

**REPUBLIC OF TURKEY
ANKARA UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
INTELLECTUAL PROPERTY RIGHTS, TECHNOLOGY POLICIES
AND INNOVATION MANAGEMENT**

BIG DATA AND INTELLECTUAL PROPERTY RIGHTS

Master Thesis Statement

RUMEYSA CANAN ULUÇAY

Ankara, 2022

**REPUBLIC OF TURKEY
ANKARA UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
INTELLECTUAL PROPERTY RIGHTS, TECHNOLOGY POLICIES
AND INNOVATION MANAGEMENT**

BIG DATA AND INTELLECTUAL PROPERTY RIGHTS

Master Thesis Statement

RUMEYSA CANAN ULUÇAY

Assist. Prof. Dr. Zehra ÖZKAN ÜNER

Ankara, 2022

**REPUBLIC OF TURKEY
ANKARA UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
INTELLECTUAL PROPERTY RIGHTS, TECHNOLOGY POLICIES
AND INNOVATION MANAGEMENT**

BIG DATA AND INTELLECTUAL PROPERTY RIGHTS

Supervisor:

Dr. Öğr. Üyesi Zehra ÖZKAN ÜNER

THESIS JURY MEMBERS

- 1-Doç.Dr. Hüseyin Can Aksoy
- 2- Assist. Prof. Dr. Selin ÖZDEN MERHACI
- 3- Assist. Prof. Dr. Zehra ÖZKAN ÜNER

Date of the Thesis Examination

19.12.2022

REPUBLIC OF TURKEY
ANKARA UNIVERSITY
To The Directorate of Graduate School of Social Sciences

I hereby declare that all information in my master thesis titled “Big Data and Intellectual Property Rights” which has been prepared under Assist. Prof. Dr. Zehra ÖZKAN ÜNER’s supervision has been gathered and submitted in compliance with academic rules and ethical conduct principles and the data obtained from other sources have been duly indicated both in the text and references. I also declare that I have acted according to scientific research and ethical rules during the study process and if otherwise come forth, I will accept all legal consequences. (..../..../.....)

Rumeysa Canan ULUÇAY

TABLE OF CONTENTS

ABBREVIATIONS	iii
INTRODUCTION	1
CHAPTER I	
CONCEPT OF BIG DATA	3
I. Term of Big Data	3
II. Sources of Big Data	4
A. IOT	5
B. Smart Cities	7
III. Value Extracted from Big Data	9
IV. Terms Related with Big Data	12
A. Data	12
B. Open data	13
C. Personal Data	14
D. Metadata	15
E. Data Set	16
F. Database	16
CHAPTER II	
LEGAL PROTECTION OF BIG DATA	19
I. Personal Data Protection	21
II. Data Ownership	25
III. Unfair Competition	28
IV. Contractual Protection	29
V. IP Rights	30
A. Patent Protection	31
B. Copyright Protection	32
C. Trade Secret Protection	33
D. Sui Generis Database Protection	35
CHAPTER III	
COPYRIGHT PROTECTION	37
I. Copyright Protection of Database in International Agreements	38
II. EU Law	41

A. The Database Directive.....	41
B. The Subject Matter of Database Directive.....	41
C. Database Authorship.....	42
D. Restricted Acts and Exceptions to Restricted Acts.....	43
E. Term of Protection	45
III. Turkish Law	45
IV. Big Data Issues for Copyright Protection	46
A. Originality.....	48
B. Human Intervention in Second Degree.....	51
CHAPTER IV	
SUI GENERIS DATABASE PROTECTION.....	53
I. EU Law.....	54
A. The Database Directive.....	54
B. Exclusive Rights and Exceptions to Those Rights	58
C. Term of Protection	60
D. Author or Maker of the Database?.....	61
E. Criticisms to the Database Directive.....	62
II. Turkish Law.....	66
III. Big Data Issues for Sui Generis Database Protection.....	67
A. Substantive Investment	70
B. Machine Generated Data and the Notion of Data Producer's Right.....	74
CONCLUSION	76
BIBLIOGRAPHY	79
ABSTRACT	88
ÖZET	88

ABBREVIATIONS

Art	Article
Database Directive	Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases
DSM Directive	Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC
EC	European Commission
ECJ	European Court of Justice/Court of Justice of the European Union
EU	European Union
GDPR	General Data Protection Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC
ICSU	International Council for Science
IoT	Internet of Things
IP	Intellectual Property
IPR	Intellectual Property Rights
p.	Page
para.	Paragraph
pp.	Pages
Trade Secret Directive	Directive (EU) 2016/943 of the European Parliament and the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use or disclosure

TRIPS	Trade-Related Aspects of Intellectual Property Rights
Law No. 5846	Turkish Law on Intellectual and Artistic Works No. 5846
UNESCO	The United Nations Educational, Scientific and Cultural Organization
US/USA	United States of America
WCT	WIPO Copyright Treaty
WIPO	World Intellectual Property Organization
WMO	World Meteorological Organization



INTRODUCTION

Data is referred as the “oil of new era” and with the rapid technological advancements, different mechanisms appear to create and collect “data”. Data is collected through various sources such as social media, online applications, e-commerce websites. Also, due to emerging technologies, internet of things and smart cities are regarded as the prominent sources of large data sets. Large data sets, which could not be processed with traditional ways, constitute big data and to extract a value from the data, big data analysis is used in the data driven industry. While elaborating the big data concept, it is crucial to comprehend the related terms with big data which are ‘open data’, ‘meta data’, ‘personal data’, ‘data set’ and ‘database’ to have precise aspect while these terms will be using in the context, and the difference from and connection with big data.

“Big data” is one of the most popular terms of digital age and its importance come from extracted value of big data analysis, the need of policies for disclosure of big data practices and promoting innovation in technologies related big data. Thus, the main discussions for big data are around about its usage, importance and effects on the economy. In this thesis, after providing a general view over big data concept, mainly the legal protection of big data will be examined.

Different legal disciplines appear in the big data concept, but IPRs will have prominent place in this thesis as the main supportive legal policy for the innovations with regard to big data. Databases with collection of independent materials systematically are utilized for the arrangement of big data to extract a value, so among IPRs, copyright, sui generis database right and trade secret protection are the most related ones with the big data protection. Trade secret is also examined in IP protection of big data, however, the minimum standards to protect undisclosed information having commercial value could

hardly exist in terms of big data. As patent in general is applied for the technological products not for the data itself, it will not be approached detailly in this thesis.

Also, data is subject to other various legal concepts such as data ownership, unfair competition and contractual obligations. Additionally, in case the big data includes personal data, data protection law shall be applicable and that intersection with other legal disciplines over big data will be provided in general framework.

Copyright is granted for original works and also provide protection for original databases. Although database is mainly protected by copyright, some legislations such as “*Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases*” and “*Turkish Law on Intellectual and Artistic Works No. 5846*” endow sui generis protection to database by evaluating the investment committed for obtaining, presenting or verification of contents in a substantive way. In that case, the various views of the scholars and also private sector in big data will be provided to discuss whether the legal protection of big data could be applicable and is necessary or the free access to database by both users and database owners is preferable to gain more benefit due to much larger datasets.

This thesis will bring the intersection between those by focusing on the terms and related legislation in European Union and Turkey in terms of big data and intellectual property rights. Could IPRs provide the sufficient protection for big data and is there a need of additional mechanisms in the related legislation to challenge with new data driven world? Instead of creating new instruments such as property rights over data, could sufficient protection be achieved due to revisions of copyright and/or sui generis right to improve database’s functionality on the present and future of the data industry? All those questions will be discussed through different perspectives within the concept of big data.

CHAPTER I

CONCEPT OF BIG DATA

I. Term of Big Data

“Big data” is identified in the Report of McKinsey as *“datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze”*¹. For the first use of term of big data, the credit is given to John Mashey² during his career at Silicon Valley. In the middle of 1990s, John Mashey, working as Chief Scientist at Silicon Graphics, prepared the 1998 slide deck with the title of “Big Data and the Next Wave of InfraStress” and contributed the phenomenon, “big data” term, in those years.³

Big data refers the so large data which is not possible to process it by using traditional methods and expressed with three terms in general as ‘Volume’, collecting huge quantity of data; ‘Velocity’, collecting data rapidly, and ‘Variety’, collecting data in all formats from various sources such as Internet browser, apps, cars etc. In addition to those, 5V is also used to explain the big data by adding the ‘Value’ and ‘Veracity’. The ‘Value’ should be revealed from big data. While the data coming from various sources in different forms with a huge speed, uncertainty about the exactness of the data could be appeared in a crucial level. Thus, ‘Veracity’, ability to trust the data, requires means and techniques to comprehend the data and use rankings for the data quality.⁴

¹ James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Byers, “Big Data: The next frontier for innovation, competition, and productivity”, *at 1, McKinsey 2*, 2011, <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation> (access date: 14.11.2022) .

²https://en.wikipedia.org/wiki/Big_data#:~:text=The%20term%20big%20data%20has,Mashey%20for%20popularizing%20the%20term.

³ Francis X. Diebold, “On the Origin(s) and Development of the Term 'Big Data'”, PIER Working Paper No. 12-037, 2012, p. 3.

⁴ Bob McQueen, *Big Data Analytics for Connected Vehicles and Smart Cities*, Artech House, 2017, <https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1825911&site=eds-live> (access date: 14.11.2022), p.42.

In the past decade of “Big Data Revolution”, creation and gathering of personal and nonpersonal digital data increased rapidly, so, extensive datasets and databases are generated for the arrangement of that data⁵. The total amount of data generated by humans and computers in 2020 is calculated as more than 64.2 zettabytes. In 2021, the estimation for the amount of data generated in the world was around 79 zettabytes (79 trillion GB) and it is predicted to rise 180 zettabytes by 2025.⁶ OECD report explains the largeness of the data generated daily in 2015 that 2.5 exabytes data generated daily equals the 167.000 times of information in all books existed in the Library of Congress of the United States.⁷

II. Sources of Big Data

Data created and gathered by humans and/or machines either consciously or unconsciously includes enormous value in data-driven economy.⁸ Two facts are stated by considering the business models and points of interests regarding data; “(i) *most data are generated by someone else, and b) the value of data increases by their use, not the restriction on their use*”.⁹ With that fact, especially for the large scaled companies creating the tools to gather and analysis data are interested with the development of data sources.

⁵Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, Singapore Academy of Law, 2020, p.1.

⁶ Statista (2022). Big data - Statistics & Facts. <https://www.statista.com/topics/1464/big-data/#dossierKeyfigures> (access date: 14.11.2022)

⁷ OECD, “Data-Driven Innovation: Big Data for Growth and Well-Being”, *OECD Publishing*, Paris, 2015, p. 20.

⁸ Peter K. Yu, “Data Producer’s Right and the Protection of Machine-Generated Data”, *Tulane Law Review*, 93(4), 2019, p. 860.

⁹ Joren De Wachter, “IP Strategy in the Age of Big Data”, *International In-House Counsel Journal*, 7(26), 2014, p. 8.

It is hard to classify the limitless sources of big data, but, mobile phones, sensors, social media, smart devices, web searches could be listed firstly to create the enormous data. Thus, although various means exist to collect the data, IoT and smart cities involving the main data gathering mechanisms such as sensors, mobile phones or smart devices are largely subject to examination in terms of big data and will be provided through general perspective with its contribution to big data constitution.

A. IOT

Kevin Ashton created the term of IoT by connecting the internet with things as the standard tool for computers to comprehend the real world to meet the people's need¹⁰. IoT is a system including interconnected devices called as smart devices in general and that system transfers data and connect one with the other due to the Internet.¹¹

OECD identifies the Internet of Things as “*IoT refers to an ecosystem in which applications and services are driven by data collected from devices that sense and interface with the physical world.*”¹². IoT raises the connection among devices tremendously and will maximize data sources with the result of increase in the analytics of big data potentially.¹³

Jacopo Ciani explains IoT as the gathering of daily smart devices including “*microchips, sensors and wireless communications capabilities*” with the connection with Internet and

¹⁰ Chana Schoenberger, “The Internet of Things”, *FORBES*, 2002, <https://www.forbes.com/global/2002/0318/092.html#33f75a713c3e> (access date: 14.11.2022).

¹¹ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p. 53.

¹² OECD, “The Internet of Things: Seizing the Benefits and Addressing the Challenges”, *OECD Digital Economy Papers*, No. 252, *OECD Publishing*, Paris, 2016, p.5.

¹³ Francesco Banterle, “The Interface Between Data Protection and IP Law: The Case of Trade Secrets and the Database sui generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis”, *Personal Data in Competition, Consumer Protection and Intellectual Property Law*, Springer Berlin Heidelberg, Vol 28, 2018, p.428.

each, thus receiving, gathering and sending large data, following activities and interacting with other devices to enable services in accordance with the users' needs and wishes. The number of IoT connected devices in the world is calculated as 11.3 billion in 2021 and estimated to reach 29.4 billion by 2030 as a yearly basis.¹⁴

One of the crucial components of the IoT is sensor data which is created by devices such as cars, smart watches, health care devices connected to system retrieving data from sensors.¹⁵ However, IoT objects involving sensor equipped and connected devices make harder to distinguish data creation from data obtaining in the concept of automatic data collection with systematic categorization of data.¹⁶

In some instances, the creator of the data could be end user instead of IoT provider such as the track of heart-rate. In such a case, the owner of the data should be the user and it is argued the owner has a right to license it for secondary uses. Also, on the other side, when the smart device creates data for example; calculating the number of passengers at the exact location, copyright issues are linked with the organization as no specific data is produced by the individual. However, because of that the originality requirement is not met for such "works" the copyright obstacle would not be overcome even without

¹⁴ Statista, "Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2021, with forecasts from 2022 to 2030", <https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/> (access date: 14.11.2022).

¹⁵ Kristina Karanikolova, Julien Chicot, Anna Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), "Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report", *Publications Office*, 2018, <https://op.europa.eu/en/publication-detail/-/publication/5e9c7a51-597c-11e8-ab41-01aa75ed71a1>, p. 29.

¹⁶ European Commission, "Evaluation of Directive 96/9/EC on the legal protection of databases", {SWD(2018) 147 final}, 2018, [https://ec.europa.eu/transparency/documents-register/detail?ref=SWD\(2018\)147&lang=en](https://ec.europa.eu/transparency/documents-register/detail?ref=SWD(2018)147&lang=en) (access date: 14.11.2022), p. 15.

considering the who or what is the creator of the data.¹⁷ Thus, the complexity of big data protection comes from its sources also.

B. Smart Cities

The crucial point of big data is to combine data sets in novel ways by utilizing algorithms and other processes of big data to present new perspective and services. That point is the same for smart cities including traffic flows, activities in social media and other sensor devices which provides a 'richer' look for the current state affairs.¹⁸

The definition of smart is written in OECD report as "*An application or service is able to learn from previous situations and to communicate the results of these situations to other devices and users*"¹⁹. However, McQueen finds an explanatory definition of the smart city as unconvincing because of that smart city technologies newly develops and are applied and, in the future, a single identification could be achieved.²⁰

In smart city examples, raw and unstructured data is collected and aggregated from lots of sources through "*CCTV and other cameras, mobile device use, sensors, smart cards in mass transit, etc*"²¹. The functions of smart cities are explained with that the data collected from various devices and sensors will be bringing together to process due to wired and wireless communication and processed to present, perfect and predict.²² Giffinger and Milanovic explains the performance of smart cities through six characteristics:

¹⁷Christine Storr and Pam Storr, "Internet of Things: Right to Data from a European Perspective", in Marcelo Corrales, Mark Fenwick, Nikolaus Forró (edit), *New Technology, Big Data and the Law*, Springer, 2017, p. 76.

¹⁸ Karl Löfgren and C. William R. Webster, "The value of Big Data in government: The case of 'smart cities.'", *Big Data & Society*, 2020, p. 8-9.

¹⁹OECD, "Building Blocks for Smart Networks", OECD Digital Economy Papers, No. 215, *OECD Publishing*, Paris, 2013, p.4.

²⁰ McQueen, "Analytics for Connected Vehicles and Smart Cities. Artech House", p. 83.

²¹ Löfgren and Webster, "The value of Big Data in government: The case of 'smart cities.'", p.5.

²² McQueen, "Analytics for Connected Vehicles and Smart Cities. Artech House", p. 86.

“environment, economy, mobility, people, living, governance”²³. Indeed, smart cities aim to provide high level quality of life by using the technological innovations, and, IoT and big data are inseparable from smart cities.

The required basic minimum principles for smart cities are aligned as follows²⁴:

“(a) the quality standards of data to be used in the smart city; (b) the ethical standards regarding privacy and data protection; (c) a clearer policy regarding the ownership of unstructured and structured data; and (d) agreed standards regarding safety and protection of storing of data”.

Thus, the policies put forward with regard to smart cities require a query about the current legal perspective for the big data protection.

With the effects of the involvement of both public service providers and enterprises as partners, the beneficiary of the value arising from using data in smart city concepts is complicated. Everyday life of persons turns into data due to the technological advancements and both private and governmental organizations desire to acquire, store and process this data.²⁵ Thus, it is complicated that whether the inspection, ownership, access to data and the privatization of value is done by commercial partners or society through public agencies in smart cities. That complexity has a deep impact on the actors to address legal protection of big data.

²³ Rudolf Giffinger, Natasa Pichler-Milanovic and et. al, “Smart Cities: Ranking of European Medium-Sized Cities” *Vienna University of Technology*, Vienna, 2007, p. 12.

²⁴ Löfgren and Webster, “The value of Big Data in government: The case of ‘smart cities.’”, p. 11.

²⁵ Leighton Evans, “The Privacy Parenthesis: Private and public spheres, smart cities and big data”, in Claudio Coletta, Leighton Evans, Liam Heaphy, Rob Kitchin (ed.), *Creating smart cities*, Taylor & Francis Books, 2019, p. 196.

III. Value Extracted from Big Data

New economy which is in transition period is explained as knowledge based. For that economy, the key entry of the competition is the knowledge in the market which is also main element for the improvement of market.²⁶ Many philosophers connect the power of knowledge with comprehension of information which attributes meaning to data.²⁷

The data is classified as “*a non-rivalrous commodity*” because of that a person’s use of it does not essentially prevent or decrease the value for another person’s or 10.000 persons’ usage. The data gathered through large datasets in use of few or many persons constitutes big data to extract gorgeous value from it. Hence, in the various definitions of big data, many scholars draw attention to decision making mechanisms for big data about its value. For example; Gartner, presenting consultancy services to its client enterprises globally over forty years, states “*Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making*”²⁸. Those advantages make big data attractive for many leading technology companies worldwide. Thus, big data is seen as the means of providing privilege in the data-driven economy. Many leading firms are utilizing big data to create value and others should discover how to do it in case they are keen to compete.²⁹

Big data is utilized for the analysis of large data sets and extract valuable information from those to arise new correlations such as finding business, political trends or

²⁶ Anirban Mazumder, *Database Law Perspectives from India*, Springer, 2016, p. 1.

²⁷ Cambridge International AS & A Level Information Technology, “Data, information and knowledge”, *Cambridge International Examinations*, 2015, <https://www.cambridgeinternational.org/images/285017-data-information-and-knowledge.pdf> (access date: 11.11.2022), p.4.

²⁸ Gartner IT glossary, Definition of Big data. <http://www.gartner.com/it-glossary/big-data> (access date: 14.11.2022).

²⁹ Manyika, Chui, Brown and et. al., “Big Data: The next frontier for innovation, competition, and productivity”.

hampering the crime.³⁰ However, the commercial processing of data requires investment in IT infrastructure, time and human resources. When we consider the data-driven economy, value of data is significantly increasing due to analysis technologies and its great capability in prediction.³¹ The perspective of public policy connects the maximization of social benefits arising from big data with the access to the databases and permission for the combination data with other datasets to get better qualified data analysis.³²

McKinsey report of 2011 exemplifies the big data's financial contribution to the various sectors such as 250 billion euros value per year to Europe public sector administration and the revenue amounted 100 billion usd+ for service providers due to global personal location data.³³ Some of the innovation companies which are the most successful ones as Google, Amazon, Facebook and eBay structured their business model on the gathering and exploiting big data.³⁴

The power of big data is also considerably observed in health issues. In 2010, some researchers used internet searches to determine the adverse effects of specific drugs. In cooperation with Microsoft, researchers analyzed the searches of users on “*Google, Bing and Yahoo*”. As a result of examination of words used in queries, they found evidence for that pravastatin and paroxetine searches were also hyperglycemia-associated words.

³⁰ Nikolaus Forgó, Stefanie Hännold, Benjamin Schütze, “The Principle of Purpose Limitation and Big Data”, in Marcelo Corrales, Mark Fenwick, Nikolaus Forgó (edit), *New Technology, Big Data and the Law*, Springer, 2017, p.17.

³¹ Banterle, “The Interface Between Data Protection and IP Law: The Case of Trade Secrets and the Database sui generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis”, p. 12

³² Man Yip, Protecting Consumer's Personal Data in the Digital World – Challenges and Changes, *Research Collection School Of Law*, 2018, https://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=4862&context=sol_research (access date: 14.11.2022), p. 116-117.

³³ Manyika, Chui, Brown and et. al., “Big Data: The next frontier for innovation, competition, and productivity”.

³⁴ Helena Ursic Vrabec and Bart Custers, “Legal barriers and enablers to big data reuse”, *European Data Protection Law Review (EDPL)*, 2(2), 2016, p. 209.

Thus, this connection appeared a hypothesis confirmed experimentally that Paxil and Pravastatin could lead to adverse side effects in case those are taken together.³⁵

Also, making more proper estimations and better decisions in potential to handle with pandemics, natural disasters or social unease is possible due to big data techniques and technologies.³⁶ United Nations Global Pulse³⁷ initiative report summarizes the big data for development as “*Big Data for Development is about turning imperfect, complex, often unstructured data into actionable information.*”³⁸ Thus, protecting and incenting intellectual creations with regard to big data should be fostered for development. Hugenholtz explains the economic potential of big data as “*machine-generated data are both input and output to innovative manufacturing processes and value-added services, and thus a major driver of economic growth.*”³⁹ The big data is only really used by creating information from those databases with a mining which is beneficial and valuable for end users⁴⁰. Therefore, the ways for selection and arrangement with regard to the big data should be advanced to acquire maximum benefit and the legal approach to provide an incentive for those will be discussed.

³⁵ Ryen W White, Nicholas P Tatonetti, Nigam H Shah, Russ B Altman, Eric Horvitz, “Web-scale pharmacovigilance: listening to signals from the crowd”, *Journal of the American Medical Informatics Association*, Volume 20, Issue 3, 2013, p. 404–408.

³⁶ Hamid Ekbia, Michael Mattioli, Inna Kouper, et al, “Big data, bigger dilemmas: A critical review”, *Journal of the Association for Information Science & Technology*, 66(8), 2015, p.1539.

³⁷ UN Global Pulse is the Secretary-General’s initiative on big data and artificial intelligence for development, humanitarian action, and peace, see <https://www.unglobalpulse.org/about/>.

³⁸ UN Global Pulse, “Big Data for Development: Challenges and Opportunities”, 2012, <https://www.unglobalpulse.org/wp-content/uploads/2012/05/BigDataforDevelopment-UNGlobalPulseMay2012.pdf> (access date: 14.11.2022), p. 6.

³⁹ P. Bernt Hugenholtz, “Against ‘Data Property’”, in H. Ullrich, P. Drahos, G. Ghidini (Eds.), *Kritika: Essays on Intellectual Property*, Vol. 3, 2018, p.14.

⁴⁰ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p.5.

IV. Terms Related with Big Data

Although the many terms related of big data are commonly used even in daily life, it is quite important to comprehend those terms: data, open data, personal data, metadata, dataset and database. Due to the clarification on these terms, legal protection of big data will be discussed clearly in the following chapters.

A. Data

Data, which can be without content or idea, are the raw form of the facts. It can be qualitative with the meaning of descriptive or quantitative being numeric as a result of count, quantification or other mathematical calculation. However, also a number can be qualitative data, in case this number is not a result of any measurement or calculation.⁴¹ Thus, data is the digitalized information and could be age, customer information or even website clicks.

Data are all kinds of facts, findings and numbers and even works which are proper to make inferences.⁴² In fact, people, generally without conscience, pay the return of digital services with their data.

Data processing includes two stages as gathering, assembling or generating data firstly and then store and control it in the system. Data is considered in general together with the

⁴¹ David T. Bourgeois, James L. Smith, Shouhong Wang and Joseph Mortati, *Information Systems for Business and Beyond*, 2019, Open Textbooks, <https://digitalcommons.biola.edu/cgi/viewcontent.cgi?article=1000&context=open-textbooks> (access date: 14.11.2022), p. 1., <https://digitalcommons.biola.edu/cgi/viewcontent.cgi?article=1000&context=open-textbooks>, p.66.

⁴² Ünal Tekinalp, *Fikrî Mülkiyet Hukuku*, Beta İstanbul, 1999, p. 130.

system which involves its, however, to put the difference between the content and the system is important with regard to IPRs.⁴³

In legal terms, Kemp states the data as funny stuff and explains it with that no rights exist in data but, in relation to data, comprehensive rights and obligations appears.⁴⁴ Thus, those rights mainly in IP will be discussed within the concept of big data.

B. Open data

Data which is available both legally and practically for use and republication by anybody is named as open data⁴⁵. The perspective on data usage is affected significantly by open data. Both civil societies and governments are in favor of open data respectively have an idea of those open data may contribute social life and increase the transparency in decision mechanism.⁴⁶

The technological design of data is the first factor that data should be available preferably online, machine readable so structured and linkable. IPRs, limiting or preventing the reuse and distribution of open data, is classified as second factor. Some initiatives in the area of open data defenses that open data should be free of IPRs although innovative IP policies are encouraged due to open data regimes in case of not restricting reuse of users and redistributing the data.⁴⁷

⁴³ Research Data Management Service Group, "Introduction to intellectual property rights in data Management", Cornell University, <https://data.research.cornell.edu/content/intellectual-property> (access date: 14.11.2022).

⁴⁴ Richard Kemp, "Legal aspects of managing Big Data", *Computer Law & Security Review: The International Journal of Technology Law and Practice*, 30(5), 2014, p. 486.

⁴⁵ Juho Lindman, Matti Rossi, Virpi Kristiina Tuunainen, "Open data services: research agenda", *46th Hawaii International Conference on System Sciences, IEEE*, 2013, p. 1239.

⁴⁶ Bridgette Wessels, Rachel Finn, Thordis Sveinsdottir and Kush Wadhwa, *Open Data and the Knowledge Society*, Amsterdam University Press, 2017, p. 63.

⁴⁷ OECD, "Data-Driven Innovation: Big Data for Growth and Well-Being", p. 39.

For some specific areas such as scientific researches, the open data policies are also supported⁴⁸ to encourage the advancement benefiting from existing ones freely. As due to the concept of open data, access, use and distribution of the content is free without applying the limitations arising from IPRs.

C. Personal Data

“General Data Protection Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC” (GDPR) identifies ‘personal data’ as *“any information relating to an identified or identifiable natural person.”*. Art. 3/a of Turkish Personal Data Protection Law No.6698 has same definition for personal data with GDPR. Personal data does not attribute property to data subject in current legislation in both EU and Turkey.

Personal data could be family name, occupation, data concerning health, ethnic origin etc. Personal data processing means any action on personal data, through automated means or not, for example; collection, structuring, combination, storing or alignment of data.⁴⁹

The difference between the knowledge of service providers and most end users is likely subject to rise due to IoT improves. Users are not aware of an extent of personal data gathered and shared through methods of embedded devices without any documentation to consult including privacy notice.⁵⁰

⁴⁸ Banterle, “The Interface Between Data Protection and IP Law: The Case of Trade Secrets and the Database sui generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis”, p. 31.

⁴⁹ GDPR, Art. 4(2), Turkish Personal Data Protection Law, Art. 3-e.

⁵⁰ Wessels, Finn, Sveinsdottir, Wadhwa, *Open Data and the Knowledge Society*, p. 60.

In recital 63 of GDPR, the application of “right to access of data subject” is limited by considering the IPRs as follows:

“That right should not adversely affect the rights or freedoms of others, including trade secrets or intellectual property and in particular the copyright protecting the software. However, the result of those considerations should not be a refusal to provide all information to the data subject.”

However, this provision is quite ambiguous; for example, what should be the scope of asserting trade secrets protection or IPRs while refusal of data portability by data controller.⁵¹

D. Metadata

Metadata and data are matched with “*a universe and DNA*” and metadata is often identified as “*data about data*”. Information describing content or an item describing another item could be named as metadata so the metadata is more than descriptive data. For example; the book occurs the metadata of any book or article in its references.⁵²

Metadata is data come from the internet embedded devices in digital universe. It is added to present data and it provides useful data which increases rapidly and proper for analysis.⁵³ The metadata management support is necessary in environments of big data in

⁵¹ Joseph Drexel, “Data Access and Control in the Era of Connected Devices”, *The European Consumer Organisation*, BEUC, Brussels, Belgium, 2018, https://www.beuc.eu/sites/default/files/publications/beuc-x-2018-121_data_access_and_control_in_the_area_of_connected_devices.pdf (access date:14.11.2022), p. 83.

⁵² John W. Warren, “Zen and the Art of Metadata Maintenance”, *Mason publishing group/george mason university press*, Volume 18, Issue 3, 2015, p. 1.

⁵³ Wessels, Finn, Sveinsdottir, Wadhwa, *Open Data and the Knowledge Society*, p. 57-58.

some cases aiming to reach a large-scale analysis from a small narrow question in business requiring validity of analytics.⁵⁴

E. Data Set

Big data includes more complex and large data sets. Cambridge Dictionary defines the data set as “*a collection of separate sets of information that is treated as a single unit by a computer*”.⁵⁵ With an increase in volume and variety of proper data sets, the ability to link various datasets rises. Thus, further information could be derived from those data.⁵⁶

Big datasets are referred as “*large, diverse, complex, longitudinal, and/or distributed datasets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources available today and in the future*”.⁵⁷ Boyd and Crawford explain the big data mentioning data set that “*Big Data is less about data that is big than it is about a capacity to search, aggregate, and cross-reference large data sets*.”⁵⁸ It is referred that big data is output of the combination of large data sets.

F. Database

“*The Directive 96/9/EC (1996) of the European Parliament and of the Council of 11 March 1996, on the legal protection of databases*” (hereinafter referred to as ‘the

⁵⁴ Roxana-Maria Holom, Katharina Rafetseder, Stefanie Kritzing, Harald Sehrschrn, “Metadata management in a big data infrastructure”, *Procedia Manufacturing*, Volume 42, ISSN 2351-9789, 2020, p. 376.

⁵⁵ Cambridge Dictionary, <https://dictionary.cambridge.org/dictionary/english/dataset> (access date: 14.11.2022).

⁵⁶ OECD, “Data-Driven Innovation: Big Data for Growth and Well-Being”, p. 219.

⁵⁷ National Science Foundation, “Core Techniques and Technologies for Advancing Big Data Science & Engineering (BIGDATA)”, PROGRAM SOLICITATION NSF 12-499, 2012, <http://www.nsf.gov/pubs/2012/nsf12499/nsf12499.pdf>.

⁵⁸ Danah Boyd and Kate Crawford, “Critical Questions for Big Data”, *Information, Communication & Society*, 15:5, 2012, p. 663.

Database Directive’) gives the meaning of database as “*a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.*”. Although the meaning is provided in the Database Directive, many scholars considers that it should be clarified⁵⁹.

Ateş simply identifies the database as products including information collected, stored, materialized, made tangible and, thus, gained the ability of “marketability”.⁶⁰ Lavenue identifies the database as “*the most valuable form of technical data.*”.⁶¹ It could be referred from the ability provided by the database to infer the necessary information easily in database by the end user.

The storage ability of databases and presentation of existing data or information to the end users are the two important aspects of electronic databases.⁶² As to talk about the existence of a database, it should be accessible. Accessibility of a database means that it should be readable, visible, hearable for persons, shortly, cognisable by a human.⁶³

The determination of a database is bound to presence of collection of independent elements which could be separated from each other without effecting its value in information, literature, music or any other.⁶⁴ In case the third party extracts geographical information utilized to produce and commercialize another map from a topographic map it is regarded as adequate informative value which is independent element of database

⁵⁹ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report”, p. (vi).

⁶⁰ Mustafa Ateş, “Veri Tabanlarının Hukukî Koruması”, *Ankara Üniversitesi Hukuk Fakültesi Dergisi*, Volume 55, issue 1, 2006, p. 48.

⁶¹ Lionel M. Lavenue, “Database rights and technical data rights: the expansion of intellectual property for the protection of databases”, *Santa Clara Law Review*, 38(1), 1997, p. 1.

⁶² Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p.9.

⁶³ Ateş, “Veri Tabanlarının Hukukî Koruması”, p. 69.

⁶⁴ Fixtures Marketing Ltd v Organismos prognostikon agonon Podosfairou AE (OPAP). Case C-444/02. Judgment of the Court (Grand Chamber) of 9 November 2004, para. 29.

according to the Art. 1(2) of Database Directive.⁶⁵ To reach that result, the Court accepted that according to Art. 1(2) of the Directive, combination of information pieces is also independent material which is not special to only a piece of information.⁶⁶

Databases are quite important for the economy, but its impact is not limited with the economy. Databases have also crucial impacts for the progress of culture and science so the investment on those is subject to protection. Mr. Coble from House of Representatives states in the Report of Collections of Information Antipiracy Act that “*Whether the focus is on financial, scientific, legal, medical, bibliographic, news, or other information, databases are essential tools for improving productivity, advancing education and training, and creating a more informed citizenry.*”⁶⁷. Thus, considering the contribution for extracting a value through big data processing and big data analysis, the database protection will be examined in detail.

⁶⁵Freistaat Bayern v Verlag Esterbauer GmbH. Case C-490/14. Judgment of the Court (Second Chamber) of 29 October 2015.

⁶⁶ Francesco Banterle, “Data Ownership in the Data Economy: A European Dilemma”, in Tatiana-Eleni Synodinou, Philippe Jougoux, Christiana Markou, Thalia Prastitou (Eds.), *EU Internet Law in the Digital Era*, Springer, 2020, p. 225.

⁶⁷ Collections of Information Antipiracy Act. H.R. Rep. No. 106–349, 1999, <https://www.congress.gov/106/crpt/hrpt349/CRPT-106hrpt349.pdf> (access date: 14.11.2022), p.6.

CHAPTER II

LEGAL PROTECTION OF BIG DATA

People searching for information would like to reach the exact information at the exact time without dealing with the confused and unrelated information so AI, other technologies in data visualization and best practices in a matter of organization will improve to be sure about the people's want regarding information⁶⁸. Extracting value from big data to provide the demanded information easily is the result of generation of database. It could be exemplified with improving your product to present most likeable one by end user or offering the most favorable products for the customers in ecommerce websites. Indeed, determining the connection and making estimations or other precious analytical results could be achieved after processing of big data. These connections and comprehensions could be used for various aims including targeted advertisement, surveillance and other possible applications.⁶⁹ Thus, the legal protection of big data is mainly discussed over the database.

By considering the extracted value from big data due to database, governments and organizations are quite interested with the investments in electronic databases. Singapore Academy of Law lists the tension among three points about big data concept as follows:

- (1) AI and big data innovations and investments should be incited by rewarding original works creators,
- (2) the investment in skill, time and exert for data governance to provide qualified data should be recognized, and

⁶⁸ Kemp, "Legal aspects of managing Big Data", p.483.

⁶⁹ Daniel Gervais, "Exploring the interfaces between big data and intellectual property law", *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 10(1), 2019, p. 4-5.

(3) maximization of the advantages of original works should be ensured through the adequate access of third parties.⁷⁰

Thus, to achieve the maximum benefit from qualified data, the incentive through granting IPRs for big data innovations is crucial item of the data economy.

Due to the network among “connected users, devices and sensors”, the generation and obtainment of data are never been easy as today. However, also the copying content of a database is much easier than creating it. Because of the easiness of reproduction of its contents by a third party, investment rates for databases could be in a low stage for organizations having little financial incentive.⁷¹ Therefore, to encourage the creation of a database, its protection should not be confined narrowly in terms of legal protection, in general, IPRs in information society could be applied for databases.⁷²

In addition to IPRs, other various disciplines are discussed for the big data protection. Burtscher and Fritz state the intersection of different legal concepts as “*data privacy, database rights, IPRs, antitrust law as well as the basic civil rights of ownership and possession are playing a role when dealing with the legal alien big data but are each only addressing bits of it.*”⁷³. Various concepts of law are on the agenda to pay attention while the big data is used in e-commerce or some traditional sectors such as retail, logistics and health care.

Digital technology has two effects on databases which are to make easier the piracy and, therefore, require a powerful protection for databases and secondly, to make the control

⁷⁰ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p.7.

⁷¹ Darly Lim, “Re-defining the rights and responsibilities of database owners under competition law”, *Singapore Academy of Law Journal*, 18(2), 2006, p. 420.

⁷² F. Wilem Grosheide, “Database Protection—The European Way”, *Washington University Journal of Law & Policy*, Volume 8, 2002, p.41

⁷³ Bertram Burtscher, Gernot Fritz, “Big data: who owns and who may use and exploit big data?”, 2015, <https://www.lexology.com/library/detail.aspx?g=77ab3ffb-8f25-469c-ad80-d14bbcee9551> (access date: 14.11.2022).

and track of access and charge of user more easily and aggravate the access to databases.⁷⁴

To incent the database creators, EU policy and Turkish law provide both copyright protection for creativity and also sui generis protection in case of substantive investment. In those chapters of copyright and sui generis protection, complexity between human intervention and machine generated data will be discussed.

Although big data has intercessions with other disciplines mainly IP law and somehow competition law, personal data is widely used in big data, thus, the issue will be discussed also approaching the personal data protection in a superficial way. Also, some scholars are supporting to provide data ownership and other believes the power of contractual protection. Additionally, unfair competition is regarded as a remedy to support the IP rights.

In this thesis, personal data protection, unfair competition, contractual protection, data ownership and IP rights including patent and trade secret will be handled in general by considering the limited applications of those for big data protection. Then, copyright protection and sui generis database protection will be elaborately discussed as main legal protection ways of big data within the frameworks of EU and Turkish Law after providing a view in international instruments. This thesis will bring a perspective for de lege lata and also de lege ferenda for big data protection.

I. Personal Data Protection

In the age of big data, persons' control over data is extremely hard. Personal data, both as raw or information having copyright protection, exist in the online complex information

⁷⁴ Mazumder, *Database Law Perspectives from India*, p. 23.

systems for social media and other platforms.⁷⁵ The primary aim of processing big data is to extract a value from the data and as the personal data is mainly processed to do that, the personal data protection law should also be considered while discussing the legal protection of big data.

Hartzog states “*In the world of big data, more is always better.*”⁷⁶ However, for enabling privacy for data subjects, enforcement of the principle of data minimization is always better. Thus, in case the personal data exists in the concrete case the issue becomes more complex.

Ownership regime over the data sets could be evaluated by taking connection between personal data protection and IP into consideration. In personal data protection law, data subject has a right to privacy concerning his/her data during data processing actions of data controllers. Also, IPRs may be applicable over those datasets. Thus, IPRs challenges with the privacy rights such as the “data portability” or “right to object to data processing” and that decreases the exact originality of existing IPRs in data set.⁷⁷

The level of “transferability or assignability of a right” specifies its category as personal right or IP right, so if it is determined that data could be assigned, then it will be emancipated from the data subject.⁷⁸ As personal data protection law allows individuals an extent of authority over their personal lives and information. Thus, it acknowledges and promote human dignity and liberty. Also, although intellectual property law contributes to an extent to dignity and liberty in terms of creators and inventors, it is

⁷⁵ Francesco Banterle, “Data Ownership in the Data Economy: A European Dilemma”, p.223.

⁷⁶ Woodrow Hartzog, *Privacy’s Blueprint: The Battle to Control the Design of New Technologies*, Harvard University Press, 2018, p. 51.

⁷⁷ Andreas Sattler, “From Personality to Property?”, in Mor Bakhoun, Mark-Oliver Mackenrodt, Beatriz Conde Gallego and Gintarė Surblytė-Namavičienė (Eds), *Personal Data in Competition, Consumer Protection and Intellectual Property Law Towards a Holistic Approach?*, MPI Studies on Intellectual Property and Competition Law, vol 28. Springer, 2018, p. 42; Kohler, J. (1880), *Das Autorrecht*, Fischer, p. 412-413.

⁷⁸ Andreas Sattler, : Kohler, J. (1880), *Das Autorrecht*, Fischer, p.74.

generally regarded with utilitarian terms of incenting innovation and access to knowledge due to exclusivity provided for IP owners against free-riders. Thus, special techniques have been developed to balance competing rights, freedoms and concerns for instance permitting access to the public domain as a counterpoint to IP or freedom of expression as an exception for privacy.⁷⁹

Big data desires to reuse the data for various purposes so data protection principles as “purpose limitation” and “data minimization” do not comply with that desire. Also, in case the outdated framework is insisted for data protection, benefits of big data could be under risk.⁸⁰ As, the significance of big data is not only related with its volume or growing speed, however, also data coming from incredible array of sources.⁸¹

GDPR provides many kinds of rights for data subjects to have the control over their data such as ‘*right of access*’, ‘*right to rectification*’ and ‘*right to erasure*’, ‘*right to restriction of processing*’ and ‘*right to data portability*’. ‘Right to data portability’ has a most related and significant impact on the database protection. As transferring data to rivals of the database maker for replacing current service contract or with a kind of legitimate aim should be up to willing of individual customers.⁸² Art. 20 (1) of GDPR clarifies the right to data portability:

“The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data

⁷⁹ Megan Richardson, “Intellectual Property and Privacy Law”, *Oxford University Press*, 2021, p. 136-137.

⁸⁰ ICO, “Big Data and Data Protection”, 2014, <https://rm.coe.int/big-data-and-data-protection-ico-information-commissioner-s-office/1680591220> (access date: 14.11.2022), p. 40.

⁸¹ Christopher Kuner, Fred H. Cate, Christopher Millard, Dan Jerker B. Svantesson, “The challenge of ‘big data’ for data protection”, *International Data Privacy Law*, Vol. 2, No. 2, 2012, p.47.

⁸² Matthias Leistner, “Protection of and Access to Data under European Law”, *Oxford University Press*, 2021, p. 393.

have been provided, where: (a) the processing is based on consent pursuant to point (a) of Article 6(1) or point (a) of Article 9(2) or on a contract pursuant to point (b) of Article 6(1); and (b) the processing is carried out by automated means”.

In accordance with the Art. 20 of the GDPR, database owner should respect the right of data subject to move his/her personal data from one database to another. Thus, data portability could be a probable regulatory means for issues in competition and concerns to access in big data.⁸³ Whether the data portability right is a form of new exogenous restraint for copyright protection while the information systems of data controller are under the protection of copyright.⁸⁴ Therefore, data portability should be provided as the exception to copyright protection of a database and the user should be allowed to transfer own personal data to another database.⁸⁵

Consequently, while benefiting from legal rights over big data, it is essential to ensure implementing principles of personal data processing or vice versa. For example; there is a support to improve public authorities’ access to private sector data especially for scientific purposes. However, by considering the data holders’ investment in collecting or adapting the data, the compensation issue should be addressed in addition to data protection principles (transparency, proportionality, purpose limitation etc.).⁸⁶ Therefore, the IPRs will be on the agenda to determine the rights of data holders and whether copyright or sui generis protection will be claimed. Data holders are in need of certainty about number of issues including compensation to recover their investments to collect or

⁸³ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p.23.

⁸⁴ Tatiana-Eleni Synodinou, “The Portability of Copyright-Protected Works in the EU”, in Tatiana-Eleni Synodinou, Philippe Jougoux, Christiana Markou, Thalia Prastitou (Eds.), *EU Internet Law : Regulation and Enforcement*, Springer, 2017, p.228.

⁸⁵ Synodinou, p.230.

⁸⁶ Communication From the Commission to the European Parliament, The Council, The European Economic and Social Committee and The Committee of the Regions, “Towards a common European data space”, COM/2018/232 final, 2018, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0232&from=en> (access date: 14.11.2022), p. 12-14.

arrange the data, but also fundamental rights of data subjects should be supervised by considering the vital importance of right to privacy.

II. Data Ownership

Big data has a huge place in the economy so the legal concerns also exist with that growing effect as the data could not be classified as ordinary asset. The data could be copied or combined with other data easily and also used by others at the same time which are special qualities by comparing the physical assets. Thus, some questions with regard to the IPRs appear such as “Who owns the data and what rights come attached with a dataset?” or “Who is responsible when an inaccurate piece of data leads to negative consequences?”. Those issues should be clarified to benefit from the whole potential of big data.⁸⁷

The legal control of data processors and/or intermediaries over the data is obvious, but for the individuals as the owners of data, the issue is more complex. Big data requires to gather data through various sources and create value through secondary use of the data due to data correlations. Thus, the potential economic value of raw data put the data ownership policy on the agenda.

Also, big data and IoT accelerate the requirement of reaching to data over various resources and secondary use of those to find out data connections. By taking the information and insights extracted from the data, even raw data has a value. Thus, the

⁸⁷ Manyika, Chui, Brown and et. al., “Big Data: The next frontier for innovation, competition, and productivity”.

arguments with regard to data ownership appears such as is raw data public or private goods or could be settle under property right.⁸⁸

As property rights are not valid for data, an individual could not sell it or claim it is stolen. Ownership of data and having sufficient control over data which makes up the database are unclear points.⁸⁹ Also, recent scandals such as intentional exploitation of the data in Facebook by Cambridge Analytica⁹⁰ increased the calls for conferral of property rights in personal data via popular media.⁹¹

Even the data ownership is attributed, it does not solve the problems as the legislation should organize some issues in intellectual property law. Primary issue is to determine what is the subject matter of protection, so, a need for clarification of ‘data’ in the structure of ‘data ownership’ exists. Also, the extent of protection including the kinds of interests and uses, exceptions and limitations should be specified.⁹²

The subheading of data ownership including the subject matter and scope of protection is not still clearly identified. Drexl explains that situation focusing the perspective of competition in terms of data analysts. “Owning data” will not provide a competitive privilege for them by comparing the cost of their rivals. However, due to the better access to big data from various sources which is not under the protection of ownership but only

⁸⁸ Banterle, “The Interface Between Data Protection and IP Law: The Case of Trade Secrets and the Database sui generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis”, p. 29.

⁸⁹ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p. 32.

⁹⁰ In the 2010s, personal data belonging to millions of Facebook users was collected without their consent by British consulting firm Cambridge Analytica, predominantly to be used for political advertising. See https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge_Analytica_data_scandal.

⁹¹ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p. 37.

⁹² Joseph Drexl, “Designing Competitive Markets for Industrial Data – Between Propertisation and Access”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 257, 2017, p. 260-261.

with the contractual relationship with the owners of those data sets they may gain more benefits for the efficiency and veracity of its big data analyses.⁹³

Private property rights for data could not be optimal solution, also, because of economic reasons. With the consumption of the person valuing rivalrous good, the maximum social welfare is achieved, however, for when the non-rivalrous goods is consumed by everyone valuing it with the maximum benefit is achieved.⁹⁴ Additionally, new property rights mostly cause conflicts and litigation in addition to ambiguity of the allocation of the right. Because of that data could be integrated and organized in various datasets it is hard to realize factual control independently over the data and so data ownership will decrease the transparency with high possibility and raise the risk of property right breaches unintentionally.⁹⁵

In any case, because of those big data situations are not systematically covered by the sui generis protection regulated in Database Directive, claims for indirect property rights on digital data could not be prevented.⁹⁶ However, in general, data ownership is not supported as a proper solution for protecting the big data analysts because access to data at others is more crucial to struggle with the velocity and veracity than claiming the breach of third parties' unauthorized use. Big data analysts focus on the technology by justifying protection of software due to copyright and probably know-how in technic instead of data ownership.⁹⁷ Therefore, data ownership still continues to have a debatable place in big data protection.

⁹³ Drexl, "Designing Competitive Markets for Industrial Data – Between Propertisation and Access", p. 274.

⁹⁴ OECD, "Data-Driven Innovation: Big Data for Growth and Well-Being", p. 196.

⁹⁵ Drexl, "Designing Competitive Markets for Industrial Data – Between Propertisation and Access", p. 275.

⁹⁶ Communication From the Commission to the European Parliament, The Council, The European Economic and Social Committee and The Committee of the Regions, "Towards a common European data space", p.6.

⁹⁷ Drexl, "Designing Competitive Markets for Industrial Data – Between Propertisation and Access", p. 273-274.

III. Unfair Competition

Art. 10bis(2) of the Paris Convention for the Protection of Industrial Property identifies the unfair competition as “*Any act of competition contrary to honest practices in industrial or commercial matters constitutes an act of unfair competition.*”. Unfair competition law protects competitors, consumers and public interest due to safeguarding competition.

Copyright or sui generis right protection for non-relational databases widely used in Big Data corpora could not be applied mostly. Therefore, misappropriation (tort-based) is prone to fill that gap particularly for AI system produced data which has eminent but nondurable value.⁹⁸ Some scholars regard not applying sui generis protection regime as the decrease on the possible monopolies of creating information and offers neighboring right or unfair competition law for a better balance.⁹⁹

On the other hand, unfair competition is regarded as unsubstantial protection base for the industrial data protection as it was not clearly and specifically identified and also is away from granting an exclusive right. Also, existence of various law schemes of unfair competition law globally and deficit of international agreements may harm the legal certainty of a legal protection.¹⁰⁰ Therefore, unfair competition does not have a broad place to be applied for big data protection.

⁹⁸ Gervais, “Exploring the interfaces between big data and intellectual property law”, p. 19.

⁹⁹ Helga Tabuchi, “International Protection Of Non-Original Databases; Studies on the Economic Impact of the Intellectual Property Protection of Non-Original Databases”, *Data Science Journal*, Volume 3, 2004, p.179.

¹⁰⁰ Andreas Wiebe, “Protection of Non-Personal Data: A New Legal Framework for Data Ownership?”, in Barczewski, M. (Ed.), *Value of Information: Intellectual Property, Privacy and Big Data*, Peter Lang GmbH, 2018, p. 13, 14.

IV. Contractual Protection

Database owner could search another way for protection in case the claim of copyright protection is rejected or the defendant achieved avoiding from copying the elements under copyright protection of database. Thus, the contractual breach could be put among the legal claims.¹⁰¹

The ECJ ruled in case the Database Directive is not applicable both for copyright and sui generis protection (Article 6(1), Art. 8 and Art. 15), without being against the applicable national law, the author of a database could benefit from contractual protection against use of own database by third parties who are lawful users of such a database authorized to use the database by the creator.¹⁰² Thus, the use of database is tried to be in balance between rights of database owner and lawful user.

On the other hand, natural persons or entities desiring to appropriate data may benefit from contractual terms introducing proprietary rights not having a ground in IP policy. That demonstrates the inadequacy of ownership concept due to contractual terms for control of competitive rights over big data.¹⁰³

Also, the contractual protection is regarded as insufficient protection by considering the unauthorized access to database. In that case because of contractual terms will be applied only to the parties of the contract, the intellectual property instrument should supplement the contract.¹⁰⁴ Thus, although data contracts could be referred as stronger than IPRs,

¹⁰¹ Jeff Langenderfer and Steven W. Kopp, "Which Way to the Revolution? The Consequences of Database Protection as a New Form of Intellectual Property", *Journal of Public Policy & Marketing*, 22(1), 2003, p.86.

¹⁰² Ryanair Ltd v PR Aviation BV. Case C-30/14. Judgment of the Court (Second Chamber) of 15 January 2015.

¹⁰³ Pamela Andanda, "Towards a Paradigm Shift in Governing Data Access and Related Intellectual Property Rights in Big Data and Health-Related Research", Springer, IIC (2019) 50, 2019, p. 1072.

¹⁰⁴ Mazumder, *Database Law Perspectives from India*, p.154.

IPRs could be claimed against everyone although contractual protection shall apply between parties of the contract.¹⁰⁵

V. IP Rights

The heart of any innovative business is connected with the intellectual capital. Benefiting it better will cause the future business success and failing to do will be resulting with lost.¹⁰⁶ Rens qualifies intellectual property law as the only strong international legal policy to control the information flow on the Internet although the increased attention is on the privacy regulation.¹⁰⁷

IPRs are seen as barriers for data reuse in case they limit the exploitation of datasets. However, they are also regarded as encouragers for data owners to share their IP protected data due to the better legal protection.¹⁰⁸

For the intangible goods the best type of protection is suggested as IPRs to provide clear ownership. It is also valid for data in particular, legal concepts such as copyright and other IPRs which are applicable to databases and also trade secrets could be utilized.¹⁰⁹ One of IPRs' major aim is to incent technological disclosure to accelerate the innovation.¹¹⁰ Although people are looking for the contractual protection as the IPR stands out as it provides enforcement against all people not just the parties as happened in the contract. Thus, IPRs could be regarded as most suitable protection way and, then, the issue of which IPRs protect big data raises.

¹⁰⁵ Kemp, "Legal aspects of managing Big Data", p. 488.

¹⁰⁶ De Wachter, "IP Strategy in the Age of Big Data", p. 1.

¹⁰⁷ Andrew Rens, "Who is in Charge Here? The Internet of Things, Governance and the Global Intellectual Property Regime", *UCLA Journal of Law & Technology*, Volume 23, Issue 2, 2019, p.18.

¹⁰⁸ Vrabec and Custers, "Legal barriers and enablers to big data reuse", p. 221.

¹⁰⁹ OECD, "Data-Driven Innovation: Big Data for Growth and Well-Being", p. 195.

¹¹⁰ Michael Mattioli, "Disclosing Big Data", *Minnesota Law Review*, 99(2), 2014, p.537.

While evaluating the IPRs, it is also stated that there is no specific solution for protection of big data in worldwide and the current intellectual property system could only enhance restricted protection for big data and not fix the issue in a comprehensive and effective way. Thus, further improvement on the de facto intellectual property system may provide an actual solution for big data protection as intellectual property by considering the challenges of big data protection.¹¹¹ Also, relation between the data and IPR is not certain as the law will continue to improve in that subject as throughout history the commercialization of the innovation has been followed by the development in IPRs. Thus, in a same way, while the big data gains value IPRs will support it.¹¹²

By considering the concerns with encouraging technological disclosure in the traditional law, intellectual property policy should be adopted to fix disclosure problem regarding big data.¹¹³ Original databases storing big data is under the protection of IPRs. Thus, mostly related IPRs on big data are copyright protection of original database and sui generis protection for database with substantive investment. Patent is not widely applicable as it could provide protection for computer software not the big data related concerns. Also, trade secret protection as a complementary to those IPRs will be examined.

A. Patent Protection

Patent protects an invention which is a new solution to existed technical problem, and provides exclusive rights to owner over his/her invention for the specified time period.

¹¹¹ Meng Lu, "Intellectual property protection of big data", J. Phys.: Conf. Ser. 1693 012012, 2020, <https://iopscience.iop.org/article/10.1088/1742-6596/1693/1/012012/pdf> (access date: 14.11.2022), p. 1.

¹¹² Kemp, "Legal aspects of managing Big Data", p. 487-488.

¹¹³ Mattioli, "Disclosing Big Data", p. 538.

Within intellectual property law, patent law is also encouraged to support the innovation, however, in terms of database, the patent may only provide protection partially for the software not the data itself. Although the big data analytical algorithms are patented successfully the value of the patent will be lost soon as the stated algorithms will rapidly change.¹¹⁴

Especially, as the concept of open source or cooperative ways is commonly and increasingly used, the patent protection could be regarded as out of topic. Also, obtaining, verifying or presenting the big data due to a database are not regarded as inventions and thus not being eligible for patent protection.

B. Copyright Protection

Copyright protects the author's own intellectual creation. The main categories of works for copyright are literary and artistic works. Books, poems, paintings, music computer and also some technological works including electronic databases could be granted with copyright protection in case of the originality. Thus, the database protection under copyright is the main subject of IP protection for big data.

Electronic database system comprises of the combination of three subsystems as “*a data storage system, a calculation or organisation system, and an information transmission system*”¹¹⁵. Bourgeois explains only aim of that design of the database as follows:

“The goal of many information systems is to transform data into information in order to generate knowledge that can be used for decision making. In order to do

¹¹⁴ De Wachter, “IP Strategy in the Age of Big Data”, p. 7.

¹¹⁵ Synodinou, “The Portability of Copyright-Protected Works in the EU”, p. 226.

this, the system must be able to take data, allow the user to put the data into context, and provide tools for aggregation and analysis”¹¹⁶.

Thus, together with the existence of originality on that design, the big data concept could be granted with the copyright protection. The detailed examination for copyright protection of database will be elaborated in next chapter.

C. Trade Secret Protection

Businesses look for protection methods for their innovations against exploitation and those are, in general, copyright, patent or design rights. Another tool of protection could be trade secret to prevent access, exploitation of valuable and not widely known information. Such confidential know-how and business information is named as ‘trade secret’. WIPO classifies trade secrets as one of the IPRs and states that “*Trade secrets are intellectual property (IP) rights on confidential information which may be sold or licensed.*”¹¹⁷.

The Trade Secret Directive refers the trade secret protection as a ‘complement’ or an ‘alternative’ to IPRs.¹¹⁸ EU adopted “*Directive (EU) 2016/943 of the European Parliament and the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use or disclosure*” (‘Trade Secret Directive’) to regulate the trade secret protection, however, the Trade Secret Directive is commented as outdated as in the time of adoption the new data economy still were not comprehended fully. Thus, the content of the Directive is not

¹¹⁶ David T. Bourgeois, James L. Smith, Shouhong Wang and Joseph Mortati, *Information Systems for Business and Beyond*, p. 68.

¹¹⁷ WIPO, Frequently Asked Questions: Trade Secrets. https://www.wipo.int/tradesecrets/en/tradesecrets_faqs.html#:~:text=Trade%20secrets%20are%20intellectual%20property,may%20be%20sold%20or%20licensed (access date: 14.11.2022).

¹¹⁸ The Trade Secret Directive, Recital 1, 2.

sufficiently clear for example about the application of trade secret protection for the data generated by smart devices.¹¹⁹ Confidential databases could be protected under trade secret protection in Turkish Commercial Code. Article 55 of Turkish Commercial Code exemplifies the unfair competition cases and one of those is written in 55(1)-d as “*disclosure of production or confidential information illegally*”.

Copyright protection shall apply only for the expressions not for ideas, so others are allowed to use the idea such as IoT solutions. Thus, for the companies, trade secret may provide better protection rather than copyright protection by comparing their investment for IoT solutions.¹²⁰ However, some scholars put trade secret protection at the second plan, for example, Drexl evaluates trade secret protection less proper way as a kind of tort liability to protect data than sui generis database protection or possible data ownership right.¹²¹ Thus, trade secret protection regarded as controversial topic for big data protection.

Art. 39 of TRIPS Agreement specifies the minimum standards for the protection of undisclosed information with three points: (i) it should be secret, (ii) it should have commercial value due to having secrecy and (iii) reasonable precautions should be taken, by the person legally in charge of control over that information to keep it as secret. Art. 2(1) of the Trade Secret Directive presents the same standards on trade secret.

Art. 2(1) of the Trade Secret Directive is hardly available for big data, because the minimum standards for the protection of undisclosed information to be secret, having commercial value and reasonable precautions taken to keep it as secret are the problematic issues in terms of big data. The various accesses to big data considering the connection

¹¹⁹ Drexl, “Designing Competitive Markets for Industrial Data – Between Propertisation and Access”, p. 268-269.

¹²⁰ Storr and Storr, “Internet of Things: Right to Data from a European Perspective”, p. 77.

¹²¹ Drexl, “Data Access and Control in the Era of Connected Devices”, p.91.

of smart devices could not meet the requirement of secrecy and the necessary precautions should be taken by the trade secret owner.¹²² Also, although the data has commercial value it could be hard to connect the value with the secrecy of the data as the value of the data appears when it is correlated with other data in the concept of big data analyses.

Trade Secrets Directive does not establish a property right system as it only brings liability for specific tortious treatment. It is stated that trade secret protection is better suited for data economy because third parties could acquire access to data, not providing exclusive protection against data use.¹²³ Also, trade secret right and sui generis right may exist together as cumulative protection accepted and applied commonly in IP. When comparison is made, sui generis protection is not valid for raw facts or data and applicable for the substantive part of a database. Trade Secret Directive protects the commercial value of trade secret resulting from its secrecy although sui generis rights even protect publicly available databases.¹²⁴

D. Sui Generis Database Protection

There is no international mechanism or norm to provide legal protection for nonoriginal databases. A new right, sui generis database right, is identified for databases whose elements are permanently changing in real time such as data gathered by cars for the intensity of the streets.¹²⁵

¹²² See Drexl, "Designing Competitive Markets for Industrial Data – Between Propertisation and Access", p. 269: Drexl gives the example of that data collected by smart cars on freely accessible roads could be collected by the cars of many manufacturers.

¹²³ Drexl, p. 269.

¹²⁴ European Commission, "Evaluation of Directive 96/9/EC on the legal protection of databases", p.44.

¹²⁵ Drexl, "Data Access and Control in the Era of Connected Devices", p.76.

Sui generis database protection requires substantive investment for obtaining, verifying or presenting the contents in a qualitative and/or quantitative way. It could be explained that although copyright preserves “original structure” with other meaning, “skeleton” of database; sui generis protection applies for the “body” of database.¹²⁶

Sui generis protection could be applied for marketing databases, sporting event record or databases of rainfall from specific locations and most probably does not cover a list of runners.¹²⁷ Thus, by comparing with the examples of possible copyrighted databases, the scope is broader to achieve sui generis database protection. Sui generis database protection will be discussed in final chapter detailly.

¹²⁶ Banterle, “Data Ownership in the Data Economy: A European Dilemma”, p. 224.

¹²⁷ Wilks and Christie (Eds.), “IP Rights in Data Handbook: Protecting and exploiting IP in data, big data and databases internationally”, p. 30.

CHAPTER III

COPYRIGHT PROTECTION

Big data requires database to analyze and extract the value from it. Databases are used largely in various activities and industries in the EU Community rather than information market as a crucial element for the development of that market.¹²⁸ Investment in production of databases can be increased due to law providing the protection for intellectual effort because together with extensive compilation of information in consequence of database protection, society's abilities to solve the problems will be enhanced.¹²⁹ Also, appropriation of a database as free of charge could cause a loss or decrease on the incentive to investment for databases so lack of databases or reduced quality of current databases should be prevented.¹³⁰

“Collection” and “independent” as keywords used generally in the definition of a database shows that database should be comprised a number of separate data.¹³¹ The organization of information in large compilations for easy reference with minimum input is achieved due to the electronic databases which are the backbones of Internet. Unlike the random information, databases provide users the opportunity of easily selection and manipulation on what they looking for. However, these beneficial qualities of databases strength its protection as the duplication and distribution of the digital information over the Internet which is quite easy.¹³² Therefore, copyright is on the agenda to enhance incentive for database makers and protect their rights. Copyright protection for a database is provided with international agreement by using the various terms for it such as “*Collections ... by*

¹²⁸ Terence Prime, “International Trade and the Rule of Law: The Case of European Protection of Computer and Databases”, *Ankara Üniversitesi SBF Dergisi*, 54-4, 2015, p.136.

¹²⁹ Lim, “Re-defining the rights and responsibilities of database owners under competition law”, p.421.

¹³⁰ Mark J. Davison, *The Legal Protection of Databases*, Cambridge University Press, 2003, p.239.

¹³¹ Grosheide, “Database Protection—The European Way”, p.51.

¹³² Langenderfer and Kopp, “Which Way to the Revolution? The Consequences of Database Protection as a New Form of Intellectual Property”, p.83.

reason of the selection and arrangement of their contents” or “*compilations of data*”. EU Law and Turkish Law by adopting EU’s strategy for database protection provide copyright protection with directly using the term of “database”.

Leistner comments that a database work including typical creativity in the advancement of AI models and applications grounded on big data through the structure and weigh of the cost functions, choose and combination of training data etc. could be granted with copyright protection under EU copyright law.¹³³ The copyright protection is regarded as a powerful incentive for database works.

On the other hand, some scholars name copyright protection for databases as “thin protection”¹³⁴, because it does not protect the facts and the data from different sources could be used by others for new compilations. Database protection is provided to encourage the innovation in this subject, however, also the access of user to that database is not prevented completely. In that point, intellectual property law regulates protection for database under copyright, but it brings also exceptions. However, the application of exceptions differs in jurisdictions.

I. Copyright Protection of Database in International Agreements

The necessity of protection of database in international level resulted with the regulations in international agreements with regard to copyright. Although there is no clear attribution to database in Berne Convention for the Protection of Literary and Artistic Works, precise inclusion of database to copyright protection in Agreement on Trade-Related Aspects of

¹³³ Leistner, “Protection of and Access to Data under European Law”, p. 386.

¹³⁴ Ray K. Harris, Susan Stone Rosenfield, “Copyright Protection for Genetic Databases”, *Jurimetrics*, Vol. 45, No. 2, American Bar Association, 2005, p. 239.

Intellectual Property Rights (TRIPS Agreement) and WIPO Copyright Treaty (WCT) provided copyright protection for databases.

Berne Convention does not explicitly list database as one of the examples for collections under copyright protection. In 1967, while having last revision on substance of Berne Convention, electronic databases open to public access or database software of mass-market were not available and so the collections referred in Convention by drafters were paper-based anthologies and encyclopedias.¹³⁵ Although the term of “database” is not used, it is accepted that Art. 2/5 includes databases as the collections of data:

“Collections of literary or artistic works such as encyclopaedias and anthologies which, by reason of the selection and arrangement of their contents, constitute intellectual creations shall be protected as such, without prejudice to the copyright in each of the works forming part of such collections.”

The most crucial point for the copyright protection of databases is originality as also stated in the Berne Convention. Although the database includes copyrighted works it has no effect on copyright protection of database.¹³⁶

Art. 10/2 of TRIPS Agreement provides the copyright protection for compilations of data, without a required form such as machine readable or not, selecting or arranging its contents should be regarded as intellectual creations. Appearance of electronic databases happened in 1990s and to achieve usable version data must be indexed and re-indexed in general. The language of TRIPS Agreement is a response of that advancement.¹³⁷

Additionally, TRIPS mentions the protection could not be extended to materials in database and not weaken contents’ own copyright protection if any. Andanda approaches

¹³⁵ Gervais, “Exploring the interfaces between big data and intellectual property law”, p. 7.

¹³⁶ Grosheide, “Database Protection—The European Way”, p.43.

¹³⁷ Gervais, “Exploring the interfaces between big data and intellectual property law”, p. 8

the copyright protection of big data in terms of technical challenges by mentioning the TRIPS Agreement and states that “Article 10(2) of the TRIPS Agreement requires the data to be compiled in machine readable or other form. Notably, the format of big data involves dynamic data sets and uses cloud computing services, which technically makes it difficult to meet this requirement.”¹³⁸.

WCT is consistent with TRIPS Agreement and provide copyright protection for compilations of data constituted as intellectual creations. WCT also explains the copyright protection of databases in Art. 5 under the heading of “*Compilations of Data (Databases)*”. The expressions for compilations in the TRIPS Agreement and WCT are adequately broad to cover automated electronic databases.¹³⁹

There was an attempt for a treaty in favor of database protection before WIPO. As the Directive does not grant protection for U.S. firms with the reason of that the U.S. law has not comparable protection for databases, international treaty for the database protection was offered to WIPO.¹⁴⁰ However, the concerns about the impact of database treaty on free flow of scientific information stated by global organizations such as World Meteorological Organization (WMO), UNESCO, International Council for Science (ICSU) or American National Research Council and international treaty including the sui generis database protection could not concluded.¹⁴¹ Thus, there is no international agreement on database protection in the current system.

¹³⁸ Andanda, “Towards a Paradigm Shift in Governing Data Access and Related Intellectual Property Rights in Big Data and Health-Related Research”, p.1066.

¹³⁹ Paul Goldstein and P. Bernt Hugenholtz, *International Copyright: Principles, Law, and Practice*, Oxford University Press, 2001, p. 175.

¹⁴⁰ Harris and Rosenfield, “Copyright Protection for Genetic Databases”, p. 250.

¹⁴¹ Davison, “The Legal Protection of Databases”, p.235.

II. EU Law

A. The Database Directive

Although the Berne Convention, TRIPS Agreement and WCT provides copyright protection for databases but only for original works, “*Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases*” (the Database Directive) applies both copyright and sui generis protection for databases. The global harmonization for the Database Directive purposes the balance the interests of “*database makers, database operators, users, authors, competitors, SEES, and the public at large*”¹⁴².

The Database Directive has three goals which are the harmonization of database protection among EU Member States, inciting to invest in database creation to advance the ability of European digital industry in the international competition and ensure the balance between interests of users and makers of database.¹⁴³

The Copyright part of the Database Directive starts with the criteria of “*originality*” continues with “*restricted acts*” and “*exceptions to restricted acts*”. The implementation of Database Directive should have been completed by Member States until 1 January 1998.

B. The Subject Matter of Database Directive

The one of major successes due to adoption of Database Directive is mentioned as harmonization among Member States over the elements for copyright protection of a

¹⁴² Jens L. Gaster, “The New EU Directive Concerning the Legal Protection of Databases”, *Fordham International Law Journal*, Volume 20, Issue 4, 1996, p. 1149.

¹⁴³ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report”, p. (ii).

database. Both nonelectronic and electronic databases are covered in the Database Directive which brings the criteria of “*by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation*”¹⁴⁴ for copyright and does not accept other condition. However, it is specially stated in the same article that copyright protection conferred upon by the Database Directive does not cover its contents and that protection shall not be prejudice to any rights existed for those contents.

The Database Directive tries to keep the balance between the interests of database author and database user. Copyright protection for database begins “*with the systematic or methodical arrangement of the collection, followed by the selection or arrangement of the contents.*”¹⁴⁵. The database’s internal structure could not be seen by the end user and surpass over other kind of data storage due to creative choices. That structure is crucial for the achievement of database in commercial way as it has an effect on the access speed to the content of database. Also, database’s external structure including the search tools, graphical signs etc. contributes the decent operation. In that case, copyright provides protection to the database made with author’s own intellectual creation through creative choices of him/her, not imposed because of functional or technical restraints.¹⁴⁶

C. Database Authorship

Art. 4 of the Database Directive uses term of “*database authorship*” and attributes the author of the database to “*the natural person or group of natural persons who created the base*”. Also, it states a discretion for member authorities to designate legal person as the rightsholder or not in its legislation. For the copyright protection of database by whom,

¹⁴⁴ The Database Directive, Art. 3.

¹⁴⁵ Grosheide, “Database Protection—The European Way”, p. 52.

¹⁴⁶ Tatiana-Eleni Synodinou, “The Portability of Copyright-Protected Works in the EU”, p. 227.

how or when the data is put in the database is insignificant as it is same for that the intellectual creation of users while generating this content is not important by evaluating the database protection.¹⁴⁷

D. Restricted Acts and Exceptions to Restricted Acts

Restricted acts are specified in Art. 5 in the Database Directive and database author has a right to exclude others from reproduction, translation, adaptation, arrangement, distribution, any communication, display or performance to the public. National law and Berne Convention shall apply for moral rights so moral rights are out of the scope of Database Directive.¹⁴⁸

The exceptions for copyright protection of databases differ from the copyright regulation of other works. Lawful user is authorized to use the part of a database and options are provided to Member States to limit the “Restricted Acts”.¹⁴⁹ Reproduction for private use is only allowed for non-electronic databases. Teaching, scientific research, public security, administrative or judiciary proceeding are listed as kinds of exemptions in Art. 6(2) of the Database Directive.

However, exceptions are found too narrow by the stakeholders. Restricting the private use exemption with the nonelectronic databases and nonexistence of text and data mining exemption are mainly criticized.¹⁵⁰ Leistner also finds those exceptions insufficient for the access to data.¹⁵¹ Thus, the “*Directive (Eu) 2019/790 Of The European Parliament*

¹⁴⁷ Tatiana-Eleni Synodinou, “The Portability of Copyright-Protected Works in the EU”, p. 228.

¹⁴⁸ The Database Directive, Recital 28.

¹⁴⁹ The Database Directive, Art. 6.

¹⁵⁰ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report”, p. 14.

¹⁵¹ Leistner, “Protection of and Access to Data under European Law”, p. 387.

And Of The Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC” (‘DSM Directive’) regulates exceptions and restrictions of copyright and related rights to prevent the exploitation of works and other subject matters.

Art. 3(1) of the DSM Directive for “*reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access.*” and Art. 4 of the DSM Directive for “*reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining.*” regulates exceptions with regard to the Art. 5(a) and Art. 7(1) of the Database Directive. Indeed, Art. 3 and Art. 4 of the DSM Directive states an obligation for member states to provide those exceptions which are the means to provide a balance between database owner and user as stated in the Art. 1 of DSM Directive. Therefore, the exceptions to database protection are extended with the DSM Directive in favor of text and data mining.

Leistner defenses that European law enables balanced and sensible legal structure by considering the copyright protection for database works. On the contrary, Hugenholtz states the requirement to take measures for “*promoting access to data and fostering data mining rather than commodification of data by creating property rights in data.*”¹⁵². Thus, with regard to exception for copyright protection, the need of reform appears as even recent exceptions for text and data mining regulated in the DSM Directive are already in an outdated position and inadequate to handle extensively with the difficulties and opportunities of the data economy.¹⁵³

¹⁵² Hugenholtz, “Against ‘Data Property’”, p.14.

¹⁵³ Leistner, “Protection of and Access to Data under European Law”, p. 399.

Especially for the scientific researches there is a concern for the access to useful databases. With regard to genetic databases, the need of balance between incenting the improvement of beneficial databases and free reach for scientific facts is stated.¹⁵⁴

E. Term of Protection

The minimum period of copyright protection is determined as at least 50 years after the author's death in both Berne Convention and TRIPS Agreement. Art. 1(1) of the "Directive 2006/116/EC of The European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (codified version)" with the heading of "Duration of Authors' Rights" brings 70 years protection in addition to author's life period. That time period is also valid for the copyright protection of databases.

III. Turkish Law

In the "Decision No 1/95 of the EC-Turkey Association Council of 22 December 1995 on implementing the final phase of the Customs Union", Turkey shall adapt the Turkish legislation for the protection of database to align it to the EU legislation¹⁵⁵.

Article 6(11) of the Law on Intellectual and Artistic Works No. 5846 identifies the database as "*obtained by the selection and compilation of data and materials according to a specific purpose and a specific plan, which are in a form that can be read by a device or in any other form*". It was also stated in a same way with the Database Directive that

¹⁵⁴ Harris and Rosenfield, "Copyright Protection for Genetic Databases", p. 250.

¹⁵⁵ Decision No 1/95 of the EC-Turkey Association Council of 22 December 1995 on implementing the final phase of the Customs Union. [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:21996D0213\(01\)&from=en](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:21996D0213(01)&from=en) (access date: 14.11.2022).

this protection shall not extend to data materials in database. The Law puts database as a work under the heading of “Adaptations and Collections” in Art. 6 and lists it as one kind of collections.

Copyright protection should be provided for original expression as a result of selecting, coordinating or arranging facts in database not for the facts themselves.¹⁵⁶ However, although the Directive accepts the selection or arrangement of database’s content for the evaluation of originality, Art. 6(11) of the Law No. 5846 requires the author’s own intellectual creation for both selection and arrangement. The Court of Cassation ruled that the characteristic of author over a database, a kind of collection work, exists through the selection and arrangement of data and materials used and it should be appeared on the entire database.¹⁵⁷ There is no special provision of exceptions or limitations stated for the database protection in Law No. 5846. Therefore, similar criticisms over the Database Directive are also valid for database protection in Turkish Law.

IV. Big Data Issues for Copyright Protection

Copyright protection for databases should not be in a lower threshold, however, databases reflecting the author’s own creativity genuinely could be granted for copyright protection. That requirement of originality significantly decreases the applicability of copyright protection over big data concept, but also may provide exclusive protection for prominent accomplishments to select and/or combine the various training data, algorithms and predominant factors.¹⁵⁸

¹⁵⁶ Cahit Suluk and Ali Orhan, *Uygulamalı Fikri Mülkiyet Hukuku*, Arıkan Yayıncılık, 2005, p. 266.

¹⁵⁷ Republic of Türkiye, Court of Cassation 11th Civil Chamber, E. 2005/3506, K.2006/4425, 13.04.2006.

¹⁵⁸ Leistner, “Protection of and Access to Data under European Law”, p. 384.

Brand has a famous quote for the dilemma of value of the information which could be commented in a same way for the big data:

*“Information wants to be free because it has become so cheap to distribute, copy, and recombine—too cheap to meter. It wants to be expensive because it can be immeasurably valuable to the recipient. That tension will not go away. It leads to endless wrenching debate about price, copyright, “intellectual property,” and the moral rightness of casual distribution, because each round of new devices makes the tension worse, not better.”*¹⁵⁹

That dilemma keeps the database protection as the subject of common arguments in data industry. For example, it is envisaged providing copyright protection for original databases could be even more prominent than sui generis protection to specify the data economy structure although it is not expected today. As original structures or precise meta-data organizing pre-existing data pools with an exact perspective may be granted for copyright.¹⁶⁰ On the other hand, because of the special technological structure of the database, copyright is evaluated as hardly applicable to address a database protection.¹⁶¹

Additionally, the other problematic way of copyright protection is explained with that copyright is improper as the database has its specific value from comprehensiveness. Lack of creativity or originality appears in case the existence of more complete database and it results with the going away from thick copyright protection.¹⁶² Therefore, while struggling for the volume of big data, the application of copyright protection could be out of place.

¹⁵⁹ Stewart Brand, *The Media Lab : Inventing the Future at M.I.T.*, Penguin Books, 1998.

¹⁶⁰ Matthies Leistner, “Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform”, 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3245937, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3245937 (access date: 14.11.2022), p. 16.

¹⁶¹ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p.14.

¹⁶² Grosheide, “Database Protection—The European Way”, p. 43.

A. Originality

The word of original is identified as “*not the same as anything or anyone else and therefore special and interesting*”.¹⁶³ Copyright protection requires being original, however it is a kind of work with the human creation. The meaning of originality is not explained in the Berne Convention. Dreier and Karnell ask the question of that “*Did the creator bring it all - or at least enough of it – out of his or her own mind, more or less unique as that may prove to be, or where did he or she take it?*”¹⁶⁴ to decide about the originality of work. The original can be understood as new, however everything new is not original although it is regarded as literary or artistic, the new should express the personality to be original.¹⁶⁵

Examples of original databases could be marketing databases having a complex structure or an anthology including poems classified by theme. Directories of telephone or record of sporting event is not listed as original databases to provide copyright protection.¹⁶⁶ Therefore, a database is subject of intellectual property law and it could be under protection of copyright in case the required originality exists.

The Database Directive requires only “originality” for the copyright protection of databases and clearly states that other condition shall not be needed. “*databases which, by reason of the selection or arrangement of their contents, constitute the author’s own intellectual creation shall be protected as such by copyright.*”. It could be subjectively new not evaluating with the rigid evaluation and may be attributed to copyright. Also, the originality mostly comes with being new in case of the existence of expressing personality

¹⁶³ Cambridge Dictionary, <https://dictionary.cambridge.org/dictionary/english/original> (access date: 14.11.2022).

¹⁶⁴ Thomas Dreier and Gunnar W. G. Karnell, “Originality of the copyrighted work: European perspective”, *Journal of the Copyright Society of the U.S.A.*, 39(4), 1992, p.291.

¹⁶⁵ Dreier and Karnell, p.290.

¹⁶⁶ John Wilks and Alec Christie (Eds.), “IP Rights in Data Handbook: Protecting and exploiting IP in data, big data and databases internationally”, *DLA Piper*, 2014, p. 30.

so being new does not mean having originality.¹⁶⁷ ECJ explains the criteria of originality for a database as follows:

*“As regards the setting up of a database, that criterion of originality is satisfied when, through the selection or arrangement of the data which it contains, its author expresses his creative ability in an original manner by making free and creative choices ... and thus stamps his ‘personal touch’”.*¹⁶⁸

Thus, the personal touch for copyright protection having originality could be the main issue for copyright protection of databases. However, that originality should exist for the selection of the data not finding out the content.

It is the same ruling in Case C-604/10 (*Football Dataco v Yahoo!*), creating data whether through an intellectual effort or skill is not acceptable for the evaluation of copyright protection of database and without originality in the data selection or arrangement of the data included in the database, the copyright protection is not applicable. ECJ explains the aim of the Database Directive in this case as following:

*“its purpose is to stimulate the creation of data storage and processing systems in order to contribute to the development of an information market against a background of exponential growth in the amount of information generated and processed annually in all sectors of activity ... and not to protect the creation of materials capable of being collected in a database.”*¹⁶⁹

Generating inferred data by selecting and combining data, making structure and weigh of cost functions or structures for systematic can be granted as “creative” database in the

¹⁶⁷ Dreier and Karnell, “Originality of the copyrighted work: European perspective”, p.290.

¹⁶⁸ *Football Dataco Ltd and Others v Yahoo! UK Ltd and Others*. Case C-604/10, 2012. Judgment of the Court (Third Chamber) of 1 March 2012, para 38.

¹⁶⁹ *Football Dataco Ltd and Others v Yahoo! UK Ltd and Others*, para. 34.

concept of AI and big data. However, higher standards for originality requires prominent selection and combination of data to pass the threshold and achieve copyright protection.¹⁷⁰

Copyright protection may attribute for deciding the data will be gathered in the database, organizing the data and associating those various data, however, that protection may have a thin layer for databases.¹⁷¹ Therefore, even the substantive investment is made for creating that database, unless required originality exists, the copyright protection shall not apply for the database.

Prof. Franz van Isacker states “*Originality does, by words of, of Belgium, does share the fate of the pink elephant, of being easy to recognize but difficult to define.*”. This wording is justified more for the evaluation of the originality with regard to copyright protection of database.¹⁷² The requirement of originality is investigated in a stricter way than being original is just not sufficient, selecting or arranging contents of database must be author’s own intellectual creation.

If the data is not processed it is hard to find the useful information for end users. McManis advises to support the analysis and refinement of data instead of simply gathering it, however the protection should be enhanced for the compilations providing novel and creative organization and so adding value to the existing facts.¹⁷³ Thus, the original selection and arrangement could be subject to copyright protection.

¹⁷⁰ Matthias Leistner and Lucie Antoine, “IPR and the use of open data and data sharing initiatives by public and private actors”, European Parliament, 2022, [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf) (access date: 15.11.2022), p. 49.

¹⁷¹ Research Data Management Service Group, “Introduction to intellectual property rights in data Management”.

¹⁷² Dreier and Karnell, “Originality of the copyrighted work: European perspective”, p.289.

¹⁷³ Charles R. McManis, “Database protection in the digital information age”, *Roger Williams University Law Review*, 7(1), 2001, p.28.

As databases usually are lack of sufficient creativity for copyright protection, there is unambiguity with the reason of necessity for the creativity level.¹⁷⁴ However, considering the present regime, technological solutions are increasingly preferred by database creators to protect themselves against the possibility of copyright protection is failed. Therefore, the implementation of encryption algorithms for the protection of data by preventing the access arises its popularity.¹⁷⁵

In reality, creativity or human intellection could not appear mostly in the extensive and systematic databases. However, for the databases achieving to pass the threshold of the originality should be granted of copyright protection to foster maximum benefit from big data analysis.

B. Human Intervention in Second Degree

Machine generated data could be personal or nonpersonal data created by computer process, application or sensor receiving information from machinery, hardware and/or software without direct human intervention.¹⁷⁶ The point of “author’s own creation” is hardly exist for machine generated data.

The ambiguous areas in database protection increases with the advancement in technology. It is still argued that the machine generated databases may be benefited from database protection or not. There are also strong views against the data protection due to

¹⁷⁴ Leonard D. DuBoff, “Introduction to computer law”, *Hastings Communications and Entertainment Law Journal* (Comm/Ent), 14(2), 1991, p. 233.

¹⁷⁵ Langenderfer and Kopp, “Which Way to the Revolution? The Consequences of Database Protection as a New Form of Intellectual Property”, p.92.

¹⁷⁶ European Commission, “Building a European Data Economy”, {SWD(2017) 2 final}, p. 9.

IPRs as the data is the demonstration of the observed and temperate facts than a human's intellectual creation.¹⁷⁷

This position, considering authors doing essential arrangements to generate computer-created works, interestingly contrasts with the extent of the European copyright law. Nimmer states that “*The more comprehensive a data collection becomes, the harder it is to protect it via copyright.*”¹⁷⁸. However, it is also hard to eliminate copyright protection in today's fast growing data economy without having the supportive power of copyright protection in terms of the innovation.

It is indisputable that machine learning puts the human intervention to the second degree for the work. Therefore, the base notions of IP law, specifically for copyright law in terms of authorship, challenges the separation between human contribution and outputs of the system to evaluate the probable IP protection.¹⁷⁹

¹⁷⁷ Taina Pihlajarinne and Rosa Maria Ballardini, “Owning Data via Intellectual Property Rights: Reality or Chimera?”, in Ballardini, R. M., Kuommamäki, P., Pitkänen, O. (Ed), *Regulating Industrial Internet Through IPR, Data Protection and Competition Law*, Kluwer Law International, 2019, p.16.

¹⁷⁸ Harris and Rosenfield, “Copyright Protection for Genetic Databases”, p. 231: Nimmer & Nimmer, supra note 29, 3.04[B][2], at 3-33.

¹⁷⁹ Gervais, “Exploring the interfaces between big data and intellectual property law”, p.5

CHAPTER IV

SUI GENERIS DATABASE PROTECTION

Although the legal protection of big data has complicated points for sui generis database right it is regarded as the best option to support advancements in databases due to substantive investments. The considerable amount of time and effort could be sufficient for the sui generis database protection even for excessive amounts of data. The most proper way with regard to the IoT database protection is considered as sui generis protection.¹⁸⁰

Davison refers arguably sui generis protection provided for a database as another name of copyright protection. As same actions could constitute both authorship for copyright and being database maker for sui generis right. Additionally, the test for sui generis protection is same with the sweat of the brow within copyright policy and, furthermore, sui generis right is granted for intellectual creativity through “selection and arrangement of a database” or coexistence of “sweat of the brow and intellectual creation”. Finally, the rights provided under sui generis protection are the same with those of copyright and exceptions are also very similar, even identical in a number of cases.¹⁸¹

Although the Database Directive adopts two-tier system for legal protection, sui generis right provides its uniqueness, as copyright protection shall not apply for most databases even the low standards for originality may exists. The difference between copyright and sui generis right provided in the Database Directive is explained by Grosheide that copyright breach appears when the structure is copied, however, breach against sui generis protection is determined in case of copying the contents without considering their

¹⁸⁰ Storr and Storr, “Internet of Things: Right to Data from a European Perspective”, p. 76.

¹⁸¹ Davison, “The Legal Protection of Databases”, p.224.

copyrightability¹⁸². Also, both copyright and sui generis protection exists together in a concrete case and those rights could be benefited independently.¹⁸³ On the other hand, the example of CD is given in the Directive for not being applicability of those two rights. CD including some music performances is not protected through Directive both in terms of copyright and sui generis right respectively because of not being original and not a result of substantive investment.¹⁸⁴

Consequently, with the adoption of Database Directive, EU leads the sui generis protection for database although it could not pass the threshold of originality, due to the substantive investment it will be granted for the protection.

I. EU Law

A. The Database Directive

U.S. Supreme Court ruling in “Feist Publications Inc. v. Rural Telephone Service Co. Inc,” (Feist decision) is the cornerstone for the database protection in both U.S. and EU as this decision promoted the perspective of sui generis protection in terms of EU Commission. U.S. Copyright Office connects the urgency of the issue with regard to database protection in the 90s with the Feist decision among the changes of legal area, technological developments and international outlook. Because of rapid advancements in technology to gather, classify, reproduce and disseminate information and EU’s actions to harmonize the legal applications of members states, in 1996 the evaluation over new

¹⁸² Grosheide, “Database Protection—The European Way”, p.39.

¹⁸³ Grosheide, p. 39.

¹⁸⁴ The Database Directive, Recital (19).

form for database protection was seen both before WIPO and U.S. Congress¹⁸⁵. In that decision the Supreme Court states as follows¹⁸⁶:

“The selection, coordination, and arrangement of Rural's white pages do not satisfy the minimum constitutional standards for copyright protection. ... Rural's white pages are entirely typical. Persons desiring telephone service in Rural's service area fill out an application and Rural issues them a telephone number. In preparing its white pages, Rural simply takes the data provided by its subscribers and lists it alphabetically by surname. The end product is a garden-variety white pages directory, devoid of even the slightest trace of creativity.”

The Court did not accept the claim for breach of copyright as the telephone directory does not provide required level of originality for copyright protection by considering the choice or setting of its contents.

“Sweat of the brow” principle had also an effect on some European countries including England and Ireland for the copyright protection of database, then it replaced with the higher standard of originality with Feist. Lower Court decision for the “Feist Publications Inc v Rural Telephone Service Company Inc” was reversed by requiring originality of selection and arrangement which is away from the standard of sweat of brow.¹⁸⁷ Also, it resulted with a new database right due to the Directive to enhance protection for investment in database creations in Europe.¹⁸⁸

¹⁸⁵ US Copyright Office, “Report on Legal Protection for Databases”, 1997, <https://www.copyright.gov/reports/db4.pdf> (access date: 14.11.2022), p. 1.

¹⁸⁶ Feist Publications, Inc. v. Rural Telephone Service Co., Inc. No. 89-1909. Supreme Court of the United States (10th Circuit), 1990, <https://cyber.harvard.edu/people/tfisher/1991%20Feist.pdf>.

¹⁸⁷ Prime, “International Trade and the Rule of Law: The Case of European Protection of Computer and Databases”, p.137.

¹⁸⁸ Marc J. Davison, “Database Protection: Lessons From Europe, Congress, and WIPO”, *Case Western Reserve Law Review*, Vol. 57:4, 2016, p. 831.

The Commission of the European Communities mentioned the need of database protection to provide harmonization among member states firstly in “Green Paper on Copyright and The Challenge of Technology” in 1988.¹⁸⁹ This Paper is published in the Official Journal of the European Communities on 20 March 1989. Chapter VI with the heading of “Databases” clearly express that database having the appropriate level of originality should be protected under copyright as compilations and taking an action to make it definite is necessary. Although this paper mentioned also the sui generis protection, Commission of the European Communities did not have an approach for sui generis protection in “Follow-up to the Green Paper on Copyright and the Challenge of Technology”.¹⁹⁰

The Commission realized that only copyright does not provide sufficient protection for the factual databases with substantive investment. As such databases could not meet the originality requirement in most of the Member States although the existence of great ability and effort for the compilation. Database makers’ objective is to generate comprehensive database not an original one especially for electronic databases so those are not the result of intellectual creation in the fact.¹⁹¹

The information generated in a huge amount in today’s world requires also an investment for the professional information processing abilities. Thus, the need of legal protection regime for such advanced systems is tried to be founded through Directive with a stable and common protection for database makers.¹⁹²

¹⁸⁹ European Commission, Green Paper on Copyright and the Challenge of Technology, Copyright Issues Requiring Immediate Action, (COM (88) 172 final) 176, 1988.

¹⁹⁰ Grosheide, “Database Protection—The European Way”, p.47.

¹⁹¹ Mark Powell, “The European Union’s Database Directive: An International Antidote to the Side Effects of Feist?”, *Fordham International Law Journal*, Volume 20, Issue 4, 1996, p.1220.

¹⁹² Directive, Recital 9,10,12.

Although the database industry in U.S. dominates others, EU took the initiative for creation of an intellectual property right for data.¹⁹³ Powell associates the sui generis protection provided in the Directive with the introduction of “sweat of the brow” doctrine in European side.¹⁹⁴ That doctrine attaches importance to the endeavor and investment made by the compiler.¹⁹⁵ Thus, it does not require creativity for the selection and creation of the information and only attach importance to the labor invested to collect, present and verify the data.¹⁹⁶

Recital 40 and Recital 41 state the object of sui generis protection as follows:

“(40) Whereas the object of this sui generis right is to ensure protection of any investment in obtaining, verifying or presenting the contents of a database for the limited duration of the right; (41) Whereas the objective of the sui generis right is to give the maker of a data-base the option of preventing the unauthorized extraction and/or re-utilization of all or a substantial part of the contents of that database,”

The Database Directive presents sui generis protection for a database maker executing substantial investment in a quality and/or quantity for either “*the obtaining, verification or presentation*” of the contents.¹⁹⁷ Gaster did not put database protection under the related rights. He identifies this new right as independent category also not interesting the organizational success as realized for the Rome Convention.¹⁹⁸

¹⁹³ McManis, “Database protection in the digital information age”, p.8.

¹⁹⁴ Powell, “The European Union’s Database Directive: An International Antidote to the Side Effects of Feist?”, p.1226.

¹⁹⁵ US Copyright Office, “Report on Legal Protection for Databases”, p. 3.

¹⁹⁶ Davison, “Database Protection: Lessons From Europe, Congress, and WIPO”, p. 833.

¹⁹⁷ The Database Directive, Art. 7(1).

¹⁹⁸ Gaster, “The New EU Directive Concerning the Legal Protection of Databases”, p. 1143.

B. Exclusive Rights and Exceptions to Those Rights

The wording used in the Database Directive to indicate the purpose of protection supports that sui generis right could be regarded as a kind of indirect property in data. Due to sui generis protection the database maker may prevent ‘extraction and/or reutilization’ of substantial part of the database content.¹⁹⁹

Directive grants database maker to prevent “extraction” and/or “reutilization”. Art. 7(2)-a defines the extraction as “*the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form*”. Reutilization is referred in subparagraph (b) as “*any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission.*”. Davison states that the right of extraction corresponds with the right of reproduction.²⁰⁰ As both of the definitions are used for the substantial part of a database, database owner could not prevent the extraction or reutilization of insubstantial part of a database. However, it does not encourage the repeatedly and systematically extraction.²⁰¹ As in case of the repetitive and systematic extraction of insignificant part of database that will harm the investment and maybe cause reconstruction of whole or substantial part of the database without the consent of database maker.²⁰² ECJ comments “substantial part evaluated quantitatively” in the Art. 7(1) of the Directive as “*the volume of data extracted from the database and/or reutilized*” and states that it should be associated with the content volume of the whole of database.²⁰³

¹⁹⁹ European Commission, “Evaluation of Directive 96/9/EC on the legal protection of databases”, p. 39.

²⁰⁰ Davison, “Database Protection: Lessons From Europe, Congress, and WIPO”, p. 833.

²⁰¹ See the Database Directive, Art. 7(5): “The repeated and systematic extraction and/or re-utilization of insubstantial parts of the contents of the database implying acts which conflict with a normal exploitation of that database or which unreasonably prejudice the legitimate interests of the maker of the database shall not be permitted.”

²⁰² Mazumder, *Database Law Perspectives from India*, p. 122.

²⁰³ The British Horseracing Board Ltd and Others v William Hill Organization Ltd. Case C-203/02. Judgment of The Court (Grand Chamber) of 9 November 2004, para. 70.

Although the Directive identifies the terms of “extraction” and “reutilization”, it is silent for the terms of “substantial”, “obtaining”, “verifying” or “presenting”. “Obtaining” is commented as the gathering of data, works or other things constituting the database, “verification” refers to the controlling, rectification and update of the current data in database and “presentation” relates to the revival and contact of the compiled data.²⁰⁴

ECJ explains the investment in “*verification*” as “*resources used, with a view to ensuring the reliability of the information contained in that database, to monitor the accuracy of the materials collected when the database was created and during its operation.*” and “*presentation*” through “*the resources used for the purpose of giving the database its function of processing information, that is to say those used for the systematic or methodical arrangement of the materials contained in that database and the organisation of their individual accessibility*”.²⁰⁵ Thus, ECJ provides clarification for those terms and relates ‘verification’ with the reliability of the information and ‘presentation’ with the methodical arrangement and organization of content.

Art. 8(1) of the Database Directive provides an access to insubstantial part to extract and/or reutilize for lawful users any contrary provision to that article in a contract will be null and void.²⁰⁶ While formation of lawful user is for insubstantial part of a database, Art. 9 regulates the exceptions to sui generis right. As stated regarding copyright protection of a database, the aims of teaching or scientific research in the case of extraction, public security or an administrative or judicial procedure in the case of the extraction and/or reutilization are also exempted from infringement of a database.²⁰⁷

²⁰⁴ P.Bernt Hugenholtz, Implementing the European Database Directive, 1998, <https://www.ivir.nl/publicaties/download/PBH-HCJ-LIB.pdf>, p.9.

²⁰⁵ Fixtures Marketing Ltd v Oy Veikkaus Ab. Case C-46/02. Judgment of the Court (Grand Chamber) of 9 November 2004, para. 37.

²⁰⁶ The Database Directive, Art. 15.

²⁰⁷ The Database Directive, Art. 9(b) and 9(c).

In addition to those exceptions regulated in the Database Directive, European Commission published “Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act)” on 23 February 2022. Art. 35 of the Data Act brings exception for the databases containing certain data for the sui generis right in Database Directive:

“In order not to hinder the exercise of the right of users to access and use such data in accordance with Article 4 of this Regulation or of the right to share such data with third parties in accordance with Article 5 of this Regulation, the sui generis right provided for in Article 7 of Directive 96/9/EC does not apply to databases containing data obtained from or generated by the use of a product or a related service.”

Although, Data Act is not still finalized, in case the Art. 35 becomes applicable, new exception of “*databases containing data obtained from or generated by the use of a product or a related service*” for the applicability of rights conferred by the Database Directive will be valid.

C. Term of Protection

The time period of protection is fifteen years and it starts from first of January following the year of conclusion of making database as stated in Art. 10 of the Database Directive. However, protection period of fifteen years under sui generis protection is criticized and offered five years considering the maximum time of datasets protection.²⁰⁸ Also, 15 years protection could be turned into perpetual protection due to substantial updates on the database as it is renewable in accordance with Art. 10(3). New substantial investment on

²⁰⁸ Pihlajarinne and Ballardini, “Owning Data via Intellectual Property Rights: Reality or Chimera?”, p.16.

a database through addition or updates of new data in current data will provide additional fifteen year protection to a database maker.²⁰⁹ However, the burden of proof is at the side of database maker to prove the existence of substantive change as a result of substantial new investment.²¹⁰

D. Author or Maker of the Database?

The Database Directive uses the term of “*maker of a database*” not author of a database for sui generis right holder as sui generis protection is not provided for sufficient level of intellectual human creation, it is a result of substantive investment. Database maker is identified as “*the maker of a database is the person who takes the initiative and the risk of investing; whereas this excludes subcontractors in particular from the definition of maker*”²¹¹. However, this definition is criticized by considering the practical situation in modern data-driven economy because that causes legal uncertainty in a high level. As the creation and collection of data will often be in larger networks such as platforms enabling data sharing by numerous players.²¹² Joint ownership is also considered, due to network functions creation of data significantly increased and so determining who is database maker is much harder than each passing day. While determining the database maker, in case the data is created by different person or entity than the obtaining it, it is easy to distinguish database maker.²¹³

Art. 11 of the Database Directive sorts the sui generis protection’s beneficiaries as “*database whose makers or rightholders are nationals of a Member State or who have*

²⁰⁹ Jerome H. Reichman and Ruth L. Okediji, “When Copyright Law and Science Collide: Empowering Digitally Integrated Research Methods on a Global Scale”, *Minnesota Law Review*, 96(4), 2012, p.1421.

²¹⁰ The Database Directive, Recital (54).

²¹¹ The Database Directive, Recital (41).

²¹² Drexler, “Data Access and Control in the Era of Connected Devices”, p.77.

²¹³ Drexler, p.71.

their habitual residence in the territory of the Community.”. Also, for the companies having “*registered office, central administration or principal place of business*” within the community benefits from the protection as stated in the second paragraph of same article. However, the condition for those companies only having registered office in the Community is to have a real link with the Member State’s current economy.

E. Criticisms to the Database Directive

There are criticisms for the Database Directive about it does not meet the requirements of advancements in the technology for aggregation of data and big data.²¹⁴ Through various big data types and cases with AI uses, EU policy acceleratingly attaches importance to role of the Database Directive in those technological advancements so imminent reform is seen in the agenda by considering especially the Commission’s Evaluation of the Database Directive and academic Evaluation Report.²¹⁵

“*Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases*” was prepared for the European Commission to investigate effects of Database Directive on business sector in data-driven economy and comprehend whether the aims of the Directive are achieved. First evaluation report is published in 2005 and did not state a clear positive impact of the Directive. According to the figures, although there is a significant growing in database industry of US, the Database Directive’s effect became in an opposite way.²¹⁶ This Study also presented a survey to the stakeholder including the

²¹⁴ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report”, p. vi.

²¹⁵ Leistner, “Protection of and Access to Data under European Law”, p. 384.

²¹⁶ European Commission, “DG Internal Market and Services Working Paper: First evaluation of Directive 96/9/EC on the legal protection of databases”, 2005, https://ec.europa.eu/info/sites/default/files/evaluation_report_legal_protection_databases_december_2005_en.pdf (access date: 14.11.2022), p. 24.

costs and benefits which they met as a result of the Database Directive. The stakeholders stated the benefits as having a legal certainty and more powerful protection for databases to exclude unauthorized users.²¹⁷

The Database Directive has also deficit points which could be listed in general perspective for the need of clarity for database definition, requirement of registration for sui generis data protection and compulsory license. The Database Directive accepts databases as “*vital tool in the development of an information market*”²¹⁸ which is used in various sectors. However, because of that the description of a database in the Directive is wide and the identification of databases in logistical way is partly difficult it is hard to take proper analysis on the production of databases to evaluate the effects of the Directive²¹⁹.

By evaluating the adverse impacts of Database Directive over the researches of scientific, educational and cultural, some measures should be taken to decrease those negative impacts. As an example; compulsory licence should be applicable in case the required data is only derived from the database as a sole source.²²⁰ Art. 20 of GDPR with regard to personal data portability could be taken as a model in a more general way for a probable novel provision on compulsory licenses.²²¹ However, the compulsory license provisions should be applied by considering the connection with privacy, personal data protection and confidentiality together with additional authorization required by those areas.²²² Thus, there is also an objection for the proposal of compulsory license whose aim is

²¹⁷ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report”, p. iii.

²¹⁸ The Database Directive, Recital (9).

²¹⁹ Davison, “Database Protection: Lessons From Europe, Congress, and WIPO”, p. 843.

²²⁰ Reichman, Okediji, “When Copyright Law and Science Collide: Empowering Digitally Integrated Research Methods on a Global Scale”, p. 1451.

²²¹ Leistner, “Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform”, p. 12.

²²² Leistner, p. 11.

regarded as the justification for receiving a payment by benefiting from IP right for the use of database which is under protection of sui generis right. Compulsory license system will cause a second stage for access regulation and de facto data control which will make harder to access to data instead of solving that problem. As it will incent a legal base for de facto data holders to try the possibility of claiming sui generis protection for database and demand for compensation.²²³

There are lots of other reservations regarding the Directive in terms of sui generis protection. Drexl explains one of those as “joint ownership”²²⁴, as in today’s world data are generated through network structures progressively and it makes much harder to determine the database maker.²²⁵

The Database Directive adopts reciprocity principle and provides sui generis protection for the database makers, not national of or having habitual residence in Member States, whose national laws present comparative protection to databases created by nationals of or having their habitual residence in the Community territory.²²⁶ For example; US law does not provide comparable protection for Member States, the database makers from US could not apply for the sui generis protection. Giving rise to sui generis protection policy globally is aimed with that condition of reciprocity for database protection. However, that intention has hailed such as U.S. is more tended to protection regime based on unfair competition and so does not provide comparable protection.²²⁷ Also, although special IP right is granted for database protection, in terms of economic, sui generis right presented

²²³ Drexl, “Data Access and Control in the Era of Connected Devices”, p. 82.

²²⁴ See the Database Directive, art. 4(3): “*In respect of a database created by a group of natural persons jointly, the exclusive rights shall be owned jointly.*”

²²⁵ Drexl, “Data Access and Control in the Era of Connected Devices”, p.68.

²²⁶ The Database Directive, Resital (56).

²²⁷ Leistner, “Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform”, p. 15.

with the Database Directive does not seem to provide enlargement for the database industry in Europe.²²⁸

Sui generis right for a database regulated in the Database Directive does not seem under the category of copyright or neighboring rights. Thus, the reciprocity principle to nationals of other countries for sui generis database protection could be evaluated acceptable. However, another conclusion could be made due to deeply inspection on the Database Directive.²²⁹

Public Consultation for the evaluation of Database Directive directed a survey to database owners. 50.8% of them replied positively that encouragement of investment in “advanced information processing systems (e.g. algorithmic processing)” and stimulation of database production is realized due to sui generis right.²³⁰ However, Singapore Academy of Law (Law Reform Committee) states that although introducing sui generis right is aimed to provide more extensive or at least more exact protection for all stages of database creation, it has achieved little to foster investment in accordance with the EU experience. On the contrary, US law not having such kind of protection demonstrated stronger growth in the creation of databases.²³¹ However, there is no convincing empirical evidence to respond that whether the American legislation is improper or the Directive exaggerate the protection provided by databases in the sense of those will be causing much restriction for the access to database or insufficient incentive for the investment in databases.²³²

²²⁸ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p. 26.

²²⁹ Davison, “The Legal Protection of Databases”, p. 223.

²³⁰ European Commission, “Evaluation of Directive 96/9/EC on the legal protection of databases”, p. 17.

²³¹ Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, p.2.

²³² U.S. Copyright Office, “Report on Legal Protection for Databases”, p.88.

II. Turkish Law

Art. 6(11) identifies the database and the arrangement of Art. 6(11) is criticized in terms of law making technic as that sub article covers both original and unoriginal databases although it exemplifies “Adaptations and Collections” granted original databases with copyright protection and does not include the sui generis database protection. Thus, this definition should have been in Art. 1(B) to also cover sui generis protection a proper systematic of the legislation.²³³

In 2004, Turkish Law is adopted the sui generis protection with the Additional Art. 8. Substantial investment in both qualitatively or quantitatively in creating, verifying or demonstrating the contents brings sui generis protection for database. In case of granting with sui generis right, subparagraphs (a) and (b) give rights to database maker, respectively, prohibiting transfer of all or substantial piece by any way and in any form, and, rental, sale, distribution or transmitting of all or substantial piece to the public in any means. Additional Art. 8 regulates 15 years protection term starting from the publicity for those databases. Law. No. 5846 allows renewal for the sui generis protection and states that *“After every qualitative and quantitative addition, removal, or modification producing a substantial change in the contents of the database and requiring a new investment, the new database resulting from such investment shall qualify for its own conditions of protection.”*. Thus, assigned term of protection and the opportunity of renewal for a database maker is applied similarly with the Database Directive.

By considering the wording of Additional Art. 8, it is free to benefit from the data by extracting not whole or insubstantial part of a database.²³⁴ Even though extracting contents of only nonelectronic database with private purposes is listed as exception in the

²³³ Ateş, “Veri Tabanlarının Hukukî Koruması”, p. 57-58.

²³⁴ Suluk, Orhan, *Uygulamalı Fikri Mülkiyet Hukuku*, p. 269.

Database Directive²³⁵, Turkish law does not make that separation and provides those exception for also electronic databases²³⁶. General exceptions stated in the Law No. 5846 and required extraction for purposes of public security, administrative and judiciary are also valid.

III. Big Data Issues for Sui Generis Database Protection

The copyright protection of database could be on the agenda with the requirement of the originality on the database contents by considering selection or arrangement of its. As an option to failing copyright protection, with the decreased level of protection sui generis protection shall applies in case of the substantive investment without the need of required level of creativity. Drexel explains that the sui generis protection may be accepted better protection way for IoT based data at a first glance, however, it also has own limitations regarding the subject matter and also protection scope.²³⁷

One of the areas in which big data analysis used considerably is customer information. With regard to profiling activities, the investment is made to create database and process data in a systematic manner rather than the creation of it so customer information constituting the dataset should be benefited from sui generis protection.²³⁸ Hence, the unwillingness of technology companies for more powerful protection in U.S. is commented as benefits of free flow of data outweighs the stronger data protection and decrease of data flows.²³⁹ Leistner also attributes sui generis protection for database as a

²³⁵ The Database Directive, Art. 9(a).

²³⁶ Law No. 5846, Art. 38.

²³⁷ Drexel, "Designing Competitive Markets for Industrial Data – Between Propertisation and Access", p. 267.

²³⁸ Banterle, "The Interface Between Data Protection and IP Law: The Case of Trade Secrets and the Database sui generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis", p.427.

²³⁹ Yu, "Data Producer's Right and the Protection of Machine-Generated Data", p. 890.

potential obstacle for the advancement of big data and AI applications. Also, he states it as a critical information and transaction expense issues in the EU case of big data and AI use regarding access to data.²⁴⁰

The socioeconomic value of right to access to data for business is questioned. Intellectual creations are awarded with the protection of IPRs due to support the investments in new knowledge. In terms of IoT and big data it is arguable that there is no need for investment in knowledge as data is created with technological inventions protected through intellectual property law.²⁴¹

Whole data sets are needed for users dealing with big data settings to produce value added new products or services. Database sui generis protection also could be entitled to the databases of private companies gathering customer or personal data in case of the existence of the conditions for it.²⁴² The extraction from the database rather than a creation of sources independently, should have been demonstrated by its owner, however it is increasingly hard to prove it in that networked environment.²⁴³ Web scraping is quite popular nowadays as one of newest techniques to get information through other resources. Web scraping is defined as “*an automated process that uses a bot to fetch web pages automatically and extract the content displayed, which is later parsed, reformatted and stored into a database.*”²⁴⁴ It is quite hard to follow these kinds of unauthorized extractions and claim sui generis database production against infringers.

²⁴⁰ Leistner, “Protection of and Access to Data under European Law”, p. 388.

²⁴¹ Storr and Storr, “Internet of Things: Right to Data from a European Perspective”, p. 80-81.

²⁴² Björn Lundqvist, “Big Data, Open Data, Privacy Regulations, Intellectual Property and Competition Law in an Internet-of-Things World: The Issue of Accessing Data”, in Mor Bakhoun, Mark-Oliver Mackenrodt, Beatriz Conde Gallego and Gintarė Surblytė-Namavičienė (Eds), *Personal Data in Competition, Consumer Protection and Intellectual Property Law Towards a Holistic Approach?*, MPI Studies on Intellectual Property and Competition Law, vol 28. Springer, 2018, p.200.

²⁴³ Wiebe, “Protection of Non-Personal Data: A New Legal Framework for Data Ownership?”, p.13.

²⁴⁴ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report”, p. 28.

It is also evaluated that ECJ's case law does not provide allowance for access to complete data sets in the publicly available database to use those for big data analysis and so by considering the functionality and dynamism of the data driven economy it is not convenient to continue with those access rules.²⁴⁵ Also, in the age of 'Industry 4.0' or the IoT, the Database Directive could not provide proper protection with 15 years protection term starting from the completion of a database by considering the current database technology as the structure of a database is not static and permanently changes through datasets and data services in real time.²⁴⁶

Reform is necessary for the Database Directive and sui generis database right. Legal uncertainty should be decreased certainly by specifying the terms of protection especially and crucially about machine generated data such as IoT, but additionally certain services.²⁴⁷ Davison states the necessity of recognition for that if users, who are also creators of data and so crucial actors, are treated to decrease their access to data, however, it could cause an adverse effect on capacity of those.²⁴⁸ Technological advancements may incent private database creators to make the access with limits or charge, but, on the other hand a significant amount of data produced privately is open to public access and it seems that the large amount of data publicly available will continue.²⁴⁹

De Wachter shows the reason of the value of data as its flow not from its sources and states that big data's focus should be on the access and utilize the flow of data rather than intellectual property rights as outdated model of artificial scarcity by considering the

²⁴⁵ Leistner, "Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform", p. 10.

²⁴⁶ Drexler, "Designing Competitive Markets for Industrial Data – Between Propertisation and Access", p. 268.

²⁴⁷ Leistner, "IPR and the use of open data and data sharing initiatives by public and private actors. European Parliament", p. 34.

²⁴⁸ Davison, "The Legal Protection of Databases", p. 252.

²⁴⁹ Langenderfer and Kopp, "Which Way to the Revolution? The Consequences of Database Protection as a New Form of Intellectual Property", p. 91-92.

exponential growth of big data.²⁵⁰ Drexl also states that *“first, many firms are producers of data and have to rely on access to data of other players at the same time.”*²⁵¹ However, database protection with regard to connected devices may cause significant barriers for free flow of data.²⁵² Similarly, OECD report use the term of “co-opetition” combined from the collaboration and competition to explain the interaction among the actors of data ecosystem. Collaboration among companies provides them a value which could not be achieved on its own.²⁵³ Therefore, the abolishment of sui generis protection to optimize the legal structure for future improvement of data economy in Europe is widely supported.²⁵⁴

A. Substantive Investment

The Database Directive requires qualitatively and/or quantitatively substantial investment to attain, authenticate or demonstrate contents. A CD including the compilations of some musical performances recordings is not subject to the sui generis database protection as the substantive investment is not regarded for the creation of works but it is required with regard to the obtaining of the database content.²⁵⁵

The Database Directive does not include the explanation of “substantive investment”. However, Recital 7 of the Database Directive explains the required investment of *“considerable human, technical and financial resources”* to make a database and Recital 40 what the investment may include:

²⁵⁰ De Wachter, “IP Strategy in the Age of Big Data”, p. 8, 12.

²⁵¹ Drexl, “Designing Competitive Markets for Industrial Data – Between Propertisation and Access”, p. 260.

²⁵² Drexl, “Data Access and Control in the Era of Connected Devices”, p. 68.

²⁵³ OECD, “Data-Driven Innovation: Big Data for Growth and Well-Being”, p. 34.

²⁵⁴ Leistner, “Protection of and Access to Data under European Law”, p. 400.

²⁵⁵ See the Database Directive, Recital (19).

*“Whereas the object of this sui generis right is to ensure protection of any investment in obtaining, verifying or presenting the contents of a database for the limited duration of the right; whereas such investment may consist in the deployment of financial resources and/or the expending of time, effort and energy;”*²⁵⁶

Thus, the sufficient scope of investment should be determined by the national courts. Law No. 5846 has also no definition of substantive investment.

Without gathering, profiling could not be achieved to have the essential processing phase so the investment for processing is common than creating the data for profiling activities. Using analytics software to collect the data efficiently could be regarded as investment. Thus, the expenditure for software, IT things, human capital or time to collect and process the data is generally exists as an investment to organize and arrange the data.²⁵⁷

Documentation of the investments and structure of database, which is equally important with building a database, is stated as top tip for database owners. The way of obtainment/verification/presentation of information and the amount of the investment for each stage of process are critical to benefit from sui generis protection.²⁵⁸

ECJ ruled that database maker could apply for sui generis protection endowed with Art. 7(1) in the Database Directive against meta search engines reutilizing the whole or substantial part of the contents of protected database and provide a form of access to protected database for the end user even different access route aimed by database

²⁵⁶ The Database Directive, Recital (40).

²⁵⁷ Banterle, “The Interface Between Data Protection and IP Law: The Case of Trade Secrets and the Database sui generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis”, p. 18.

²⁵⁸ Wilks and Christie (Eds.), “IP Rights in Data Handbook: Protecting and exploiting IP in data, big data and databases internationally”, p. 33.

maker.²⁵⁹ Also, new database, constituted by adding, removing or changing in terms of quality or quantity caused a significant change on the content of the database and required substantial investment, deserves protection with its own provisions. Exclusions stated in the Law and required because of public security, administration and judiciary proceedings are the points in which sui generis database protection shall not apply.

Art. 7(1) of the Directive does not apply for the resources of verification used while creating data or other materials mainly gathered in a database as those could not be considered for the assessment of substantial investment.²⁶⁰ In the case of *Fixtures Marketing Ltd v Oy Veikkaus Ab*, ECJ states that “*Obtaining the contents of a football fixture list thus does not require any investment independent of that required for the creation of the data contained in that list.*”²⁶¹ Thus, in that case substantive investment for sui generis protection does not cover the creation of materials within database. In the Case C-203/02, ECJ explained two different subjects as creating data and collecting it by linking the database protection with adequate substantive investment for obtainment, verification or its presentation not connecting the resources used to create those materials.²⁶²

Similarly, the Executive Summary in 2018 Database Directive Final Evaluation Report clearly states the exclusion of investment made to create data as follows:

“In the current context, it seems that the Database Directive does not apply to the databases generated with the means of machines, sensors and other new technologies (such as the Internet of Things or artificial intelligence). In fact, the

²⁵⁹*Innoweb v Wegener*. Case C-202/12, 2013. Judgment of the Court (Fifth Chamber) of 19 December 2013.

²⁶⁰ Pihlajarinne and Ballardini, “Owning Data via Intellectual Property Rights: Reality or Chimera?”, p.8.

²⁶¹ *Fixtures Marketing Ltd v Oy Veikkaus Ab.*, para. 44.

²⁶² *Fixtures Marketing Ltd v Organismos prognostikon agonon Podosfairou AE (OPAP)*, para. 45.

generation of these databases is closely interlinked with the creation of their content (i.e. data). However, case law indisputably excludes investments in data creation from the scope of the sui generis right”²⁶³.

The Database Directive’s aim to present sui generis protection is the encouragement of storage and process in practice to demonstrate information not creating content in a database.²⁶⁴ Thus, both in *Fixtures Marketing* and *British Horseracing*, ECJ ruled to limit the scope of protected databases.

Although the ECJ did not provide exact answer for the question of distinction between creating and obtaining of the data, national courts are determined for that point. The German Federal Supreme Court decided that the private entity, Toll Collect, may claim sui generis right for the dynamic database utilized to bill individual operators in the *Autobahnmaut* case. The justification for the attribution of sui generis protection is the investment done by Toll Collect in terminals which register lorries preferring the motorways. The Court also put the difference of *Autobahnmaut* case from *British Horseracing* with that data existed independently from investment of Toll Collect so not created by database maker. Thus, sui generis database right would also apply for the cases with machine generated data. Data gathered during registration were combined with other data including the day, density of motorway use, plate numbers of lorries and operators’ credit card numbers to form an editing systematically or methodologically.²⁶⁵

²⁶³ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report”, p. (ii).

²⁶⁴ *The British Horseracing Board Ltd and Others v William Hill Organization Ltd*, para. 31.

²⁶⁵ Drexler, “Data Access and Control in the Era of Connected Devices”, pp. 72-74: Federal Supreme Court (Bundesgerichtshof) of 25 March 2010, Case I ZR 47/08 *Autobahnmaut* [2010] *Gewerblicher Rechtsschutz und Urheberrecht* 1004, <http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=en&nr=53335&pos=0&anz=1>, para. 15, p. 18,19.

Thus, although the substantial investment is accepted for collecting the data to endow sui generis protection, it will not be applicable for data generation. However, it is stated that the advancement in model of open source caused an investment for collecting data and curation works not for creating database.²⁶⁶ Thus, most databases created under the data generation category through a sole source or an investment in the creation of data could cause a decrease in the number of databases protected via sui generis protection.

B. Machine Generated Data and the Notion of Data Producer's Right

On the contrary of German Federal Supreme Court decision, Andanda excludes machine-generated databases and big data collecting from various sources from sui generis protection as sui generis protection requires a substantial investment to obtain, verify or present the contents of a database.²⁶⁷ Although the contexts of IoT, AI, sensor/machine generated data, big data gains crucial importance, the regulation about those is not clear such as current identification of a database covers those or they may benefit from sui generis protection.²⁶⁸ The thresholds of substantive investment should be elaborately reconsidered and regulated by taking technological advancements on machine generated data into consideration. Therefore, highly usage of big data analytics and increase in IoT devices highlight the importance of large and comprehensive datasets.²⁶⁹ Data producer's right is seen as a way making access to data and its transfer much easier.

²⁶⁶ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), "Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report", p. (v).

²⁶⁷ Andanda, "Towards a Paradigm Shift in Governing Data Access and Related Intellectual Property Rights in Big Data and Health-Related Research", p.1068

²⁶⁸ Karanikolova, Chicot, Gkogka, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology), "Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report", p. (v).

²⁶⁹ Yu, "Data Producer's Right and the Protection of Machine-Generated Data", p. 866.

Sui generis database protection only may be applicable in existence of substantive investment for obtaining, verifying or presenting the contents of a database. Raw machine generated data could not provide those requirements in general and European Commission put the data producer's rights in the agenda with "Building a European Data Economy" proposal. Data Producer's right is identified as "*A right to use and authorise the use of non-personal data could be granted to the 'data producer', i.e. the owner or long-term user (i.e. the lessee) of the device.*" and this approach's aim is to clarify the legal situation and providing more option to the data producer for machine generated data.²⁷⁰ Data producer right focuses on non-personal data.

However, introduction of a new right for data producer is still under discussion. In that case, who will have the control over the data? For a smart car, car manufacturer, car owner and software producer are the actors who could claim a right on machine generated data. Although joint ownership is stated as a solution, next question for specific rights for each actor will appear.²⁷¹

Machine generated data has significant importance due to increased use in analytics of big data and rapidly growth of IoT devices and the caution from many counterarguments appeared against a presentation for new data producer's right.²⁷²

²⁷⁰ European Commission, "Building a European Data Economy", p. 13.

²⁷¹ Yu, "Data Producer's Right and the Protection of Machine-Generated Data", p. 907.

²⁷² Yu, p. 888.

CONCLUSION

Data is created and transferred in a rapid pace. Due to velocity, large datasets constitute big data with a high volume and variety. Turning those data into value requires human intellectual creation and/or investment. Because of that well created database provides big data analysis with extraction of value, the database protection is discussed with regard to big data concept.

Peculation is quite easy and very fast in database industry, it is crucial to provide efficient legal framework for database protection to prevent attacks and apply legal ways in case of infringement. Legal frameworks enforceable to database, which is most related subject matter for big data regarding IP protection and engaged in big data analysis with decision making mechanisms, are mainly copyright and sui generis database protection. Trade secret protection is also stated as applicable to confidential databases. However, when the nature of data is considered, the value of big data extracts by correlating the data with other data so the secrecy could not be achieved while desiring a big data analysis.

Although, other legal concepts about big data protection such as unfair competition or contractual protection is discussed, the limited applicability of those provides the intensity for IPRs applicability. Additionally, the necessary legal policies and regulations should be adopted to spur the innovation with regard to big data, however, the fundamental rights should not be inferior to the innovation as in case the big data includes personal data, applying the rules of data protection law is essential.

Considering the IPRs in EU Law, the Database Directive provides two-tier protection. Sui generis protection for databases alongside the copyright protection has a significant impact on the data access. For the copyright protection, “originality” is required with regard to data selection or arrangement in both European Union law and Turkish law. By

considering the comprehensive structure of big data, the existence of sufficient originality encountering the requirement of copyright protection could be the main problem to protect it through IPRs. In case, the factual manner is used to present the information or data sets, copyright protection shall not apply for big data, as copyright protection for big data requires originality so the existence of author's own personality. Thus, in general, the big data is not granted with copyright protection because databases are generally regarded as not the expression of author but just machine created.

Although the base of copyright protection is original database works, the Database Directive brought new sui generis right for database based on a substantive investment to get, verify or demonstrate the contents. For the introduce of sui generis protection, it is also suggested that the cost of it outweighed its benefits by considering the advantages of data free flow on innovations.

All those legal approaches exist to incent the advancement of database productions and provide a balance between database owner/maker and database user. Although some scholars are in favor of introducing sui generis right for such databases others shifts the perspective that free access to data underpins the growth. The right in data is close to impossible, but the legal rights related to the data should be under examination by considering the rapid technological advances. Also, offered data ownership as a new property right over the data is quite ambiguous. As the control over data is almost impossible by considering the rapid flow of data through various datasets and possible unintentional data right breaches.

It is clear that the rapid technological advancements require updates in the related legislation of big data protection and crucial to be aware of the interests of all actors in data industry while both creating new rules or modernizing the current legislation. Consequently, the balance is required to incent the creation of a database while not

aggravating the access to data through providing limitations and exceptions. IPR should continue to support intellectual creations with regard to big data. Thus, the legal protection could be provided under IPRs, mainly copyright and sui generis protection, also including the machine generated data with human intervention in second degree and if those are not applicable the combination of contractual provisions or technical precautions may leverage the advancements in database industry within the concept of big data.



BIBLIOGRAPHY

Books / e-Books

- Bourgeois, David T., Smith, James L., Wang, Shouhong, Mortati, Joseph. *Information Systems for Business and Beyond*, 2019, *Open Textbooks*. 1, <https://digitalcommons.biola.edu/cgi/viewcontent.cgi?article=1000&context=open-textbooks> (access date: 14.11.2022).
- Brand, Stewart. *The Media Lab : Inventing the Future at M.I.T.*, Penguin Books, 1998.
- Ciani J. A Competition-Law-Oriented Look at the Application of Data Protection and IP Law to the Internet of Things: Towards a Wider ‘Holistic Approach.’ Vol 28. Springer Berlin Heidelberg; 2018.
- Davison, Marc J. *The Legal Protection of Databases*, 2003, Cambridge University Press.
- Goldstein, Paul and Hugenholtz, P. Bernt. *International Copyright: Principles, Law, and Practice*, Oxford University Press, 2001.
- Mazumder, Anirban. *Database Law Perspectives from India*, Springer, 2016.
- McQueen, Bob. *Big Data Analytics for Connected Vehicles and Smart Cities*, Artech House, 2017.
<https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1825911&site=eds-live> (access date: 14.11.2022)
- Richardson, Megan. *Intellectual Property and Privacy Law*, Oxford University Press, 2021.
- Suluk, Cahit and Orhan, Ali. *Uygulamalı Fikri Mülkiyet Hukuku*, Arıkan Yayıncılık, 2005.
- Tekinalp, Ünal. *Fikrî Mülkiyet Hukuku*, Beta İstanbul, 1999.
- Wessels, Bridgette, Finn, Rachel, Sveinsdottir, Thordis and Wadhwa, Kush. *Open Data and the Knowledge Society*, Amsterdam University Press, 2017.
- Hartzog, Woodrow. *Privacy’s Blueprint: The Battle to Control the Design of New Technologies*, Harvard University Press, 2018.

Articles

- Andanda, Pamela. “Towards a Paradigm Shift in Governing Data Access and Related Intellectual Property Rights in Big Data and Health-Related Research”, *Springer, IIC* 50, 2019, pp. 1052–1081.
- Ateş, Mustafa. “Veri Tabanlarının Hukukî Koruması”, *Ankara Üniversitesi Hukuk Fakültesi Dergisi*, Cilt 55, issue 1., 2006, pp. 47-84.
- Banterle, Francesco. “Data Ownership in the Data Economy: A European Dilemma”, in Tatiana-Eleni Synodinou, Philippe Jougoux, Christiana Markou, Thalia Prastitou (Eds.), *EU Internet Law in the Digital Era*, Springer, 2020, pp. 199-225.
- Banterle, Francesco. “The Interface Between Data Protection and IP Law: The Case of Trade Secrets and the Database sui generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis”, *Personal Data in Competition, Consumer*

- Protection and Intellectual Property Law*, Springer Berlin Heidelberg, Vol 28, 2018, pp. 411-443.
- Boyd, Danah and Crawford, Kate. "Critical Questions for Big Data", *Information, Communication & Society*, 15:5, 2012, pp. 662-679.
- Burtscher, Bertram and Fritz, Gernot. "Big data: who owns and who may use and exploit big data?", 2015, <https://www.lexology.com/library/detail.aspx?g=77ab3ffb-8f25-469c-ad80-d14bbcee9551> (access date: 14.11.2022).
- Cambridge International AS & A Level Information Technology, "Data, information and knowledge", *Cambridge International Examinations*, 2015, <https://www.cambridgeinternational.org/images/285017-data-information-and-knowledge.pdf> (access date: 11.11.2022).
- Davison, Marc J.. *Database Protection: Lessons From Europe, Congress, and WIPO*, Case Western Reserve Law Review, Vol. 57:4, 2016, pp. 829-854.
- De Wachter, Joren. "IP Strategy in the Age of Big Data", *International In-House Counsel Journal*, 7(26), 2014, pp. 1-13.
- Dreier, Thomas and Karnell, Gunnar W. G. "Originality of the copyrighted work: European perspective", *Journal of the Copyright Society of the U.S.A.*, 39(4), 1992, pp. 289-302.
- Drexler, Joseph. "Designing Competitive Markets for Industrial Data – Between Propertisation and Access", *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 257, 2017, pp. 257-292.
- DuBoff, Leonard D. "Introduction to computer law", *Hastings Communications and Entertainment Law Journal (Comm/Ent)*, 14(2), 1991, pp. 215-234.
- Ekbja, Hamid, Mattioli, Michael, Kouper, Inna, et al. "Big data, bigger dilemmas: A critical review", *Journal of the Association for Information Science & Technology*, 66(8), 2015.
- Evans, Leighton. "The Privacy Parenthesis: Private and public spheres, smart cities and big data", Claudio Coletta, Leighton Evans, Liam Heaphy, Rob Kitchin (ed.), *Creating smart cities*, Taylor & Francis Books, 2019.
- Forgó, Nikolaus, Händel, Stefanie, Schütze, Benjamin. "The Principle of Purpose Limitation and Big Data", in Marcelo Corrales, Mark Fenwick, Nikolaus Forgó (edit), *New Technology, Big Data and the Law*, Springer, 2017, pp. 17-42.
- Gaster, Jens L. "The New EU Directive Concerning the Legal Protection of Databases", *Fordham International Law Journal*, Volume 20, Issue 4, 1996, pp. 1129-1150.
- Gervais, Daniel. "Exploring the interfaces between big data and intellectual property law", *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 10(1), 2019, pp. 3-19.
- Giffinger, Rudolf, Pichler-Milanovic, Natasa and et. al. "Smart Cities: Ranking of European Medium-Sized Cities", *Vienna University of Technology*, Vienna, 2007.
- Grosheide, F. Wilm. "Database Protection—The European Way", *Washington University Journal of Law & Policy*, Volume 8, 2002, pp. 39-77.
- Harris, Ray K., Rosenfield, Susan Stone. "Copyright Protection for Genetic Databases", *Jurimetrics*, Vol. 45, No. 2, American Bar Association, 2005, pp.225-250.

- Holom, Roxana-Maria, Rafetseder, Katharina, Kritzinger, Stefanie, Sehrschön, Harald, “Metadata management in a big data infrastructure”, *Procedia Manufacturing*, Volume 42, ISSN 2351-9789, 2020, pp. 375-382.
- Hugenholtz, P. Bernt. “Against 'Data Property' ”, in H. Ullrich, P. Drahos, & G. Ghidini (Eds.), *Kritika: Essays on Intellectual Property*, Vol. 3, 2018, pp. 48-71.
- Kemp, Richard. “Legal aspects of managing Big Data”, *Computer Law & Security Review: The International Journal of Technology Law and Practice*, 30(5), 2014, pp. 482-491.
- Kuner, Christopher, H. Cate, Fred, Millard, Christopher, Jerker B. Svantesson, Dan. “The challenge of ‘big data’ for data protection”, *International Data Privacy Law*, Vol. 2, No. 2, 2012, pp. 47-49.
- Langenderfer, Jeff, Kopp, Steven W. “Which Way to the Revolution? The Consequences of Database Protection as a New Form of Intellectual Property”, *Journal of Public Policy & Marketing*, 22(1), 2003, pp. 83–95.
- Lavenue, Lionel M. “Database rights and technical data rights: the expansion of intellectual property for the protection of databases”, *Santa Clara Law Review*, 38(1), 1997, pp. 1-64.
- Law Reform Committee, “Rethinking Database Rights and Data Ownership in an AI World”, *Singapore Academy of Law*, 2020.
- Leistner, Matthies. Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform, 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3245937 (access date: 14.11.2022).
- Leistner, Matthias. “Protection of and Access to Data under European Law”, *Oxford University Press*, 2021, pp. 383-400.
- Lim, Darly. “Re-defining the rights and responsibilities of database owners under competition law”, *Singapore Academy of Law Journal*, 18(2), 2006, pp. 418-472.
- Lindman, Juho, Rossi, Matti. Kristiina Tuunainen, Virpi. “Open data services: research agenda”, *46th Hawaii International Conference on System Sciences*, IEEE, 2013, pp. 1239-1246.
- Löfgren, Karl and Webster, C. William R.. “The value of Big Data in government: The case of ‘smart cities.’”, *Big Data & Society*, 2020.
- Lundqvist, Björn. “Big Data, Open Data, Privacy Regulations, Intellectual Property and Competition Law in an Internet-of-Things World: The Issue of Accessing Data”, in Mor Bakhoun, Mark-Oliver Mackenrodt, Beatriz Conde Gallego and Gintarė Surblytė-Namavičienė (Eds), *Personal Data in Competition, Consumer Protection and Intellectual Property Law Towards a Holistic Approach?*, MPI Studies on Intellectual Property and Competition Law, vol 28. Springer, 2018, pp. 191-214.
- Mattioli, Michael. “Disclosing Big Data”, *Minnesota Law Review*, 99(2), 2014, pp. 535–584.
- McManis, Charles R. “Database protection in the digital information age”, *Roger Williams University Law Review*, 7(1), 2001, pp. 7-46.
- Pihlajarinne, T. and Ballardini, R. M. “Owning Data via Intellectual Property Rights: Reality or Chimera?”, in Ballardini, R. M., Kuomamäki, P., Pitkänen, O. (Ed), *Regulating*

Industrial Internet Through IPR, Data Protection and Competition Law, Kluwer Law International, 2019.

Powell, Marc. “The European Union’s Database Directive: An International Antidote to the Side Effects of Feist?”, *Fordham International Law Journal*, Volume 20, Issue 4, 1996, pp. 1215-1250.

Prime, Terence. International Trade and the Rule of Law: The Case of European Protection of Computer and Databases, *Ankara Üniversitesi Sosyal Bilimler Fakültesi Dergisi*, 54-4, 2015, pp. 121-143.

Reichman, Jerome H., Okediji, Ruth L. “When Copyright Law and Science Collide: Empowering Digitally Integrated Research Methods on a Global Scale”, *Minnesota Law Review*, 96(4), 2012, pp. 1362–1481.

Rens, Andrew. “Who is in Charge Here? The Internet of Things, Governance and the Global Intellectual Property Regime”, *UCLA Journal of Law & Technology*, Volume 23, Issue 2, 2019.

Sattler, Andreas. “From Personality to Property?”, in Mor Bakhoun, Mark-Oliver Mackenrodt, Beatriz Conde Gallego and Gintarė Surblytė-Namavičienė (Eds), *Personal Data in Competition, Consumer Protection and Intellectual Property Law Towards a Holistic Approach?*, MPI Studies on Intellectual Property and Competition Law, vol 28. Springer, 2018, pp. 27-54.

Storr, Christine and Storr, Pam. “Internet of Things: Right to Data from a European Perspective”, in Marcelo Corrales, Mark Fenwick, Nikolaus Forgó (edit), *New Technology, Big Data and the Law*, Springer, 2017, pp. 65-96.

Synodinou, Tatiana-Eleni. “The Portability of Copyright-Protected Works in the EU”, in Tatiana-Eleni Synodinou, Philippe Jougoux, Christiana Markou, Thalia Prastitou (Eds.), *EU Internet Law : Regulation and Enforcement*, Springer, 2017, pp. 217-266.

Tabuchi, Helga. “International Protection Of Non-Original Databases; Studies on the Economic Impact of the Intellectual Property Protection of Non-Original Databases”, *Data Science Journal*, Volume 3, 2004, pp. 175-181.

Ursic, Vrabec H., Bart, Custers. “Legal barriers and enablers to big data reuse”, *European Data Protection Law Review (EDPL)*, 2(2), 2016, pp. 209-221.

Warren, John W. “Zen and the Art of Metadata Maintenance”, *Mason Publishing Group/George Mason University Press*, Volume 18, Issue 3, 2015.

W White, Ryen, P Tatonetti, Nicholas, Shah, Nigam H, B Altman, Russ and Horvitz, Eric. “Web-scale pharmacovigilance: listening to signals from the crowd”, *Journal of the American Medical Informatics Association*, Volume 20, Issue 3, 2013.

Wiebe, Andreas. “Protection of Non-Personal Data: A New Legal Framework for Data Ownership?”, in Barczewski, M. (Ed.), *Value of Information: Intellectual Property, Privacy and Big Data*, Peter Lang GmbH, 2018, pp. 9-26.

Yip, Man. “Protecting Consumer’s Personal Data in the Digital World – Challenges and Changes”, *Research Collection School Of Law*, 2018, pp. 104-117.

Yu, Peter K. Data Producer's Right and the Protection of Machine-Generated Data. *Tulane Law Review*, 93(4), 2019, pp. 859–930.

Reports, Studies and Papers

Collections of Information Antipiracy Act. H.R. Rep. No. 106–349, 1999, <https://www.congress.gov/106/crpt/hrpt349/CRPT-106hrpt349.pdf> (access date: 14.11.2022).

Communication From the Commission to the European Parliament, The Council, The European Economic and Social Committee and The Committee of the Regions. Towards a common European data space. COM/2018/232 final, 2018, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0232&from=en> (access date: 14.11.2022).

Diebold, Francis X., “On the Origin(s) and Development of the Term 'Big Data'”, PIER Working Paper No. 12-037, 2012 (access date: 14.11.2022).

Drexler, Joseph. “Data Access and Control in the Era of Connected Devices”, The European Consumer Organisation, BEUC, Brussels, Belgium, 2018, https://www.beuc.eu/sites/default/files/publications/beuc-x-2018-121_data_access_and_control_in_the_area_of_connected_devices.pdf (access date: 14.11.2022).

European Commission, “Evaluation of Directive 96/9/EC on the legal protection of databases”, {SWD(2018) 147 final}, 2018, [https://ec.europa.eu/transparency/documents-register/detail?ref=SWD\(2018\)147&lang=en](https://ec.europa.eu/transparency/documents-register/detail?ref=SWD(2018)147&lang=en) (access date: 14.11.2022)

Hugenholtz, P. Bernt. “Implementing the European Database Directive”, 1998, <https://www.ivir.nl/publicaties/download/PBH-HCJ-LIB.pdf> (access date: 14.11.2022).

Karanikolova, Kristina, Chicot, Julien, Gkogka, Anna, et al. (for European Commission, Directorate-General for Communications Networks, Content and Technology). “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : Final Report”, *Publications Office*, 2018, <https://op.europa.eu/en/publication-detail/-/publication/5e9c7a51-597c-11e8-ab41-01aa75ed71a1> (access date: 14.11.2022).

European Commission. “Building a European Data Economy”, {SWD(2017) 2 final}, 2017.

European Commission. “DG Internal Market and Services Working Paper: First evaluation of Directive 96/9/EC on the legal protection of databases”, 2005, https://ec.europa.eu/info/sites/default/files/evaluation_report_legal_protection_databases_december_2005_en.pdf (access date: 14.11.2022).

European Commission. “Green Paper on Copyright and the Challenge of Technology, Copyright Issues Requiring Immediate Action”, (COM (88) 172 final) 176, 1988.

ICO. “Big Data and Data Protection”, 2014, <https://rm.coe.int/big-data-and-data-protection-ico-information-commissioner-s-office/1680591220> (access date: 14.11.2022).

- Leistner, Matthias and Antoine, Lucie. “IPR and the use of open data and data sharing initiatives by public and private actors”, European Parliament, 2022, [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf) (access date: 15.11.2022).
- Lu, Meng. “Intellectual property protection of big data”, J. Phys.: Conf. Ser. 1693 012012, 2020, <https://iopscience.iop.org/article/10.1088/1742-6596/1693/1/012012/pdf> (access date: 14.11.2022).
- National Science Foundation, “Core Techniques and Technologies for Advancing Big Data Science & Engineering (BIGDATA)”, PROGRAM SOLICITATION NSF 12-499, 2012.
- OECD. “Building Blocks for Smart Networks”, OECD Digital Economy Papers, No. 215, *OECD Publishing*, 2013.
- OECD. “Data-Driven Innovation: Big Data for Growth and Well-Being”, *OECD Publishing*, Paris, 2015.
- OECD. “OECD Digital Economy Outlook” 2020, *OECD Publishing*, Paris, 2020.
- OECD. “The Internet of Things: Seizing the Benefits and Addressing the Challenges”, OECD Digital Economy Papers, No. 252, *OECD Publishing*, Paris, 2016.
- UN Global Pulse. Big Data for Development: Challenges and Opportunities, 2012, <https://www.unglobalpulse.org/wp-content/uploads/2012/05/BigDataforDevelopment-UNGlobaIPulseMay2012.pdf> (access date: 14.11.2022).
- US Copyright Office. “Report on Legal Protection for Databases”, 1997, <https://www.copyright.gov/reports/db4.pdf> (access date: 14.11.2022).
- Wilks, J. (Ed). IPRs in Data Handbook: Protecting and exploiting IP in data, big data and databases internationally. DLA Piper, 2014.

Legislation

- Agreement On Trade-Related Aspects of Intellectual Property Rights, 1995
https://www.wto.org/english/docs_e/legal_e/27-trips.pdf (access date: 14.11.2022).
- Decision No 1/95 of the EC-Turkey Association Council of 22 December 1995 on implementing the final phase of the Customs Union. [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:21996D0213\(01\)&from=en](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:21996D0213(01)&from=en) (access date: 14.11.2022).
- Directive 2006/116/EC of The European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (codified version), <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02006L0116-20111031&from=EN> (access date: 14.11.2022).

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:01996L0009-20190606&from=EN> (access date: 14.11.2022).

Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L0790&from=EN> (access date: 14.11.2022).

Law No. 5846 of Literary and Artistic Works,
<http://www.mevzuat.gov.tr/MevzuatMetin/1.3.5846.pdf> (access date: 14.11.2022).

Law No. 6102 Turkish Commercial Law,
<https://www.mevzuat.gov.tr/mevzuatmetin/1.5.6102.pdf> (access date: 14.11.2022).

Law No.6698 Turkish Personal Data Protection Law,
<https://www.mevzuat.gov.tr/mevzuatmetin/1.5.6698.pdf> (access date: 14.11.2022).

Paris Convention for the Protection of Industrial Property, 1883,
https://www.wipo.int/edocs/lexdocs/treaties/en/paris/trt_paris_001en.pdf (access date: 14.11.2022).

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679> (access date: 14.11.2022).

The Berne Convention for the Protection of Literary and Artistic Works, 1979,
<https://wipolex.wipo.int/en/text/283693> (access date: 14.11.2022).

WIPO Copyright Treaty, 1996, <https://wipolex.wipo.int/en/text/295157> (access date: 14.11.2022).

Court Cases

Feist Publications, Inc. v. Rural Telephone Service Co., Inc. No. 89-1909. Supreme Court of the United States (10th Circuit), 1990.
<https://cyber.harvard.edu/people/tfisher/1991%20Feist.pdf> (access date: 14.11.2022).

Freistaat Bayern v Verlag Esterbauer GmbH. Case C-490/14. Judgment of the Court (Second Chamber) of 29 October 2015.
<https://curia.europa.eu/juris/document/document.jsf?text=&docid=170741&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=14590916> (access date: 14.11.2022).

Fixtures Marketing Ltd v Organismos prognostikon agonon Podosfairou AE (OPAP). Case C-444/02. Judgment of the Court (Grand Chamber) of 9 November 2004.
<https://curia.europa.eu/juris/showPdf.jsf?text=&docid=49635&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=2207895> (access date: 14.11.2022).

Fixtures Marketing Ltd v Oy Veikkaus Ab. Case C-46/02. Judgment of the Court (Grand Chamber) of 9 November 2004.
<https://curia.europa.eu/juris/showPdf.jsf?text=&docid=49636&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=2216743> (access date: 14.11.2022).

Football Dataco Ltd and Others v Yahoo! UK Ltd and Others. Case C-604/10, 2012. Judgment of the Court (Third Chamber) of 1 March 2012.
<https://curia.europa.eu/juris/document/document.jsf?text=&docid=119904&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=2208564> (access date: 14.11.2022).

Innoweb v Wegener. Case C-202/12, 2013. Judgment of the Court (Fifth Chamber) of 19 December 2013.
<https://curia.europa.eu/juris/document/document.jsf?text=&docid=145914&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=2209259> (access date: 14.11.2022).

Republic of Türkiye, Court of Cassation 11th Civil Chamber, E. 2005/3506, K.2006/4425, 13.04.2006.

Ryanair Ltd v PR Aviation BV. Case C-30/14. Judgment of the Court (Second Chamber) of 15 January 2015. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62014CJ0030&from=EN> (access date: 14.11.2022).

The British Horseracing Board Ltd and Others v William Hill Organization Ltd. Case C-203/02. Judgment of The Court (Grand Chamber) of 9 November 2004.
<https://curia.europa.eu/juris/showPdf.jsf?jsessionid=A51202BECC18BDEB34D4A5C2FD007E10?text=&docid=49633&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=2206682> (access date: 14.11.2022).

Internet Sources

Cambridge Dictionary, <https://dictionary.cambridge.org/dictionary/english/dataset> (access date: 14.11.2022).

Cambridge Dictionary, <https://dictionary.cambridge.org/dictionary/english/original> (access date: 14.11.2022).

Research Data Management Service Group. “Introduction to intellectual property rights in data Management”, *Cornell University*, <https://data.research.cornell.edu/content/intellectual-property> (access date: 14.11.2022).

Gartner IT glossary, Definition of Big data. <http://www.gartner.com/it-glossary/big-data> (access date: 14.11.2022).

Manyika, James, Chui, Michael, Brown, Brad and et. al. “Big Data: The next frontier for innovation, competition, and productivity”, at 1, *McKinsey*, 2011, <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation> (access date: 14.11.2022).

Schoenberger, Chana R. “The Internet of Things”, *FORBES*, 2022, <https://www.forbes.com/global/2002/0318/092.html#33f75a713c3e> (access date: 14.11.2022).

Statista. Big data - Statistics & Facts, 2022, <https://www.statista.com/topics/1464/big-data/#dossierKeyfigures> (access date: 14.11.2022).

Statista. Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2021, with forecasts from 2022 to 2030, 2022, <https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/> (access date: 14.11.2022).

WIPO, Frequently Asked Questions: Trade Secrets. https://www.wipo.int/tradesecrets/en/tradesecrets_faqs.html#:~:text=Trade%20secrets%20are%20intellectual%20property,may%20be%20sold%20or%20licensed (access date: 14.11.2022).

ABSTRACT

New technologies raised some legal questions for also in intellectual property law such as IPRs applicability and scope of the protection for big data existing in almost all sectors with a buzzword status. Author/maker of the database should have been granted legal protection to acquire economic benefit from its creation and so promote the innovations in this subject. What is firstly presented in this assignment is “big data concept” with its sources of Internet of Things and smart cities. Arguments for different protection ways for big data in IP are mainly about copyright, sui generis database protection and trade secret protection. Contractual protection and unfair competition will also have a place with its limited applicability. Data ownership is widely discussed issue about data regime although the challenges of control over data make that legal remedy fictional. The thesis will provide a general guide in terms of EU Law and Turkish Law for the IP protection of big data by evaluating the constitution of elements of big data and a database benefited to extract a value from big data. Thus, the application of copyright protection will be discussed for big data mainly in terms of originality and sui generis protection for substantive investment together with the most related IP instruments’ role and required reform for those.

Keywords: Big data, copyright, sui generis database right, trade secret, data ownership.

ÖZET

Yeni teknolojiler bazı hukuki soruları fikri mülkiyet hukukunda da gündeme getirdi, örneğin; hemen hemen tüm sektörlerde mevcut olan ve kullanımı klişe haline gelmiş büyük veri için fikri mülkiyet haklarının uygulanabilirliği ve korumanın kapsamı. Veri tabanının yazarı/yapımcısına, oluşturulmasından ekonomik fayda elde etmek ve böylece

bu konudaki yenilikleri teşvik etmek için yasal koruma sağlanmış olmalıydı. Bu tezde ilk olarak sunulan şey, “büyük veri kavramı” ile Nesnelerin İnterneti ve akıllı şehirler olmak üzere büyük veri kaynaklarıdır. Büyük veri için fikri mülkiyet alanında farklı koruma yöntemlerine yönelik argümanlar temel olarak telif hakkı, sui generis veri tabanı koruması ve ticari sır koruması ile ilgilidir. Sözleşmesel koruma ve haksız rekabet de kısıtlı uygulama alanı olmasına rağmen yer alacaktır. Veriler üzerindeki kontrolün zorlukları bu yasal çareyi kurgusal kılmasına rağmen, veri mülkiyeti, veri rejimi hakkında geniş çapta tartışılan bir konudur. Büyük veri oluşumunun elementleri ve büyük veriden bir değer elde etmek için yararlanılan veri tabanı değerlendirilerek büyük verinin fikri mülkiyet koruması için AB Hukuku ve Türk Hukuku açısından genel bir kılavuz sağlanacaktır. Böylelikle, en ilgili fikri mülkiyet araçlarının rolü ve bunlar için gerekli görülen reform ile birlikte telif hakkı koruması temel olarak orijinallik ve sui generis koruma ise esaslı bir nispet dahilinde yatırım açısından tartışılacaktır.

Anahtar kelimeler: Büyük veri, telif hakkı, sui generis veri tabanı hakkı, ticari sır, veri sahipliği.