no potential conflict of interest.
Financial support:	The authors received no financial support for the research.
Ethics Board Approval:	The authors declare no need for ethical approval for the research.

References

Andrejevic, M. (2014). Big data, big questions: The big data divide. *International Journal of Communication*, *8*(17), 1673-1689.

Andrejevic, M., & Gates, K. (2014). Big data surveillance: Introduction. *Surveillance & Society*, *12*(2), 185-196.

Bauman, Z., & Lyon, D. (2020). Akışkan Gözetim. (E. Yılmaz, Trans.). Ayrıntı Yayınları.

Bentham, J., & Božovič, M. (1995). *The panopticon writings*. Verso Trade.

Blevins, J. L. (2017). Panoptic missorts and the hegemony of US data privacy policy. *The Political Economy of Communication*, *4*(2), 18-33.

Bogard, W. (2006). Surveillance assemblages and lines of flight. In *Theorizing surveillance* (pp.111-136). Willan.

Bogard, W. (2006). Welcome to the society of control: The simulation of surveillance revisited. In Kevin D. Haggerty, and Richard V. Ericson (Eds.), *The new politics of surveillance and visibility*, (pp. 55-78). University of Toronto Press.

Bogard, W. (2012). Simulation and post-panopticism. In *Routledge handbook of surveillance studies* (pp. 30-37). Routledge.

Brayne, S. (2017). Big data surveillance: The case of policing. *American sociological review*, *82*(5), 977-1008.

Bridle, J. (2020). Yeni karanlık çağ: Teknoloji ve geleceğin sonu. (K. Güleç, Trans.). Metis Yayınları.

Campbell, J. E., & Carlson, M. (2002). Panopticon.com: Online surveillance and the commodification of privacy. *Journal of Broadcasting & Electronic Media*, *46*(4), 586-606.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile networks and applications*, *19*(2), 171-209.

Finn, J. (2014). Liquid surveillance: A conversation. *Canadian Journal of Communication*, *39*(3), 497.

Galič, M., Timan, T., & Koops, B. J. (2017). Bentham, Deleuze and beyond: An overview of surveillance theories from the panopticon to participation. *Philosophy & Technology*, *30*(1), 9-37.

Gandy Jr, O. H. (1996). Coming to Terms with the Panoptic Sort Oscar H. Gandy Jr. *Computers, surveillance, and privacy*, 132.

Gandy Jr, O. H. (2003). Public opinion surveys and the formation of privacy policy. *Journal of social issues*, *59*(2), 283-299.

Gandy Jr, O. H. (2021). *The panoptic sort: A political economy of personal information*. Oxford University Press.

Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *The British Journal of Sociology*, *51*(4), 605-622.

Haggerty, K. D., & Ericson, R. V. (2019). The new politics of surveillance and visibility (pp. 3-26). University of Toronto Press.

Lerman, J. (2013). Big data and its exclusions. *Stan. L. Rev. Online*, 66, 55.

Lokke, E. (2020). Mahremiyet: Dijital Toplumda Özel Hayat (D. Başak, Trans.). Küy Yayınları.

Lyon, D. (2010). Liquid surveillance: The contribution of Zygmunt Bauman to surveillance studies. *International political sociology*, *4*(4), 325-338.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard Business Review*, *90*(10), 60-68.

Özuz, E. (2018). Dijital sosyoloji perspektifinden toplumsal değerlerin değişim sürecinde YouTuberlar: İlköğretim öğrencileri örneği. [Yüksek Lisans Tezi]. Hacettepe Üniversitesi, Sosyoloji Bölümü

Penn Today. (2021). The panoptic sort: Surveillance Q&A with Oscar Gandy. Retrieved from The Panoptic Sort: Surveillance Q&A with Oscar Gandy | Penn Today (upenn.edu)

Pero, R. (2015). Liquid surveillance: A conversation. New Media & Society, *17*(3), 478–480.

Schermann, M., Hemsen, H., Buchmüller, C., Bitter, T., Krcmar, H., Markl, V., & Hoeren, T. (2014). Big data. *Business & Information Systems Engineering*, *6*(5), 261-266.

Subudhi, B. N., Rout, D. K., & Ghosh, A. (2019). Big data analytics for video surveillance. *Multimedia Tools and Applications*, *78*(18), 26129-26162.

Sullivan, J. (2014). Uncovering the data panopticon: The urgent need for critical scholarship in an era of corporate and government surveillance. *The Political Economy of Communication*, *1*(2).

Tankard, C. (2012). Big data security. *Network Security*, *2012*(7), 5-8.

Tufekci, Z. (2014). Engineering the public: Big data, surveillance and computational politics. *First Monday*, *19*(7).