

EXCEPTION FOR TEXT AND DATA MINING FOR THE PURPOSES OF SCIENTIFIC RESEARCH IN THE CONTEXT OF LIBRARIES AND REPOSITORIES

Jakub Míšek

Jakub.Misek@law.muni.cz

Institute of Law and Technology, Masaryk University, Czech Republic

This paper is licensed under the Creative Commons license: CC BY-ND 4.0 (<http://creativecommons.org/licenses/by-nd/4.0/>).

Abstract

The Copyright Directive in the Digital Single Market aims to update copyright legislation to meet the needs of the information society. This paper will deal with its Articles 3 and 4, which introduce a mandatory exception to copyright and the sui generis database right, allowing text and data mining. Article 3 specifically serves the purposes of scientific research and Article 4 covers a general, albeit limited, exception. The first part of the paper will introduce this new institute and the context of its adoption. The second part will critically evaluate it and analyse its strengths and weaknesses. The third part will then cover the activities of libraries and repositories, including the context of public sector information publication.

Keywords

Text and data mining, copyright, database protection, public sector information, grey literature

Introduction¹

Text and data mining (hereinafter referred to as “TDM”) collectively identifies techniques by which a large amount of information² can be gathered from a large number of documents by revealing the links between various sources.³ In the context of grey literature,⁴ i.e. documents from various originators available online and easily processed by automated means, this is an extremely useful technological procedure that can be widely applied.⁵ As Truyens and Van Eeck point out, typical TDM activities include sorting text and data into one or more categories, grouping text and data together, extracting a common concept (for example, finding what the documents say), analysing document sentiment and modelling relationships between the entities contained in the text and data.⁶ Compared to standard data analysis, TDM practices differ in that data are not used to test prepared models or hypotheses but rather to discover new, currently hidden, contexts, patterns and models.⁷ Traditionally, the TDM process is carried out in three steps: accessing the analysed content, mining or copying the analysed content, and finally analysing the content itself.⁸

Apart from the fact that processed documents may be protected by special information rights such as protection of personal data, confidential information or trade secrets, intellectual property law constitutes a fundamental obstacle to TDM. Even if the data analysed or acquired during analysis are not protected in any way,⁹ TDM may infringe the author’s rights in the case of copyrighted works that are part of the analysed database or may infringe the database

¹ The author gratefully acknowledges the aid from the Masaryk University under grant No. MUNI/A/1006/2018.

² Information in this context means information in a semantic sense, that is data put into a particular context within which they create knowledge (cf. e.g. ADRIAANS, Pieter. Information. In: ZALTA, Edward N. (ed.). *The Stanford Encyclopaedia of Philosophy* [online]. Stanford: Metaphysics Research Lab, Stanford University, 2013 [Accessed 30 June 2019]. Available from: <https://plato.stanford.edu/archives/fall2013/entries/information/>; FLORIDI, Luciano. *Information: a very short introduction*. Oxford, New York: Oxford University Press, 2010, 130 p. Very short introductions, 225. ISBN 978-0-19-955137-8.

³ See ROSATI, Eleonora. Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity. *Asia Pacific Law Review* [online]. Rochester, NY: Social Science Research Network (SSRN), 2019, p. 2 [Accessed 8 October 2019]. Available from: <https://papers.ssrn.com/abstract=3452376>.

⁴ The paper works with the concept of grey literature in the sense of the so-called Prague Definition, chosen because of its timeliness and comprehensive nature. See SCHÖPFEL, Joachim. Towards a Prague Definition of Grey Literature. *Twelfth International Conference on Grey Literature: Transparency in Grey Literature: Grey Tech Approaches to High Tech Issues*. 2010, pp. 11-26.

⁵ Cf. MYŠKA, Matěj. Text and Data Mining of Grey Literature for the Purpose of Scientific Research. *The Grey Journal*. 2017, Vol. 13, p. 32.

⁶ TRUYENS, Maarten; VAN EECKE, Patrick. Legal aspects of text mining. *Computer Law & Security Review* [online]. 2014, Vol. 30, No. 2, p. 153 [Accessed 8 October 2019]. ISSN 0267-3649. Available from: <https://doi.org/10.1016/j.clsr.2014.01.009>

⁷ See LUPAȘCU, Monica. Text and Data Mining Exception - Technology into Our Lives. *International Conference: CKS - Challenges of the Knowledge Society*. 2019, p. 906. ISSN 2068-7796.

⁸ See ROSATI, Eleonora. Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity. *Asia Pacific Law Review* [online]. Rochester, NY: Social Science Research Network (SSRN), 2019, p. 8-15 [Accessed 8 October 2019]. Available from: <https://papers.ssrn.com/abstract=3452376>

⁹ Cf. e.g. KOŠČÍK, M. Duševní vlastnictví k výzkumným datům. In: KOŠČÍK, Michal et al. *Výzkumná data a výzkumné databáze. Právní rámec zpracování a sdílení vědeckých poznatků*. Prague: Wolters Kluwer ČR, 2018, p. 34. ISBN 978-80-7552-952-7.

copyright.¹⁰ It may also affect sui generis database right within the meaning of Article 7 et seq. of Directive 96/9/EC.¹¹ The main problem for TDM thus lies in the wide scope in which are interpreted the concepts of reproduction (in the context of copyrighted works),¹² and mining¹³ (in the context of the sui generis database right of the database maker).¹⁴

In order for a modern society to fully benefit from the application of TDM techniques, it is essential to ensure the comprehensibility and clarity of the relevant legislation. It must enable this analytical procedure to be carried out without the risk of infringement of intellectual property rights and the consequent legal sanctions. Senseney and Koehl, for example, argue that in the context of the use of TDM for scientific research, researchers experience a 'chilling effect'¹⁵ when exposed to legal uncertainty about the consequences of their actions.¹⁶ An effective system of exceptions and limitations to copyright and the sui generis database right is thus essential for the effective application of TDM techniques.

This paper introduces the applicable exceptions and limitations to copyright and the sui generis database right of the database maker that may apply to TDM, in particular in the context of the new EU Directive 2019/790 on copyright and related rights in the Digital Single Market (hereinafter referred to as the "DSM Directive"), which newly introduced exceptions for the implementation of TDM. It also critically assesses the practical applicability of this exception and, in its third part, places the new exception in the context of the activities of libraries and repositories, in conjunction with the legislation on the re-use of public sector information.¹⁷

Exceptions for TDM

The possible application of the exceptions and limitations of copyright resulting from Directive 2001/29/EC and of the exceptions and limitations of sui generis database right resulting from Directive 96/9/EC has already been addressed by a number of authors in the context of TDM implementation.¹⁸ In the context of copyright, consideration has been given in particular to the

¹⁰ A database is protected by copyright in accordance with Article 3 of Directive 96/9/EC if it constitutes the author's own intellectual creation by way of selection or arrangement of the content.

¹¹ See TRUYENS, Maarten; VAN EECKE, Patrick. Legal aspects of text mining. *Computer Law & Security Review* [online]. 2014, Vol. 30, No. 2, p. 163-164 [Accessed 8 October 2019]. ISSN 0267-3649. Available from: <https://doi.org/10.1016/j.clsr.2014.01.009>

¹² Within the meaning of Article 2 of Directive 2001/29/EC.

¹³ Within the meaning of Article 7(1) of Directive 96/9/EC.

¹⁴ Cf. DUCATO, Rossana; STROWEL, Alain. Limitations to Text and Data Mining and Consumer Empowerment: Making the Case for a Right to "Machine Legibility". *IIC - International Review of Intellectual Property & Competition Law*. 2019, Vol. 50, No. 6, p. 658. ISSN 0018-9855.; MYŠKA, Matěj; HARAŠTA, Jakub. Omezení autorského práva a zvláštních práv pořizovatele databáze v případě datové analýzy. *Časopis pro právní vědu a praxi*. 2016, Vol. 23, No. 4, pp. 378. ISSN 1805-2789.

¹⁵ This is an undesirable form of self-censorship, where one would rather not take certain steps or activities to make sure that any negative effect, such as a legal sanction, will be avoided. For more information on the chilling effect, see e.g. PENNEY, J.W. Chilling Effects: Online Surveillance and Wikipedia Use. *Berkeley Technology Law Journal* [online]. 2016, Vol. 31, No. 1, pp. 117–182 [Accessed 8 October 2019]. ISSN 1086-3818. Available from: <https://doi.org/10.15779/Z38SS13>

¹⁶ See SENSENEY, Megan; DICKSON KOEHL, Eleanor. Text data mining beyond the open data paradigm: Perspectives at the intersection of intellectual property and ethics. *Proceedings of the Association for Information Science & Technology* [online]. 2018, Vol. 55, No. 1, p. 891 [Accessed 30 June 2019]. ISSN 2373-9231. Available from: <https://doi.org/10.1002/pr2.2018.14505501162>

¹⁷ In particular Directive 2003/98/EC, as amended by Directive 2013/37/EC, which will soon be replaced by EU Directive 2019/1024.

¹⁸ See e.g. BROOK, Michelle; MURRAY-RUST, Peter; OPPENHEIM, Charles. The social, political and legal aspects of text and data mining (TDM). *D-Lib Magazine* [online]. 2014, Vol. 20, No. 11–12 [Accessed 30 June 2019]. ISSN 1082-9873. Available from: <https://doi.org/10.1045/november14-brook>; CASPERS, Marco; GUIBAULT, Lucie. A right to 'read' for machines: Assessing a black-box analysis exception for data mining. *Proceedings of the Association for Information Science & Technology*

limitation of copyright for temporary copies (based on Article 5(1)) and to the limitation of the right of reproduction for scientific purposes.¹⁹ In the context of the sui generis database right, it was then possible to consider the rights of authorized users who could mine non-essential parts of the database,²⁰ and finally an exception allowing the mining of a publicly accessible database for scientific research purposes.²¹ However, as the texts cited in the introduction to this section have already extensively demonstrated, neither of these exceptions and limitations is broad and flexible enough to act as a suitable tool for implementing TDM.

In the context of grey literature and Czech law, we can also rely on an exception for an official work (at least in part of the examined documents), both in the case of copyright protection²² and, appropriately, in the context of the protection of the sui generis database right.²³ However, even these exceptions are not broad enough to handle a large amount of analysed data efficiently, as there are also many copyrighted works in the context of grey literature.

The factual usefulness of the implementation of TDM and the shortcomings mentioned²⁴ have inspired the European legislator to introduce, into the DSM Directive, exceptions to copyright and database rights to allow TDM. The TDM exceptions appear in the new directive as two variants, enshrined in Articles 3 and 4. Both exceptions are mandatory for Member States and must therefore be transposed into national law. Article 3 establishes the exception of text mining for scientific research. Article 4 then provides for a general exception allowing TDM, but also contains a huge number of limitations that make it difficult to apply.

Critical Evaluation of the New TDM Exception

Ducato and Strowel argue that the creation of a new comprehensive TDM exception was not appropriate and, in their view, the European legislature should have set out to formulate a new right to “electronic reading of documents” to cover TDM with sufficient flexibility.²⁵ However, if we continue to move within the current intellectual property rights paradigm, which includes a broad interpretation of the concepts of reproduction and mining, and accept that there is no right to automated text reading, the existence of an exception to copyright and the sui generis

[online]. 2016, Vol. 53, No. 1, pp. 1-5 [Accessed 30 June 2019]. ISSN 2373-9231. Available from: <https://doi.org/10.1002/pr2.2016.14505301017>; GEIGER, Christophe; FROSIO, Giancarlo; BULAYENKO, Oleksandr. Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data?: Legal Analysis and Policy Recommendations. *IIC International Review of Intellectual Property and Competition Law* [online]. 2018, Vol. 49, No. 7, pp. 814-844 [Accessed 8 October 2019]. ISSN 2195-0237. Available from: <https://doi.org/10.1007/s40319-018-0722-2>; HANNAY, William M. Legally Speaking - Of Mindfields and Minefields: Legal Issues in Text and Data Mining. *Against the Grain* [online]. 2014, Vol. 26, No. 1, pp. 52-55 [Accessed 30 June 2019]. ISSN 10432094. Available from: <https://doi.org/10.7771/2380-176X.6663>; MYŠKA, Matěj. Text and Data Mining of Grey Literature for the Purpose of Scientific Research. *The Grey Journal*. 2017, Vol. 13, pp. 32-37. ISSN 1574-1796.; MYŠKA, Matěj; HARAŠTA, Jakub. Omezení autorského práva a zvláštních práv pořizovatele databáze v případě datové analýzy. *Časopis pro právní vědu a praxi*. 2016, Vol. 23, No. 4, pp. 375-384. ISSN 1805-2789.; TRIAILLE, Jean-Paul et al. *Study on the legal framework of text and data mining (TDM)*. Luxembourg: Publications Office, 2014, p. 41. ISBN 978-92-79-31976-1., et seq.

¹⁹ See Article 5(3)a) of Directive 2001/29/EC.

²⁰ See Article 8 of Directive 96/9/EC.

²¹ *Ibid.*, Article 9b).

²² See Section 3a) of Act No 121/2000

²³ *Ibid.*, Section 94.

²⁴ These issues are analyzed in detail by e.g., TRIAILLE, Jean-Paul et al. *Study on the legal framework of text and data mining (TDM)*. Luxembourg: Publications Office, 2014. ISBN 978-92-79-31976-1.

²⁵ *Ibid.*

database right is necessary for the effective application of TDM procedures. Unfortunately, the inclusion of this exception in the DSM Directive was not without some critical moments.

The fundamental problem of Article 3 of the DSM Directive is the very narrow scope of its application resulting from the necessary simultaneous fulfilment of conditions relating to the entity carrying out the TDM (research organizations²⁶ and cultural heritage institutions), conditions relating to the purpose of the TDM (only scientific research purposes) and the necessity of legal access to the protected content.²⁷ Another problem lies in the third paragraph of Article 3, which states that rightsholders “shall be allowed to apply measures to ensure the security and integrity of the networks and databases where the works or other subject matter are hosted”, while these measures “shall not go beyond what is necessary to achieve that objective”. This is a very unclear provision whose interpretation is likely to be difficult. In practice, for example, a provider of information (datasets, copyrighted works) may limit the number of accesses to its database, justifying this by the necessity to ensure technical integrity. In fact, however, this will make it technically impossible to implement TDM and thus apply this exception. A third and fundamental problem is that the exception laid down in Article 3 of the DSM Directive concerns only the right to reproduce works or mine databases but not any other rights of the author or database maker.

In assessing the TDM exception enshrined in Article 4, we must conclude that despite the benefits resulting from the exception being ultimately mandatory for Member States,²⁸ and that the exception applies to any TDM implementation regardless of purpose or executor (assuming it has legitimate access to the protected rights), the applicability of this exception will be very problematic for two reasons. Firstly, the DSM Directive considerably limits the possible period of reproduction and extraction retention to the time required directly for the purpose of performing TDM. This greatly limits the possibility of future analyses or, for example, checks on the repeatability of the activity. This may be particularly problematic in the case of purely private research which does not benefit from the exception laid down in Article 3 of the DSM Directive. The second major problem then lies in Article 4(3), which allows rightsholders to exclude the application of this exception.²⁹

These shortcomings show that although the new TDM exception has the potential to help in certain narrow areas, it still suffers from a number of problems. In legal terms, these make the implementation of TDM only a little more certain compared to the previous legislation. This view is further supported by the fact that when performing TDM, interference with the original work is minimal. There is no further publication of the work or any parts of it. What is extracted from the work for further use is new information that would otherwise remain hidden and that is not copyrighted.³⁰ In view of this, I am in favour of the viewpoint of Ducato and Strowel that, instead of this new exception, the European legislator should have set out to formulate a new

²⁶ Recital 11 of the DSM Directive states that companies in the private sector may also be exempted if they do research in cooperation with the public sector.

²⁷ See also DUCATO, Rossana; STROWEL, Alain. Limitations to Text and Data Mining and Consumer Empowerment: Making the Case for a Right to “Machine Legibility”. *IIC - International Review of Intellectual Property & Competition Law*. 2019, Vol. 50, No. 6, p. 665. ISSN 0018-9855.

²⁸ The Council and Parliament proposals included a version which made the exception optional. Cf. *ibid*.

²⁹ See also ROSATI, Eleonora. Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity. *Asia Pacific Law Review* [online]. Rochester, NY: Social Science Research Network (SSRN), 2019, p. 20-21 [Accessed 8 October 2019]. Available from: <https://papers.ssrn.com/abstract=3452376>

³⁰ Copyright protects only the specific expression of the work, not the ideas behind it.

“electronic document reading” right that would cover TDM adequately and at the same time correspond more to the technical and legal reality.³¹

TDM, Libraries, Repositories and PSI

In the context of the activities of libraries and repositories, two options need to be considered in terms of how the new TDM exception can affect them. The first option is that the library or repository, as an institution, carries out its own TDM analysis of the documents it manages. Although many of these documents will be covered by the official work exception, repositories and libraries still function essentially as hosting ISPs for a number of documents that are copyrighted and for which the library does not need to have a direct licence to use the work further.³² However, if the TDM analysis of managed documents is carried out for scientific research purposes, libraries and repositories, as cultural heritage institutions,³³ can rely on the TDM exception enshrined in Article 3 of the DSM Directive. Moreover, in this case there would be no problems related to the sui generis database right or copyright protection of the database, since in the vast majority of cases these rights are held by libraries and repositories.³⁴

The second, more interesting, option is where an external entity wants to process resources that the library or repository manages using TDM techniques. In this case, as libraries and repositories are often set up by the public sector, we are dealing with the field of legislation concerning the provision and re-use of public sector information (PSI). This is expressed at European level by Directive 2003/98/EC as amended by Directive 2013/37/EC (hereinafter referred to as the “PSI Directive”), which will be replaced in 2021 by the recently adopted EU Directive 2019/1024 on open data and the re-use of public sector information (hereinafter referred to as the “Open Data Directive”).³⁵ In general, the obligations under the PSI Directive can be summarized in such a way that if any information (in terms of documents)³⁶ is provided, it must be done in such a way as to enable and facilitate the subsequent re-use of that information to the greatest extent possible. Of course, TDM is also an example of re-use.

From the point of view of libraries and repositories, the crucial question is whether the library can publish its databases and effectively enable TDM to be performed. Again, the main issue will be the copyrighted works that are part of the collection because the sui generis database right and database copyright can be licensed by the library or repository to allow TDM without the need for an exception. In Article 1(6), the Open Data Directive even forbids the regulated

³¹ See DUCATO, Rossana; STROWEL, Alain. Limitations to Text and Data Mining and Consumer Empowerment: Making the Case for a Right to “Machine Legibility”. *IIC - International Review of Intellectual Property & Competition Law*. 2019, Vol. 50, No. 6, pp. 649-684. ISSN 0018-9855.

³² For more information on a grey literature repository like an ISP, see MYŠKA, Matěj; ŠAVELKA, Jaromír. A Model Framework for publishing Grey Literature in Open Access. *Journal of Intellectual Property, Information Technology and E-Commerce Law*. 2013, Vol. 4, p. 109. ISSN 2190-3387.

³³ Cf. recital 13 of the DSM Directive.

³⁴ A similar case would be where a library or repository wishes to use the resources of other libraries (or, in general, resources held by third parties) to carry out TDM for scientific purposes. Again, given the nature of libraries as cultural heritage institutions, it would generally be possible to carry out a TDM analysis provided that the conditions set out in Article 3 of the DSM Directive are met.

³⁵ The Open Data Directive builds on current legislation in basic procedures and principles, and only introduces a few new obligations that *de facto* only reinforce them.

³⁶ The PSI legislation works with the concept of information in the sense of a “document that records content”, which corresponds to Buckland’s concept of information. See BUCKLAND, Michael Keeble. Information as a Thing. *Journal of the American Society for Information Science and Technology*. 1991, Vol. 42, No. 5, pp. 351-360. ISSN 2330-1643.

entities to exercise the sui generis database rights if this prevents the re-use of documents. The PSI Directive does not affect documents protected by copyright (and related rights) of third parties.³⁷ However, this conclusion must be interpreted as meaning only cases where the provision of the works in question would infringe the copyrights of third parties.³⁸ If the library or repository is licensed to distribute the work and communicate it to the public, it may make the works available online. In that case, the obligations arising from the PSI or the Open Data Directive will be fully applied.

However, in addition to publication, it is necessary to consider the possibility of the obligation to provide information upon request under the Act on Free Access to Information³⁹ for libraries and repositories that are also regulated subjects under that Act. The second possibility for disseminating copyrighted works and other resources for TDM is therefore to comply with a request for that information. This can not only provide a copy of the database with copyrighted works but also provide direct access to the database of works through the API.⁴⁰ In the context of Czech law, in certain cases it is possible to request information pursuant to Act No 106/1999 Sb. and provide third-party copyrighted documents even without a license from the author or rights executor. This is due to the statutory license to use the work for official use enshrined in Section 34 of Act No 121/2000 Sb., the Copyright Act.⁴¹ However, a necessary requirement for using the statutory license is to pass a three-step test.⁴² The third step of the three-step test is that the use of the work must not unduly affect the legitimate interests of the author. In the context of providing a copyrighted database of works for the purpose of TDM, the legal situation of the information applicant can be taken into account when evaluating the third step of the three-step test. We will therefore examine whether it qualifies for the possible application of the TDM exception and, if so, it can be argued that the provision of the works in question is not contrary to copyright protection because the rights of the author are not unduly affected by this process. In conclusion, however, it must be pointed out that the TDM exception itself is not sufficient as legal title for the publication of copyrighted works.

In the context of the evaluation of the possibility of subsequent use of works through TDM, this will undoubtedly depend on the entity in question, the purpose of the TDM and, where relevant, whether copyright executors have restricted the possibility of applying the TDM exception within the meaning of Article 4(3) of the DSM Directive. If such a decision has been made, it is essential that this information is sufficiently accessible to other potential users, for example through a metadata record attached to the work in question. An important conclusion, however, is that access to works deposited in the repository, whether the works have been published in accordance with a license or under a statutory official license pursuant to Section 34 of the

³⁷ See Article 1(2)b) of the PSI Directive.

³⁸ See also Czech legislation, namely Section 11(2)c) of Act No 106/1999, on Free Access to Information, as amended.

³⁹ Typically, in the context of Czech law, these will be 'public institutions', i.e. institutions established by the state or self-governing units or fully owned by them. For more information, see Ruling of the Constitutional Court of 24 January 2007, file No I. ÚS 260/06, No N 10/44 SbNU 129 [available from <http://nalus.usoud.cz>, accessed 10 November 2019].

⁴⁰ For more information on access to and re-use of public sector information, in particular copyrighted works, see e.g. EECHOUD, Mireille Van; JANSSEN, Kathleen. Rights of Access to Public Sector Information. *Masaryk University Journal of Law and Technology*. 2013, Vol. 6, No. 3, pp. 471-499. ISSN 1802-5951.; GILCHRIST, John. Accessing and Reusing Copyright Government Records. *Queensland University of Technology Law & Justice Journal*. 2010, Vol. 10, No. 2, pp. 213-232. ISSN 1445-6230.

⁴¹ Cf. TELEČ, Ivo; TŮMA, Pavel. *Copyright Act - Commentary*. 2nd edition. Prague: C. H. Beck, 2019, p. 414. ISBN 978-80-7400-748-4.

⁴² See Article 5(5) of Directive 2001/29/EC. Cf. also e.g. GEIGER, Christophe; GERVAIS, Daniel; SENFTLEBEN, Martin. The Three-Step Test Revisited: How to Use the Test's Flexibility in National Copyright Law. *American University International Law Review*. 2013, Vol. 29, No. 3, pp. 581-626. ISSN 1520-460X.

Copyright Act, will constitute legitimate access to the works, which is a prerequisite for the application of TDM exceptions under the DSM Directive.

Conclusion

With the DSM Directive and the exception allowing the implementation of TDM, the European legislator has embarked on a journey, the success and feasibility of which will only be decided over time. As the preliminary analyses show, the two formulated exceptions for TDM are applicable only to a very narrow number of cases. This is either because of the narrow focus, as in Article 3, or because of the factual option of the rightsholder to exclude the application of that exception, as in Article 4. The exception will have no far-reaching impact for libraries and repositories. It can be expected that this exception will facilitate the work, legal status and certainty of libraries wishing to conduct TDM analysis of the documents they manage. For external entities interested in carrying out TDM with respect to grey literature documents in depositories and libraries, there is an interesting synergy with the PSI legislation, as the TDM exception establishes a specific way to legally handle copyrighted documents and also slightly facilitates access to such documents.

References

ADRIAANS, Pieter. Information. In: ZALTA, Edward N. (ed.). *The Stanford Encyclopaedia of Philosophy* [online]. Stanford: Metaphysics Research Lab, Stanford University, 2013 [Accessed 30 June 2019]. Available from:

<https://plato.stanford.edu/archives/fall2013/entries/information/>

BROOK, Michelle; MURRAY-RUST, Peter; OPPENHEIM, Charles. The social, political and legal aspects of text and data mining (TDM). *D-Lib Magazine* [online]. 2014, Vol. 20, No. 11–12 [Accessed 30 June 2019]. ISSN 1082-9873. Available from:

<https://doi.org/10.1045/november14-brook>

BUCKLAND, Michael Keeble. Information as a Thing. *Journal of the American Society for Information Science and Technology*. 1991, Vol. 42, No. 5, pp. 351-360. ISSN 2330-1643.

CASPERS, Marco; GUIBAULT, Lucie. A right to ‘read’ for machines: Assessing a black-box analysis exception for data mining. *Proceedings of the Association for Information Science & Technology* [online]. 2016, Vol. 53, No. 1, pp. 1-5 [Accessed 30 June 2019]. ISSN 2373-9231. Available from: <https://doi.org/10.1002/pr2.2016.14505301017>

DUCATO, Rossana; STROWEL, Alain. Limitations to Text and Data Mining and Consumer Empowerment: Making the Case for a Right to “Machine Legibility”. *IIC - International Review of Intellectual Property & Competition Law*. 2019, Vol. 50, No. 6, pp. 649-684. ISSN 0018-9855.

EECHOUD, Mireille Van; JANSSEN, Katleen. Rights of Access to Public Sector Information. *Masaryk University Journal of Law and Technology*. 2013, Vol. 6, No. 3, pp. 471-499. ISSN 1802-5951.

FLORIDI, Luciano. *Information: a very short introduction*. Oxford; New York: Oxford University Press, 2010, 130 p. Very short introductions, 225. ISBN 978-0-19-955137-8.

GEIGER, Christophe; FROSIO, Giancarlo; BULAYENKO, Oleksandr. Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data?: Legal Analysis and Policy Recommendations. *IIC International Review of Intellectual Property and Competition Law* [online]. 2018, Vol. 49, No. 7, pp. 814-844 [Accessed 8 October 2019]. ISSN 2195-0237. Available from: <https://doi.org/10.1007/s40319-018-0722-2>

GEIGER, Christophe; GERVAIS, Daniel; SENFTLEBEN, Martin. The Three-Step Test Revisited: How to Use the Test's Flexibility in National Copyright Law. *American University International Law Review*. 2013, Vol. 29, No. 3, pp. 581-626. ISSN 1520-460X.

GILCHRIST, John. Accessing and Reusing Copyright Government Records. *Queensland University of Technology Law & Justice Journal*. 2010, Vol. 10, No. 2, pp. 213-232. ISSN 1445-6230.

HANNAY, William M. Legally Speaking - Of Mindfields and Minefields: Legal Issues in Text and Data Mining. *Against the Grain* [online]. 2014, Vol. 26, No. 1, pp. 52-55 [Accessed 30 June 2019]. ISSN 10432094. Available from: <https://doi.org/10.7771/2380-176X.6663>

KOŠČÍK, Michal et al. Výzkumná data a výzkumné databáze. *Právní rámec zpracování a sdílení vědeckých poznatků*. Prague: Wolters Kluwer ČR, 2018, 180 p. ISBN 978-80-7552-952-7.

LUPAȘCU, Monica. Text and Data Mining Exception - Technology into Our Lives. *International Conference: CKS - Challenges of the Knowledge Society*. 2019, pp. 905–914. ISSN 2068-7796.

MYŠKA, Matěj. Text and Data Mining of Grey Literature for the Purpose of Scientific Research. *The Grey Journal*. 2017, Vol. 13, pp. 32-37. ISSN 1574-1796.

MYŠKA, Matěj; HARAŠTA, Jakub. Omezení autorského práva a zvláštních práv pořizovatele databáze v případě datové analýzy. *Časopis pro právní vědu a praxi*. 2016, Vol. 23, No. 4, pp. 375-384. ISSN 1805-2789.

MYŠKA, Matěj; ŠAVELKA, Jaromír. A Model Framework for publishing Grey Literature in Open Access. *Journal of Intellectual Property, Information Technology and E-Commerce Law*. 2013, Vol. 4, pp. 104-115. ISSN 2190-3387.

PENNEY, J.W. Chilling Effects: Online Surveillance and Wikipedia Use. *Berkeley Technology Law Journal* [online]. 2016, Vol. 31, No. 1, pp. 117–182 [Accessed 8 October 2019]. ISSN 1086-3818. Available from: <https://doi.org/10.15779/Z38SS13>

ROSATI, Eleonora. Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity. *Asia Pacific Law Review* [online]. Rochester, NY: Social Science Research Network (SSRN), 2019 [Accessed 8 October 2019]. Available from: <https://papers.ssrn.com/abstract=3452376>

SENSENEY, Megan; DICKSON KOEHL, Eleanor. Text data mining beyond the open data paradigm: Perspectives at the intersection of intellectual property and ethics. *Proceedings of the Association for Information Science & Technology* [online]. 2018, Vol. 55, No. 1, pp. 890-891 [Accessed 30 June 2019]. ISSN 2373-9231. Available from: <https://doi.org/10.1002/pr2.2018.14505501162>

SCHÖPFEL, Joachim. Towards a Prague Definition of Grey Literature. *Twelfth International Conference on Grey Literature: Transparency in Grey Literature: Grey Tech Approaches to High Tech Issues*. 2010, pp. 11-26.

TELEC, Ivo; TŮMA, Pavel. *Copyright Act - Commentary*. 2nd edition. Prague: C. H. Beck, 2019, 1295 p. ISBN 978-80-7400-748-4.

TRIAILLE, Jean-Paul et al. *Study on the legal framework of text and data mining (TDM)*. Luxembourg: Publications Office, 2014. ISBN 978-92-79-31976-1.

TRUYENS, Maarten; VAN EECKE, Patrick. Legal aspects of text mining. *Computer Law & Security Review* [online]. 2014, Vol. 30, No. 2, pp. 153-170 [Accessed 8 October 2019]. ISSN 0267-3649. Available from: <https://doi.org/10.1016/j.clsr.2014.01.009>