Mugla Journal of Science and Technology

OPEN DATA AND BIG DATA MODEL FOR HIGHER EDUCATION AND POSSIBLE OUTCOMES

Sümeyye KAYNAK*, Department of Computer Engineering, Sakarya University, Türkiye, sumeyye@sakarya.edu.tr
(https://orcid.org/0000-0002-7500-4001)
Baran KAYNAK, Information Systems Engineering, Sakarya University, Türkiye, kaynak@sakarya.edu.tr
(https://orcid.org/0000-0002-9004-2639)
Ahmet ÖZMEN, Software Engineering, Sakarya University, Türkiye, ozmen@sakarya.edu.tr
(https://orcid.org/0000-0003-2267-2206)

Received: 09.11.2022, Accepted: 22.03.2023 Research Article
*Corresponding author
DOI: 10.22531/muglajsci.1201726

Abstract

The basic outputs of universities can be listed as education, research-development and service to society. Managerial software systems at universities generate large amount of open data during daily operations. The data generated by these systems contain valuable public institutional performance information along with critical private information. These public data can be classified, collected and processed by using big data approaches for performance monitoring. In this study, an open data platform is modelled, and issues are discussed related how open data is collected, stored and processed using big data approaches to extract interested performance information. It is shown that institutional performance information can be presented according to a wide variety of metrics from the collected data. Scientific studies that can be carried out in higher education using big data are examined under 4 headings: Creating an open data directive for universities, development of open data platform, institutional accreditation service, creating a digital twin. This platform can be used for online institutional evaluation either by university management or accreditation agencies.
Keywords: Open data, Big data, Higher education

YÜKSEK ÖĞRENİMDE AÇIK VERİ VE BÜYÜK VERİ MODELİ VE OLASI SONUÇLARI

Özet

Bir üniversitenin temel çıktıları eğitim, araştırma-geliştirme ve topluma hizmet olarak listelenebilir. Üniversitelerdeki yönetim yazılım sistemleri günlük işlemler sırasında büyük miktarda açık veri üretir. Bu sistemler tarafından üretilen veriler, kritik özel bilgilerle birlikte değerli kamu kurumsal performans bilgilerini içerir. Bu kamuya açık veriler, performans izleme için büyük veri yaklaşımları kullanılarak sınıflandırılabilir, toplanabilir ve işlenebilir. Bu çalışmada, bir açık veri platformu modellenmiş ve ilgili performans bilgilerinin çıkarılması için büyük veri yaklaşımları kullanılarak açık verilerin nasıl toplandığı, saklandığı ve işlendiği ile ilgili konular tartışılmıştır. Toplanan verilerden kurumsal performans bilgilerinin çok çeşitli ölçütlere göre sunulabileceği gösterilmiştir. Yükseköğretimde büyük veri kullanılarak yapılabilecek bilimsel çalışmalar 4 başlık altında incelenmektedir: Üniversiteler için açık veri yönergesi oluşturulması, açık veri platformunun geliştirilmesi, kurumsal akreditasyon hizmeti, dijital ikiz oluşturulması. Bu platform, gerek üniversite yönetimi gerekse akreditasyon kurumları tarafından çevrimiçi kurumsal değerlendirme için kullanılabilir.
Anahtar Kelimeler: Açık veri, Büyük veri, Yüksek öğretim

Cite

Kaynak, S., Kaynak B., Özmen A., (2023). "Open Data and Big Data Model for Higher Education and Possible Outcomes", Mugla Journal of Science and Technology, 9(1), 7-15.

1. Introduction

A university ecosystem consists of units (institute, faculty, department, rectorate, dean etc.), entities (students, instructors, academicians, administrative staff etc.), collaborations (public and private companies etc.), internal and external relations. These components are independent and interrelated with each other. In the university ecosystem, data is produced at different speed, structures, and diversity. The development and dissemination of information and communication technologies has provided a great increase in the data volume, diversity, and production speed of educational institutions. In traditional information systems, the data is collected, processed, and stored by monolithic or micro-service architecture systems. In open data platform, however, the volume of data and the speed of data arrival is unknown by nature, hence different approaches are necessary to handle such as distributed and fault tolerant components must be employed.

Processing data of different structures and types taken from different data sources will positively increase the quality of useful information that can be extracted from the data. Combining the data/information of the education ecosystem with the up-to-date data/information from other ecosystems associated with this ecosystem can increase the scope of the information obtained, the accuracy of the prediction, and enable the discovery of new patterns that were not known to exist before. Big data technologies are used for these purposes.

The idea of sharing non-personal data produced from experiments made during scientific studies or produced by corporate systems without any preconditions for consumption is called "open data". Sharing of data increases transparency. Easy access to data enables further research and accelerates the emergence of new discoveries and scientific publications. In the last decade, open data initiatives around the world have gained momentum. It has become a part of the primary digital policies of countries [1]. Some countries have prepared open data action plans at various levels [2], [3], [4]. Türkiye has an e-government strategy and action plans in this context, even though it is at the beginning of the adoption of the open data concept and the transformation of this philosophy into action [5].

Open Data comes mostly from government, education, and demographic and social data sets. Private and public institutions, enterprises, and individuals also feed the open data [6]. Schools, universities, and other educational institutions contain high volumes of data in different varieties and types. Making this dataset produced by universities, schools, and other educational institutions available to everyone makes a great contribution to research with some problems. This study focuses on universities. What are the open data sources in universities is an issue that needs to be investigated? Determining open data sources with a scientific study according to the type of institution and transforming them into a directive will minimize the data security problem, while the opportunities it will offer to researchers who benefit from open data will increase.

In this study:

- An example open data platform is proposed for universities.
- The architectural structure of the open data platform is explained in detail.
- This platform can be used by both university management and accreditation agencies.
- Possible open data sources that are part of the platform are presented.
- Possible solutions for collecting data from open data sources and processing them with big data methods are presented.

2. Background

2.1. Big Data

Big data represents data that has the characteristics of diversity, volume, speed, accuracy and value. The volume property of big data refers to the amount of data available. The speed property refers to how quickly data is generated and moved. The diversity property refers to the variety of data types. Accuracy refers to the precision and correctness of the data. The value feature, on the other hand, refers to the added value that big data can provide. In order to obtain information, such data must be stored, processed and analyzed. Collecting, storing and processing data of different types and volumes from different sources cannot be done with traditional information processing and storage technologies. At this point, big data technologies come into play.

Educational institutions contain many unstructured data. Traditional databases are based on a fixed schema and can only work with structured data. NoSQL, The Apache Hadoop Distributed File System (HDFS), Data Lakes, Blockchain, RainStar, in-memory database, cloud technologies are technologies used for storing big data. Recently, it is seen that cloud and blockchain technologies with a distributed architecture are frequently preferred in the storage of data of educational institutions [7], [8], [9].

In order to benefit from the advantages of big data in educational institutions, the collected and stored big data must be analyzed. With big data analysis, more in-depth analyzes can be made than traditional methods. In traditional analytics, structure data in megabytes and gigabytes are analyzed. In big data analytics, larger volumes (quintillion bytes etc.) of structural/semi-structural/non-structural data can be analyzed. While answers to previously defined questions are sought in traditional analytics, there is no clear predefined question in big data analytics, the aim is to reveal unexpected new findings and facts [10]. Ensemble analysis, association analysis, high-dimensional analysis, deep analysis, precision analysis, divide-and-conquer analysis are widely used big data analytics techniques [11].

2.2. Open Data

Open data is data that can be openly accessed, used, and shared by anyone for any purpose, even for commercial purposes, without being bound by patents or other control mechanisms. The Open Definition project, a Knowledge International project, presented the key features of openness in 3 items [12]:

- Availability and access: The data must be available as a whole and downloadable at no more than the cost of one-time replication. Data must also be available in appropriate and modifiable form.
- Reuse and redistribution: Data must be provided under conditions that allow reuse and redistribution. Data must be machine readable.

- Universal participation: Everyone should be able to use, reuse and redistribute. (There should be no discrimination against workplaces, groups or individuals. Data can also be used for commercial purposes).

Neelie Kroes, vice-president of the open data European commission, in the opening speech of the EU's "Open Data strategy" stated that data is the gold of the digital age [13]. By making data open; transparency, democratic control, self-empowerment, innovation, efficiency increase [14].

3. Open Data Sources at Universities

Countries provide support for the development of data-related technologies, the development of education programs on data literacy and open data. Nowadays there is increasing pressure to open data generated public institutes/organization and scientific system [15].

Open data is a free resource for everyone. Organizations use open data mostly for optimization purposes. The next most common use of open data is research and development. The research and development goal are followed by the development of products or services and advocacy goals. The education sector contains all these goals. There has been some progress in making open and using training data that is relevant, valid, reliable, timeless, punctual, transparent, comparable, accessible, cost-effective, consistent and with potential for disaggregation, but major gaps remain [16].

There is little research to suggest how and why open data should be used in education. Atenas et al. prepared a survey on the use of open data in education. It was seen that 26 people who participated in this survey did not actually use open data [17]. Determining open data sources that are free of personal data and comply with the open data principle for universities is an issue that needs to be studied. Atenas and Havemann has examined the intersection of open data and education sector at 3 points: Education Landscape, education indicators and educational resources [16]. Education landscape represents the use of open data by policy makers and organizers of action plans. Education indicators represents open data for students and parents. Educational resources represent open data for use by educators. In the study [16], the outcomes of open data in education are evaluated in terms of teacher, students, parents, and administrators.

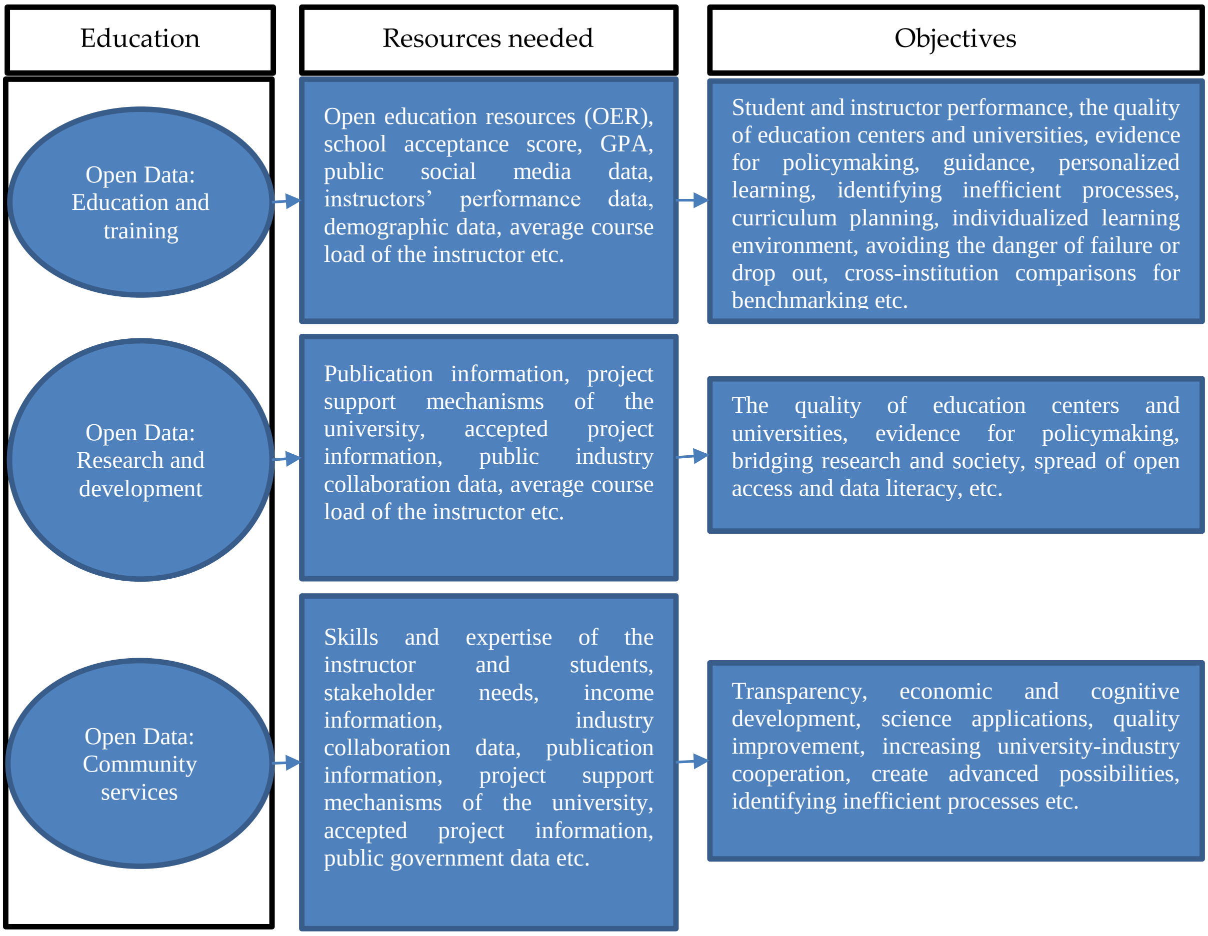

Figure 1. Open data sources and objectives in higher education [18], [19], [20], [21], [6],[22].

In this study, in order to make big data analytics easier, the intersection of open data with education is presented with a different conceptual model. As seen in Fig. 1., the outputs of a university can be classified as education-training, research-development and service to society. In this study, the intersection of open data and higher education is examined according to university outputs. With this conceptual framework, the ways of using open data in education are explained.

At the education-training, open education resources (OER), student and instructor academic performance, educational indicators are examined. The intersection of the outputs of education and training processes with open data affects 3 main education stakeholder groups (policymakers, parents and students and instructors). At the research-development, the academic aspect of university is examined using open data. The academy develops with the accumulation and sharing of knowledge. The concept of open access means that scientific studies can be accessed, read, saved, copied, printed, scanned, indexed completely free of charge, without financial, legal and technical barriers, by targeting the sharing and accumulation of knowledge. The concept of "open access" has been on the agenda of the world intensively for the last 10 years. Open access is supported by many scientists, publishers, and researchers. The intersection of the outputs of research and development processes with open data directly affects policymakers, instructors, and society.

The widespread use of open data in higher education increases transparency. With big data analytics, processes, human resources can be improved, this provides economic and cognitive development. This improvement and development bring well-trained individuals to the society. With open data, a university offers parents, students, external institutions a preview of the university. Open data can help parents and students choose the most suitable school for them. Also, students can discover the educational programs that suit them best.

With the spread of open data in higher education and institutions, a balance of labor demand and supply can be established. It increases the cooperation between university and industry. Graduate students (human resources) can be offered more productive jobs suitable for their abilities.

There are at least 2 important points to consider when trying to use education data in higher schools [16]. Firstly, to protect the privacy of educators, students and university stakeholders in the use and sharing of personal data. Later, protecting and increasing the capacity to create and use data.

3. Proposed Approach

3.1. Open Data Platform

In this study, a university level open data platform has been proposed as a repository area so that big data methods can be applied on the data collected from university sources. Fig. 2 shows the model and internal components such as faculties, research centers and administrative units which run under a university management software system.

Personal information should be filtered out, and a copy of the remaining part can be moved to a public platform. The data owner is responsible for extracting personal information. Data collection should not include personal information. This system only accepts data in a specific format. It is suggested to refer to a directive prepared together with the lawyers for the separation of personal data from open data. Approval of this directive by the Council of Higher Education will make its operation legal by all universities. Agent software to be added to the information systems in universities will be implemented in accordance with this directive. In accordance with the directive from the university information systems, personal data will be filtered by the agent software and sent to the open data platform.

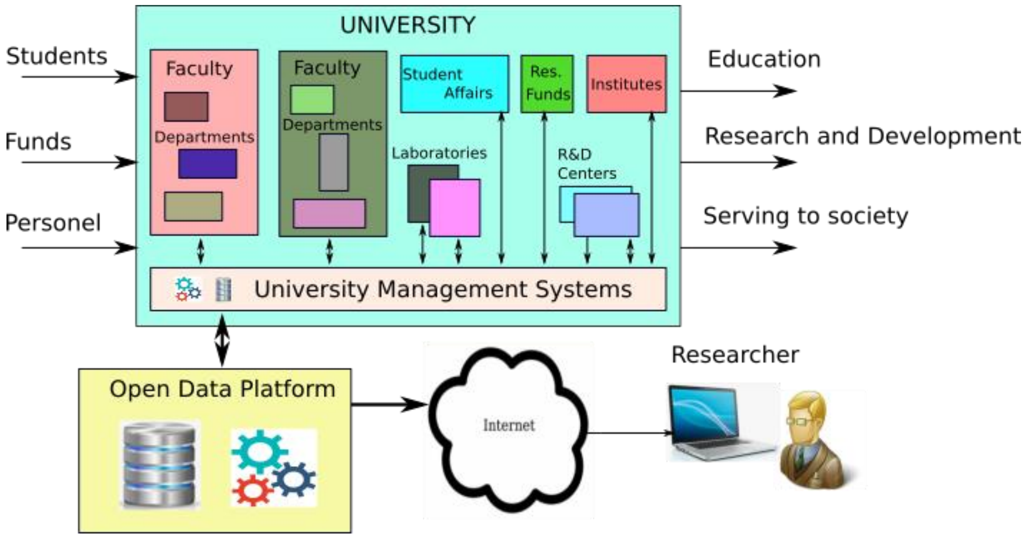

Figure 2. Block diagram showing the inputs and outputs of a university and the general structure of the open data platform.

In this context, determination of the open data sources of universities is a problem that needs to be investigated. Open data can be considered for any business or institution. The scope of this study is limited to universities. If a comprehensive study is effective, it will also set a model for other business lines or institutions.

Independent identification and irregular presentation of open data sources threatens data integrity and negatively affects the trust to the data. In this respect, open data sources should be determined by scientific research and how data will be collected and labeled from these sources should be transformed into a legal directive. The directive should also define the systematics to ensure that the data is stored uniquely. On the other hand, open data sources should have characteristics of being findable, accessible, interoperable, and reusable (FAIR) [23]. Different data quality standards such as "FAIR", "COLD", "W3C", "DCAT" exist. The FAIR data quality standard focuses on the use and applicability of the data in real life, while other data quality standards generally focus on issues such as accessibility, transparency, or use of linked data. The FAIR data quality standard has been adopted based on its characteristics and focus.

The general architecture of the open data platform is shown in Fig. 3. The agency software that collects data from the information systems of the universities and the gate software that meets the data coming to the open data platform (gateway) constitute the data collection part of the system. The communication between the agency software and the gateway is thought to be in the REST (Representational State Transfer) architecture [24], [25]. Gateway has a multi-channel structure and will convert the incoming data into a format suitable for the queue system and send it immediately. The queue system is designed to be in a distributed architecture and a scalable structure that can easily expand with the increase in data resources. There are different applications (such as LinkedIn, WhatsApp, Twitter, Facebook) in the literature and in the field of application. There are alternative solutions such as open source RabbitMQ and Apache-Kafka for the queuing system that will form the backbone of the platform, and comparative preliminary studies on this subject are ongoing [26], [27], [28].

The data coming to the queue system is saved into the appropriate data stores. The software that manages the queuing system decides whether to keep the incoming data in the associated database (SQL), document-based database (NoSQL) or cache (such as Redis) by looking at the formats and lifetimes of the incoming data. At the output part of the platform there are interfaces that can present data in various formats to researchers. For example, REST-API can be used for machine-based receivers. Users and researchers can fetch data with a web browser or download data as a file in CSV format. In some cases, special services may be performed with the micro-service units. For example, institutional accreditation reports can be provided by a special micro-service. The monitoring module contains the log services of the open data platform. Through this interface, authorized users and platform administrators monitor data traffic and see the usage statistics of the platform.

Institutional performance can be measured by establishing a relationship between the inputs and outputs of universities. In this respect, the data that represents both input and output sizes should be determined. For example, university outcomes can be classified as education, research and development, and community service. Inputs can be determined as students placed in departments with entrance exams, income from various channels and employed personnel. Collecting and processing such a large volume of data requires big data approaches. Performing big data operations over collected open data reveals important online performance information about the universities. This information is important for university management to identify improvement processes.

3.2. Data Collection

Since living systems constantly generate data, storage and processing of collected open data creates problems. For this purpose, a copy of the corporate open data can be served from the platform. Hosting of many universities' data on the same platform becomes a big data problem. However, efficiency will increase as information resources will be shared, and costs will decrease for universities.

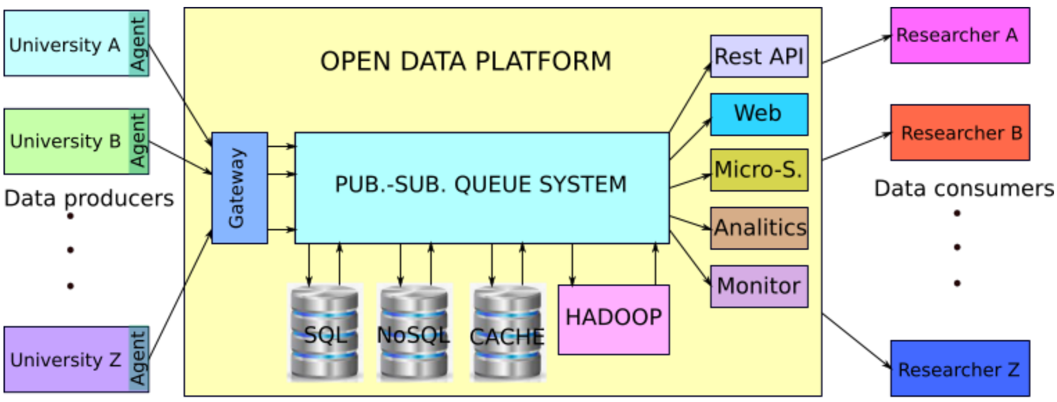

Figure 3. Open Data Platform architecture block diagram

The platform becomes a meeting point for researchers who are allowed downloading data freely with various interfaces. In this context, the open data platform can be built on Tübitak-Ulakbim IT resources with cloud technologies.

While filtering and collecting relevant data from university management systems, the data logging strategy may be slowing down the systems by increasing the load over the systems. The data collection policy has to be selected carefully, otherwise the data collection mechanisms to be added to university management systems can affect the behavior of the system and cause it to slow down significantly. A careless design in the data collection software can even cause a change in the size to be measured. For example, while the system is able to enroll 1000 students per day during student enrollment, this number should not drop to 800 after open data collection is activated. In this regard, open data collection from institutions should be done in a way that causes minimal disturbance to the institution.

Data collection policies can be event-driven or time-driven. Event-driven policy creates a data record as soon as an interested event occurs. For example, when a new student registered an event record is generated. On the other hand, time-driven policy works with timers, a new record is created periodically when the timer expires. In this policy, the sampling period can be either fixed or dynamic. In this approach, data can be collected daily bases where the system does not operate to reduce the intrusion. For example, logging the registered student count daily can be implemented in this method by setting up the timer period 24 hours. In this approach, some information loss occurs on behalf of data reduction. For example, the student count changes during the daytime will be lost in this approach.

The time complexity of the 24-hours period time-driven approach will be $O(k)$ for k universities. The time complexity of the event-driven approach will be $O(kn)$for k universities and n events. Investigations showed that the time-driven approach with a dynamic period of 24 hours as default creates a better solution for the proposed system. In case more precision is desired, the sampling period can be reduced to obtain frequent sampling. This approach results in; 1) easy to implementation, 2) less data generation, and 3) variable sampling period to improve precision when needed.

3.3. Data Labeling and Uniqueness

Collected data must be stored in databases for later retrievals. Hence, each data must be classified and labeled uniquely with a metadata, and then stored on the platform for a certain period of time. The data here can be just a positive integer number such as student count or a document such as a report. Variety in data types creates a necessity for employing different database systems in the platform such as relational or non-relational database systems. The meta data must be systematic and self-describing, so that the researchers easily retrieve the data they want. The structure should help the users to easily find the same type of data of a different university be able to conduct comparative studies between institutions. Fig. 4 shows an example labeling structure for this study.

Fig. 4 shows an example of a partially formed tree structure for the student numbers of a university. The design is self-descriptive and scalable. In this design, the tree representation creates the labels for the values. N-ary tree structure (i.e., hierarchical representation) is selected to make the data findable for humans, and the depth is fixed to four to make it faster for queries (for computers). New classes can be created in the same level by selecting the most adjacent class. For example, a new education model (i.e., weekend education) can be created just by adding a new leaf in the same level of 1. session, 2. session and distance education. In this way the labeling format will be preserved, and this requires no changes in the system. Also, it has easy to learn structure: For example, a query string can easily be obtained by visiting the tree nodes for "What is the number of normal education students at University A at this date?". Similarly, it is possible to analyze the data by looking at the label (metadata). The metadata format is fixed for open data stored in the system, and the sample structure of the format is shown in Fig. 5. Since this data format is consistent for all universities, access to specific data is very easy.

3.4. Data Processing

It can benefit from digital twin (DT) technology during the processing of data. Interest in digital twin is growing exponentially and is a promising technology. The concept of DT is confused with many different terminologies. DT is not a computer model or simulation. These are just a part of DT. DT is also different from the concepts of Digital Model, Digital Shadow, Product Avatar, Digital Thread. A general definition of DT, including its features that differ from different terminologies, was made by Singh et al. as follows [29]:

“A Digital Twin is a dynamic and self-evolving digital/virtual model or simulation of a real-life subject or object (part, machine, process, human, etc.) representing the exact state of its physical twin at any given point of time via exchanging the real-time data as well as keeping the historical data. It is not just the Digital Twin which mimics its physical twin but any changes in the Digital Twin are mimicked by the physical twin too.”

Digital twin technology was initially used in the aerospace industry, then its use began to become widespread in other industries [30]. Design/planning, optimization, maintenance, security, training, validation, analysis are examples of different applications of digital twin.

In this study, the creation of a university's digital twin from the collected data was examined. It is possible to create a digital replica of a university from the data collected on the open data platform. It is possible to access details about the output of interest of the system on the digital twin, compare it with data from previous years and reveal weak performing points.

The digital twin can also be used for future predictions of university outcomes, rather than analysis of snapshots. For example, by evaluating the graduate information, an estimate can be made for the department quotas of the next years. Answers can be sought to the questions of what kind of cooperation should be made in order to increase the R&D outputs of the university and to make more publications.

3.5. Data Retrieval

Either raw data or processed data can be queried through the platform. Raw data can be queried either by web API or downloaded in the form of Excel format. Beyond this, a user may want to download a performance report generated by the system. A well-defined report can be generated by a micro-service which is placed on the platform to meet the demands in a fast and scalable manner. For example, most of the information required for institutional evaluation can be generated by a specifically designed micro-service for all universities in the platform.

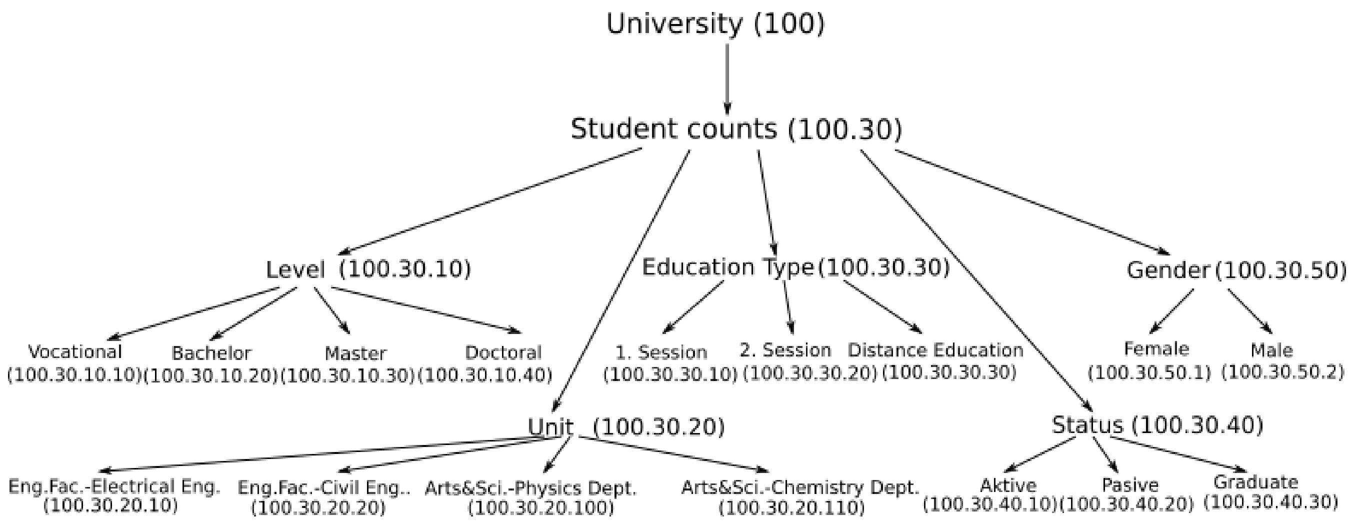

Figure 4. Labeling structure for the data collected from universities. The student numbers of a university are shown in a 4-level hierarchy as an example

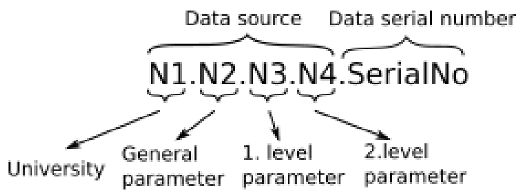

Figure 5. Metadata (label) format that identifies the data in the platform uniquely.

4. Results and Discussion

In this preliminary research study, it has been understood that it will be possible to produce useful information from open data collected from the systems of universities by using big data methods. In this context, the following scientific studies should be carried out:

1. Creating an Open Data Directive for Universities:

Determining the data and data sources to be collected in the scope of open data for universities is the first step of big data studies. Anonymous data sources should be identified and labeled with metadata to make comparisons between universities.

2. Development of Open Data Platform:

Storage, management and processing of collected open data is a challenging problem for every individual university. For this purpose, copying the corporate open data to a stable platform emerges as the most correct solution. The platform can host many universities open data which increases the efficiency as the computing infrastructures are shared. The platform becomes a meeting point for the researchers who seek open data to study on. Open data platform can be created with cloud technologies on Tübitak-Ulakbim IT resources.

3. Institutional accreditation service:

New services can be provided with micro-service software architecture for different purposes on the open data platform. For example, a university's institutional evaluation service can be developed with a micro service architecture. Expanding the institutional evaluation or accreditation service to other universities on the platform can be facilitated by the advantages offered by the architectural approach.

In a significant part of the accreditation processes, performance data is analyzed within the framework of previously determined metrics. In the next phase, the evidence is examined through face-to-face interviews and the results of the analysis are verified. In this context, the accuracy and reliability of the data presented is extremely important. Unreliable or inconsistent data mislead its users and cause harm. Hence, the data sources must be determined by a third party and the systems that collect data must be approved (certified) by the authorized bodies. It is important to conduct an exemplary study on this subject before institutional open data portals become widespread in our country.

4. Creating a digital twin:

The concept of digital twin refers to the creation of a digital copy with the data collected from a physical system [31]. With the appropriate structuring of open data, it is possible to create a digital twin of an organization in various perspectives. Through the digital twin, the processes of large and complex institutions such as universities can be monitored online and unexpected outputs can be detected in a short time. On the other hand, future predictions can be made for interesting perspectives of the institution using modern data processing algorithms such as artificial intelligence on the digital twin.

5. Conclusion

In this study, the determination of open data sources in universities, the creation of an open data platform and the benefits that can be obtained from these data are discussed. Not only researchers, but also institutions that make their data available to the public benefit greatly from this practice. Because the results of the scientific research on the shared data returns valuable outputs for possible improvements. Therefore, open data is becoming widespread all over the world, especially in government institutions created with public funds. It is estimated that many institutions in our country will open their anonymous data to share with public in the near future. It is seen that municipalities are in pioneering studies in this regard.

As such, the article analyzes the issues of how to collect, store and evaluate open data for universities, beyond producing a solution to a specific sub-problem. Analysis results will primarily be of interest to researchers working on open data or big data.

6. References

[1] Thorsby, J., Stowers, G.N.L., Wolslegel, K., and Tumbuan, E., "Understanding the content and features of open data portals in American cities", Government Information Quarterly, vol. 34, no. 1, pp. 53-61, 2017.

[2] Muenchen.de “Willkommen beim Open-Data-Portal München”, *muenchen.de*, Online. Available: https://opendata.muenchen.de/, Accessed: 05.10.2022.

[3] NYC.gov, “Open Data for All New Yorkers”, *NYC Open Data*, Online. Available: https://opendata.cityofnewyork.us/, Accessed: 05.10.2022.

[4] data.europa.eu, “Open Data Best Practices in Europe”, Online.Available: https://data.europa.eu/en/datastories/open-data-best-practices-europe, Accessed: 02.09.2022.

[5] Sağıroğlu, Ş., Bensghir, T. and et. al. (Sağıroğlu, Ş., and Koç, O.), “Büyük Veri ve Açık Veri Analitiği: Yöntemler ve Uygulamalar”, Grafiker Yayınları, Ankara, 2017.

[6] Drigas, A. S. and Leliopoulos, P., "The Use of Big Data in Education", IJCSI International Journal of Computer Science Issues, vol. 11, no. 5, 2014.

[7] Bughin, J., Chui, M., and Manyika, J., "Clouds, big data, and smart assets: Ten tech-enabled business trends to watch", McKinsey Quarterly, vol. 56, pp. 75-86, 2010.

[8] Kataev, M. and Bulysheva, L., "Blockchain system in the higher education: Storing academical students' records and achievements accumulated in the educational process", Systems Research and Behavioral Science, vol. 39, no. 3, pp. 589-596, 2022.

[9] Park, J., "Promises and challenges of Blockchain in education", Smart Learning Environments, vol. 8, no. 1, pp. 33, 2021.

[10] Lukić, J., "The Impact of Big Data Technologies on Competitive Advantage of Companies", Facta Universitatis Series: Economics and Organization, vol. 14, no. 3, pp. 255-264, 2017.

[11] Shu, H., "Big data analytics: six techniques", Geo-spatial Information Science, 19(2), 119-128, 2016.

[12] Open Foundation Knowledge, “What is open?” Available: https://okfn.org/opendata/, Accessed: 05.10.2022.

[13] Gürdal, G., "Açık veri–açık erişim: Dünya ve Türkiye", Dünya ve Türkiye Bilişim Dergisi, pp. 104-109, 2014.

[14] Akdamar E., “Akıllı Kent İdealine Ulaşmada Açık Verinin Rolü,” Sosyal Bilimler Araştırma Dergisi, vol. 6, no. 1, pp. 45–52, 2017.

[15] Raffaghelli, J. E., "Open Data for Learning: A case study in Higher Education", Exploring the Micro, Meso and Macro Proceedings of the European Distance and E-Learning Network 2018 Annual Conference, 2018.

[16] Atenas, J. and Havemann, L., "Open Data Sectors and Communities: Education", in The State of Open Data: Histories and Horizons, edited by Davies, T., Walker, S.B., Rubinstein, M., and Perini, F., African Minds, South Africa, 2019.

[17] Atenas, J. and Havemann, L., "Open Data as Open Educational Resources: Towards Transversal Skills and Global Citizenship", Open Praxis, Vol. 7, pp. 377-389, 2015.

[18] Coughlan, T., "The use of open data as a material for learning", Educational Technology Research and Development, vol. 68, no. 1, pp. 383-411, 2020.

[19] Toptaş, O. and Şen, A., "Eğitimde Ölçme Değerlendirmede Büyük Verinin Önemi," Düşünce ve Toplum Sosyal Bilimler Dergisi, no. 4, pp. 223–243, 2021.

[20] West D. M., “Big Data for Education Data Mining, Data Analytics and Web Dashboards”, 2012.

[21] Siemens, G., and Gasevic, D., "Guest editorial - learning and knowledge analytics," Educational Technology and Society, vol. 15, no. 3, pp. 1-2, 2012.

[22] Schalkwyk F. van, “Open data on universities – New fuel for transformation”, *University world news*, Online. Available: https://www.universityworldnews.com/post.php?story=20170710104034491, Accessed: 05.10.2022.

[23] Dumontier M. and Wesley K., "Advancing Discovery Science with FAIR Data Stewardship: Findable, Accessible, Interoperable, Reusable", The Serials Librarian, vol. 74, no. 1-4, pp. 39-48, 2018.
[24] F. Haupt, F. Leymann, A. Scherer and K. Vukojevic-Haupt, "A Framework for the Structural Analysis of REST APIs," 2017 IEEE International Conference on Software Architecture (ICSA), Gothenburg, Sweden, 2017, pp. 55-58.
[25] Halili F. and Ramadani E., "Web Services: A Comparison of Soap and Rest Services," Modern Applied Science, vol. 12, no. 3, pp. 175-182, 2018.
[26] Sharvari, T., and Sowmya Nag, K., "A study on Modern Messaging Systems- Kafka, RabbitMQ and NATS Streaming", 2019.
[27] Fleischer H., "Applying event-driven architecture in the jvalue ods," *Master Thesis*, 2020.
[28] Hiraman, B. R., Viresh M. C., and Abhijeet K. C., "A Study of Apache Kafka in Big Data Stream Processing," 2018 International Conference on Information, Communication, Engineering and Technology (ICICET), Pune, India, 2018, pp. 1-3.
[29] Singh, M., Fuenmayor, E., Hinchy, E. P., Qiao, Y., Murray, N., and Devine, D., "Digital Twin: Origin to Future", *Applied System Innovation*, vol. 4, no. 2, pp. 1–20, 2021.
[30] Gabrielli, S., Rizzi, S., Mayora, O., More, S., Pérez Baun, J.C., and Vandevelde, W., "Multidimensional Study on Users' Evaluation of the KRAKEN Personal Data Sharing Platform", Applied Sciences, vol. 12, no. 7, pp. 3270, 2022.
[31] Fuller, A., Fan, Z., Day, C., and Barlow, C., "Digital Twin: Enabling Technologies, Challenges and Open Research", IEEE Access, vol. 8, pp. 108952-108971, 2020.