



**UNIVERSITY  
OF TURKU**

This is a pre-copy-edited, author-produced version of an article accepted for publication in *European Law Review* following peer review. The definitive published version (*Maryna Manteghi 2023, In Search of Balance: Text, Data Mining and Copyright in the Digital Single Market Directive from a Fundamental Rights Perspective. European Law Review 48 (4), 443-457*) is available online on Westlaw UK or from Thomson Reuters DocDel service.

AUTHOR	Maryna Manteghi
TITLE	In Search of Balance: Text, Data Mining and Copyright in the Digital Single Market Directive from a Fundamental Rights Perspective
YEAR	2023
VERSION	Author's accepted manuscript
LICENSE	CC BY-NC

# **In Search of Balance: Text, Data Mining and Copyright in the Digital Single Market Directive from a Fundamental Rights Perspective**

**Maryna Manteghi**

*University of Turku*

Copyright; Data mining; Digital technology; EU law; Freedom of information; Fundamental rights; Research and development

## **Abstract**

*Text and data mining (TDM) is integral to data-driven innovation. This advanced analytical technique allows users to generate new knowledge from the plethora of digital data. The EU regulation of TDM is based on mandatory exceptions provided under arts. 3 and 4 of the Directive on Copyright in the Digital Single Market (DSM Directive). Even though these provisions have improved the regulation of TDM compared to the legal uncertainty that existed before the adoption of the DSM Directive, their effectiveness could be challenged in terms of fundamental rights. This article argues that the exceptions fail to take into account the freedom of information and the freedom of research in a balanced way. These fundamental freedoms have been only addressed superficially. The study raises constitutional concerns regarding the purpose limitation of art. 3 and its narrow scope of beneficiaries, the “lawful access” requirement, the storage conditions, the “opt-out” mechanism of art. 4, and the application of technological protection measures. In this regard, the paper proposes amendments to the regulation of TDM so that the balance between the fundamental rights to information and research and the fundamental right to intellectual property would be realized better than in the current law. The findings of this article are crucial for developing efficient and consistent legal grounds for carrying out TDM at EU-wide level, which would enhance research and innovation across the EU Member States.*

## **Introduction**

Data-driven innovation has become a primary engine of progress and growth in the 21<sup>st</sup> century. Advanced types of data usage trigger the emergence of new actors and the development of novel

business strategies in the digital economy.<sup>1</sup> In 2019, the EU legislator adopted the Directive on Copyright in the Digital Single Market (the DSM Directive)<sup>2</sup> to address rapid technological developments and eliminate legal uncertainty for both rightholders and users concerning certain uses of works in the digital environment.<sup>3</sup> Text and data mining (TDM) has been recognised as a significant advancement in the scientific field and a research tool which is essential for promoting innovation.<sup>4</sup> This analytical technique is employed to generate new knowledge from existing structured and unstructured digital data<sup>5</sup> and thereby harness a massive growth of digital information. TDM is one of the most significant recent developments in research and the production of new knowledge. This analytical technique is indispensable for the development and functioning of innovative artificial intelligence (AI) technologies, which require broader access to large masses of information.<sup>6</sup> For instance, TDM is crucial for the operation of search engines which would not be capable to explore huge repositories of digital data, available on the Internet, without relevant algorithms and machine learning.

However, the EU copyright and *sui generis* database protection could restrict a substantial part of TDM research. To mine, a computer is required to create copies of data which could be protected under these legal regimes. Consequently, TDM may potentially infringe the copyright holders' exclusive right to reproduction<sup>7</sup> or the database holders' *sui generis* right granted to those actors who made a substantial investment in the creation of a database.<sup>8</sup> To clarify the extent to which research organisations and the private sector could lawfully perform TDM, the DSM Directive has introduced two mandatory exceptions under arts. 3 and 4. Even though these provisions have improved the regulation of TDM compared to the legal uncertainty that existed before the adoption

---

<sup>1</sup> Jose Ramon Saura, Domingo Ribeiro-Soriano and Daniel Palacios-Marques, "From User-generated Data to Data-driven Innovation: A Research Agenda to Understand User Privacy in Digital Markets" (2021) 60 IJIM 1, 1.

<sup>2</sup> Directive 2019/790 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L 130/92.

<sup>3</sup> DSM Directive, recital 3.

<sup>4</sup> DSM Directive, recital 8.

<sup>5</sup> DSM Directive, art. 2 (2) defines TDM as "any automated analytical technique aimed at analyzing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations."

<sup>6</sup> Rossana Ducato and Alain Strowel, "Limitations to Text and data Mining and Consumer Empowerment: Making the Case for a Right to "Machine Legibility" (2019) 50 (6) IIC 649, 651; Sean M. Fiil-Flynn et al, "Legal reform to enhance global text and data mining research" (2022) 378 (6623) Science 951, 951.

<sup>7</sup> Directive 2001/29 on the harmonisation of certain aspects of copyright and related rights in the information society [2001] OJ L 167/10 (InfoSoc Directive), art. 2.

<sup>8</sup> Directive 96/9 on the legal protection of databases [1996] OJ L 77/20 (Database Directive) art. 7 (1).

of the DSM Directive,<sup>9</sup> their effectiveness could be challenged in terms of fundamental rights. The idea behind this legislative initiative was to address the advancement of the EU's competitive position as a research area,<sup>10</sup> but in practice, both exceptions have introduced significant limitations on the right to TDM<sup>11</sup>. In particular, art. 3 only addresses scientific research and benefits merely research organisations and cultural heritage institutions.<sup>12</sup> Although art. 4 seems to be broader as it covers all purposes for the benefit of any (public or private) entity, the possibility to reserve the use of works by their rightholders (a so-called "opt-out" mechanism) significantly limits the scope of this exception.<sup>13</sup> Moreover, both provisions require that the users have "lawful access" to mined works. In this case, TDM research would depend on the will of the rightholder, who may allow, deny or restrict, through licensing or subscription, the access to content.<sup>14</sup> Further, even though arts. 3 and 4 grant the beneficiaries the right to store copies of works made in the course of TDM, many aspects remain unclear regarding the storage conditions, the issues of responsibility, and access to retained copies. And finally, the application of technological protection measures (TPMs), which is explicitly allowed under art. 3, has not been properly clarified. Moreover, no safeguards have been introduced to reduce the risk of erroneous or unjustified blocking of lawful access.

The limitations of arts. 3 and 4 has raised constitutional concerns as the restrictions on TDM are, at the same time, restrictions on the right to information protected by art. 11 of the EU Charter of

---

<sup>9</sup> It is argued that the exceptions and limitations provided under art. 5 of the InfoSoc Directive and arts. 6, 8 and 9 of the Database Directive are not applicable to TDM, see for instance Thomas Margoni and Martin Kretschmer, "The Text and Data Mining Exception in the Proposal for a Directive on Copyright in the Digital Single Market: Why it is not What EU Copyright Law Needs" (2018) <https://www.create.ac.uk/blog/2018/04/25/why-tdm-exception-copyright-directive-digital-single-market-not-what-eu-copyright-needs/> accessed 24 August 2022; Caterina Sganga, "A New Era for EU Copyright Exceptions and Limitations? Judicial Flexibility and Legislative Discretion in the Aftermath of the CDSM Directive and the Trio of the Grand Chamber of the CJEU" (2020) 21 (2) ERA Forum 311, 314.

<sup>10</sup> DSM Directive, recital 10.

<sup>11</sup> It is claimed that the provisions are narrow and normatively unambitious for regulating knowledge production performed by means of innovative digital analytics, see for instance Thomas Margoni and Martin Kretschmer, "A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology" (2022) 71 (8) GRUR International 685, 686.

<sup>12</sup> DSM Directive, art. 3 and recitals 12-13.

<sup>13</sup> DSM Directive, art. 4 and recital 18.

<sup>14</sup> Christophe Geiger et al., "The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects" (2018) Centre for International Intellectual Property Studies Research Paper No. 2018-02, 20 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3160586](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3160586) accessed 20 October 2022, p. 22.

Fundamental Rights (EUCFR)<sup>15</sup> and the freedom of research derived from arts. 11<sup>16</sup> and art. 13 of the EUCFR<sup>17</sup>. This has been clearly reflected in the Finnish implementation of the DSM Directive when the Parliament's Constitutional Law Committee rejected the government's draft implementation of the DSM Directive on the grounds that it is not in line with the Finnish constitution.<sup>18</sup> In particular, the Committee found that the proposed TDM exceptions have not been developed in such a way that the freedom of science<sup>19</sup> has been taken into account in a balanced way.<sup>20</sup> Therefore, instead of establishing a fair balance between copyright and the right to research, the TDM exceptions have introduced unnecessary limitations to the mining of text and data in the EU.<sup>21</sup>

---

<sup>15</sup> Charter of Fundamental Rights of the European Union [2012] OJ C 326/391, art. 11 provides the freedom of expression which includes freedom to hold opinions and to receive and impart information. See also art.10 (1) of the European Convention on Human Rights (ECHR) Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14, 4 [1950], ETS 5 [https://www.echr.coe.int/documents/convention\\_eng.pdf](https://www.echr.coe.int/documents/convention_eng.pdf) accessed 12 October 2022. The European Court of Human Rights (ECtHR) has broadened the scope of the right to freedom of expression and information, also including the right of access to information in *Társaság a Szabadságjogokért v. Hungary* (2009) no 37374/05; *Kenedi v. Hungary* (2009) no 31475/05; *Youth Initiative for Human Rights v. Serbia* (2013) no 48135/06, at [20 and 24]; *Timpul Info-Magazin and Anghel v. Moldova* (2007) no 42864/05, at [31]; *WVgrzynowski and Smolczewski v. Poland* (2013) no 33846/07, at [57]; *Sdružení Jihočeské Matky c. la République tchèque* (2006) no 19101/03. In this sense, see also the Court of Justice of the European Union (ECJ) case *Stichting Greenpeace Nederland and Pesticide Action Network Europe (PAN Europe) v European Commission* (T-545/11) EU: T:2013:523.

<sup>16</sup> The right to research is based on the right to information. For a more detailed analysis of the scope of the right to information see, for instance, Christophe Geiger and Bernd Justin Jütte, "Conceptualizing a 'Right to Research' and Its Implications for Copyright Law: An International and European Perspective" (2022) Joint PIJIP/TLS Research Paper Series No. 7-2022, 1 <https://digitalcommons.wcl.american.edu/cgi/viewcontent.cgi?article=1079&context=research> accessed 27 December 2022, pp. 29-35; Christophe Geiger and Bernd Justin Jütte, "The Right to Research as Guarantor for Sustainability, Innovation and Justice in EU Copyright Law" (2022) forthcoming in: "Intellectual Property Rights in the Post Pandemic World: An Integrated Framework of Sustainability, Innovation and Global Justice", Taina E. Pihlajarinne, Jukka Mähönen and Pratyush Upreti (eds.), Cheltenham, UK / Northampton, MA, Edward Elgar Publishing (2022) < <https://ssrn.com/abstract=4140627> > accessed 27 December 2022, pp. 23-25.

<sup>17</sup> Art. 13 of the EUCFR requires that the scientific research should be free of constraint. This freedom is generally considered as a part of freedom of expression see Steve Peers et al. (ed), *The EU Charter of Fundamental Rights: a Commentary* (Oxford, Portland Or.: Hart, 2021) 336, 405. In this sense see Opinion of AG Szpunar in *Pelham GmbH and Others v Ralf Hütter and Florian Schneider-Esleben* (C-476/17) EU:C: 2019:624 at [91] which states that while the EUCFR recognizes the freedom of the arts separately, in the ECHR it must be inferred from the freedom of expression enshrined in Article 10 of ECHR; *EIT Nordisk Film & TV A/S v Denmark* App no 40485/02 (ECtHR, 8 December 2005); *Dammann v Switzerland* App no 77551/01 (ECtHR, 25 April 2006) at [52].

<sup>18</sup> The Parliament's Constitutional Law Committee statement PeVL 58/2022 vp HE 43/2022 vp (14.11.2022) [https://www.eduskunta.fi/FI/vaski/Lausunto/Sivut/PeVL\\_58+2022.aspx](https://www.eduskunta.fi/FI/vaski/Lausunto/Sivut/PeVL_58+2022.aspx) accessed 23 November 2022.

<sup>19</sup> The Constitution of Finland, section 16 subsection 3 <https://www.finlex.fi/en/laki/kaannokset/1999/en19990731> accessed 23 November 2022.

<sup>20</sup> The Parliament's Constitutional Law Committee statement (see fn. 18).

<sup>21</sup> Martin Senftleben, "Compliance of National TDM Rules with International Copyright Law: An Overrated Nonissue?" (2022) 53 IIC 1477, 1478.

Against this background, the aim of this study is to discuss the DSM Directive's TDM exceptions in terms of fundamental rights. The article argues that the exceptions fail to take into account the freedom of information and the freedom of research in a balanced way. These fundamental freedoms have been only addressed superficially. In this regard, the paper proposes amendments to the regulation so that the balance between the fundamental rights to information and research and the fundamental right to intellectual property (IP)<sup>22</sup> would be realized better than in the current law on TDM. The structure of this article is as follows. After the introduction, the paper will discuss how the TDM exceptions improve the situation in terms of fundamental rights compared to the time before the DSM Directive. After that, the article will evaluate the limitations of arts. 3 and 4 and propose changes to these provisions to ensure that the rights to information and research are duly taken into account in the current regulation. Finally, the article will briefly summarise the findings reached under this study and provide concluding remarks.

### **The DSM Directive's TDM exceptions: a new hope?**

The exceptions under arts. 3 and 4 of the DSM Directive have emerged as a new hope for the better regulation of TDM in the EU. The provisions provide more legal certainty compared to a situation that existed before the adoption of the DSM Directive. As Geiger and Jütte emphasize, they clarify the lawfulness of TDM in principle.<sup>23</sup> The TDM exceptions are mandatory, implying that the EU Member States would transpose them into their national laws. The harmonised framework for TDM will promote innovation at European level, especially through cross-border research projects.<sup>24</sup> The mandatory solution to the regulation of TDM would provide normalised and consistent legal grounds for carrying out TDM at EU-wide level which would enhance European

---

<sup>22</sup> The authors' right as a part of IP is guaranteed as a human right under art. 1 of Protocol No. 1 of the ECHR: "Every natural or legal person is entitled to the peaceful enjoyment of his possessions. No one shall be deprived of his possessions except in the public interest and subject to the conditions provided for by law and by the general principles of international law" and as a fundamental right under art. 17 (2) of the EUCFR: "Intellectual property shall be protected." Copyright as a part of IP is also guaranteed at the international level by art. 27 of the Universal Declaration of Human Rights and art. 15 of the International Covenant on Economic, Social and Cultural Rights.

<sup>23</sup> Geiger and Jütte, "Conceptualizing a 'Right to Research' and Its Implications for Copyright Law: An International and European Perspective" (see fn. 16) 55.

<sup>24</sup> Geiger et al., "The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects" (see fn. 14); Geiger and Jütte, "The Right to Research as Guarantor for Sustainability, Innovation and Justice in EU Copyright Law" (n 16).

research collaboration.<sup>25</sup> This would be in line with the EU's goal to create a European research area, where new knowledge disseminates freely<sup>26</sup> and the freedom of research, which supports cross-border collaboration, including the exchange of information between researchers, which is essential for advancing scientific knowledge and addressing pressing public issues.<sup>27</sup> Therefore, having clear guidance on using TDM within the EU would encourage users to mine text and data that would facilitate the functioning of the digital single market (DSM).<sup>28</sup>

However, not only the mandatory nature of these exceptions could be appraised for contributing to the realisation of the fundamental rights to information and research in the context of data analysis. In this regard, the exception of art. 3 is also welcomed as it provides more legal certainty for researchers outlining the conditions under which they could perform TDM.<sup>29</sup> This would increase the research competitiveness of the EU and stimulate future innovation. Moreover, the fact that the exception cannot be prevented by contract<sup>30</sup> would strengthen the positions of beneficiaries of art. 3 and ensure the practical application of this exception within EU Member States. This would protect TDM research from possible restrictions imposed by rightholders under the licensing terms and help research organizations prove the lawfulness of their access. Further, the exception exempts beneficiaries from the obligation to indicate the source and the author's name of mined works which is one of the requirements of the exceptions for scientific research provided under art. 5 (3) (a) of the InfoSoc Directive and arts. 6 (2) (b) and 9 (b) of the Database Directive. Such exemption is necessary in the context of TDM as mining implies the processing of huge amounts of digital data, which makes it almost impossible for users to attribute all mined sources. Moreover, the final output produced as a result of TDM does not contain any excerpts of protected materials utilised during the mining process; it only introduces new insights and patterns

---

<sup>25</sup> Geiger et al., "The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects" (fn. 14) 20.

<sup>26</sup> See art. 179 (1) of the Consolidated version of the Treaty on the Functioning of the European Union (TFEU) and Commission, "Better Careers and More Mobility: A European Partnership for Researchers" (Communication) COM (2008) 317 final.

<sup>27</sup> Geiger and Jütte, "Conceptualizing a 'Right to Research' and Its Implications for Copyright Law: An International and European Perspective" (fn. 16) 54.

<sup>28</sup> Rossana Ducato and Alain Strowel, "Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out" (2021) CRIDES Working Paper Series no. 1/2021, 20 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3829858](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3829858) accessed 12 August 2022.

<sup>29</sup> Art. 3 of the DSM Directive benefits "universities, research institutions, libraries, museums, and any other entity conducting scientific research on a nonprofit basis or by reinvesting all the profits in their scientific research or pursuant to a public interest mission", DSM Directive, art. 2 (1) and (3).

<sup>30</sup> DSM Directive, art. 7 (1).

generated from pre-existing data. Therefore, the listing of all mined works may be burdensome for users and provide no benefits for authors.

By adopting the specific TDM exception for the purpose of scientific research under art. 3, the EU has recognised the importance of using new technologies in the field of research. Although the DSM Directive refers to the commonly used term in copyright law to indicate the purpose of this exception, the wording of art. 3 is broader than in research exceptions provided under art. 5 (3) (a) of the InfoSoc Directive and art. 6 (2) (b) of the Database Directive. Art. 3 does not limit the application of the exception “for the sole purpose” of scientific research; it uses a wider formulation “for the purpose” of scientific research. The latter term is likely to cover research projects which could indirectly pursue other objectives in addition to the primary purpose of scientific research. Furthermore, the exception provided under art. 3 has a broader application than merely “illustrate scientific research” as it covers the purpose of scientific research as such. The exception of art. 3 is very promising as it does not limit the exception to non-commercial research purposes. This would facilitate innovation and would be justified from the view of the freedom of scientific research, which protects all types of research activities.<sup>31</sup> In this sense, it is both commercial and non-commercial research that produces valuable, informative outputs, which are crucial for economic and technological developments. Moreover, research findings could be accessed more easily when commercialised<sup>32</sup> that would also serve society and stimulate progress in the EU.<sup>33</sup>

The second exception introduced under art. 4 of the DSM Directive has also introduced substantial improvements to the regulation of TDM in the EU. The provision has been adopted to recognise the significance of TDM tools not only for scientific research but also for other purposes in both public and private fields.<sup>34</sup> The exception covers all types of beneficiaries and applies for any purpose of TDM activities. The broad scope of this exception could allow any entity to carry out TDM, which is crucial in terms of fundamental rights to information and research. Indeed, it is not

---

<sup>31</sup> Peers et al. (ed), *The EU Charter of Fundamental Rights: a Commentary*, (fn. 17) 420-421.

<sup>32</sup> LIBER, “Myths and Misunderstandings about Text and Data Mining in the Copyright Reform” (2017) <http://eare.eu/assets/uploads/2017/11/Myths-and-misunderstandings-about-TDM.pdf> accessed 23 November 2022.

<sup>33</sup> LIBER, “A Copyright Exception for Text and Data Mining” (*Libereurope*) <https://libereurope.eu/wp-content/uploads/2020/11/TDM-Copyright-Exception.pdf> accessed 24 November 2022.

<sup>34</sup> DSM Directive, recital 18.

only universities and similar research organisations that carry out research and thus promote innovation in different fields. In this sense, it is crucial to encourage companies to use this advanced analytical tool to build and develop their business models and strategies in the EU. This would strengthen the European economy and increase its competitiveness.

Both TDM exceptions grant their beneficiaries a right to store copies made in the course of TDM, however, under different conditions. Art. 3 (2) of the DSM Directive allows the storage of such copies for the “purposes of scientific research, including for the verification of research results”. The provision does not limit the purpose of scientific research to current scientific activities.<sup>35</sup> This would stimulate future scientific research and thus benefit the fundamental right to research. Moreover, allowing beneficiaries to store and retain copies of mined works also for the verification of research results would ensure and promote “good science”, which requires that the hypothesis is tested before approval. In the realm of scientific research, the verifiability of research outputs may determine the quality of research which is a crucial element of knowledge construction.<sup>36</sup> As Geiger states, access to copyrighted materials and the storage of results for future research or the validation of outcomes are indispensable elements of the right to research.<sup>37</sup> Art. 4 (2) also allows its beneficiaries to retain the copies of works made in the course of TDM but only “for as long as is necessary for the purposes of text and data mining.”<sup>38</sup>

Therefore, it is noticeable that the DSM Directive’s TDM exceptions introduce some elements which are crucial for the efficient realization of the fundamental rights to information and research. The situation is better compared to the time before the adoption of these provisions. The main reason is that the statutory exceptions and limitations provided under art. 5 of the InfoSoc

---

<sup>35</sup> Eleonora Rosati, *Copyright in the Digital Single Market: Article-by-Article Commentary to the Provisions of Directive 2019/790* (Oxford University Press 2021), 57.

<sup>36</sup> Reto Hilty, “Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules Part B Exceptions and Limitations (Art. 3 Text and Data Mining)” (2017) Max Planck Institute for Innovation and Competition Research Paper No. 17-12, 11 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2900110](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2900110) accessed 21 August 2022.

<sup>37</sup> Geiger and Jütte, “The Right to Research as Guarantor for Sustainability, Innovation and Justice in EU Copyright Law” (fn. 16) 21.

<sup>38</sup> DSM Directive, art. 4 (2) and recital 18.

Directive<sup>39</sup> and arts. 6, 8 and 9 of the Database Directive<sup>40</sup> appear to be non-applicable to advanced digital tools such as TDM. Regarding the latter provisions, Caterina Sganga fairly claims that the closed and exhaustive list of mostly optional exceptions, designed during the initial phase of the Digital Era, would fail to properly regulate advanced digital technologies.<sup>41</sup> In this regard, the DSM Directive's TDM exceptions should be welcomed as a positive legislative initiative, which gives new hope for balancing copyright and the users' fundamental rights to information and research.

### **The DSM Directive's TDM exceptions: the great frustration?**

Despite many positive nuances, which have been discussed in the previous section, the DSM Directive's TDM exceptions have introduced certain limitations which could be problematic from the fundamental rights perspective. This section will discuss these limitations in detail and propose possible ways out.

#### *The limitation on the scope of beneficiaries of art. 3*

Despite the intention to benefit a research community, art. 3 restricts the beneficiaries to those conducting research in the public sector only. Therefore, private actors such as journalists, individual researchers, small and medium-sized enterprises (SMEs), start-ups and many other commercial organisations would not have the possibility to conduct TDM research under this exception. This limitation is problematic from the perspective of the freedoms of information and research, which apply to all scientific activities: the protection should not be limited to scientific research carried out only by universities or similar institutions.<sup>42</sup> The European Court of Human Rights (ECtHR) has specifically recognised that access to data for legitimate research purposes is

---

<sup>39</sup> InfoSoc Directive, art. 5 (1) provides the exception for temporary acts of reproduction and art. 5 (3) (a) provides the exception for "the sole purpose of illustration for teaching or scientific research".

<sup>40</sup> Database Directive, arts. 6 (2) (b) and 9 (b) provide the exception for scientific research, art. 6 (1) provides the "normal use" exception and art. 8 (1) provides the exception for extraction and/or re-utilization of insubstantial parts of databases.

<sup>41</sup> Sganga, "A New Era for EU Copyright Exceptions and Limitations? Judicial Flexibility and Legislative Discretion in the Aftermath of the CDSM Directive and the Trio of the Grand Chamber of the CJEU" (fn. 9) 314.

<sup>42</sup> Peers et al. (ed), *The EU Charter of Fundamental Rights: a Commentary* (fn. 17) 344-345; Margoni and Kretschmer, "A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology" (fn. 11) 695.

an essential element of users' (e.g., journalists, individual researchers etc.) right to freedom of expression, which also covers the rights to information and research.<sup>43</sup>

The limitation on the scope of beneficiaries of art. 3 could prevent research in many fields of study. Knowledge is generated not only by research institutions but also by individuals, public entities, and businesses.<sup>44</sup> For instance, AI research conducted now by many start-ups and SMEs could advance various areas and sectors (e.g., health care, transportation, banking, education etc.) by transforming the way they operate. In light of recent development and the emergence of cutting-edge AI applications, it is apparent that TDM will become an essential technique allowing new types of information-based services and applications.<sup>45</sup> The exclusion of private actors from the scope of the exception would deprive the EU of a crucial source of innovation and research as a huge amount of public interest research is conducted by commercial actors, especially in the field of medicine and AI.<sup>46</sup> In this regard, AI developers and other private organisations intended to carry out research based on TDM would be required to negotiate with all rightholders over protected materials. The process is extremely complex, taking into account the large amounts of data used for mining. This would place start-ups and SMEs in a less favourable position than large commercial companies or Internet giants, which are able to pay high licensing fees.<sup>47</sup>

Although art. 3 benefits only not-for-profit research organizations and cultural heritage institutions, the DSM Directive supports research in the framework of public-private partnerships.<sup>48</sup> However, it highlights that if the beneficiaries of this exception rely on their private

---

<sup>43</sup> *Társaság a Szabadságjogokért v. Hungary* (fn. 15) at [35-39], *Kenedi v. Hungary* (fn. 15) and *Youth Initiative for Human Rights v Serbia* (fn. 15).

<sup>44</sup> Geiger and Jütte, "Conceptualizing a 'Right to Research' and Its Implications for Copyright Law: An International and European Perspective" (fn. 16) 44.

<sup>45</sup> Ducato and Strowel, "Limitations to Text and data Mining and Consumer Empowerment: Making the Case for a Right to "Machine Legibility"" (fn. 6) 651.

<sup>46</sup> Ducato and Strowel, "Limitations to Text and data Mining and Consumer Empowerment: Making the Case for a Right to "Machine Legibility"" (fn. 6) 651; Geiger and Jütte (fn. 16) 13. For more discussions on TDM exceptions and AI see, for instance, Maryna Manteghi, "The Insufficiency of the EU's Text and Data Mining Exceptions for Using Artificial Intelligence" (2022) 44 (11) E.I.P.R. 651; Sean Flynn et al., "Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action" (2020) Joint PIJIP/TLS Research Paper Series No 48

[https://digitalcommons.wcl.american.edu/research/48/?utm\\_source=digitalcommons.wcl.american.edu%2Fresearch%2F48&utm\\_medium=PDF&utm\\_campaign=PDFCoverPages](https://digitalcommons.wcl.american.edu/research/48/?utm_source=digitalcommons.wcl.american.edu%2Fresearch%2F48&utm_medium=PDF&utm_campaign=PDFCoverPages) accessed 23 December 2022, 9.

<sup>47</sup> Manteghi, "The Insufficiency of the EU's Text and Data Mining Exceptions for Using Artificial Intelligence" (fn. 46) 654.

<sup>48</sup> DSM Directive, recital 11.

partners for carrying out TDM (e.g., by using their technological tools), they should be the only beneficiaries of art. 3.<sup>49</sup> Further, the DSM Directive clarifies that actors upon which commercial entities have “a decisive influence...because of structural situations”<sup>50</sup> should not be regarded as beneficiaries of this exception.<sup>51</sup> Nevertheless, it may be practically difficult to distinguish between commercial and non-commercial activities in public-private partnerships.<sup>52</sup> The situation can also be problematic when the research outcome of TDM can be “found subsequently to be commercially exploitable” or when research is conducted on the grounds of donations from private entities.<sup>53</sup>

Therefore, it appears crucial to provide broader access to mined data for research purposes to improve the EU’s research competitiveness and encourage new AI applications in various fields. A possible solution could be extending the scope of art. 3 to permit any person or entity with “lawful access” to content to carry out TDM for scientific research including persons associated with them that could be necessary for the verification of research results or in the case of peer-review. These changes to the scope of beneficiaries of Art. 3 would strengthen the EU’s innovative policy and increase research outcomes within the DSM.

### *The purpose limitation of art. 3*

The DSM Directive restricts the scope of art. 3 to “scientific research” only, which may create legal uncertainty and lead to a narrow interpretation of this provision. The concept of scientific research does not represent a distinct category of research, but rather defines a research process based on established scientific methods and techniques. However, to qualify as “scientific research,” new knowledge resulting from methodological and systematic data analysis should contribute to the advancement of science in either practical or theoretical domains.<sup>54</sup> The DSM

---

<sup>49</sup> DSM Directive, recital 11.

<sup>50</sup> DSM Directive, recital 12.

<sup>51</sup> DSM Directive, art. 2 (1) (b) and recital 12.

<sup>52</sup> Maria Bottis et al., “Text and Data Mining in the EU Acquis Communautaire Tinkering with TDM & Digital Legal Deposit” (2019) 2 Erasmus Law Review 190, 193; Jonathan Griffiths et al., “Comment of the European Copyright Society Addressing Selected Aspects of the Implementation of Articles 3 to 7 of the Directive (EU) 2019/790 on Copyright in the Digital Single Market” (2023) 72 (1) GRUR International 22, 37.

<sup>53</sup> Griffiths et al., “Comment of the European Copyright Society Addressing Selected Aspects of the Implementation of Articles 3 to 7 of the Directive (EU) 2019/790 on Copyright in the Digital Single Market” (fn. 52) 37.

<sup>54</sup> Peers et al. (ed), *The EU Charter of Fundamental Rights: a Commentary* (fn. 17) 420-421.

Directive does not clarify the notion of “scientific research”. It merely states that such research should cover “both the natural sciences and the human sciences.”<sup>55</sup>

Therefore, users would be required to prove that their projects based on data analysis adhere to scientific characteristics to enjoy the exception of art. 3. The process is complex because scientific research is characterised by different approaches, categories and styles. For instance, the classification methods employed to categorise the fields of science may cover other independent branches along with the natural sciences and humanities (e.g., engineering and technology, medical and health sciences, agricultural science)<sup>56</sup> or recognise alternative groups of main branches (e.g., humanities and social sciences).<sup>57</sup> Further, the scientific methods employed to conduct scientific research should satisfy many characteristics such as replicability, precision, falsifiability, objectivity, reliability, parsimony and others.<sup>58</sup> These characteristics enable the research method to test and crosscheck the main aspects of sciences. In this sense, research is not considered scientific if it is unable to observe data using these features.<sup>59</sup> As has been mentioned, the results of scientific research should contribute to the current state of science; however, some branches of knowledge (e.g., arts, music, literature, theology, etc.) cannot be considered “science” because they cannot be tested using scientific methods.<sup>60</sup>

Therefore, it is crucial to use the broader term “research” to identify the purpose of art. 3. This more general formulation would cover a creative and systematic investigation in different fields of study, undertaken to solve problems and create new reliable and valuable knowledge.<sup>61</sup> This

---

<sup>55</sup> DSM Directive, recital 12.

<sup>56</sup> See, OECD, Working Party of National Experts on Science and Technology Indicators, “Revised field of science and technology (FOS) classification in the Frascati Manual” (2007) DSTI/EAS/STP/NESTI (2006)19/FINAL <https://www.oecd.org/science/inno/38235147.pdf> accessed 2 December 2022.

<sup>57</sup> See, for instance, the ERC panels’ classification (2020) [https://erc.europa.eu/sites/default/files/document/file/ERC\\_Panel\\_structure\\_2020.pdf](https://erc.europa.eu/sites/default/files/document/file/ERC_Panel_structure_2020.pdf) accessed 2 December 2022.

<sup>58</sup> Anol Bhattacharjee, “Social Science research: Principles, Methods, and Practices” (Textbooks Collection 3, University of South Florida 2012) <https://courses.lumenlearning.com/suny-hccc-research-methods/chapter/chapter-1-science-and-scientific-research/> accessed 2 December 2022.

<sup>59</sup> Brian Kennett, *Planning and Managing Scientific Research: A guide for the beginning researcher* (ANU Press 2014) 2.

<sup>60</sup> For more discussions on what can be considered “science” see Bhattacharjee, “Social Science research: Principles, Methods, and Practices” (fn. 58).

<sup>61</sup> Peers et al. (ed), *The EU Charter of Fundamental Rights: a Commentary* (fn. 17) 420.

would be justified from the perspective of the fundamental freedom of research as it would cover a broader scope of studies in various fields.

### *The limitation on access in arts. 3 and 4*

Art. 3 of the DSM Directive requires that the beneficiaries of this exception have “lawful access” to mined works.<sup>62</sup> The concept is defined in the text of the DSM Directive as “covering access to content based on an open access policy or through contractual arrangements between rightholders and research organisations or cultural heritage institutions, such as subscriptions, or through other lawful means”.<sup>63</sup> In addition, the concept of “lawful access” should also include access to materials that are freely available online.<sup>64</sup> The “lawful access” requirement also applies to art. 4. Recital 18 of the DSM Directive indicates the conditions under which this exception should apply. First, the provision outlines that the exception “should only apply where the work...is accessed lawfully by beneficiaries, including when it has been made available to the public online”.<sup>65</sup> This is in line with art. 3, meaning that access should be deemed lawful when it is authorised by rightholders (e.g., through open access policy or contractual arrangements, etc.) or it is not restricted by the applicable law.<sup>66</sup> And second, the exception should apply when the use of protected works has not been reserved by its rightholders in an appropriate manner.<sup>67</sup>

The requirement of “lawful access” demonstrates a strict approach to the application of these exceptions as it would depend on the will of rightholders to allow access to content, deny such access for certain users or grant it on a conditional basis only.<sup>68</sup> For instance, rightholders may prevent research organisations from carrying out TDM by not issuing them licenses or raising the

---

<sup>62</sup> DSM Directive, art. 3 (1).

<sup>63</sup> DSM Directive, recital 14.

<sup>64</sup> DSM Directive, recital 14.

<sup>65</sup> DSM Directive, recital 18.

<sup>66</sup> *Stichting Brein v Jack Frederik Wullems* (C-527/15) EU:C: 2017:300 at [65]; *Football Association Premier League and Others* (C-403/08 and C-429/08) ECR I-09083 at [168]; *Infopaq International A/S v Danske Dagblades Forening* (C-302/10) EU: C: 2012:16 at [42].

<sup>67</sup> DSM Directive, recital 18.

<sup>68</sup> Reto Hilty and Valentina Moscon, “Modernisation of the EU Copyright Rules Position Statement of the Max Planck Institute for Innovation and Competition” (2017) Max Planck Institute for Innovation and Competition Research Paper No. 17-12, 4 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3036787](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3036787) accessed 18 November 2022. Geiger et al., “The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects” (fn. 14) 22.

licensing or subscription fees.<sup>69</sup> The excessive licensing costs could also restrict the scope of research based on TDM since this analytical tool requires the processing of a huge amount of data to achieve the pursued goal. According to Ducato and Strowel, if the sources are incomplete (e.g. when TDM analysis relies only on data freely available online), the final result could be scarce and insecure.<sup>70</sup> Moreover, rightholders may limit or prohibit TDM on data that is freely available online, for instance, in the terms and conditions of their websites or by employing TPMs.<sup>71</sup> In other words, subjecting TDM research to private ordering might undermine the users' rights to information and research as it could "discriminate research according to research organizations' market power".<sup>72</sup> This will aggravate the scientific and innovative "gap" between more and less funded research institutions.<sup>73</sup>

Although the requirement of "lawful access" limits the users' fundamental rights to information and research,<sup>74</sup> the condition cannot be abolished since it is an essential element in balancing competing interests. The "lawful access" requirement serves as a safeguard for authors guaranteeing that they receive compensation for the public use of their creative and original works. Even though scientific research must not be restricted, the benefits of increased access to data should be balanced with the harms to art. 17 (2) of the EUCFR caused by such access. Therefore, even if the "lawful access" requirement in art. 3 can be regarded as an interference with the rights established under arts. 11 and 13 of the EUCFR, its effects should be justified by the need to protect the rightholder's right to IP.<sup>75</sup>

The European Court of Justice (ECJ) should undertake a flexible approach to interpret the notion of "lawful access" in the context of TDM to strike a "fair balance" between those who have rights

---

<sup>69</sup> Geiger et al., "The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects" (fn. 14) 22.

<sup>70</sup> Ducato and Strowel, "Limitations to Text and data Mining and Consumer Empowerment: Making the Case for a Right to "Machine Legibility"" (fn. 6) 668.

<sup>71</sup> For more discussion on TPMs and TDM exceptions see the last subsection of this article.

<sup>72</sup> Geiger et al., "The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects" (fn. 14) 22.

<sup>73</sup> Geiger et al., "The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects" (fn. 14) 22.

<sup>74</sup> Margoni and Kretschmer, "A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology" (fn. 11) 697.

<sup>75</sup> In this sense, see *Metronome Music v. Music Point Hokamp* (C-200/96) ECLI:EU:C:1998:172 and *Laserdisken ApS v. Kulturministeriet* (C-479/04) ECLI:EU:C:2006:549.

over mined works and the beneficiaries of art. 3.<sup>76</sup> The term should not be understood in light of the “lawful source” requirement since it may limit TDM research carried out in the online environment, where the legality of sources is usually obscure.<sup>77</sup> This would also contradict the wording of recital 14 of the DSM Directive, which provides that access to content that is freely available online should be viewed as lawful.<sup>78</sup> Moreover, it would be unreasonable in terms of fundamental rights to information and research to allow automated analytical tools to verify the lawfulness of the mined materials.<sup>79</sup> Therefore, if the protected materials are available online without any technological restrictions, TDM users should not be liable for mining “unlawful sources”.<sup>80</sup> In this sense, the notion of “lawfulness” should not cover access to or use of data authorized by rightholders only.

The concept should be interpreted in the light of the term “lawful use” to ensure that the harm that the “lawful access” requirement might cause to the rights to information and research, in the context of TDM, is proportional to the benefits that art. 3 is likely to provide in terms of copyright protection. In this regard, as Synodinou argues, “lawful access” should acquire a broad meaning in the case of TDM exceptions.<sup>81</sup> This would imply that the lawfulness of access would not only depend on contractual arrangements but also would cover access to data which is not restricted by law.<sup>82</sup> Therefore, when research organizations or cultural heritage institutions have obtained “lawful access” to data through the subscriptions or licenses already excluding or limiting TDM

---

<sup>76</sup> Griffiths et al., “Comment of the European Copyright Society Addressing Selected Aspects of the Implementation of Articles 3 to 7 of the Directive (EU) 2019/790 on Copyright in the Digital Single Market” (fn. 52) 24.

<sup>77</sup> Christophe Geiger, “The Missing Goal-Scorers in the Artificial Intelligence Team: Of Big Data, the Fundamental Right to Research and the failed Text and Data Mining limitations in the CSDM Directive” (2021) PIJIP/TLS Research Paper Series no. 66, 7 <https://digitalcommons.wcl.american.edu/research/66> accessed 23 November 2022.

<sup>78</sup> DSM Directive, recital 14.

<sup>79</sup> In this sense, see *Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)* (C-70/10) ECR 2011-00000 at [52] and *Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM) v Netlog NV* (C-360/10) EU:C:2012:85 at [50].

<sup>80</sup> In this sense, see Directive 2013/40 on attacks against information systems and replacing Council Framework Decision 2005/222/JHA [2013] OJ L 218/8, art. 3; *Nils Svensson and Others v Retriever Sverige AB* (C-466/12) EU: C:2014:76; *BestWater International GmbH v Michael Mebes and Stefan Potsch* (C-348/13) EU: C:2014:2315 and *Marc Soulier and Sara Doke v Premier ministre and Ministre de la Culture et de la Communication* (C-301/15) EU: C:2016:878.

<sup>81</sup> Tatiana Eleni Synodinou, “Lawfulness for Users in European Copyright Law: Acquis and Perspectives” (2019) 10 JIPITEC 20 para 1, 26.

<sup>82</sup> Synodinou, “Lawfulness for Users in European Copyright Law: Acquis and Perspectives” (fn. 81) para 1, 26; *Football Association Premier League and Others* (C-403/08 and C-429/08) ECR I-09083 at [168]; *Infopaq International A/S v Danske Dagblades Forening* (C-302/10) EU:C:2012:16 at [42].

(e.g., for educational purposes only) in their terms, it should not affect the further application of TDM for the purpose of scientific research. Although the use is not authorized by rightholders, it cannot be deemed to be unlawful since the use is not prohibited by law and may rely on a specific copyright exception.<sup>83</sup>

Therefore, the interpretation of the “lawful access” requirement in light of the term “lawful use” would comply with the fundamental rights to information and research and help avoid confusion. This would also bring more flexibility to the definition of the TDM exceptions and broaden their scope.<sup>84</sup> Furthermore, assessing the “lawful access” condition using the fundamental rights analysis could help to strike a “fair balance” of competing interests that would strengthen the efficiency of arts. 3 and 4 of the DSM Directive.

#### *The “opt-out” mechanism of art.4*

Although Art. 4 provides the exception which applies to any type of beneficiary and for any purpose, it restricts the right to TDM by providing a so-called “opt-out” mechanism. This condition allows rightholders to expressly reserve the use of their materials in an “appropriate manner”.<sup>85</sup> However, the DSM Directive fails to provide clear and detailed rules for controlling the functioning of this mechanism. Recital 18 only provides that those who have rights over works publicly available online could only reserve their rights by using machine-readable means, including metadata and terms and conditions of a website or a service.<sup>86</sup> Moreover, in other circumstances, it should be appropriate to reserve the rights by other means, such as contractual agreements or unilateral declarations.<sup>87</sup> Considering the mining of content available online, the “opt-out” mechanism could be interpreted in light of the ECJ’s findings in *VG Bild-Kunst* and *GS Media*, where the court held that rightholders should be only able to reserve the use of their works by means of effective technological measures. Otherwise, it could be difficult for users to

---

<sup>83</sup> *Infopaq International A/S v Danske Dagblades Forening* (C-302/10) EU:C:2012:16 at [44 and 45]; *Land Nordrhein-Westfalen v Dirk Renckhoff* (C-161/17) EU: C:2018:279.

<sup>84</sup> In this sense, see Bernt Hugenhlotz, “Flexible Copyright, Can the Author’s Rights Accommodate Fair Use?” in Ruth L. Okediji (ed), *Copyright Law in an Age of Limitations and Exceptions* (Cambridge University Press 2017).

<sup>85</sup> DSM Directive, art. 4 (3) and recital 18.

<sup>86</sup> DSM Directive, recital 18.

<sup>87</sup> DSM Directive, recital 18.

determine whether the use is reserved or not.<sup>88</sup> In the absence of such measures, for instance, users intending to carry out TDM on the content available online would be obliged to check the terms and conditions of each website to ensure that mining is not prohibited. This would be problematic considering that a huge amount of data is needed for TDM analysis.

Admittedly, the “opt-out” mechanism has been introduced to ensure efficient protection of rightholders. In particular, the aim is to prevent the over-exploitation of the protected works which could be impacted by the specific TDM exception. However, many academics argue that giving too much control to rightholders over their intellectual “products” may suppress the ability of users to access, share, or express protected content across the public and private spheres.<sup>89</sup> Hugenholtz underlines that the possibility of reserving use “legitimizes a derivative market in TDM, enabling right holders to control, license or entirely prohibit such activities”.<sup>90</sup> Further, the exception of art. 4 can be overridden by a contract preventing many users with limited financial resources to conduct TDM research. This would mean that only very large commercial companies or Internet giants would benefit from this TDM exception. In this regard, users have to pay twice to be able to mine –first to acquire “lawful access” to works and a second time to read and analyse digital data.<sup>91</sup> This could deprive legitimate users of the possibility to fully and effectively exercise their fundamental rights to information and research.

To balance the competing rights at stake, it is necessary to clarify some elements of the “opt-out” mechanism. First, the ECJ should properly interpret the requirements of the reservation right that the use should be reserved “expressly” and “in an appropriate manner”. These two conditions could significantly contribute to the promotion and protection of the users’ fundamental rights to information and research in the context of TDM. Moreover, the ECJ should specifically clarify how the use could be reserved in a machine-readable format in practice. The clarification of these

---

<sup>88</sup> *VG Bild-Kunst v Stiftung Preußischer Kulturbesitz* (C-392/19) EU:C: 2021:181 at [46]; In this sense see also *GS Media BV v Sanoma Media Netherlands BV and Others* (C-160/15) EU:C:2016:644 at [46].

<sup>89</sup> Andrew Tyner, “The EU Copyright Directive: “Fit for the Digital Age or Finishing It?” (2020) 26 (2) J. INTELL. PROP. L. 275, 281.

<sup>90</sup> Bernt Hugenholtz, “The New Copyright Directive: Text and Data Mining (Articles 3 and 4)” (*Kluwer Copyright Blog*, 24 July 2019) <http://copyrightblog.kluweriplaw.com/2019/07/24/the-new-copyright-directive-text-and-data-mining-articles-3-and-4/> accessed 12 April 2022.

<sup>91</sup> LIBER, “Myths and Misunderstandings about Text and Data Mining in the Copyright Reform” (fn. 32).

aspects would create more legal certainty for TDM users in both public and private sectors and thus promote European developments, especially in the field of AI.

### *The right to store copies*

As discussed above, art. 3 (2) of the DSM Directive allows the storage of copies created in the course of mining for the “purposes of scientific research, including for the verification of research results”. Therefore, the right to store copies under art. 3 is more favourable than that in art. 4, which limits the storage to a TDM process as such.<sup>92</sup> However, many aspects concerning the storage conditions and the status of copies are unclear under the exception for scientific research. First, the DSM Directive fails to clarify if the third parties (other than the beneficiaries of Art. 3) may access datasets containing copies of mined works for the purpose of scientific research including for the further verification of research results.<sup>93</sup> Second, the exception does not cover activities such as scientific peer review and joint research which may constitute an important part of the research process. Recital 15 merely indicates that these activities may be covered by the exception for scientific research provided under art. 5 (3) (a) of the InfoSoc Directive. However, the optional and unspecified nature of this exception may call into question the legality of such uses across the EU. This may impede cross-border research activities and even decrease the quality of research results.

Art. 3, unlike art. 4, requires that the copies of mined works should be stored with “an appropriate level of security”.<sup>94</sup> It seems that this condition intends to strike a balance between copyright and the fundamental rights to information and research as it fairly restricts the “storage” right of TDM users to a needed licit extent. However, the DSM Directive does not clarify the conditions needed for satisfying such a requirement. Particularly, the DSM Directive is silent on the time limitations

---

<sup>92</sup> DSM Directive, art. 4 (2) and recital 18.

<sup>93</sup> Rossana Ducato and Alain Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out” (2021) CRIDES Working Paper Series no. 1/2021, 20 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3829858](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3829858) accessed 12 August 2022, 12; Margoni and Kretschmer, “A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology” (fn. 11) 697.

<sup>94</sup> DSM Directive, art. 3 (2).

and specific conditions under which the mined materials would be stored.<sup>95</sup> Likewise, the issue of responsibility for the storage of such copies remains unclear. Recital 15 stipulates that “the copies should be stored in a secure environment. Member States should be free to decide, at the national level and after discussions with relevant stakeholders, on further specific arrangements for retaining the copies, including the ability to appoint trusted bodies for the purpose of storing such copies.”<sup>96</sup> However, the possibility of applying various storage conditions in every EU Member State may create legal uncertainty and particularly impede the advance of cross-border research activities.<sup>97</sup> Moreover, many research organisations claim against involving “trusted intermediaries” in the process of storage. For instance, the UK Libraries and Archives Copyright Alliance (LACA) and the Network of University Libraries in Spain (REBIUN ) argue that third parties should not be trusted to hold derived data that results from TDM. They allege that the “storage” right should be enjoyed by research organisations, such as universities and libraries, which spend huge sums of money on subscriptions and also preserve copyrighted works.<sup>98</sup>

Further, art. 4 (2) also allows its beneficiaries to retain copies of works made in the course of TDM. However, the provision indicates that such copies could be retained only “for as long as is necessary for the purposes of text and data mining.”<sup>99</sup> This implies that the storage of reproductions and extractions at hand is limited to a TDM process as such. In this sense, once the mining is completed, the copies should be removed. Thus, the right to store copies under art. 4 of the DSM Directive appears to be more restrictive than that under art. 3. The former provision could limit TDM research as users would not be able to store copies for the verification of research results. Therefore, even though the right to store copies under arts. 3 and 4 of the DSM Directive benefit

---

<sup>95</sup> Maria-Daphne Papadopoulou et al., “The Exception of Text and Data Mining from the Academic Libraries Standpoint” (2021) 9 *Open Journal of Social Sciences* 502, 512 [https://www.scirp.org/pdf/jss\\_2021052414155129.pdf](https://www.scirp.org/pdf/jss_2021052414155129.pdf) accessed 14 October 2022.

<sup>96</sup> DSM Directive, recital 15.

<sup>97</sup> REBIUN, “Text and Data Mining: (Articles 3 and 4 of the EU-DSM) by REBIUN’s Copyright Working Group” (*Blogs Ifla*, 10 June 2020) <https://blogs.ifla.org/lpa/2020/06/10/text-and-data-mining-articles-3-and-4-of-the-eu-dsm/> accessed 12 December 2022.

<sup>98</sup> REBIUN, “Text and Data Mining: (Articles 3 and 4 of the EU-DSM) by REBIUN’s Copyright Working Group” (fn. 97); LACA, “The Right to Read is the Right to Mine: But Not When Blocked by Technical Protection Measures” (*Uklaca*, 1 August 2019) <https://uklaca.org/2019/08/01/the-right-to-read-is-the-right-to-mine-but-not-when-blocked-by-technical-protection-measures/> accessed 17 October 2022.

<sup>99</sup> DSM Directive, art. 4 (2) and recital 18.

TDM users and contribute to the promotion and protection of the freedoms of information and research, it includes many aspects which should be clarified by the ECJ.

### *The restrictive effect of TPMs*

Due to a potentially high number of access requests to their materials, rightholders are permitted under art. 3 (3) of the DSM Directive to apply measures to ensure the security and integrity of their systems and databases. Such measures could, for instance, be applied to ensure that only users with lawful access to mined data can access it, including through IP address validation or user authentication.<sup>100</sup> The InfoSoc Directive provides protection against the circumvention of any effective technological measures,<sup>101</sup> which is also applicable to art. 3 of the DSM Directive.<sup>102</sup> However, the DSM Directive clarifies that the TPMs at issue should be proportionate to the risks involved, should not go beyond what is necessary to achieve the objective of ensuring the security and integrity of the networks and databases and should not prevent the effective application of the exception.<sup>103</sup> Moreover, recital 7 of the DSM Directive underlines that even though the protection of TPMs remains essential for safeguarding rightholders' exclusive rights, rightholders should apply voluntary measures to ensure that technological measures do not prevent the enjoyment of the exception.<sup>104</sup> Although art. 4 is silent in relation to the use of TPMs, it appears that the provisions concerning the application of such measures in the DSM Directive are more broadly relevant to this exception unless the use has been reserved.<sup>105</sup>

Therefore, TPMs adopted by rightholders should be strictly targeted, in the sense that they should ensure the security and integrity of the system without affecting TDM users' possibility to lawfully access protected works.<sup>106</sup> Otherwise, the rightholder's interference with the users' rights to

---

<sup>100</sup> DSM Directive, recital 16.

<sup>101</sup> InfoSoc Directive, art. 6.

<sup>102</sup> DSM Directive, art. 7 (2).

<sup>103</sup> DSM Directive, art. 3 (3) and recital 16.

<sup>104</sup> DSM Directive, recital 7.

<sup>105</sup> Rosati, *Copyright in the Digital Single Market: Article-by-Article Commentary to the Provisions of Directive 2019/790* (fn. 35) 91.

<sup>106</sup> In this sense, see for instance *Republic of Poland v European Parliament and Council of the European Union* (C-401/19) EU:C:2022:297 at [81]; *UPC Telekabel Wien GmbH v Constantin Film Verleih GmbH and Wega Filmproduktionsgesellschaft mbH* (C-314/12) EU:C: 2014:192 at [55 and 56].

information and research would be unjustified in the light of the objective pursued.<sup>107</sup> Moreover, the measures taken by the rightholders should be “sufficiently effective to ensure genuine protection of the fundamental rights at issue” in the sense that they should be able to prevent unauthorised access to works.<sup>108</sup> The examples of TPMs provided by recital 16, such as IP address validation or user authentication, do not undermine the possibility of TDM users to lawfully access information.<sup>109</sup> Such measures could be deemed to be necessary to secure the effective protection of the fundamental right to IP.<sup>110</sup>

However, rightholders may also use automated technological tools to block access to their data (such as IP address blocking or domain name server blocking), which may result in overblocking (blocking of lawful access). Such measures, as Synodinou claims, would fail the effectiveness test and would lead to an imbalance between means and ends.<sup>111</sup> The ECtHR and the ECJ have underlined that the system cannot properly distinguish between lawful and unlawful access.<sup>112</sup> As Reda notes, automated filtering technologies cannot perform the difficult legal and factual analysis that is needed to ascertain whether access to data is infringing or covered by relevant exceptions or limitations.<sup>113</sup> Therefore, the “lock-outs” may occur frequently and take lots of time to resolve. This may limit the benefits of art. 3 and lead to undesirable consequences. For instance, libraries that spend millions of euros for access to electronic resources would have to wait weeks or even months to be reconnected to digital content services.<sup>114</sup> The unreasonable delays in accessing data for research purposes could create a risk that such data would lose its informative value and thereby

---

<sup>107</sup> *UPC Telekabel Wien* (C-314/12) (fn. 106) at [56]; *Tobias Mc Fadden v Sony Music Entertainment Germany GmbH* (C-484/14) EU:C: 2016:689 at [93].

<sup>108</sup> *UPC Telekabel Wien* (C-314/12) (fn. 106) at [62].

<sup>109</sup> In relation to password-protection, see *McFadden* (C-484/14) (fn. 107) at [94].

<sup>110</sup> In this sense, see *McFadden* (C-484/14) (fn. 107) at [98-99].

<sup>111</sup> Tatiana-Eleni Synodinou, ““Intermediaries” Liability for Copyright Infringement in the EU: Evolutions and Confusions” (2015) 31 *Computer Law & Security Review* 57, 62.

<sup>112</sup> *SABAM* (C-360/10) (fn. 79) at [50 and 52]; *Scarlet Extended* (C-70/10) (fn. 79) at [50]; *Ahmet Yıldırım v Turkey* (2012) no 3111/10 at [66-68]; *Kharitonov v Russia* App no 10795/14 (ECtHR, 23 June 2020) at [38-40]; *Republic of Poland v European Parliament and Council of the European Union* (C-401/19) (fn. 106) at [86].

<sup>113</sup> Julia Reda et al., “Article 17 of the Directive on Copyright in the Digital Single Market: A Fundamental Rights Assessment” (2020), 26-27 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3732223](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3732223) accessed 1 December 2022.

<sup>114</sup> LIBER, “Europe’s TDM Exception for Research: Will It Be Undermined By Technical Blocking From Publishers?” (*Libereurope*, 10 March 2020) <https://libereurope.eu/article/tdm-technical-protection-measures/> accessed 30 November 2022.

become outdated.<sup>115</sup> Moreover, the mere fact that access is even temporarily unavailable restricts the users' rights to information and research.<sup>116</sup>

Therefore, it is necessary to clarify the application of TPMs and introduce sufficient safeguards to reduce the risk of erroneous or unjustified blocking of lawful access.<sup>117</sup> This would ensure that the protection of copyright and the protection of the users' fundamental rights affected by TPMs are counterbalanced.<sup>118</sup> Art. 17 (2) of the EUCFR is not an absolute right and thus it should not be used to justify ineffective TPMs which may lead to the overblocking of protected content. This implies that the enjoyment of the fundamental right to IP cannot be performed at any cost. In this regard, it could be reasonable to develop a legal mechanism which would work where lawful access to data is disabled.

## Conclusion

The exceptions under arts. 3 and 4 of the DSM Directive have been adopted to reconcile the discord between copyright protection and the users' right to TDM. Although the provisions have improved the regulation of this analytical technique, they fail to take into account the freedom of information and the freedom of research in a balanced way. These fundamental freedoms have been only addressed superficially. In this regard, the paper proposes amendments to the normative framework of arts. 3 and 4 so that the balance between competing fundamental rights would be realized better than in the current regulation of TDM. First, it would be reasonable to extend the scope of beneficiaries of art. 3 to permit any person or entity with "lawful access" to content to carry out TDM for scientific research. This would improve EU's research competitiveness and encourage new AI applications in various fields. Second, it would be appropriate to use the broader term "research" instead of "scientific research" to identify the purpose of art. 3. This could eliminate any uncertainties caused by different approaches, categories and styles used in the field

---

<sup>115</sup> *Republic of Poland v European Parliament and Council of the European Union* (C-401/19) (fn. 106) at [60].

<sup>116</sup> ECtHR, *Observer and Guardian v. the United Kingdom*, Appl. no. 13585/88, 26 November 1991 at [59(b)], with further references summarizing its prior case law; Series A no. 216, and *Guerra and Others v. Italy*, App. No. 14967/89, 19 February 1998 at [53].

<sup>117</sup> *Data Protection Commissioner v Facebook Ireland Limited and Maximilian Schrems* (C-311/18) EU:C:2020:559 at [176]; *Republic of Poland v European Parliament and Council of the European Union* (C-401/19) (fn. 106) at [67]; *Grigoriy Nikolayevich Kablis v Russia* App no 59663/17 (ECtHR, 30 July 2017) at [92].

<sup>118</sup> *Scarlet Extended* (C-70/10) (fn. 79) at [45]; *UPC Telekabel Wien* (C-314/12) (fn. 106) at [47]; *Tobias Mc Fadden* (C-484/14) (fn. 107) at [100].

of science. Third, the concept of “lawful access” should be interpreted in the light of the term “lawful use”. This would imply that the lawfulness of access would not only depend on contractual arrangements but would also cover access to data which is not restricted by law. This change could ensure that the harm that the “lawful access” requirement might cause to the rights to information and research is proportional to the benefits that arts. 3 and 4 are likely to provide in terms of copyright protection. Fourth, the “opt-out” mechanism of art. 4 should be clarified to understand how the use could be reserved in a machine-readable format in practice and when it is deemed to be reserved “expressly” and “in an appropriate manner”. This would provide more legal certainty for private actors that constitute a crucial source of innovation in the EU. Fifth, it would be reasonable to create the possibility to share datasets, containing copies of mined works, with other researchers for the verification of research results including for the purpose of scientific peer review and joint research. Moreover, it is needed to determine the time limitations and specific conditions under which the mined materials would be stored. Sixth, it is necessary to clarify the application of TPMs and introduce sufficient safeguards to reduce the risk of erroneous or unjustified blocking of lawful access. This would ensure that the protection of copyright and the protection of the users’ fundamental rights affected by TPMs are counterbalanced. The proposed solutions are likely to provide more legal certainty and ensure a more balanced regulatory framework for TDM in the EU. The findings indicate the need to reconsider the TDM exceptions in light of fundamental rights to ensure that research and innovation are not impeded by the excessive restrictions of copyright and database protection.