https://doi.org/10.55843/ivisum2382113k

TEXT AND DATA MINING IN TERMS OF INTELLECTUAL PROPERTY

Canan Küçükali, page 113-133

ABSTRACT

With the intense penetration of the internet into our lives, sharing has increased worldwide, whether in intellectual property or industrial property, as in other fields. Text and data mining has become an area for supporting experts working in the field of intellectual property, facilitating the work of decision- makers, and helping them in solving problems.

While data mining activities are involved different issues such as copyright, database rights, and data protection, text and data mining activities that are not regulated in TRIPS and international agreements are discussed as to whether they cause violations or not.

The problem in this matter is tried to be solved differently in each legal system by being evaluated within the scope of fair use in US law and by being included in the scope of the Directive in EU law. However, both EU law and other laws have areas that need to be renewed in the face of technological developments

Keywords: Information Technology Law, Internet, Text and Data mining, Intellectual Property, Industrial property, EU. Directive.

Assoc. Prof. Dr. Canan Küçükali

Doğuş University, Faculty of Law, Commercial Law Department

e-mail: ckucukali @dogus.edu.tr

UDK: 004.6:004. 451.5:347.77(560)

Date of received: 23.07.2023

Date of acceptance: 20.08.2023

Declaration of interest: The authors reported no conflict of interest related to this article.

INTRODUCTION

The development of countries takes place with the provision of cultural exchange between nations and the training of well-equipped and trained scientists. Information law, which is an area where many legal fields intersect, is a science that examines the processing of data and the transfer of the result of the transaction. Informatics is intimately tied to preserving social memory for future generations and the collecting data about it. However, while raw data is worthless on its own, when these data are processed for a specific purpose, they become information and gain value (Kalıkov, 2006: 1).

"Data mining" also known as information discovery from databases in the field of informatics, has emerged as a way to take the necessary and accurate information out of the data that the masses have collected in the databases, analyze it, and interpret it in to make predictions about the future based on the past data. Data analysis processes “digging”,and “mining” to find the correct value that satisfies an equation by consuming electricity and computer power and trying Numerical values; Those who do this operation are called “miners” (Güçlütürk, 2018: 16).

With this activity, wide-ranging techniques are used in different areas such as increasing revenues, reducing costs, improving customer relations, and reducing risks. Text and data mining is an activity that not only helps researchers find the correct answers but also helps them find the right questions (Carroll, 2019: 903).

Data mining activities in Turkey have just been realized and started to be used in intellectual property, as in many other fields.Text and data mining is the process of transforming raw data into information, examining the data, and discovering confidential and valuable information. These transactions are; Statistics is made by using mathematical disciplines, modeling techniques, database technology, and various computer programs (Akpınar, 2000, cited in Kalikov, 2006:7; Eker, 2008; Baykal, 2006: 3). Data mining, which is a reliable business model utilized in many disciplines, is the activity of detecting and using hidden dependencies and patterns from seemingly meaningless data stacks to obtain a valid result, and the process of finding anomalies, patterns and correlations in large data sets to predict outcomes (Kalikov, 2006:17; Carroll, 2019: 895; Quantais, 2020: 7; Arslan, 2008: 3).

Based on data mining activities, there are tools and methods to explain information, classify such information, index, and depict such information, which has advanced greatly at the rate of scientific advancement (Margoni, 2020: 1; Feng and Mao, 2022; Iglesias et al. 2021: 10). With this method, it is possible to discover relationships between massive amounts of data, and the connections between them, draw information from databases that have been hidden by computer programs, and use these relationships and rules to predict future trends and behavioral patterns. These data are then transformed into information. (Dongsong and Zhou, 2004; Rosati, 2018; 6,7; He et al. 2019; Baykal, 2006:3; Kalikov, 2006 :1; Piatetsky-Shapiro et al. 1991; Linoff et al.; Akpınar, 2000: 2).

Data mining; is a field that includes multiple disciplines such as artificial intelligence, machine learning, statistics, database Technologies, and visualization and uses their methods. Machine learning and artificial intelligence are scientific disciplines that can use standard tools to contribute to this field. This field, which has a wide application area and has a lot to do with statistics and machine learning, processes data with automatic or semi-automatic methods to find patterns among the data.

On the other hand, it is argued that legal practices will have little impact on law practice since they require advanced cognitive abilities, and the current artificial intelligence capacity does not have such a high level of cognition even though it has made significant improvements (Okamoto, 2009: 83). Data mining activities in Turkey are an area that has just been realized and started to be used in the field of intellectual property as in many other fields.

I. DATA MINING APPLICATIONS

Data mining methods are used in different fields. For example, areas such as increasing efficiency in public health administration, creating new recipes in the kitchen, and preventing crime, including cyber attacks, can be counted on without increasing productivity in the workplace (Rosati, 2018: 3).

Data mining is becoming more significant in the business world for issues like market basket analysis, customer relationship management, and customer evaluation because it boosts customer loyalty, increases marketing campaign returns, identifies customer buying patterns,

forecasts sales, identifies connections between customer demographics, raises response rates to mail campaigns, and assists companies in acquiring and retaining customers. In the era of computers, transactions are tracked, and these data are examined by professionals so that the best decisions may be made. Numerous technological techniques, including clustering, data summarization, change analysis, and deviation detection, are used to develop the algorithms for these analyses (Kalikov, 2006: 2). Data mining has evolved as a result of several elements, including computer networks, hardware, mathematical calculations, and market movements (Savaş, 2012: 6).

he problem is identified and the hypothesis is described, correspondingl, in the data mining process. This model, which was developed through the collection of data, preparation, development, and evaluation of the model, is used and tracked (Kalikov, 2006: 39; Matheus, et al., 1993 as referenced in Arslan, 2008:24, 25; Savaş et al., 2012: 8; Han and Kamber, 2001: 45–53).

Data mining techniques are used to define and prepare the data on the database where the information is stored, and a model is produced by moving through the steps of data preparation, gathering, merging, cleaning, and transformation.

The best model is found for the problem's solution after the program has been installed (Kalikov, 2006: 40, 41).In contrast to conventional data analysis techniques, if the information is obtained by mining activities, it is the first unknown, then an efficient approach, and finally a practical method (Adekitan et al. 2019:1–23).

II. USING DATA MINING IN INTELLECTUAL PROPERTY

Big data is increasingly used by intellectual property owners in the fields of production and operations, which provides the opportunity to discover new generation innovative technologies and enable better decision-making. On the other hand, the statistical use of data mining in the field of intellectual property is limited to how much technology has been learned by academics to use it (Margoni, 2020: 18).

Data mining activities in Turkey are an area that has just been realized and started to be used in the field of intellectual property as well as in many other fields. The lack of adequate national legislation and EU copyright

legislation to improve data mining activities, uncertainties, a lack of awareness and skill, and infrastructure issues are some of the reasons why data mining in Europe lags behind that in America and China, even though those two countries are making progress in the field (Rosati, 2018 : 2). On the basis of countries, the reason why data mining in Europe lags behind America and China, although America and China are making progress in data mining, can be counted as the inadequacy of national legislation and EU copyright legislation to improve data mining activities, uncertainties, lack of awareness and skills, and infrastructure problems (Rosati, 2018: 2).

Data mining is a technique that can be used to support intellectual proper ty experts, simplify the job of lawyers in court, and help solve issues. In intellectual property research, the application of data-driven analytical research and unique methodological classification and computational methods helps in solving the problems that arise in this field. The application of data mining and predictive analysis in copyright, research creates some advantages in terms of automation. Likewise, data mining methods can be used in the marketing of products subject to industrial property.
The goal of a significant amount of applied data in the practical applicati ons of data mining technology is to conduct patent technology research o f various businesses in the patent field of particular technologies and pro ducts, perform intelligence analysis, integrate it, then make statistical an alyses and technical groups text mining, combination theory, patent map s, and other
technologies are making statistical maps, collecting and reporting these maps (Liying Li, 2019:140). Data about intellectual property is analyzed, generated, and acquired during the analysis process (Wei Z. 2019: 17–18).

Thanks to the data mining activity in the field of intellectual property, in the studies to be carried out on the registered trademarks and patent designs in the databases of the Turkish Patent and Trademark Office, the consumer profiles of the companies that own intellectual property will be able to be determined to determine where consumers will prefer brands, patent designs or other products subject to intellectual property.

Intellectual property analytics is the data science that aids decision-making by analyzing large amounts of intellectual property information to

discover relationships, trends, and patterns among data. Artificial intelligence, machine learning, and deep learning methods can be used to analyze state-of-the-art data in intellectual property analytics.

Data mining is a method that can be used to support experts working in the field of intellectual property, facilitate the work of practitioners in courts, and to help solve issues.

A lot of work has been done to create data analysis for intellectual property with data mining technology. As a matter of fact, in a study, the application of newly developing data mining and related modern technologies in the field of intellectual property was investigated, and as a result, it was suggested to make a systematic classification in this field with the view that there are terminological uncertainties (Margoni, 2006: 2). A study argues that along with data-driven analytical tools, text and data mining can achieve its autonomous methodological classification through computational legal methods. With the study, it has been determined that text and data mining provide some advantages in terms of automation with predictive analytics in the field of intellectual property (Margoni, 2006: 2). In this study, it is aimed to examine the diversity of uses to express license terms and conditions in legal documents in order to "guess" non-standard licenses and terms of service, and to identify the expressions used in licenses to define the most popular licensing terms, phrases and license concepts with a set of labels. The goal of the research was to use specialized software either locally or as a component of an internet service to turn the data they collected into text and data mining. The project has created a static matrix or ontology to determine the degree of compatibility of various licenses within the sa me layer, based on what situation and how frequently the relevant licens es form. This is done in light of the multi-
layered compatibility problem.

Researchers manually select the relevant licenses from a drop-down menu, but in the future, the tool will be able to automatically obtain licenses from the metadata of the content and services used. While using the service requires acceptance of the terms of service specified by the service provider, these documents have never been standardized.

Since the current techniques for analyzing patents are inadequate and immature, data mining addresses the issue of the patent examination

approach. In the future, generalized backward neural networks (GRNN) and support vector machines (SVM) algorithms will be more effective than a conventional single algorithm because the problem cannot always be solved by a single approach. Research is possible.

Another study evaluated research and analysis techniques as well as patent research in the area of intellectual property for data mining. Due to the technical nature of data mining, the most popular very thorough and feed-forward BP back propagation neural network algorithm and regression neural network algorithm were used. In this study, a comparison of the target algorithm based on the GRNN network algorithm and the SVM control algorithm with the experience of intellectual property data mining was made to confirm the time sufficiency of the algorithm and the accuracy of the classification and to find the clever algorithm that can be applied in different situations. With a higher sample size, it converges on the regression surface and estimation. The fusion of the two algorithms not only creates a novel approach with potential for widespread use in the future, but it also paves the way for more in-depth studies on intellectual property (Li, 2019:149). Web-based music streaming services that give unlimited music to their customers through a subscription model include Spotify, Deezer, Fizzy, and YouTube Music. These services use data mining to match the member's favourite music with the daily lists, emphasize the appropriate music genres, assist the user in making a decision, and even provide guidance. According to our opinion, data mining operations will be beneficial in targeting the consumer and growing buyer potential in this copyright-related intellectual property market. Another compelling study using data mining was on maqam recognition in Turkish music. The conversion of works to digital media has led to the beginning of the employment of computer programs in musicology research. Scientific study uses musical works as data, so the discipline of computational music is expanding quickly. Turkish music works on the maqam system, studied in data mining, machine learning, and categorization. An experimental study was conducted to estimate maqam using a machine learning algorithm on the data file, which is presumed to be obtained via note identification from a sound file and contains only the note sequences of 1261 Turkish music works.The software was developed to use MusicXML format works in machine learning application, and four different derived data fields were added to the original dataset in order to increase the performance in machine

learning with this software. As a result, 89.7% success was achieved in recognizing the maqam with the 'Random Forest' algorithm (Abidin et al. 2017: 1221-1232).

The MusicXML format was developed to transfer the information to be collected from musical works to different platforms by transferring them to digital media (Good, 2001:113-124).

Microsoft developed a "singing model" using text-to-speech technology, which can only function by using examples from the internet, and "Song training data" obtained from various music websites through data mining called "Deep Singer." He has created artificial intelligence that is able to communicate in English, Chinese, and Cantonese. Although writing lyrics and synchronizing them with the notes of a song is a human-specific skill, the article from Zhejiang University in China points out that the new technology's distinguishing traits are its use of data from the web environment and the delay between the transition from words to performance. The benefits are highlighted, including the removal of lengthy and expensive arrangements, the ability to play in multiple languages, the separation of the singers without the need for complicated acoustic models, and more.

In the first study, excellent results were obtained in processing the 92-hour performance data of 89 singers in three languages while maintaining accurate pitch and genuine sound quality. However, in our opinion, in the face of these and similar facilities, besides the advantages of data mining, there is also the suspicion that the authors of the work can put aside their long-lasting labor and intellectual efforts to create a work and make it worthless.

According to Turkish Copyright Code No. 5846's Article 42/A, the license fees for professional associations are decided upon. The prices of goods in the relevant sector, the proportion of that sector's GDP that is devoted to that sector, the frequency of use and/or transmission of works, performances, phonograms, productions, and broadcasts, unit price and lump-sum payment, payment schedule, and similar considerations are all used to determine the tariff (Erverdi, 2011: 78). It tracks the financial rights of member authors resulting from the public offering of their works, and it identifies all unlicensed users in broadcasting organizations, interactive uses, public places, concerts, and live events within the

parameters of the authorization received from the members of the Professional Associations. Consistent and reliable data should be obtained in order to compile data-based information on intellectual property rights and to develop strategies and policies for this.

In our opinion, data mining can be used to increase demand by determining factors like interest in the works, frequency of use, user age, occupation, social position, and geographic location from data on the use of works protected by intellectual property, as well as the segments that constitute violations. Similar sales graphics according to the occupation and interest distribution of the users in the system can be utilized in data mining and industrial property by identifying those who purchase a particular product.

IV. COPYRIGHT PROBLEM IN DATA MINING ACTIVITIES

The majority of text and data mining research is based on previously published works. The existence of copyright in the copies of works utilized for data mining is up for debate. Copyright is typically granted to the holder of a variety of exclusive economic rights, enabling them to regulate how their work is used and barring unauthorized usage by others (Küçükali, 2017: 157). The best protection of the rights holder's interests is ensured when the author's economic rights are balanced, with the exception of copyright. Thus, extensive protection provided by intellectual property laws ensures the author's rights in the event that his work is shared and utilized by third parties. The copyright laws.

The importance of developing records, data, and national statistics is crucial in order to track the cultural industries that are covered by copyright and continue to grow in importance as well as to develop policy. The professional associations tasked with defending the financial interests of the data, artifacts, and related right holders should also develop data sets that will enable the measurement of the economic dimension of the relevant sectors. According to the Tenth Development Plan (2014-2018), Ministry of Development, Ankara (2013), a number of systematic studies should be carried out, including the advantages of data exchange between them and the matching of professional association databases. These opinions were reiterated in the Eleventh Development Plan, which also recommended that analytical budget data in fields like the cultural

economy be reformed to provide more effective and local planning data (The Eleventh Development Plan 2019–2023).

Text mining and data mining may infringe the rights of the authors when using intellectual property. On the other hand, this type of data processing is crucial for scientific study and cultural advancement. Data is typically covered by copyright protection. Unless there is an exemption, it is legal to duplicate data in a protected database either temporarily or permanently, to remove a significant portion of the data, or to reuse the data.

Globally, in 2014 the UK introduced new restrictions on copyright law for non-commercial text and works used in data mining activities. Likewise, Japan has amended its copyright law to enable text and data mining research, and the European Union has accepted that with the latest copyright law amendment, members must implicitly allow this competition in certain forms of use in text and data mining research (Carroll, 2019: 895).

Some people claimed that data mining activities have nothing to do with copyright because there is no exploitation of the work involved in these activities in the discussions that followed the development of data minin g (European Copyright Society, 2017; Hilty and Richter, 2016: 593; Sen ftleben: 2017), even though the majority of data mining activities fall un der the purview of exceptions. They claim that consent from the right ow ner must be obtained in order to enter (Rosati, 2018: 9; Guibault, 2016). In order to resolve this ambiguity, the EU Commission in the European Union proposed a directive in 2016. This directive, known as the Directive on Copyright and Related Rights in the European Union in the Market (2019/790, of the European Parliament and of the Council of April 17, 2019 on Copyright and Related Rights in the Digital Single Market and "Digital Single Amending Directives 96/9/EC and 2001/29/EC, 2019 O.J. (L 130) 92, Amending Directive 2019/790 and 96/, has been criticized by commentators for its detrimental effects on the competitive climate in Europe, particularly its exclusion of startups and independent researchers.

It is stated that although the exception provisions of the Directive are supported by a strategic innovation policy objective, they are conceptually incorrect, theoretically flawed, and normatively unpretentious. New exceptions have been introduced to copyright binding member states in

order to access more content beyond the borders of the directive, for research, cultural heritage and education purposes, to use copyright elements more and to operate the copyright market better (Quintais, 2020: 7).

Article 5(1) of the InfoSoc Directive or Infosoc regarding reproductions for scientific research. Since the regulations in article 5(3) (a) are ambiguous for data mining, it has become mandatory to apply Articles 3 and 4 of the Digital Single Market Directive by passing the member states into their own legislation (Strowel and Ducato, 2019: 310).

Data mining is also a key tool in recent artificial intelligence inventions (Strowel and Ducato, 2021: 299-316). The innovations brought by this type of technology create the need for regulation in-laws. The European legislator has introduced two specific exceptions to copyright and related rights in the Digital single market. The first of these exceptions is related to text and data mining activities, especially for scientific research purposes, which are included in Article 3 of the Directive (Quintais, 2020: 8). In article 2, the Directive defines text and data mining as any automated analytical method used to analyze text and data. In article 8, it is defined as the automatic computational analysis of information in digital forms such as sounds, images, or data enabled by new technologies (Margoni, Kretschmer, 2021: 6). The exception in article 3 of the directive is the reproduction of copyrighted works, the use of all or a significant part of the databases in optional press releases. Research institutions and cultural heritage organizations benefit from this exception, provided that there is no commercial use (Strowel and Ducato, 2021: 301; Margoni, Kretschmer, 2021: 4).

Article 3 of the Directive has limited/exempted the copyrights of the authors of reproduction and dissemination so that member states, research organizations, and cultural heritage organizations can conduct text and data mining on works and other materials to which they have legal access rights, for the purpose of scientific research (Strowel et al. Ducato, 2019, 301; Quintais, 2020: 8). Article 3 of the Directive has excluded the text and data mining of research institutions and cultural heritage institutions for research purposes. Universities are excluded from this scope. Use by politicians to test draft policies and for regulations on new legal interventions, or by journalists to check the authenticity of news is excluded (Strowel and Ducato, 2019: 303).

Copies created within the scope of an activity carried out in accordance with the first paragraph in accordance with Article 3/2 of the Directive can be stored and kept for scientific purposes, provided that they are in a secure environment. Pursuant to Article 3/3 of the Directive, right holders will also be able to take the necessary measures to protect the security of networks or databases containing data mining works or related material. Rights holders will be allowed to implement measures to ensure the security and integrity of networks and databases where works and other subjects are hosted. (Nuredin, 2022) Such measures shall not go beyond what is necessary to achieve this aim. Member States shall encourage beneficiaries, research institutions, and cultural heritage institutions to identify commonly accepted best practices for the implementation of the obligation and measures referred to in paragraphs 2 and 3 of Article 3 of the Directive, respectively.

While leaving EU-based especially small and medium-sized businesses and startups at a level, multinational companies such as Google, Facebook, Amazon, and Microsoft will be able to engage in such activities without prior permission and will significantly reduce the costs of artificial intelligence developments. Another disadvantage is that they do not fall within the scope of the exceptions in Articles 3 and 4 from accessing online archives for questioning and verification in journalism activities and fighting fake news. (Margoni, Kretschmer, 2021: 10).

The 2019/790 EU directive has tried to accept the competitive benefit provided by the fair use clause in the US copyright law and has tried to balance the interests between the authors and the beneficiaries of data mining for scientific research purposes.

Article 4 of the Directive regulates that data mining can be done by anyone for any purpose, but it can be reserved by the right holders (Margoni, Kretschmer, 2021: 4). More broadly than Article 3 of the Directive, the rights of reproduction and processing also cover computer programs, regardless of whether they are for profit or not.

Article 5(1) of the Infosoc Directive relates to the reproduction and processing of images that are accessed legally for the purpose of data mining activity, which is not covered by the temporary copying exception. This reservation shall not affect the application of the data mining exception for scientific purposes specified in Article 3. However, as noted by many researchers, the exceptions are narrow and limited in the directive (Strowel

and Ducato, 2019; Geiger et al. 2018-2). However, the public interest must come first. It is argued that the limitations imposed on databases and computer programs in the field of copyright in the directive are still not sufficient for text and data mining activities, especially for data-oriented technologies such as machine learning and artificial intelligence ((Strowel and Ducato, 2021: 301).

According to article 2/1 of the directive, a research institution; A non-profit university, which uses all its profits for scientific studies or for the benefit of a public interest recognized by a member state, whose primary purpose is to carry out scientific research studies or educational activities, including the conduct of scientific research, and within these scope university libraries, research institutes or another institution. If there is an undertaking that has a decisive influence on the research institution and if this institution can benefit primarily from the research results, this research institution is not within the scope of the Directive. A cultural heritage institution is also defined in Article 2/3 of the Directive as a library or museum, archive or film or sound heritage institution open to public access. Those operating under these organizations will also benefit from the exemption granted to these organizations.

By means of open access or by providing access within the framework of a contract between the right holders and the research or cultural heritage institution, access to the content of those affiliated with these institutions and covered by the subscription will provide the lawful access condition sought by the Directive.

In the 4th article of the Directive, unlike the 3rd article, the text and data mining exception limited the will of the right holder (Strowel and Ducato, 2019: 302; Çonkar, 2020: 686-688). According to the regulation made by Articles 3 and 4 of the Directive, text and data mining activity does not necessarily constitute an infringement of copyrights or neighboring rights. Temporary reproductions of copyrighted material may be exempted.

Legal regulation has been brought for these acts that do not meet the temporary copy exception conditions in InfoSoc Directive 5(1). The new exception limits the rights of the rightholder in use through machine-readable tools made available to the public online, for example, in the case of content through the terms and conditions of a service and website and the use of metadata. This Directive shall not affect the scientific uses in article 3. Exceptions brought to the Directive are criticized for being narrow-scoped.

Article 5 of the Directive provides an exception for the use of digital works across borders for educational purposes only. These libraries exclude other purely commercial institutions such as museums (Strowel and Ducato: 2019).

In the evaluation of the violation of the right to reproduce databases in accordance with the economic structure of copyright, text and data mining research may be prohibited depending on the technique used in text and data mining, except for the 2009/24 EC software directive 96/9 EC Database directive. These; Reproduction of copyright content is the extraction of a substantial part of the database, and the reproduction and processing of computer programs. In addition, the text and data mining exceptions introduced in the Directive should be transferred to the legislation of the countries.

The US Copyright Law regulates that these uses remain within the fair use doctrine as long as the act does not constitute infringement. Article 107 of the American Copyright Act grant states that the reproduction of research copies is not an infringement of copyright (Carroll, 2019 ;909).

In computational data analysis activity, US copyright law provides a competitive advantage to researchers. This flexibility does not exist as EU law brings limited exceptions to text and data mining activities. Considering together the limitation on copies listed in Section 106 (a) of the US Copyright Act and the fair use principle in Article 107, they may legally conduct computational research on data or scientific papers they have access to, as long as they have a durable output of research that does not contain more original wording than is permitted. In addition, these researchers may retain full-text articles, expand the analysis of their computational research, or reproduce their research results for others to access.

In addition, these copies cannot be used as an article backup source for readers. If computational copies of a scientific article are taken from an infringing source, it does not pose a problem and falls within the scope of fair use according to American copyright law. This analysis is still valid even if the researcher buys the scientific article from a site that violates rights. Because computational copies made from a computer are not considered such a resource and fall within the scope of fair use. Reproduction of them will not affect markets under US Copyright Law (Carroll, 2019: 963).

In the age of big data, text and data mining is still a new tool, especially for research conducted by research institutions such as universities or institutes. It is thought that the narrowness of the EU directives to those who benefit from copyright exceptions will prevent the development of text and data mining applications in Europe.

Comprehensive regulation has been introduced in this Directive, which is not among the restrictions and limitations adopted by the Digital Single Market Directive and international agreements such as TRIPS, which has accepted limitations in the field of copyright to allow EU law computational research in carrying out text and data mining activity.

It is important to keep in mind the fact that text and data mining is not about changing the existing content of data, but about extracting more information. Due to the nature of text and data mining, it does not have a competitive structure by distorting data content or disrupting existing models (Rosati, 2018: 430). In order to make the EU fit for the age of big data, a reform has been proposed to expand the scope of text and data mining (Geiger et al. (2018: 814–844). The custodian who creates temporary copies of the data to be analyzed should be shared with other researchers, and copies of these copies should be distributed to other researchers. uses the reproduction rights of the copyright holders.

While the way of cultural development is opened with the processing of ideas and the presentation of works, benefiting from these works or other intellectual products is also an element that ensures development. The fact that the restrictions do not have any exceptions or are very limited prevents the public from accessing and benefiting from intellectual products, so development is hindered, while the exaggerated application of the legal restrictions and exceptions may cause the efforts of the intellectual product owners to be unrewarded. It is a fact that these exceptions should be updated and re-evaluated as a result of technological developments.

CONCLUSION

In the field of informatics; "Data mining" has emerged as a way, a data analysis technique, in order to extract the necessary and correct data from the heaps of data collected in databases, to analyze and interpret information, to make future predictions from past data, and to facilitate this.

Technological developments all over the world necessitate data mining applications in our country. Interest in text and data mining is increasing day by day and its usage areas are expanding. Data mining activities are an open potential area that increases efficiency in the fields of research and the business world. Data mining studies carried out in our country are carried out in every field such as education, trade, insurance, engineering, banking, stock market, medicine and telecom. Most of companies focus on how to increase their sales and how to attract attention with customer/user analysis. Like the use of data mining in many areas, it will provide great convenience and opportunity to intellectual property owners in the field of intellectual property.

I believe that adopting data mining techniques will result in a more equitable distribution of the salaries and other fees that the professional association will be required to collect from its members.

A significant and enthusiastic group of legal annotators with specialized training in copyright licenses, data and text mining, and a long-term period can be accessed through data mining from among numerous common and generic copyright licenses. Computational legal procedures have a lot of potentials, but for this to happen, lawyers need to be very tech- savvy and stay up to date with technology. The extent to which lawyers can use technology will determine how much they can profit from data mining.

Researchers in the field of interdisciplinary methodology are also required to have doctoral-level training proficient in applying statistical methods for legal analysis, and data and text mining and machine learning. In legal research, the application of techniques such as data mining and machine learning, text and data analytics is a new methodological approach. It is necessary to include undergraduate and graduate doctorate courses on data mining in the curriculum of law faculties, and to provide training on computational law methods (Margoni, 2020: 3).

On the other hand, in the face of the advantages of innovations brought by technology such as data mining, artificial intelligence and deep learning, it also raises the suspicion that it will be a development that can devalue and discard the long-term effort and intellectual efforts of the authors to create works.

The storing researcher, who creates temporary copies of the data to be analyzed, uses the copyright holders' right of reproduction to share these copies with other researchers and distribute their copies. The US Copyright Law regulates that these uses remain within the fair use doctrine as long as they do not constitute infringement.

While many countries in the world are making arrangements in their laws on text and data mining, EU countries have added this to the limitations in copyright laws in this area with the EU Directive No. 2019/790.

In parallel with technological developments in the field of copyright, some arrangements and updates are made. With the Digital Single Market EU Directive in Europe, exceptions and limitations regarding data mining have been introduced. These limitations in data mining activities have made it easier for researchers. On the other hand, exaggerating and expanding these limitations may prevent and discourage intellectual property owners from creating intellectual products in different sectors. The balance between these two must be maintained very well.

As can be seen, regarding the rapidly increasing text and data mining activities, the EU Directive has introduced exceptions that are not found in Law No. 5846.

It is necessary to balance the public interest in these activities with the interests of the authors against the new limitations in favor of text and data mining activities that violate the reproduction and transmission rights of the authors and database owners. In this context, it is necessary to train lawyers who can choose the most appropriate technique for the problem.

REFERENCES

ABİDİN, D., ÖZTÜRK, Ö. ve ÖZACAR, T. Ö. (2017) Klasik Türk müziğinde makam tanıma için veri madenciliği kullanımı, Gazi Üniversitesi Mühendislik Mimarlık Dergisi, 32(4),1221-1232.

ADEKİTAN, A. I., ABOLADE, J. ve SHOBAYO, O. (2019) Data mining approach for predicting the daily Internet data traffic of a smart university. J Big Data 6(11), 1–23.

AKPINAR, H. (2000). Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği, İstanbul Üniversitesi işletme Fakültesi Dergisi, Nisan 2000, 29(1), 1-22.

ARSLAN, H., (2008). Sakarya Üniversitesi Web Sitesi Erişim Kayıtlarının Web Madenciliği ile Analizi" Sakarya Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi.

BAYKAL, A.(2006). Veri Madenciliği Uygulama Alanları, D.Ü. Ziya Gökalp Eğitim Fakültesi Dergisi, 95-107.

CARROLL, M. W.,(2019). Copyright and the Progress of Science: Why Text and Data Mining Is Lawful. 53 UC Davis Law Review 893-964, American University, WCL Research Paper No. 2020-15, https://ssrn.com/abstract=3531231, Erişim Tarihi:22.05.2022.

ÇONKAR, M. H.(2020). *Dijital Tek Pazarda Telif Hakkı ve Bağlantılı Haklar*, Tütk Hukukunun Avrupa Birliği Hukukuna Uyumu Özel Hukuk Kitap Bölümü, (Editörler: Arslan Kaya, Baki İlkay Engin, Ali Paslı, Şehriban İpek Aşıkoğlu, Elif Oğuz) İstanbul Üniversitesi Yayınları.

DUCATO, R. ve STROWEL, A."Limitations to Text and Data Mining and Consumer Empowerment. Making the Case for a Right to 'Machine Legibility'" (19 March 2019), Kluwer Copyright Blog, http://copyrightblog.kluweriplaw.com/2019/03/19/limitations-to-text-and-data-mining-and-consumerempowerment-making-the-case-for-a-right-to-machine-legibility/

DONGSONG, Z. ve ZHOU, L. (2004)."Discovering Golden Nuggets: Data Mining in Financial Application, IEEE Transactions on Systems, Applications and Reviews, 34(4), 513-515.

EKER, H.,(2008) "Veri Madenciliği veya Bilgi Keşfi", http://www.bilgiyonetimi.org/cm/pages/mkl_gos.php?nt=538.

ERVERDİ, Z. (2011). Eser Sahipleri Meslek Birlikleri ile Üyeleri Arasındaki İlişki Müzik Meslek Birlikleri Üzerinden Bir Değerlendirme, İstanbul Bilgi Üniversitesi Sosyal Bilimler Enstitüsü Hukuk Yüksek Lisans Tezi.

GEIGER, C., FROSIO, G. ve BULAYENKO, O. (2018). Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data?. *IIC* **49,** 814–844. https://doi.org/10.1007/s40319-018-0722-2

GEIGER, C. , FROSIO , G. ve BULAYENKO, O., (2018-2) "The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in

the Digital Single Market - Legal Aspects", Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2018-02.

GOOD M., MusicXML for Notation and Anaylsis, Computing in Musicology, 12, 113-124, 2001. https://michaelgood.info/publications/music/musicxml-for-notation-and-analysis/

GÜÇLÜTÜRK, O. G.(2018). "Blockchain: A Trustless Network or a Technologically Disguised Shift of Trust?", https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3440044

FENG, K., Mao K. (2022). "Intellectual Property Legal Problem Based on the Big Data Algorithm", The 4th International Conference on Cyber Security Intelligence and Analytics (CSIA 2022), In: Xu, Z., Alrabaee, S., Loyola-González, O., Zhang, X., Cahyani, N.D.W., Ab Rahman, N.H. (eds), Lecture Notes on Data Engineering and Communications Technologies, vol 125. Springer, Cham. https://doi.org/10.1007/978-3-030-97874-7_18 (145-150)

HE J., HAN, M. , JI, S., Du, T. ve Li. Z. (2019) Spreading social influence with both positive and negative opinions in online networks. Big Data Min Anal, Vol. 2, 100–117.

IGLESIAS, M., SHAMUILIA, S. ve ANDERBERG, A. (2021). Intellectual Property and Artificial Intelligence, A Literature review, Luxembourg: Publications Office of the European Union.

HILTY, R.M., RICHTER, H., 'Text and Data Mining (Article 3 COM(2016) 593 final)', in Modernisation of the EU Copyright Rules — Position Statement of the Max Planck Institute for Innovation and Competition, Max Planck Institute for Innovation and Competition, Munich, 2017, http://pubman.mpdl.mpg.de/pubman/item/escidoc:2470998:12/component/escidoc:2479390/E-Book%20- %20Hilty%20-%20Moscon%20-%2018.09.2017.pdf.

KALIKOV, A. (2006), Veri Madenciliği ve Bir E-Ticaret Uygulaması, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü.

KÜÇÜKALİ, C.(2017). *Fikri Hakların İhlalinden Kaynaklanan Tazminat Davası*, (1. Baskı) Ankara, Seçkin Yayınevi.

LINOFF, G. S. ve Berry, M. J.A, "Data Mining Techniques: For Marketing. Sales and Customer Relationship Management" Chapter 1 "What is Data Mining and Why Do It?" http://www.data-miners.com/companion/BDMTnotes.pdf.

LIYING L. (2019). Analysis and data mining of intellectual property using GRNN and SVM, Published online: 18 November 2019 # Springer-Verlag London Ltd., part of Springer Nature 2019, Personal and Ubiquitous Computing (2020) 24:139–150 https://doi.org/10.1007/s00779-019-01344-8

MARGONI, T. (2020). Text and Data Mining in Intellectual Property Law: Towards an Autonomous Classification of Computational Legal Methods.

NUREDIN, A. (2022). Uluslararası İnsan Hakları Hukuku. International Vision University.

OKAMOTO, K.(2009). Teaching Transactional Lawyering, 1 Drexel L. Review 69(83). https://drexel.edu/~/media/Files/law/law%20review/inaugural/Okamoto.ashx

PIATETSKY-SHAPIRO, G., FRAWLEY, W. J., ve MATHEUS, C. J., "Knowledge Discovery in Databases", AAAI/MIT Pres, 1991.

ROSATI, E. (2018) An EU text and data mining exception for the few: Would it make sence?, Editorial Journal of Intellectual Property Law &Practice, 13(6), 429-430. http://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_BRI%282018%29604942

ROSATI, E. (2018). The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market-Technical Aspects. https://www.europarl.europa.eu/RegData/etudes/BRIE/2018/604942/IPOL_BRI(2018)604942_EN.pdf

SAVAŞ, S., TOPALOĞLU, N. ve YILMAZ, M. "Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri, İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi 11 (21) Bahar 2012, 1-23.

STROWEL, A., DUCATO, R., Artificial Intelligence and Text and Data Mining, A copyright carol, 22 April 2021, The Routledge Handbook of Copyright Law, London. pp. 299-316.

STROWEL, A., and DUCATO, R. (2019). "Limitations to Text and Data Mining and Consumer Empowerment. Making the Case for a Right to 'Machine Legibility" (19 March 2019), Kluwer Copyright Blog, http://copyrightblog.kluweriplaw.com/2019/03/19/limitations-to-text-and-data-mining-and consumerempowerment-making-the-case-for-a-right-to-machine-legibility/

SAMUELSON, P. (2018). "The EU's Controversial Digital Single Market Directive – Part II: Why the Proposed Mandatory Text- and Data-Mining Exception Is Too Restrictive" (12 July 2018) Kluwer Copyright Blog, http://copyrightblog.kluweriplaw.com/2018/07/12/euscontroversial-digital-single-market-directive-part-ii-proposed-mandatory-text-data-mining-exception-restrictive/

SENFTLEBEN, M.(2017). 'EU Copyright Reform and Startups — Shedding Light on Potential Threats in the Political Black Box', http://www.innovatorsact.eu/wp-content/uploads/2017/03/Issues-Paper-CopyrightDirective-2.pdf

QUINTAIS, J. P., "The New Copyright in the Digital Single Market Directive: A Critical Look", EIPR, 2020, 42(1), (1-24) https://pure.uva.nl/ws/files/47491008/SSRN_id3424770.pdf.

European Copyright Society (ECS), 'General Opinion on the EU Copyright Reform Package', 24 January 2017, https://europeancopyrightsocietydotorg.files.wordpress.com/2015/12/ecs-opinion-on-eucopyright-reform-def.pdf.

WEI, Z. (2019) Application analysis of data mining technology in software engineering. Modern Inf Technol 3(15):17–18.

https://www.sbb.gov.tr/wp-content/uploads/2020/04/FikriHaklarOzelIhtisasKomisyonuRaporu.pdf