Journal of Research in Business : 5 (2), 147-159 December • Dezember • Aralık: 2020 • ISSN: 2630-6255
DOI: 10.35333/JRB.2020.278

RESEARCH ARTICLE • FORSCHUNGSARTIKEL • ARAŞTIRMA MAKALESİ

BIG DATA AND OPEN DATA ANALYTICS: THE APPLICATIONS AND THE PATH TAKEN IN TURKEY

BÜYÜK VERİ VE AÇIK VERİ ANALİTİĞİ: TÜRKİYE'DE ALINAN YOL VE UYGULAMALAR

Ezgi ERDOĞAN[*]
Juan S. BEJARANO[**]

Abstract

Having data is becoming more and more important for companies and countries. This is because data is very useful for generating added value and providing transparent government policies. Achieving these benefits depends on the success of data analysis, and therefore investments in data science are constantly increasing. The fact that the data sources are diverse today and the data flow is continuous and very fast also brings the data to be very large and too complex to be processed. In this study, the concepts of big data and open data are introduced technically and the importance and gains of open data that have social and economic effects are mentioned with international examples. In addition, topics covered in data analysis and the use of machine learning in data analysis are given. Following the introduction of services provided using data analysis in Turkey, the study is completed presenting data portals and government policies of Turkey and obstacles and problems hindering the development of open data in Turkey.

Keywords: Data portal, open data, big data, data science, machine learning
JEL Classification: Y1, Y2

Öz

Veriye sahip olmak hem şirketler hem de ülkeler için gün geçtikçe daha önemli hale gelmektedir. Bunun sebebi, verinin katma değer üretmek ve şeffaf devlet politikaları sunmak için oldukça faydalı olmasıdır. Bu faydaların elde edilmesi veri analizinin başarısına bağlıdır ve bu nedenle veri bilimine yapılan yatırımlar sürekli artmaktadır. Günümüzde veri kaynaklarının çeşitli olması, veri akışının sürekli ve çok hızlı olması verinin çok büyük ve işlenmek için çok karmaşık olmasını da beraberinde getirmektedir. Bu çalışmada, büyük veri ve açık veri kavramları teknik olarak tanıtılmış, açık verinin sosyal ve ekonomik etkileri olan

* Dr., Marmara University, Faculty of Arts and Sciences, Department of Mathematics, ezgi.erdogan@marmara.edu.tr
** Dr., Universidad Politécnica de Valencia, Escuela Técnica Superior de Ingeniería Informática, Industrias de la Comunicación y Culturales, juabebal@upv.es.

To cite this article: Erdoğan, E. & Bejarano, J. S. (2020). Big data and open data analytics: The applications and the path taken in Turkey. *Journal of Research in Business*, 5(2), 147-159.

"Çalışmada Etik Kurul izni gerekmemektedir."

Submitted: 08.06.2020 **Revised:** 08.08.2020
Accepted: 12.08.2020 **Published Online:** 29.12.2020

önem ve kazanımlarından uluslararası örneklerle bahsedilmiştir. Bunlara ek olarak, veri analizinde ele alınan konular ve makine öğrenmesinin veri analizinde kullanımı verilmiştir. Çalışma, Türkiye'de veri analizi kullanılarak sunulan hizmetlerin verilmesinin ardından, Türkiye'deki veri portalları ve devlet politikaları ile veri analizinin Türkiye'de gelişmesini yavaşlatan engel ve problemler sunularak tamamlanmıştır.
Anahtar Kelimeler: Veri portalı, açık veri, büyük veri, veri bilimi, makine öğrenmesi.
JEL Sınıflaması: Y1, Y2**.**

1. Introduction

As electronic environments develop and become widespread, our perspectives on the evaluation of data change and develop: the size of the data increases, it becomes easier to process different data together. As the size of the data increases, there are also changes in both the techniques and technologies and methods and solutions used. Under these developments, a field as a new branch of science called *data scienc*e is appeared. The aim of the researchers working in this area is to obtain the products having high added-value after processing the collected data. As stated in the following table (Table 1), the size of the data is constantly increasing and new sizing scales or new data dimension measures are used. Today, data scientists are max out at yottabytes, but soon, brontobytes will measure the volume of sensor data generated by the Internet of Things (IoT).

Table 1: Size of Data

Unit	Equal to	Size in bytes
Byte (B)	1	1
Kilobyte (KB)	1,024 bytes	1,024
Megabyte (MB)	1,024 kilobytes	1,048,576
Gigabyte (GB)	1,024 megabytes	1,073,741,824
Terrabyte (TB)	1,024 gigabytes	1,099,511,627,776
Petabyte (PB)	1,024 terrabytes	1,125,899,906,842,624
Exabyte (EB)	1,024 petabytes	1,152,921,504,606,846,976
Zettabyte (ZB)	1,024 exabytes	1,180,591,620,717,411,303,424
Yottabyte (YB)	1,024 zettabytes	1,208,925,819,614,629,174,706,176
Brontobyte (BB)	1,024 yottabytes	1,237,940,039,285,380,274,899,124,224

To understand the size of the data, the examples given below are striking;

- Every day 500 Terabytes of data is added on Facebook.
- The CERN's Large Hadron Collider generates 1 Petabyte per second.
- Now, every day 1 Exabyte of data is created on the internet. (This is equal to 250 million DVDs.)

As the amount, size, or diversity of data increases, new concepts emerge, or concepts change over time. The importance of data does not revolve around how much data you have, but what you do with it. The data is no longer simply being a data but transforming into products that provide

added-value. Today, the most advanced and useful data are "smart" data containing intelligent factors and semantic structures.

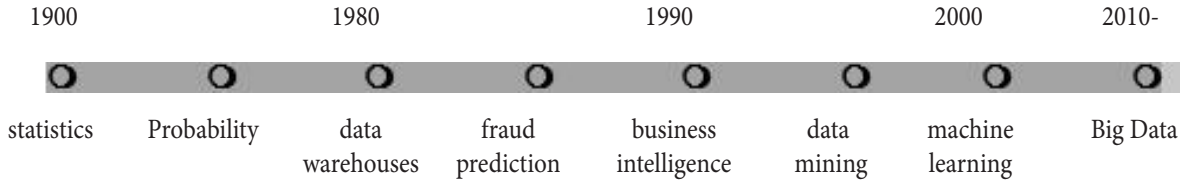

Figure 1: A Brief History of Data

Figure 1 shows the historical development of the data (Urszula, 2018). Thanks to new technologies, the concept of big data has entered our lives in a very recent past, and due to its many benefits in different fields, investments and postgraduate educations in data analysis are increasing rapidly. Now, let us look more detailed at the notions of big data and open data.

2. Big Data and Open Data : What They Are, Why They Matter

In this section, we will separately introduce the notions of big data and open data with both technically and the economic and social effects.

2.1. Big Data

The term "Big Data" refers to data that is so large, fast or complex that it's difficult or impossible to process using traditional methods. It is produced in different formats, quickly and in large volumes. As can be seen from the definition of big data, it also brings new problems in addition to its benefits.

The data produced is obtained in three different formats given below:

1. *Structured:* Organised data format with a fixed schema. Relational database management system (RDBMS) can be given as an example. RDBMS is a data storage system where data is kept in rows and columns in tables and has a high data consistency. It is a widely used management system because it offers a combination of both system performance and ease of implementation, and provides a reliable method for storing and retrieving large amounts of data.
2. *Semi-Structured:* Partially organised data that does not have a fixed format. Ex: eXtensible Markup Language (XML), Javascript Object Notation (JSON). XML is a language that allows the structuring and tagging of documents, it is not a language in itself, but a system that allows defining languages according to the needs. JSON based on a javascript sub-language that allows its easily interpretation and generation is a simple text format for data exchanges.
3. *Unstructured:* Unorganised data with an unknown schema. Ex: Audio, video files etc.

Big data is not only a collection of datasets with different formats but also it is an important asset that can be used to obtain enumerable benefits. Companies take advantage of the big data accumulated in their systems to improve operations, present better customer service and personalized marketing campaigns based on specific customer preferences, ultimately, increase profitability. There are some of the many sources of big data such as Sensors/meters and activity records from electronic devices, social interactions, business transactions, electronic files, Broadcastings. You can take data from any source and when you combine big data with high-powered analytics, you can accomplish find answers that enable c*ost reductions, time reductions, new product development and optimized offerings*, and *smart decision making*.

Big data is often characterized by six components (volume, variety, velocity, veracity, value, variability), briefly called 6 Vs of big data (Table 2) which are explained as follow.

i. *Volume:* The volume appears as one of the most important known features and problems of big data. With big data, you'll have to process high volumes of low-density, unstructured data.
ii. *Variety:* The diversity of data is due to the diversity of data sources. The produced data can be seen in structural, semi-structural or non-structural formats. If the data are used together in these different structures, this diversity in the data creates problems.
iii. *Velocity:* The term "velocity" refers to the speed at which big data is generated and (perhaps) processed. With the growth in the Internet of Things (IoT), data streams into businesses at an unprecedented speed and must be handled in a timely manner. The fact that the data produced in a flow can be analyzed and managed in real-time is another problem of big data.
iv. *Veracity:* The degree of reliability of big data is important to keep under control the data coming from so many different sources that is difficult to link, match, cleanse and transform data across systems. It must be ensured the accuracy of the data and its suitability for analysis with technologies specific to big data.

Table 2: 6V's of Big Data

Volume	Variety	Velocity	Veracity	Value	Variability
Size	Diversity	Speed	Accuracy	Usability	Instability
How much data?	*What type of data?*	*How fast can data be accessed?*	*Is data reliable?*	*Is data useful?*	*How diverse is data*

v. *Value:* It is the value produced from data that makes data analysis important by various organizations. Any data that cannot produce value is nonsense. The value produced depends on factors such as the content of the data, the purpose of production, the field of application.
vi. *Variability:* This term ascribes the ways in which the big data can be used and formatted. It is challenging that data flows are unpredictable – changing often and varying greatly– but companies need to know when something is trending, and how to manage daily, seasonal and event-triggered peak data loads.

If the existing data have the above features, it is very difficult to generate value from this data by traditional methods. Therefore, the need to handle this data with a big data perspective and to analyze it with big data technologies has arisen. We list some good-to-know terms and most popular technologies as below.

- ***Cloud*** is a widely used approach in big data and delivers on-demand computing resources on a pay-for-use basis.
- ***Hadoop*** is a framework used for distributed storage of huge amounts of data (its HDFS component) and parallel data processing (Hadoop MapReduce). Hadoop often means the ecosystem that includes multiple big data technologies, such as Apache Hive, Apache HBase, Apache Zookeeper and Apache Oozie.
- ***Apache Spark*** is a framework used for in-memory parallel data processing, which makes real-time big data analytics possible. For example, an analytical system can determine that a visitor spends quite a long time on certain product pages, but has not yet added to the cart. To motivate a purchase, the system can offer a discount coupon for that product.

2.2. Open Data

Before giving the technical information, we would like to mention the importance and gains of open data that have economic and social effects.

Since data has enormous potential for communities and economies, Governments are the biggest sources for open data. Open data can encourage economic growth, better decision-making, more transparency and efficiency of governments, higher quality of life. Numerous studies in the world have measured the impact of data for economic, political and social development, and some of them have identified the importance of "open growth" for economic growth. For these reasons, an increasing number of government and organizational data are being opened for use. Open data portals of some countries and organizations have been given below.

- European Data Portal https://www.europeandataportal.eu/en
- USA Open Data, https://www.data.gov/
- UK Open Data, https://www.data.gov.uk
- Germany Open Data, https://www.data.gov.de
- Iran Open Data https://iranopendata.org/en/
- Zimbabwe Open Data https://zimbabwe.opendataforafrica.org/
- World Bank Data, http://datacatalog.worldbank.org/
- USA Public Health Data sets, http://phpartners.org/health_stats.html
- USA National Institute of Health, http://projectreporter.nih.gov/reporter.cfm
- The Development Data Hub, http://data.devinit.org
- UN Data, http://data.un.org/Explorer.aspx

To understand economic and social gains of open data it will be very useful to look the 2020 report of European Data Portal. In this report, the value to be created by open data in the next five years in Europe is estimated (Huyer & van Knippenberg, 2020). In the report, the impact is exemplified by efficiency gains and cost savings due to open data. Since the estimated numerical values given in the report are very dramatic, we would like to share a few of them in this study. Recall that these numbers are for European Union and annual.

- *Approximately 54 – 202 thousand lives can be saved in the EU because in the event of an accident, emergency services can arrive at the scene of an incident 1 minute faster.* Open maps data and open government data help to make the time shorter by the following steps. First, it is easier to find local emergency number even one is on a trip. Second, the identification of the caller's location is optimized when calling an emergency numbers. Third, the routing of an ambulance to the scene of an incident can be quicker with a traffic-data. Saving lives is a very important achievement, but in addition, there is an economic profit in this case. *Approximately €312 – €400 thousand can be saved healthcare costs due to faster first aid by bystanders.*
- *Approximately 27 million hours can potentially be saved annually for train users in Europe.* In an ideal open data-empowered world all the people are aware of the delay before they leave home because they can find that out thanks to an app on their phone and this will reduce the waiting time in stops and stations. This model can be applied to other public transports as well and as a result, open data can enhance the willingness of people to use public transport as a more eco-friendly alternative to a private one. *In addition to saved hours, open data can potentially save €739.8 million due to time saved in public transport.*
- *Open data can aid the development of machine translations resulting in a potential cost saving of €1.1 billion in translations on European Data Portal.*

In addition to these benefits, open data provides material and moral gains in many other areas such as energy savings and increased production (see the report in (Huyer & van Knippenberg, 2020) for more details).

- Both the public and the private sector create the value of open data. *The open data market size could in an optimistic grow scenario to €334.20 billion in 2025 in Europe.* The economic impact of open data can – in addition to its market size – also be calculated in terms of the number of people who are employed due to open data. *In according to the report, now 1.09 million people are employed in the open data market. It is estimated that there will be 1.12 – 1.97 million employees according to 2025 estimates based on the baseline (0.5%) or optimistic (10.4%) growth rates.*

Now let us look at the open data technically. Open data is data that is available and free for you to use. This means that someone or an organization has already taken care of the technical and legal considerations that are needed to allow the data to be freely used, re-used, shared by anyone, anytime and anywhere.

Typically, open data is structured. This means that the data is already organized so that computer software can read it easily. Most open data is in a machine-readable format. There are some common formats including followings:

- Comma separated Value Files or CSV files
- JavaScript Object Notation or JSON,
- Application Programming Interface or APIs
- Shapefiles
- eXtensible Markup Language or XML.

Occasionally, sometimes open data is presented in other formats such as PDFs and HTML pages. These formats make it difficult and sometimes impossible for a computer to read the data. With machine-readable data, computer software can re-purpose, synthesize and model the data to analyze and generate insights.

A lot of people and organizations use open data to develop apps. Open data covers a range of subject areas including the following and much more.

- *Art and culture:* data about collections held by galleries, libraries, archives, and museums.
- *Science:* data generated from scientific research from agriculture to zoology.
- *Finance:* data on financial markets including stocks, shares and bonds…. as well data on government expenditure and Revenue.
- *Statistics:* data produced by government statistical offices including census, labour market surveys, Social Services.
- *Environment:* data generated from the environment including the level of pollutants, levels of precipitation, temperatures.

Making the data accessible to everyone offers easier access to information and content. It enables many parties to analyze the same data and generate new projects and reports using this data. Also in the field of economy, it contributes to the development of innovative services and the creation of new business models.

3. Big Data and Open Data with Deep Learning

With the development of technology, accessing data has become easier today. However, there are some difficulties caused by this convenience, the first of which is that the amount, diversity, and flow rate of the collected data are constantly increasing. Considering that data sharing is increasing day by day, processing and analyzing this constantly growing and changing data in the fastest and most accurate way has become much more important for companies to gain competitive advantage. The processing cycle of big data is described in Figure 2.

Figure 2: Big Data Processing Cycle

After the data is collected and processed, the inference from this data with high added-value is called data analysis. We can group the main topics that data analysis researchers are working on as follows:

- Machine Learning
- Data Mining
- Text Mining
- Predictive Analytics
- Statistical Analysis.

Machine learning is one of the most used methods for data analysis and Artificial Neural Networks is one of the most preferred machine learning models. These models have been used for the solution of various regression problems since the 1980s and give successful results in many real life problems. Deep learning is an advanced model of Artificial Neural Networks. Although it is not a new model, its success occurred when the amount of data increased and technology capable of fast processing appeared and it became one of the best methods in Big / Open Data analysis. Big data obtained from real world problems are capable of feeding deep learning models, and these models are able of tacking real world problems thanks to the fact that these models can reduce big data to meaningful pieces.

Some of the industrial applications of the deep learning model are given below.

-Artificial Intelligence-based Super Computers: *IBM Watson, Microsoft Oxford, Google Deep Mind,*

-Image Classification: *Face Recognition, Environmental recognition in autonomous cars,*

-IMDB Movie Review,

-Emotion Analysis,

-Language translation: *Google Translate,*

-Automatic Speech Recognition with Deep Learning: *Amazon echoDot – Alexa, Apple Siri,*

-Face Recognition with Deep Learning: *DeepFace (Facebook face recognition app),*

-Deep Learning in High Energy Physics: *Data analysis in CERN experiments.*

In addition, Artificial Intelligence has been used in medicine for a long time. Expert Systems used in the 1970s are the first concrete use of artificial intelligence in medicine. These systems are rule-based systems that consist of many "if ... if ..." rules arranged based on knowledge bases compiled from

leading experts of a particular field. The system, which receives patient registration information, laboratory results and symptom queries as an input, can perform diagnosis, prescribing and treatment planning. Until the mid '90s, researches using Expert Systems in medicine have increased gradually.

Other artificial intelligence methods such as Artificial Neural Networks, Fuzzy Logic, and Genetic Algorithms have started to increase their use in the field of medicine since the '90s. The use of Artificial Neural Networks, especially in cancer screening and diagnosis, multivariate data analysis, and time series analysis such as ECG and EEG, have yielded successful results. Unlike Expert Systems, there is no need for a rule base or mathematical model for Artificial Neural Networks but labeled and quality data is required (Bilge (2007)).

Various types of health data, such as electronic health records, medical images, genomic data, sensor data, text, are usually in a complex, heterogeneous, and unstructured format. With traditional data mining and statistical learning approaches, it is almost impossible to draw meaningful conclusions from these complex health data. The development of deep learning technologies is successful in analyzing these complex data (Miotto et al. (2017)).

Early diagnosis of cancerous tissues is very important for the patient but difficult for the pathologist. Therefore, the diagnosis of the early stages is unfortunately not always possible at this limited time (Stumpe (2017)). Recently, it has been seen that it is possible to classify mitotic cells in high-resolution breast tissue histopathological images with a performance of over 96 percent by developing Evolutionary Artificial Neural Network models (see (Albayrak & Bilgin (2016)) and (Cruz-Roa et al. (2017)).

4. Works on Open Data and Big Data in Turkey

Information Technologies are developing rapidly in Turkey as it is all over the world. Informatics is among the priority areas today and the spread of R&D centers and technoparks are encouraged in this area. However, data applications are not at the desired level since open data policies are not widespread in Turkey. With the approval of the Prime Ministry on 23 August 2013, it was aimed that the portals of seffaflik.gov.tr and data.gov.tr will be established and it was aimed to spread it, but it has not been done yet. To achieve this, the following goals and practices are aimed at the 2016-2019 National e-Government Strategy and Action Plan:

E4.2.1: Creating an Open Data Sharing Portal

E4.2.2: Converting and Sharing Public Data into Open Data.

Public awareness was raised on big data and universities started to offer postgraduate education in these fields. The Scientific and Technological Research Council of Turkey (TÜBİTAK) has opened the call for specific projects in this regard, to follow and support the developments in the area. These developments are essential and must be supported, but the publications and projects must not be only aims of Turkey, but also should be to produce technology and benefit.

On the other hand, private companies have understood well that the value can be obtained from the data and they try to improve the quality of the products owned and increase their competitiveness. Technological communication operator companies Turkcell, Vodafone and Türk Telekom are good examples.

The following table indicates the lacks of data analysis in Turkey with solutions (Sağıroğlu, 2019).

Table 3: Problems and Solutions of Data Analysis in Turkey

Step	Problem	Solution
Technology	•lack of data sharing platforms •lack of data storage environments	•creating common standards that can be worked together •making contracts to access data •to develop common platforms for data sharing
Data	•formats that are difficult to use •inconsistent data •different standards •lack of common identifiers •lack of open data and no data to use	•saving data in formats that machines can read •consistent data collection and retention •using common standards •making non-personal data public
Prohibitions	•the Law on the Protection of Personal Data is not well known •lack of data sharing examples	•learn The Law on the Protection of Personal Data well •regulating the law, being aware that we have to share data for research
Organization	•no time and resources for data sharing and collaboration •managers' lack of understanding of data and its strength •lack of data sharing environments	•to create environments where values can be obtained from data •supporting data processing projects not only in official but also in any environment •extending the data-based management and decision making approach

It should be well understood that industry-specific initiatives and collaboration between the private and public sectors encourage value creation. Besides, combining open data with personal, shared, or mass-sourced data is vital to further expand the open data market.

Despite these problems and deficiencies mentioned above, there are some good applications and portals for big data and open data in Turkey. We take a look at them in the following paragraphs.

e-Pulse Personal Health System *(https://enabiz.gov.tr/)*

One of the most successful big data applications in Turkey is e-Pulse Personal Health System. The e-Pulse System was launched in 2015 with the target of 'Ensuring the active participation of individuals in their health-related decisions' stated in the 2013-2017 Strategic Plan of the Ministry of Health. The system is a national personal health record application, where people can access their lab results, medical images, prescription and medication details, emergency information, diagnosis details, reports and health records that contains all the details concerning the examinations via desktop and mobile platforms. In addition to being able to record health data such as step, pulse, calorie, blood pressure, sugar, which are obtained by using wearable technologies and mobile

applications, individuals can also report organ donation and blood donation, and share their position and emergency health data in emergency situations with emergency buttons. There are some awards this application has received worldwide (Beyhan (2017)):

Best Health Application, Health & Well Being, World Summit Awards, 2016.

Best Change Management Project, International Data Corporation (IDC), CIO SUMMIT, 2017.

Istanbul Metropolitan Municipality Open Data Portal (*https://data.ibb.gov.tr/en *)

IMM Open Data Portal, launched by Istanbul Metropolitan Municipality, has been opened recently. The Open Data Portal, which presents data published by the IMM and its environmental organizations, aims to increase public transparency and citizen participation. Data sets collected on the portal are shared in many categories for use by academics, researchers, entrepreneurs and citizens. The portal, where open data sets collected in the areas of Mobility, Life, Environment, Energy, People, Disaster Management and Security are shared for now, is aimed to be enriched with a transparent management approach. In addition, individuals can request the data they want to see on the portal from the municipality. The purpose of the portal is to increase the quality of city services, performance and interaction with citizens, to reduce public services costs and resource consumption, and to strengthen communication between citizens and management.

***Official Statistics Portal** (http://www.resmiistatistik.gov.tr)*

Official Statistics Portal was opened in order to determine the basic principles and standards regarding the production and publishing of official statistics and to provide up-to-date, reliable, transparent and impartial data in the fields needed at the national and international levels. The portal opened to the service of users in March 2014 provides open access to all official statistics.

***Open Data Portal of the Ministry of Agriculture and Forestry** (http://veri.tarimorman.gov.tr)*

It is a portal prepared to provide fast and easy access to the data managed and produced by the Ministry. Data sets are divided into *borders., land and cover, protected areas and water.*

***TÜBİTAK Open Archive Aperta (**https://acikveri.ulakbim.gov.tr)*

TÜBİTAK Open Archive is created by The Scientific and Technological Research Council of Turkey (TÜBİTAK) to store research data and publications produced from projects carried out or supported by the institution, thereby providing open access to these data in accordance with "TÜBİTAK Open Science Policy".

Presidency Digital Transformation Office (https://cbddo.gov.tr)

Lastly, it can be added that in line with developing technologies, social demands and reform trends in the public sector, Presidency Digital Transformation Office was established in order to gather works related to digital transformation (e-Government), cyber security, national technologies, big data and artificial intelligence, which are carried out separately under different institutions. No open data has been shared so far, but it is seen that efforts to create an open data portal have been carried out.

5. Conclusion and Discussion

The importance of open data is becoming more and more noticeable by various governments, administrations and companies. Open data portals are created, investments are made, and postgraduate educations are increased by companies and governments all over the world.

Turkey has also cared about the data subject, and attempts have been made on this issue in Turkey. Unfortunately, neither the training nor the portals reached the desired level. It should be underlined that the data is used very effectively to create products with high added-value and to increase the understanding of the transparent state. For these reasons, more projects related to data science should be created, data portals should be opened, personal data protection law should be better understood, and data sharing should be encouraged.

Author Contribution

Authors made literature review together. Ezgi Erdoğan has contributed the translation of Turkish reports and articles, besides web pages of open data portals in Turkey. The authors made equal contributions to the technical part. Ezgi Erdoğan has a contribution to the Conclusion and Discussion section.

Conflict of Interest

The authors declared no potential conflicts of interest.

Financial Support

The authors have not received any financial support for this study.

References

Albayrak, A., & Bilgin, G. (2016). Mitosis detection using convolutional neural network based

features. *In Computational Intelligence and Informatics (CINTI), 2016 IEEE 17th International Symposium (pp. 335-340)*

Beyhan O., Büyük Veri Çalışmaları, Sağlık Bilgi Sistemleri Genel Müdürlüğü, Türkiye Bilişim Derneği, Kamu Bilişim Platformu-19, 26-29 October 2017, Retrieved June 8, 2020. Retrieved June 8, 2020. http://www.kamu-bib.org.tr/kamubib-19/sunumlar/ozan-beyhan-kamu-bib19.pdf.

Bilge, U. (2007). Tıpta Yapay Zeka ve Uzman Sistemler, *IV. National Medical Informatics Congress, pp. 113-118*

Cruz-Roa, A., Gilmore, H., Basavanhally, A., Feldman, M., Ganesan, S., Shih, N. N. & Madabhushi, A. (2017). Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. *Scientific Reports, 7, 46450*

Huyer, E., & van Knippenberg, L., (2020) The economic impact of open data: Opportunities for value creation in Europe, *Capgemini Invent, European, Data Portal, 26.01.2020,* doi: 10.2830/63132.

Miotto, R., Wang, F., Wang, S., Jiang, X., Dudley, J. T., (2017). Deep learning for healthcare: review, opportunities and challenges, Briefings in Bioinformatics, pp. 1-11.

Özkan Ö. (Eds) (2019). *Open data,* law, regulation and public relations study group report, blockchain platform of Turkey, August 2019, Retrieved June 8, 2020. https://Bctr.Org/Dokumanlar/Acik_Veri.Pdf.

Sağıroğlu Ş. (2019) Büyük ve Açık Veri Türkiye Uygulamaları, Big Data Conference, 26 June 2019.

Stumpe, M., (2017), Assisting pathologists in detecting cancer with deep learning. Google research group, Retrieved June 8, 2020. https://research.googleblog.com/2017/03/ assistingpathologists-in-detecting.html.

UrszulaEC1307560, (2018). 2. Historical development of big data [Blog post], Retrieved June 8, 2020. http://ec1307560.blogspot.com/2018/10/2-historical-development-of-big-data.html.

Resume

Ezgi Erdoğan is a Research Assistant of Mathematics at the Faculty of Arts and Sciences, Marmara University. She has a Ph.D. in Mathematics from Yıldız Technical University, Istanbul. Her main work topic is functional analysis, a discipline in which she has published several articles in impact journals. Pure mathematics as her main research interest allows her to collaborate in the most theoretical part of multidisciplinary activities, such as models for the construction of control indices in economics and information sciences, and in Big Data.

Juan S. Bejarano is a PhD student in Communication and Cultural Industries at the Polytechnic University of Valencia, Spain. He has a Master's degree in Information Management from the Universidad Politécnica de Valencia. He is a computer engineer from the Universidad San Buenaventura in Bogota, Colombia. His research focuses on data transparency in the public sector, taking as a reference new data processing systems such as blockchain.