

# The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects

Legal Affairs



Policy Department for Citizens' Rights and Constitutional Affairs  
Directorate General for Internal Policies of the Union  
PE 604.941- February 2018





**DIRECTORATE GENERAL FOR INTERNAL POLICIES**

**POLICY DEPARTMENT FOR CITIZENS' RIGHTS AND  
CONSTITUTIONAL AFFAIRS**

**LEGAL AFFAIRS**

# **The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects**

**IN-DEPTH ANALYSIS**

## **Abstract**

This in-depth analysis, commissioned by the European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs at the request of the Committee on Legal Affairs (JURI-Committee), is a contribution to the workshop on "Text and data mining" held on 22 February 2018 in Brussels. It provides an analysis of the Commission's Proposal (which introduces in Article 3 a mandatory exception to copyright allowing to carry out text and data mining of protected works), assesses its positive and negative impacts and provides some suggestions for possible improvements. Advantages of introducing an "open clause" on top of an enumerated list of exceptions to address some of the related problems are also reviewed.

## **ABOUT THE PUBLICATION**

This research paper was requested by the European Parliament's Committee on Legal Affairs and commissioned, overseen and published by the Policy Department for Citizens' Rights and Constitutional Affairs.

Policy Departments provide independent expertise, both in-house and externally, to support European Parliament committees and other parliamentary bodies in shaping legislation and exercising democratic scrutiny over EU external and internal policies.

To contact the Policy Department for Citizens' Rights and Constitutional Affairs or to subscribe to its newsletter please write to: [poldep-citizens@europarl.europa.eu](mailto:poldep-citizens@europarl.europa.eu)

## **RESPONSIBLE RESEARCH ADMINISTRATOR**

Udo BUX  
Policy Department for Citizens' Rights and Constitutional Affairs  
European Parliament  
B-1047 Brussels  
E-mail: [poldep-citizens@europarl.europa.eu](mailto:poldep-citizens@europarl.europa.eu)

## **EDITORIAL ASSISTANT**

Monika Laura LAZARUK

## **AUTHORS**

Christophe GEIGER, Professor of Law, Director General and Director of the Research Department, Center for International Intellectual Property Studies (CEIPI), University of Strasbourg

Giancarlo FROSIO, Senior Lecturer and Researcher, CEIPI, University of Strasbourg

Oleksandr BULAYENKO, Researcher, CEIPI, University of Strasbourg

## **LINGUISTIC VERSION**

Original: EN

Manuscript completed in February 2018

© European Union, 2018

This document is available on the internet at:

<http://www.europarl.europa.eu/supporting-analyses>

## **DISCLAIMER**

The opinions expressed in this document are the sole responsibility of the author and do not necessarily represent the official position of the European Parliament.

Reproduction and translation for non-commercial purposes are authorised, provided the source is acknowledged and the publisher is given prior notice and sent a copy.

## **CONTENTS**

<b>CONTENTS</b>	<b>3</b>
<b>EXECUTIVE SUMMARY</b>	<b>4</b>
<b>1. INTRODUCTION</b>	<b>5</b>
<b>2. INTELLECTUAL PROPERTY PROTECTION AND TDM</b>	<b>5</b>
<b>3. COPYRIGHT EXCEPTIONS AND LIMITATIONS AND TDM</b>	<b>7</b>
Do existing exceptions and limitations apply to TDM?	8
Which problems researchers face in applying existing exceptions to TDM?	12
Would an “opening clause” serve public interest?	15
<b>4. THE TDM LIMITATION IN THE MEMBER STATES AND THE EU</b>	<b>17</b>
<b>5. THE COMMISSION’S PROPOSAL</b>	<b>18</b>
Positive and Negative Impacts	20
<b>6. CONCLUSIONS</b>	<b>24</b>

## EXECUTIVE SUMMARY

- One of the basic and fundamental principles of copyright law is that data is as such not protected; copyright only protects the creative form not the information incorporated in the protected work. Thus, Text and Data Mining (**TDM**) should in principle not be a use covered by any exclusive intellectual property (IP) rights (IPRs), both copyright and other *sui generis* rights. It could even be argued that this activity is outside the scope of exclusive rights and that any restriction would amount to undermine the underlying rationales of copyright protection and result in an inadmissible restriction of freedom of expression and information as protected by e.g. the European Court of Human Rights (ECHR) and the Charter of Fundamental Rights of the European Union. However, at some point, during the chain of activities enabling TDM research, technically some IPR relevant actions are necessary so that in the absence of a specific permission within the legal framework, TDM can lead to an infringement.
- Given the uncertainties that researchers face in applying present exceptions and limitations to TDM and the great necessity to create the best framework for European research and development, a **new limitation is urgently needed** in order to drive innovation and bridge the gap with other jurisdictions, permitting TDM activities.
- Given the lack of harmonisation in the implementation of exceptions and limitations by the Member States, the **mandatory character** of the TDM exception will facilitate cross-border research cooperation in the EU and the functioning of the Digital Single Market.
- Application to **commercial and non-commercial uses** and prohibition of **contractual override** is necessary to guarantee that the limitation achieves **effective results**.
- Given that dominant market players customarily override exceptions by imposing both contractual and technological measures—depriving users of the enjoyment of exceptions and lawful uses—limitations to **technological blocking** should be **introduced as well** by clearly spelling out that both Technological Protection Measures (TPMs) and network security and integrity measures should not undermine the effective application of the exception. Accordingly, protection against contractual and technological override should be also clearly extended to TDM mining materials not protected by IPRs, including those made available in a database.
- Moreover, the **extension of the scope of the TDM exception** beyond research organisations to all those enjoying lawful access to underlying mined materials, notably to start-ups and journalists, will increase the positive impact of the legislative action on research and innovation.
- In light of the increasing **research focus on the quality and verifiability**, a TDM exception should enable **storing and communication of research files** created for TDM.

## 1. INTRODUCTION

On 14 September 2016, the European Commission published a Proposal for a Directive on copyright in the Digital Single Market ("DSM Draft Directive").<sup>1</sup> Inter alia, the upcoming copyright reform would like to improve access to protected works across borders within the Digital Single Market (DSM) to boost research and innovation. To this end, the DSM proposal includes a set of new mandatory exceptions and limitations.<sup>2</sup> In particular, the proposal would like to introduce a **specific limitation for TDM**.

In this briefing, the authors would like to assess the proposal for a mandatory TDM limitation against the international and European framework of copyright exceptions and limitations by considering the need for such an exception, the positive and negative impacts of the proposal as currently drafted, and room for possible improvement to the end of fostering European research and innovation in the DSM.<sup>3</sup>

## 2. INTELLECTUAL PROPERTY PROTECTION AND TDM

TDM refers to a research technique to collect information from large amounts of digital data through automated software tools.<sup>4</sup> It works by

1. identifying input materials to be analysed, such as works, or data individually collected or organised in a pre-existing database;
2. copying substantial quantities of materials—which encompasses
  - a. pre-processing materials by turning them into a machine-readable format compatible with the technology to be deployed for the TDM so that structured data can be extracted and
  - b. possibly, but not necessarily, uploading the pre-processed materials on a platform, depending on the TDM technique to be deployed;
3. extracting the data; and
4. recombining it to identify patterns into the final output.<sup>5</sup>

---

<sup>1</sup> Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market, 14 September 2016, COM(2016) 593 final, 2016/0280 (Text with EEA relevance).

<sup>2</sup> *ibid.*, Art. 3-6, 7-9.

<sup>3</sup> Please note that some of the commentary presented in this briefing draws on materials included in Christophe Geiger, Giancarlo Frosio and Oleksandr Bulayenko (2017), *CEIPI Opinion on the European Commission's Proposal to Reform Copyright Limitations and Exceptions in the European Union*, CEIPI Research Paper No. 2017-9; (2018) 40 *European Intellectual Property Review*, Vol. 40, No. 1, pp. -154. For critical comments on the proposal, see also European Copyright Society (2017), *General Opinion on the EU Copyright Reform Package*, 24 January 2017; Max Planck Institute for innovation and Competition, Position Statement on the Proposed Modernisation of European Copyright Rules, available at: <http://www.ip.mpg.de/en/research/intellectual-property-and-competition-law/position-statement-modernization-of-european-copyright-rules.html>.

<sup>4</sup> See, e.g., for a reference to possible multiple TDM and data analysis techniques, Jiawei Han, Micheline Kamber and Jian Pei (2011), *Data Mining: Concept and Techniques*, 3<sup>rd</sup>. ed., Waltham, USA: Morgan Kaufmann.

<sup>5</sup> See, e.g., Diane McDonald and Ursula Kelly (2012), *The Value and Benefit of Text Mining to UK Further and Higher Education. Digital Infrastructure*, JISC, available at: <http://bit.ly/jisc-textm>; Jean-Paul Triaille, Jérôme de

Obviously, there is a tension between intellectual property protection and TDM techniques. TDM, as such, is not a use covered by any exclusive rights granted by copyright law. However, at some point, during the multiple activities enabling TDM research, IPRs might be infringed. In particular, TDM can involve some activities encroaching on the exclusive rights provided by Directive 2001/29/EC and Directive 1996/9/EC. In this respect, the TDM process might become relevant for intellectual property depending on the use of the existing sources, technical tools and the extent of the mining process.

**TDM usually involves some copying**, which even in case of limited excerpt might infringe the right of reproduction.<sup>6</sup> TDM activities can concern text or data, which both can be covered by intellectual property protection, both copyrights and database *sui generis* rights, or be outside the scope of protection (e.g. lacking originality or being in the public domain). Basically, IPRs can be affected whenever mining involves IP protected subject matters.

Only TDM tools involving minimal copying of few words or crawling through data and processing each item separately could be operated **without running into potential liability** for copyright infringement. This follows from the fact that copyright law does not protect data but only original expressions within copyright protected subject matters. In this respect, the Commission Proposal clarifies that “[t]ext and data mining may also be carried out in relation to mere facts or data which are not protected by copyright and in such instances no authorisation would be required”.<sup>7</sup> Obviously, although the proposal fails to specifically mention that, also works and other subject matter not protected by copyright or *sui generis* rights can be freely mined.

Instead, any **reproductions** resulting in the creation of a copy of a protected work along the chain of TDM activities might **trigger copyright infringement**. In this respect, pre-processing to standardize materials into machine-readable formats might trigger infringement of the right of reproduction.<sup>8</sup> Likewise, the uploading of the pre-processed material on a platform—which might occur or not depending on whether the TDM technique adopted makes use of a TDM software crawling data to be analysed directly from the source<sup>9</sup>—might also violate the right of reproduction. Mining—that stage of the TDM process where data is finally extracted—can also infringe upon the right of reproduction depending on the mining software deployed and the character of the extraction. For example, there are extraction techniques that would reproduce parts of the work so minimal to fall below the threshold of copyright infringement.<sup>10</sup>

---

Meeûs d’Argenteuil and Amélie de Francquen (2014), *Study of the Legal Framework of Text and Data Mining (TDM)*, prepared for the European Commission, p. 28, available at: [http://ec.europa.eu/internal\\_market/copyright/docs/studies/1403\\_study2\\_en.pdf](http://ec.europa.eu/internal_market/copyright/docs/studies/1403_study2_en.pdf); Sholom Weiss, Nitin Indurkha, and Tong Zhang (2010), *Fundamentals of Predictive Text Mining, Texts in Computer Sciences*, Amsterdam, Netherlands: Springer, p. 15 (noting how the popular pdf format does need to be adapted into more TDM friendly formats before TDM analysis can be carried out).

<sup>6</sup> See CJEU, C-5/08, *Infopaq International A/S v. Danske Dagblades Forening* (16 July 2009) ECLI:EU:C:2009:465, §§ 54-55 (finding that even an excerpt of 11 words might be protected). See also André Lucas, Henri-Jacques Lucas and Agnès Lucas-Schloetter (2012), *Traité de la propriété littéraire et artistique*, 4<sup>th</sup> ed., Paris, France: LexisNexis, pp. 349-353.

<sup>7</sup> Proposal, *supra* note 1, Recital 8.

<sup>8</sup> See Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, 2000 O.J. (L 167) 10-19, Art. 2.

<sup>9</sup> See Triaille, de Meeûs d’Argenteuil and de Francquen (2014), *supra* 5, p. 28.

<sup>10</sup> See *ibid.*, p. 31 (noting that “[t]his conclusion might not apply in certain exceptional circumstances, i.e. if it happens that, through the process of data analysis, the software only “crawls” through the text or the data, and processes them “one by one”, without copying the whole text but only one data or word or just a few of them at a time. In that case, even if one does not ignore that the ECJ in its *Infopaq* case accepted that an excerpt of a



Again, TDM might involve the reproduction, translation, adaptation, arrangement, and any other alteration of a **database protected by copyright**, which means the original selection and arrangement of the database's content.<sup>11</sup> For example, pre-processing for extraction might cleanse from a database portions and data that are irrelevant for data analysis. In this respect, pre-processing might violate both the right of reproduction and the right to make adaptations and arrangements.<sup>12</sup>

Moreover, TDM might **infringe sui generis database rights**, in particular the extraction—and to a minor extent the re-utilization—of substantial parts of a database. In this context, even if extraction does occur without reproduction of the original materials, extraction itself would infringe upon the exclusive rights provided to the database owner.<sup>13</sup> In this regard, the Court of Justice of the European Union (CJEU) has provided that the transfer of data from one medium to another and its integration into the new medium constitutes an act of extraction.<sup>14</sup>

Finally, it is to be noted that the TDM output should not infringe any exclusive rights as it merely reports on the results of the TDM quantitative analysis, typically not including parts or extracts of the mined materials.<sup>15</sup> However, it is worth highlighting that contemporary research practices, striving for verifiability of TDM research results, require the ability of researchers to store source materials and to communicate them at least to their peers. From a legal perspective, this conduct could most likely trigger the infringement of the right of **communication to the public**.

### 3. COPYRIGHT EXCEPTIONS AND LIMITATIONS AND TDM

The concept of exceptions and imitations to copyright serve a critical role in the so-called intellectual property—and copyright—paradox. The **paradox of intellectual property lies in a "system that promotes, or at least, aspires to promote knowledge [...] by restricting it"**.<sup>16</sup> The paradox explicates the tension between access and protection—or

---

sentence made of 11 words could be protected by copyright, it would be our understanding that, in such a case where the software only "swallows" one or two words or pixels or data or sounds at a time and then goes on to the next ones without keeping a copy of them but just "counting" the number of occurrences of, say, the word "malaria", then we would argue that no copying relevant in terms of copyright takes place and that such activity does not require the consent of the rightholder – and thus no exception is needed").

<sup>11</sup> See Directive 1996/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, 1996 O.J. (L 77) 20, Art. 5(a-b). See also Irini Stamatoudi (2016), 'Text and Data Mining', in Irini Stamatoudi (ed.), *New Developments in EU and International Copyright Law*, Leiden, Netherlands: Kluwer Law International, pp. 264-265.

<sup>12</sup> See Directive 1996/9/EC, *supra* note 11, Art. 5(b). Actually, although Directive 2001/29/EC does not harmonise the right to make adaptations and arrangements, its application has been claimed in EU law as an expression of the more general right of reproduction. Maria Lilla Montagnani and Giorgio Aime (2018), 'Il text and data mining e il diritto d'autore', *Annali Italiani di Diritto d'Autore*, Vol. 26 (forthcoming), fn. 15; Stamatoudi (2016), *supra* 11, pp. 262; Triaille, de Meeùs d'Argenteuil and de Francquen (2014), *supra* 5, p. 34.

<sup>13</sup> See Directive 1996/9/EC, Art. 7.

<sup>14</sup> *ibid.*, Arts. 2(a), 7(1) and 7(2)(b). See also CJEU, C-203/02, *The British Horseracing Board Ltd and Others v. William Hill Organization Ltd* (9 November 2004); Stamatoudi (2016), *supra* 11, p. 267 (noting that "although extraction is likely to take place on most occasions, this is not the case for re-utilization").

<sup>15</sup> *ibid.*, p. 33; Irini Stamatoudi (2016), *supra* 11, p. 262; Montagnani and Aime (2017), *supra* 12, fn. 20.

<sup>16</sup> P Bernt Hugenholtz, *Owning Science: Intellectual Property Rights as Impediments to Knowledge Sharing*, 2<sup>nd</sup> COMMUNIA Conference, Turin, Italy, 29 June 2001, available at: <http://www.communia-project.eu>. See also Universal Declaration of Human Rights, GA Res 217 (III) A, UN Doc A/RES/217(III) art 27 (10 December 1948) (acknowledging the intellectual property paradox by contrasting a right to access to knowledge in the first paragraph with the right to the protection of the moral and material interests of the author in the second paragraph).

public and private interest—which is intrinsic in intellectual property policy.<sup>17</sup> In this respect, exceptions and limitations allow for copyrighted works to be used without a license from the copyright owner because that use serves some important public interest and guarantees the safeguard of the core values of the European Union protected by fundamental rights such as freedom of expression and information.<sup>18</sup>

In EU law, exceptions and limitations are usually implemented by Member States under a voluntary scheme, with very few exceptions provided as mandatory. The application of exceptions and limitations in EU law—and elsewhere—does occur according to the general principles set out in international law by the “three-step test”. According to this test, exceptions and limitations would be permitted (1) in certain special cases (2) which do not conflict with a normal exploitation of the work or other subject-matter and (3) do not unreasonably prejudice the legitimate interests of the rightholder.<sup>19</sup> The three-step test has been increasingly interpreted in an expansive manner especially in order to accommodate public interest in the broader possible manner. According to Hugenholtz and Okediji “limitations and exceptions that (1) are not overly broad, (2) do not rob right holders of a real or potential source of income that is substantive, and (3) do not do disproportional harm to the right holders, will pass the test”.<sup>20</sup> Furthermore, Geiger, Griffiths and Hilty spearheaded a balanced approach of the three-step test, so that all three components of the test should be considered together in a “comprehensive overall assessment” taking into account the threats that excessive levels of copyright protection pose to, *inter alia*, public interests, notably in scientific progress and cultural, social, or economic development, while providing copyright holders with fair compensation.<sup>21</sup>

### **Do existing exceptions and limitations apply to TDM?**

Obviously, application of exceptions and limitations to TDM techniques when they are invasive enough to trigger intellectual property infringement—either copyright or *sui generis* database rights—have been repeatedly claimed. A balanced approach to the implementation of the three-step test would especially support such application of exceptions and limitations to TDM to foster “public interest, notably in scientific progress [...] and economic development”.<sup>22</sup>

In Europe, several exceptions within the mandatory and voluntary list provided by Directive 2001/29/EC have been selected as possible candidates to screen TDM from intellectual

---

<sup>17</sup> See Christophe Geiger, *The Future of Copyright in Europe: Striking a Fair Balance between Protection and Access to Information / L'avenir du droit d'auteur en Europe : Vers un juste équilibre entre protection et accès à l'information*, Report for the Committee on Culture, Science and Education – Parliamentary Assembly, Council of Europe, Strasbourg, July 2009 (Revised and updated in October 2009), [www.ceipi.edu](http://www.ceipi.edu) (short version published in French in: (2009) *La Semaine Juridique*, Ed G., No. 48, pp. 50-57; extended version published in English in: (2010) *Intellectual Property Quarterly*, Vol. 14, No. 1, pp. 1-14).

<sup>18</sup> See Christophe Geiger (2009), ‘Copyright’s Fundamental Rights Dimension at EU Level’, in Estelle Derclaye (ed.), *Research Handbook on the Future of EU Copyright*, Cheltenham, UK: Edward Elgar, p. 27.

<sup>19</sup> See Directive 2001/29/EC, *supra* 8, Art. 5(5). For the three-step test as enacted in international instruments, see Berne Convention, Art. 9(2); Trade-Related Aspects of Intellectual Property Rights (TRIPS) Agreement, Annex 1C, Marrakesh Agreement establishing the World Trade Organization, 15 April 1994, Art. 13; WIPO Copyright Treaty, Art. 10. See detailed on this legal instrument, Christophe Geiger, *The Role of the Three-Step Test in the Adaptation of Copyright Law to the Information Society*, UNESCO e-Copyright Bulletin January-March 2007.

<sup>20</sup> Bernt Hugenholtz and Ruth Okediji (2012), *Conceiving an International Instrument on Limitations and Exceptions to Copyright*, Amsterdam Law School Legal Studies Research Paper No. 2012-43, 6 March 2012, p. 3.

<sup>21</sup> See Christophe Geiger, Jonathan Griffiths, and Reto Hilty (2008), ‘Towards a Balanced Interpretation of the “Three-Step Test” in Copyright Law’, *EIPR*, Vol. 4, pp. 489-496.

<sup>22</sup> See, most recently, Joao Pedro Quintais (2017), Rethinking Normal Exploitation: Enabling Online Limitations in EU Copyright Law, *AMI* No. 6, pp. 197-205.

property infringement. Elsewhere, opening clauses or fair use models have been deployed to the same effect. In the United States, for example, TDM either belongs to the ontological or the functional public domain.<sup>23</sup> Starting with *Baker v. Selden*, courts argued that protected subject matter can be used when it “must necessarily be used as an incident to” using unprotected materials.<sup>24</sup> In the U.S., *Baker’s* reasoning has been applied to software reverse engineering.<sup>25</sup> Once applied to TDM, this case law would imply that in order to mine text and data—which are itself unprotected—a user might lawfully reproduce protected materials. In *Google Books*, more recently, TDM the entire corpus of human knowledge in order to create a relational database was found a transformative use, hence fair under § 107 of the US Copyright Act.<sup>26</sup>

Under current EU copyright law, TDM might be possibly covered by exceptions and limitations available, however their **application is uncertain**.

(1) The mandatory exception for temporary acts of reproduction might apply to limited TDM techniques.<sup>27</sup> Recital 10 of the DSM Draft Directive itself clarifies that this exception still applies but its application would be limited to TDM techniques which involve only the making of temporary reproductions transient or incidental<sup>28</sup> to an integral and essential part of a technological process which enables a lawful use with no independent economic significance.<sup>29</sup> Doubts have been repeatedly casted on whether all these requirements are fulfilled by reproductions done for TDM purposes, especially whether these reproductions are transient and have no independent economic relevance.<sup>30</sup> Application of temporary reproduction exception remain limited to residual cases for the large number of specific requirements that must be fulfilled, apparently in a cumulative manner according to the CJEU.<sup>31</sup> Also, the CJEU has reaffirmed that, being an exception, Art. 5(1) of Directive 2001/29/EC must be interpreted restrictively.<sup>32</sup> All in all, the sole mandatory exception

---

<sup>23</sup> See Giancarlo Frosio (2011), *COMMUNIA Final Report on the Digital Public Domain*, report prepared for the European Commission on behalf of the COMMUNIA Network and the NEXA Center, pp. 65-68. Similar conclusions can be extended to Canada featuring a very inclusive fair dealing exception. See Michael Geist (2013), ‘Fairness Found: How Canada Quietly Shifted from Fair Dealing to Fair Use’, in Michael Geist (ed.), *The Copyright Pentology*, Ottawa, Canada: University of Ottawa Press, pp. 157-186.

<sup>24</sup> See *Baker v. Selden*, 101 U.S. 99, 104 (1880).

<sup>25</sup> See, e.g., Pamela Samuelson (2005), ‘The Story of Baker v. Selden: Sharpening the Distinction between Authorship and Invention’, in Rochelle Dreyfuss and Jane Ginsburg (eds.), *Intellectual Property Stories*, New York, USA: Foundation Press, pp. 159-193.

<sup>26</sup> See *Authors Guild v. Google*, 804 F.3d 202 (2<sup>nd</sup> Cir. 2015). See also *Authors Guild v. HathiTrust*, 755 F.3d 87 (2<sup>nd</sup> Cir. 2014).

<sup>27</sup> See Directive 2001/29/EC, *supra* 8, Art. 5(1). See also, for a commentary, Christophe Geiger and Franciska Schönherr, Irini Stamatoudi and Paul Torremans (2014), ‘The Information Society Directive’, in Irini Stamatoudi and Paul Torremans (eds), *EU Copyright Law: A Commentary*, Cheltenham, UK: Edward Elgar, pp. 403-404 and 447-448 and Michel M. Walter and Silke von Lewinski (2010), ‘Information Society Directive’, in Michel M. Walter and Silke von Lewinski (eds.), *European Copyright Law: A Commentary*, New York, USA: Oxford University Press, pp. 968-969 and 1024-1027.

<sup>28</sup> See CJEU, C-360-13, *Public Relations Consultants Association* (5 June 2014), ECLI:EU:C:2014:1195, §§ 43 and 50 (noting that an act of reproduction is “incidental” “if it *neither exists independently of, nor has a purpose independent of, the technological process of which it forms part*”) (emphasis added).

<sup>29</sup> See Proposal, *supra* note 1, Recital 10.

<sup>30</sup> See C-5/08, *supra* 6, § 64 (noting that an act is “transient” “only if its duration is limited to what *is necessary for the proper completion of the technological process* in question, it being understood that that *process must be automated so that it deletes that act automatically, without human intervention*, once its function of enabling the completion of such a process has come to an end”) (emphasis added). See also CJEU, C-360/13, *supra* 28, §§ 40 and 48.

<sup>31</sup> See CJEU, C-302/10, *Infopaq International A/S v. Danske Dagblades Forening* (17 January 2012) ECLI:EU:C:2012:16, §§ 26.

<sup>32</sup> See C-5/08, *supra* 6, § 56.

available to researchers in Directive 2001/29/EC has a very limited scope that would hardly cover most TDM activities and offers insufficient legal certainty.

(2) In some Member States, depending on whether it is implemented or not, the **research exception** may cover some TDM-relevant acts infringing upon intellectual property rights.<sup>33</sup> However, the use of this exception is marred by legal uncertainty regarding its scope and application to TDM in different Member States.<sup>34</sup> In any case, the existing EU law limits the national research exceptions to non-commercial purposes and of course to the “sole purpose of *illustration* for teaching or scientific research”.<sup>35</sup> All TDM projects that do not qualify as scientific research and have a commercial purpose, both direct or indirect economic or commercial advance,<sup>36</sup> will be excluded from the outset from the application of the exception. In addition, the exception would apply as long as the source is indicated, “unless this turns out to be impossible”. Actually, this caveat would make easier the application of this exception to TDM. The massive amount of materials to be mined would make practically impossible to fulfill this requirement, thus triggering the application of its exemption.<sup>37</sup>

As per the **copyright in a database**, the same **research exemption** would also be available.<sup>38</sup> This scenario would refer to the case where a TDM process does reproduce the whole or substantial parts of a database, thus possibly infringing a copyright in its original arrangement. All limitations mentioned regarding the research exception in Directive 2001/29/EC would apply to this case as well as the possibility of claiming an exemption from the obligation of citing the source.<sup>39</sup> A **research exception** is also provided for the ***sui generis* database right**.<sup>40</sup> This exception would apply to a lawful user of a database for extracting or re-utilizing a substantial part of the database’s contents for non-commercial scientific research purposes, as long as the source is indicated.<sup>41</sup> The same assumptions on whether the indication of the source—which in this case would be the maker of the database and where the database is found—would be exempted in case of TDM given the principle that *impossibilium nulla est obligatio* should apply to this provision as well. However, this exception has seen multiple and diverging national implementation which might have provided for additional requirements.

(3) Art. 5(2)b of Directive 2001/29/EC could also be claimed as an exception applying to reproductions done for TDM purposes: “reproductions on any medium made by a natural person for private use and for ends that are neither directly nor indirectly commercial, on condition that the rightholders receive fair compensation [...]”.<sup>42</sup> The **private copy**

---

<sup>33</sup> See Directive 2001/29/EC, *supra* 8, Art. 5(3)(a).

<sup>34</sup> See Stamatoudi (2016), *supra* 11, p. 272-273 (noting that the exception’s scope of application varies greatly from country to country and listing relevant references).

<sup>35</sup> Directive 2001/29/EC, *supra* 8, Art. 5(3)(a). See also Stamatoudi (2016), *supra* 11, p. 273 (noting that it is disputed whether the word “illustration” refers only to “teaching” or both “teaching and scientific research” and concluding that the leading view is that “illustration is logically linked to teaching, where one needs to illustrate, while scientific research would require a more in-depth analysis and therefore “illustration” would be unrelated to its purpose or needs).

<sup>36</sup> See Stamatoudi (2016), *supra* 11, p. 273.

<sup>37</sup> See Montagnani and Aime (2017), *supra* 12, § 3; Triaille, de Meeûs d’Argenteuil and de Francquen (2014), *supra* 5, p. 63; Estelle Derclaye and Paul Torremans (2005), ‘La Lettre d’Angleterre’, *Revue Juridique Thémis*, Vol. 39, p. 519.

<sup>38</sup> Directive 1996/9/EC, *supra* 11, Art. 6(2)(b).

<sup>39</sup> See Stamatoudi (2016), *supra* 11, p. 275.

<sup>40</sup> See Directive 1996/9/EC, *supra* 11, Art. 9(b).

<sup>41</sup> *ibid.*

<sup>42</sup> See Directive 2001/29/EC, *supra* 8, Art. 5(2)b.

**exception** might potentially cover some uses by individual researchers, in particular in Member States that do not have a research exemption, thus allowing reproductions done for research purposes. However, the wording of the exception poses multiple challenges to its application to TDM. First, no direct or indirect commercial uses will be covered, leaving out most TDM research, which even if done by research organisation can have at least an indirect commercial end. Second, researchers might have to face the argument that the use for TDM might not be private if the use is not strictly for his own purpose, for example if the results are used by a collective group of researchers or by his institution. Again, the application of this exception implies fair remuneration to be given. This remuneration must be calculated on the basis of the criterion of the harm caused to rightholders,<sup>43</sup> which would need to be demonstrated in the context of TDM. Finally, the exception is voluntary implemented by Member States, thus exposing researchers again to inconsistent national implementation, limited legal certainty and high transaction costs.

(4) A potential candidate, but rather unsuccessful, for serving as a limitation for TDM is the so-called “**normal use of a database**”. This is the only mandatory exception included in Directive 96/9/EC. It provides that a lawful user can perform any reproductions of the database without additional authorization from the rightholder, if they are necessary for accessing the contents of a database and making a normal use of them.<sup>44</sup> Obviously, the limitation does apply to TDM in a very limited manner, only as long as the mining does constitute the normal use of the database and all other requirements are fulfilled. Again, the notion of “normal use” might receive multiple interpretations according to the Member States in which it is applied, therefore limiting researchers’ legal certainty. In this respect, German national law has extended the notion of normal use of a database to include TDM in connection with its recent enactment of a specific TDM exception, which as such would not cover databases. Harmonization here is highly suggested by including TDM in the EU law notion of “normal use” of a database.

(5) Finally, the right of the lawful user to **extract and/or re-utilize** without the authorization of the rightholder **insubstantial parts from a database** protected by *sui generis* right might also narrowly apply to TDM.<sup>45</sup> According to this provision, extraction and/or re-utilization might be done for any purpose whatsoever.<sup>46</sup> The notion of insubstantiality of a part must be evaluated through quantitative and qualitative criteria.<sup>47</sup> The CJEU has clarified that this assessment must consider the investment in the creation of the database and the prejudice that the extraction or re-utilization cause to that investment.<sup>48</sup> Put it bluntly, harming the investment implies the infringement of the *sui generis* database right.<sup>49</sup> Apparently, if TDM does not harm that investment, and as such it should not, Art. 8(1) of Directive 1996/9/EC might serve to avoid liability for extractions from a database for TDM purposes. In addition, the repeated and systematic extraction of insubstantial parts would still be lawful as long as it does not reconstitute the whole or substantial parts of the database, therefore damaging the investment of the rightholder.<sup>50</sup> Apparently, researchers could extract repeatedly and systematically insubstantial parts of database for TDM purposes, if that extraction does not reconstitute the whole or substantial

---

<sup>43</sup> See CJEU, C-467/08, Padawan v SGAE (21 October 2008) ECLI:EU:C:2010:620.

<sup>44</sup> See Directive 1996/9/EC, *supra* 11, Art. 6(1).

<sup>45</sup> *ibid.*, 8(1).

<sup>46</sup> *ibid.*

<sup>47</sup> *ibid.*

<sup>48</sup> See CJEU, C-203/02, *supra* 14, § 73.

<sup>49</sup> See Stamatoudi (2016), *supra* 15, p. 279; Estelle Derclaye (2014), ‘The Database Directive’, in Irini Stamatoudi and Paul Torremans (eds.), *EU Copyright Law*, Cheltenham, UK: Edward Elgar, p. 111.

<sup>50</sup> See CJEU, C-203/02, *supra* 14, § 73.

parts of the database or, more generally, harm the investment for constituting the database. However, again, the reach of this provision will be actually limited as it applies only to lawful users, running TDM on databases, using specific TDM techniques crawling the database rather than making a copy of it, and can be contractually overridden.<sup>51</sup>

In sum, existing exceptions and limitations might not offer a stable legal framework to safely engage in TDM research projects and invest considerable resources. The legal doctrine appears to consistently support this position. According to Stamatoudi, for example, available exceptions would hardly be fit for purpose.<sup>52</sup> All in all, the possibility of relying on existing provisions, including temporary acts of reproduction, scientific research, normal use of a database, and extraction of “insubstantial parts” from a database protected by the *sui generis* right, without adoption of additional interpretative norms or judgements of high instances is doubtful.

In addition, it should be noted that all mentioned exceptions and limitations that could apply to TDM—but the exception of temporary acts of reproduction—are implemented by Member States on a voluntary basis. Voluntary implementation makes even less predictable whether existing exceptions and limitations can be applied to TDM projects, especially those of cross border nature. This would be a substantial loophole in the present EU legal framework in light of the DSM Draft Directive’s goal of stimulating the creation of a DSM for research and innovation.

### **Which problems researchers face in applying existing exceptions to TDM?**

Uncertainty in the application of existing exception and limitation—coupled with the legal fragmentation caused by the voluntary implementation of some of the exceptions that might be claimed to allow TDM activities—take the center stage when it comes to problems that researchers face in applying present exceptions and limitations to TDM. As also reported by the Impact Assessment, “**researchers are faced with legal uncertainty** with regard to whether and under which conditions they can carry out TDM on content they have lawful access to”.<sup>53</sup>

First, as seen, it is highly uncertain whether existing exceptions and limitations, both mandatory and voluntary, would apply to TDM. In addition, the **unharmonised EU legal framework for exceptions and limitations** further constraints potentials for an EU DSM for research and innovation. Actually, the voluntary nature of the exceptions that might possibly apply to TDM further affects cross-border collaborations as researchers would be unaware—or face high transaction costs for clearance—of whether TDM would be lawful across all EU jurisdictions involved in the research collaboration. While, the exception for temporary acts of reproduction is mandatory across all the Member States, this is not the case for the research exception, for example. Member States, therefore, are free to determine whether and how to implement all voluntary exceptions that might apply to TDM. Those exceptions then, if implemented, will be construed by national courts according to the peculiarities of that particular implementation and the specificities of the national legal system. A striking example of how national inconsistencies in implementing exceptions theoretically applicable to TDM would then make it impracticable to rely on them

---

<sup>51</sup> Cf. Stamatoudi (2016), *supra* 11, p. 280.

<sup>52</sup> See Stamatoudi (2016), *supra* 11, p. 252. For an updated commentary of the application of exceptions and limitations to TDM also in the context of the proposed reform, see also Montagnani and Aime (2017), *supra* 12.

<sup>53</sup> European Commission (2016), Commission Staff Working Document, Impact Assessment on the modernisation of EU copyright rules, 14 September 2016, SWD(2016) 301 final, Part 3/3, p. 94 (emphasis added).

for cross-border projects is the Italian implementation of the research exemption. This implementation does not include the caveat “unless [citing the source] turns out to be impossible” but apparently always requires to mention work, author and publisher of the materials used, therefore making it practically impossible to claim the application of the exception.<sup>54</sup> In this context European researcher face grave constraints to their research activities, especially given the **cross-border nature of European research projects** today. Actually, most research consortiums leading innovation in Europe and seeking eligibility for EU funding are required to involve partners from different Member States (e.g. under the Horizon 2020 programmes).

Furthermore, researchers have to cope with the rent-seeking behaviour of economic actors controlling content. This creates further barriers for TDM. Combinations of **contractual and technical measures** are frequently used to create **insurmountable hurdles for researchers** engaging in TDM projects. Research and database providers often contractually override exceptions and limitations. “Take it or leave it” contractual conditions make access conditional upon accepting providers’ terms, while obtaining specific permission from various publishers to carry out TDM research can be extremely complex.<sup>55</sup> In addition, technical measures, commonly used to ensure security and integrity of databases, also constitute a barrier for application of TDM techniques<sup>56</sup>. Finally, contractual limitations to TDM activities do not necessarily rely on copyright exclusive rights. Contractual and technological override might also prevent TDM of materials not protected by IPRs. Actually, the CJEU has stated that if a database is not protected either by copyright or by the sui generis right, the author is not precluded for laying down any contractual limitations on its use by third parties, apparently including also those preventing TDM.<sup>57</sup> In order to promote TDM research, the EU legislator should clarify that protection against contractual and technological override should also be always extended to TDM mining both materials protected and not protected by IPRs, including those made available in a database.

Finally, it should also be noted that when researchers come to publish, they might find that they lack the right copyright permissions.<sup>58</sup> Publishers can put specific clauses in their licences that rule out mining and gaining permission to mine content from various publishers can be hugely complex.

In sum, existing research exceptions in the EU law are not fully adapted to TDM, diversity of licencing practices generates high transaction costs, and there is a fragmentation of rules in the single market.<sup>59</sup>

### **Will a new mandatory TDM exception be able to drive innovation?**

The value of unlocking TDM research and unburdening it from intellectual property blocking effects, has been broadly highlighted. In the big data era, new data is created by the

---

<sup>54</sup> See Legge diritto d’autore, Art. 70(1). See also Montagnani and Aime (2017), *supra* 12, § 3.

<sup>55</sup> See, in particular on this issue, Lucie Guibault (2002), *Copyright Limitations and Contracts: An Analysis of the Contractual Overridability of Limitations on Copyright*, The Hague, Netherlands: Kluwer Law International.

<sup>56</sup> This is a general concern with regard to the functioning of exceptions and limitations provided in EU law, see Christophe Geiger (2008), ‘The Answer to the Machine should not be the Machine’, *EIPR*, Vol. 4, pp. 121-129.

<sup>57</sup> See CJEU, C-30/12, Ryanair Ltd v. PR Aviation BN (15 January 2015), ECLI:EU:C:2015:10.

<sup>58</sup> See Rita Morais, Julian Bauer and Lidia Borrell-Damian (2017), *Open Access: 2016-2016 EUA Survey Results*, June 2017, p. 27.

<sup>59</sup> See Commission (2016), *supra* 53,, p. 95.

quintillions of bytes every day.<sup>60</sup> Therefore, TDM becomes increasingly essential by enabling a much faster processing of huge amounts of text and data, a challenging task given the booming quantity of publications and information made available online. Actually, TDM is a ground-breaking tool for research of all kind both for-profit and non-profit.<sup>61</sup> Some studies have estimated that it could create value in excess of hundreds of billions for Europe if data can be used more effectively.<sup>62</sup> According to a JISC Report, TDM might bring beneficial effects in terms of “efficiency; unlocking hidden information and developing new knowledge; exploring new horizons; improved research and evidence base; and improving the research process and quality”.<sup>63</sup> Again, TDM would bring also broader social and economic positive externalities, such as “cost savings and productivity gains, innovative new service development, new business models and new medical treatments”.<sup>64</sup>

The introduction of **mandatory exceptions** and limitations would certainly be helpful in unlocking value and innovation that TDM research might generate. Actually, the optional list included in the InfoSoc Directive failed to harmonise the EU legislative framework in this regard.<sup>65</sup> The Commission’s plan of moving away from the voluntary arrangement of Directive 2001/29/EC should in itself support innovation in Europe. By promoting harmonisation, mandatory exceptions and limitations foster the DSM, which will be a driver for coordinated and cheaper innovation processes in Europe.

As earlier already noted, voluntary implementation of exceptions and limitations at the national level can be rated high among the critical issues for researchers in performing TDM research in Europe and applying existing exception and limitations. A recent study—done by CEIPI researchers for the European Union Intellectual Property Office (EUIPO)—involving an analysis of the copyright legislation in 28 Member States revealed that there are major uncertainties for consumers to know what is permitted or not with regard to exempted uses in copyright law, as “‘everyday’ uses of copyrighted works in the online world currently still lack a clear and straightforward answer as regards their legality”.<sup>66</sup> The study continues by noting that “[w]hile international and EU law have approximated the different copyright traditions to a certain extent, a closer look reveals that divergences still prevail [...] even if a few common basic principles can certainly be identified, the exceptions to these principles as well as their implementation vary significantly”.<sup>67</sup> In particular, the study lists exceptions

---

<sup>60</sup> See Xindong Wu, Xingquan Zhu, Gong-Qing Wu and Wei Ding (2014), ‘Data mining with big data’, *IEEE transactions on knowledge and data engineering*, Vol. 26, p. 97; Jeff Schultz, How Much Data Is Created on the Internet Each Day?, *DZone*, 16 October 2017, <https://dzone.com/articles/how-much-data-is-created-on-the-internet-each-day>.

<sup>61</sup> See European Commission (2014), Standardisation in the area of innovation and technological development, notably in the field of Text and Data Mining: Report from the Expert Group, April 2014, p. 3. See also Stamatoudi (2016), *supra* 11, pp. 252-282.

<sup>62</sup> See James Manyika et al (2011), *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, San Francisco, USA: McKinsey Global Institute.

<sup>63</sup> McDonald and Kelly (2012), *supra* 5, pp. 17-21.

<sup>64</sup> *ibid.*

<sup>65</sup> See Christophe Geiger and Franciska Schönherr (2014), ‘Limitations to Copyright in the Digital Age’, in Andrej Savin and Jan Trzaskowski (eds.), *Research Handbook on EU Internet Law*, Cheltenham, UK: Edward Elgar, pp. 110-142; Christophe Geiger and Franciska Schönherr (2012), ‘Defining the Scope of Protection of Copyright in the EU: The Need to Reconsider the Acquis regarding Limitations and Exceptions’, in Tatiana-Eleni Synodinou (ed.), *Codification of European Copyright Law, Challenges and Perspectives*, Amsterdam, Netherlands: Kluwer Law International, pp. 133-167; Jonathan Griffiths, Christophe Geiger, Martin Senftleben, Raquel Xalabarder and Lionel Bently (2015), ‘Limitations and Exceptions as Key Elements of the Legal Framework for Copyright in the European Union – Opinion on the Judgment of the CJEU in Case C-201/13 Deckmyn’, *EIPR*, Vol. 45, No. 1, pp. 93-101.

<sup>66</sup> Christophe Geiger and Franciska Schönherr (2017), *Frequently Asked Questions (FAQ) of Consumers in relation to Copyright, Summary Report*, a project commissioned by the European Union Intellectual Property Office, p. 8, available at: <https://euipo.europa.eu/ohimportal/web/observatory/observatory-publications>.

<sup>67</sup> *ibid.*, p. 6.



and limitations to copyright as one of the areas of major divergence in national copyright law.

From the finding above, it follows, as already highlighted in the past, “the importance of declaring that limitations and exceptions justified by the public interest be mandatory.”<sup>68</sup> This applies here with specific emphasis to the TDM exception but the mandatory nature should be extended to all limitations and exceptions of the EU acquis in order to achieve true harmonisation and legal security for all players in the creative process, from authors to exploiters and users of copyrighted works.<sup>69</sup> A unified and mandatory approach is especially crucial in the “digital environment as the internet involves uses that, most of the time, affect several copyright legislations, leading to a major insecurity regarding what is allowed.”<sup>70</sup> TDM would stand as a quintessential example of this conundrum that digitisation has brought about and lack of adequate harmonization in EU law fails to address.

Finally, as also highlighted by these authors elsewhere,<sup>71</sup> a TDM exception should serve to bridge the **gap with other jurisdictions**, such as the United States, where apparently TDM would be beyond the reach of the copyright holders’ exclusive rights.<sup>72</sup>

### **Would an “opening clause” serve the public interest?**

As the DSM Draft Directive would like to implement “measures to adapt exceptions and limitations to the digital and cross-border environment”,<sup>73</sup> other solutions could be adopted on top of the introduction a specific mandatory TDM exception to reach that goal. In the discussion leading to the DSM Draft Directive, other exceptions have been considered as

---

<sup>68</sup> Stéphanie Carre, Christophe Geiger, Jean Lapousterle, Franck Macrez, Adrien Bouvel, Théo Hassler, Xavier Seuba, Oleksandr Bulayenko, Franciska Schönherr and Marie Hemmerle-Zemp (2014), *Response of the CEIPI to the Public Consultation of the European Commission on the Review of the European Union Copyright Rules*, CEIPI Research Paper No. 2014-01, p. 6, available at: <https://ssrn.com/abstract=2971029>. Christophe Geiger, Frank Macrez, Adrien Bouvel, Stéphanie Carre, Théo Hassler and Joanna Schmidt-Salewski (2009), ‘What Limitations to Copyright in the Information Society? A Comment on the European Commission’s Green Paper ‘Copyright in the Knowledge Economy’’, *IIC*, Vol. 40, pp. 412-433.

<sup>69</sup> Cf. Geiger, *supra* 17.

<sup>70</sup> Christophe Geiger, Oleksandr Bulayenko, Théo Hassler, Elena Izyumenko, Franciska Schönherr and Xavier Seuba (2015), *Reaction of CEIPI to The Resolution on the Implementation of Directive 2001/29/EC on the Harmonisation of Copyright in the Information Society Adopted by the European Parliament on the 9th July 2015*, CEIPI Research Paper No. 2015-01, p. 17, available at: <https://papers.ssrn.com/sol3/papers=2970507>. This paper was published with the title “The Resolution of the European Parliament of 9 July 2015: Paving the Way (finally) for a Copyright Reform in the European Union?” in: *EIPR* 2015, pp. 683-701. For a comprehensive analysis of the merit and legality of this approach, see Evangelia Loli, *Copyright Exceptions and Limitations in the European Union: Exploring the Flexibilities*, CEIPI Master thesis, Strasbourg September 2017.

<sup>71</sup> See Geiger, Bulayenko, Hassler, Izyumenko, Schönherr and Seuba (2015), *supra* 70, p. 24.

<sup>72</sup> See Commission (2014), *supra* 61, pp. 12-13, 43. See also Christian Handke, Lucie Guibault and Joan-Josep Vallbé (2015), ‘Is Europe Falling Behind in Data Mining? Copyright’s Impact on Data Mining in Academic Research’, in Brigit Schmidt and Milena Dobrova (eds.), *New Avenues for Electronic Publishing in the Age of Infinite Collections and Citizen Science: Scale, Openness and Trust*, Amsterdam, Netherlands: IOS Press, pp. 120-130.

<sup>73</sup> Commission (2016), *supra* 1, Title II. See also Francisco Javier Cabrera Blázquez, Maja Cappello, Gilles Fontaine and Sophie Valais (2017), *Exceptions and Limitations to Copyright*, Strasbourg, France: IRIS Plus, European Audiovisual Observatory, pp. 61-73.

possible candidates for the reform but finally discarded.<sup>74</sup> In particular, the introduction of an open norm—or general exception—similar to US fair use has long been considered in EU legal scholarship and policy debate.<sup>75</sup>

An enumerated list of exceptions and limitations has shown little flexibility in adapting to evolving market and technological conditions, whereas an open fair use or fair dealing norm alone would be little predictable, increasing transaction costs and favouring economically stronger market players.<sup>76</sup> An “opening clause” should address uses that are not yet covered by existing exceptions and limitations but are justified by important public interest rationales and fundamental rights such as freedom of expression and the right to information.<sup>77</sup> A balanced approach—melding together the two options—might overcome the limitations of each of the two alternatives and boost the European DSM’s international competitiveness, in compliance with the “three-step test”. In this regard, CEIPI, for example, has already endorsed in the past a policy option that would guarantee legal certainty through a list of further harmonised or unified exceptions and limitations, but to combine it with a certain dose of *flexibility* of the EU legal framework, in order to ensure its capacity to adapt to a rapidly changing environment. This limited “opening” of the list of exceptions and limitations could have possibly been based on the “three-step test”.<sup>78</sup>

Among endless applications, an opening clause would make TDM research possible according to the three-step test and the impact of the use on the legitimate interest of the users, being limited by any prejudice on the potential market of the rightholder, if any exists. In light of a balanced approach to the “three-step test”, an opening clause might allow rightholders’ fair compensation if any prejudice to rightholders’ interest or conflict with the normal exploitation of the work might occur.

---

<sup>74</sup> For example, despite massive mobilization and a public consultation, it seems that freedom of panorama will not be part of the directive. See European Commission (2016), *Public consultation on the role of publishers in the copyright value chain and on the ‘panorama exception’*, 23 March 2016, available at: <https://ec.europa.eu/digital-single-market/en/news/public-consultation-role-publishers-copyright-value-chain-and-panorama-exception>. Also, discussions surround the introduction of an exception for user-generated content. See also Judith Blijden, ‘Keeping an eye on the fine print: the UGC exception and the JURI committee’, *COMMUNIA Association*, 15 June 2017, <http://www.communia-association.org/2017/06/15/keeping-eye-fine-print-ugc-exception-juri-committee>.

<sup>75</sup> See Martin Senftleben (2017), ‘The Perfect Match – Civil Law Judges and Open-Ended Fair Use Provisions’, *Am. U. Int’l L. Rev.*, Vol. 33, pp. 231-286; P. Bernt Hugenholtz (2016), *Flexible Copyright: Can EU Author’s Rights Accommodate Fair Use?*, in Irini Stamatoudi (ed.), *New Developments in EU and International Copyright Law*, Leiden, Netherlands: Kluwer Law International, pp. 417-433; Martin Senftleben (2014), ‘Comparative Approaches to Fair Use: An Important Impulse for Reforms in EU Copyright Law’, in Graeme Dinwoodie (ed.), *Methods and Perspectives in Intellectual Property*, Cheltenham, UK: Edward Elgar; Antony Dnes (2013), ‘Should the UK Move to a Fair-Use Copyright Exception’, *IIC*, Vol. 44, No. 4, pp. 418-444; Rob Van der Noll, Stef van Gompel, Lucie Guibault, Jarst Weda, Joost Poort, Ilan Akker and Kelly Breemen (2012), *Flexible Copyright: The Law and Economics of Introducing an Open Norm in the Netherlands*, SEO Economic Research Report N. 2012-60; Christophe Geiger (2008), ‘Flexibilising Copyright – Remedies to the Privatisation of Information by Copyright Law’, *IIC*, Vol. 39, No. 2, pp. 178-197.

<sup>76</sup> See, e.g., P. Bernt Hugenholtz (2017), ‘Flexible Copyright. Can EU Author’s Right Accommodate Fair Use?’ in Ruth Okediji (ed.), *Copyright Law in an Age of Limitations and Exceptions*, Cambridge, UK: Cambridge University Press, pp. 275-291; Pamela Samuelson (2017), ‘Justifications for Copyright Limitations and Exceptions’, in Ruth Okediji (ed.), *Copyright Law in an Age of Limitations and Exceptions*, Cambridge, UK: Cambridge University Press, pp. 12-59; P. Bernt Hugenholtz and Martin Senftleben (2012), *Fair Use in Europe: in Search of Flexibilities*, Institute for Information Law Research Paper No. 2012-39, available at: <https://papers.ssrn.com/sol3/papers=2013239>.

<sup>77</sup> As the Explanatory Memorandum to the directive Proposal admits, the Directive proposal will only have a “limited impact [...] on the freedom of expression and information, as recognised respectively by Articles 16 and 11 of the Charter, due to the mitigation measures put in place and a balanced approach to the obligations set on the relevant stakeholders”.

<sup>78</sup> See Geiger, Bulayenko, Hassler, Izyumenko, Schönherr and Seuba (2015), *supra* 70, pp. 20-21.

## 4. THE TDM LIMITATION IN THE MEMBER STATES AND THE EU

As earlier explained, the uncertain application of existing exceptions and limitation to TDM techniques causes major hindrances to researchers willing to carry out TDM research. In addition, exceptions and limitations that might potentially apply to TDM techniques have been implemented in peculiar ways by each Member State. This situation indeed frustrates researchers' attempts to find a consistent legal framework within which to operate.

**TDM limitations** have been long under consideration in Europe.<sup>79</sup> Such limitations have been already introduced in some Member States and other jurisdictions.<sup>80</sup> This has so far led to a **fragmented legal environment in the Member States**. In particular, after several studies assessing costs and benefits of a copyright reform,<sup>81</sup> the **UK** was the first Member State to adopt a TDM exception on **19 May 2014**.<sup>82</sup> The UK Copyright, Designs and Patents Act 1988 has provided for a text and data analysis exception for non-commercial research.<sup>83</sup> According to UK law, the making of a copy of a work by a person who has lawful access to that work does not infringe copyright if it is made so that that person can carry out a computational analysis of anything included in that work for non-commercial research purposes.<sup>84</sup> The exception does not cover reproduction of databases.

On **7 October 2016**, by Law No. 2016-1231 for a Digital Republic (*Loi pour une République numérique*), **France** introduced TDM exceptions both applying to works (Article L122-5, 10 of the CPI) and databases (Article L342-3, 5 of the CPI).<sup>85</sup> French exceptions cover acts of reproduction from "lawful sources" (materials lawfully made available with the consent of the rightholders) for TDM as well as storage and communication of files created in the course of TDM research activities. An actualization decree should regulate conditions under which TDM can be undertaken as well as the modalities for storing and communicating research files that were created for TDM purposes. The exceptions are

---

<sup>79</sup> See, e.g., Thomas Margoni and Giulia Dore (2016), 'Why We Need a Text and Data Mining Exception (But it is Not Enough)', 1 April 2016, available at: <https://zenodo.org/record/248048#.WXdf2oiGNEY>; Handke, Guibault and Vallbé (2015), *supra* 72, pp. 120-130; Michelle Brook, Peter Murray-Rust and Charles Oppenheim (2014), 'The Social, Political and Legal Aspects of Text and Data Mining (TDM)', *D-Lib Magazine*, Vol. 20, No. 11/12, § 4, <http://openaccess.city.ac.uk/4784>; Triaille, de Meeûs d'Argenteuil and de Francquen (2014), *supra* 5, pp. 96-118. Cf. Nils Dietrich, Lucie Guibault, Thomas Margoni, Krzysztof Siewicz, Gerald Spindler, and Andreas Wiebe (2013), *Safe to Be Open: Study on the Protection of Research Data and Recommendations for Access and Usage*, Göttingen, Germany: Universitätsverlag Göttingen, available at: <http://eprints.gla.ac.uk/129335>.

<sup>80</sup> See, e.g., Copyright Act, Sec. 5, Art. 47-7 (Japan). For a comment on the Japanese TDM exception, which is not limited neither by non-commercial purpose nor scientific research purposes, see Future TDM (2016), Baseline report of policies and barriers of TDM in Europe, extended version of D3.3, pp. 75-76, available at: <http://www.futuretdm.eu/knowledge-library/?b5-file=4588&b5-folder=2227> (providing an English translation of the Japanese provision).

<sup>81</sup> See Ian Hargreaves (2011), *Digital Opportunity: A Review of Intellectual Property and Growth*, Independent Report commissioned by the Prime Minister of the UK, pp. 4, 8, 41, 47, 51 and 99 (recommending to adopt a TDM exception), available at: <https://www.gov.uk/government/publications/digital-opportunity-review-of-intellectual-property-and-growth> and IPO (2012), *Exception for copying of works for use by text and data analytics*, Impact Assessment, No. BIS0312, available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/308738/ia-exception-dataanalytics.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/308738/ia-exception-dataanalytics.pdf).

<sup>82</sup> See Regulation 3 of the Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014, No. 1372, adding Article 29A to the Copyright, Designs and Patents Act 1988. The Regulations came into force on 1 June 2014.

<sup>83</sup> See Copyright, Designs and Patents Act 1988, § 29A (UK).

<sup>84</sup> *ibid.*, § 29(1).

<sup>85</sup> Art. 38 of the Law No. 2016-1231 of for a Digital Republic added paragraph 10 to Art. L122-5 and paragraph 5 to Art. L342-3 of the Intellectual Property Code (*Code de la propriété intellectuelle*) (CPI).

restricted to “text and data included or associated with scientific writings”.<sup>86</sup> TDM cannot pursue commercial objectives and should be limited to the needs of (public) research.<sup>87</sup>

An **Estonian** TDM exception entered into force on **1 January 2017**.<sup>88</sup> With regard to copyrighted works, Estonian law (*Autoriõiguse seadus*) allows: processing of materials covered by exclusive rights “for the purposes of text and data mining, provided that such use does not have a commercial purpose”. Likewise the UK exception, the Estonian exception seems to be limited to the right of reproduction of copyrighted works, excluding post-TDM communication for the purpose of quality assurance. While the exception does not contain “lawful access” or similar requirement, it demands mention of “the name of the author of the work, if it appears thereon, the name of the work and the source publication”.

**Germany** adopted its TDM exception on 30 June 2017. This amendment of the Law on Copyright and Related Rights (*Urheberrechtsgesetz*) will enter into force on **1 March 2018**. Article 60d enables TDM for scientific research absent a commercial purpose. The exception covers the acts of reproduction necessary for undertaking TDM and the making available of a “corpus” (e.g., source materials that were normalised, structured and categorised) to a “specifically limited circle of persons for their joint scientific research, as well as to individual third persons” for quality assurance. Once the TDM project is completed, the “corpus” can be sent to institutions designated by law for long term storage. Any other copy made should be deleted. It is worth noting that German law does not impose a “lawfully accessed source” requirement as France does. Also, it does not limit the source materials that can be mined to “text and data included or associated with scientific writings”. With regard to databases, their reproduction is being qualified as constituting “normal use”. In this respect, as mentioned already, it would be worth considering to extend the same qualification to EU law at large.

At a quick glance, the inconsistency of existing TDM exceptions so far adopted by Member States illustrates how fragmented the legal landscape will rapidly become if no EU-wide action is undertaken.

## 5. THE COMMISSION’S PROPOSAL

In May 2015, the European Commission issued its Digital Single Market Strategy (DSMS), in which it plans to provide “greater legal certainty for the cross-border use of content for specific purposes (e.g., research, education, text and data mining, etc.) through harmonised exceptions”.<sup>89</sup> Following up on this plan, the DSM Draft Directive includes the following exception to copyright and database *sui generis* right:

Member States shall provide for an exception to the rights provided for in Article 2 of Directive 2001/29/EC, Articles 5(a) and 7(1) of Directive 96/9/EC and Article 11(1) of this Directive for reproductions and extractions made by research

---

<sup>86</sup> *ibid.*

<sup>87</sup> Art. L122-5, 10 of the CPI refers to the “needs of public research” and Art. L342-3, 5 to “research”.

<sup>88</sup> See Estonian Copyright Act, Art. 19(3).

<sup>89</sup> European Commission (2015), Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A Digital Single Market Strategy for Europe, COM(2015) 192 final, 6 May 2015, § 2.4. Apparently, the Commission dropped previous plans to explore licencing arrangements to deal with TDM. See, e.g., European Commission (2012), Communication on Content in the Digital Single Market, COM(2012) 789 final, 18 December 2012, p. 4.

organisations in order to carry out text and data mining of works or other subject-matter to which they have lawful access for the purposes of scientific research.<sup>90</sup>

The proposal provides a TDM exception to the right of reproduction of copyright protected subject matters and the *sui generis* database extraction right. In addition, the TDM exception would apply to the new right over digital uses of press publication that the DSM Draft Directive has proposed.<sup>91</sup>

Contrary to the general structure of Article 5 of Directive 2001/29/EC—the new TDM exception would be **mandatory**. As mentioned, other than the exception for temporary copies, Directive 2001/29/EC provides an optional list of exceptions and limitations that Member States are free to implement. In addition, Articles 6 and 9 of Directive 96/9/EC, apart from the exception for the “normal use” of a database, also provide only for voluntary exceptions and limitations.

Several limitations would apply to the TDM exception.

1. First, TDM exception’s **beneficiaries are limited to research organisations**, meaning “any organisation the primary goal of which is to conduct scientific research or to conduct scientific research and provide educational services”.<sup>92</sup>
2. Second, the exception allows only purpose-specific lawful access for TDM, namely “for the **purpose of scientific research**”.<sup>93</sup>
3. Third, the exception applies only to works or other subject-matter to which research organisations “have **lawful access**”.
4. Finally, a further limitation is provided by Article 3(3) and Recital 12 allowing rightholders to introduce measures to protect the “**security and integrity**” of **their networks** and databases where works are hosted. However, such measures shall not go beyond what is necessary to achieve that objective.<sup>94</sup>

Within these limitations, the TDM exception’s scope is very inclusive as it applies both to **commercial and non-commercial uses and**—very importantly—**cannot be overridden by contract**.

Any contractual provision contrary to the exception provided for in paragraph 1 shall be unenforceable.<sup>95</sup>

---

<sup>90</sup> Proposal, *supra* note 1, Art. 3(1).

<sup>91</sup> *ibid.*, Art. 11(1). See also for a discussion new neighbouring right proposal, Christophe Geiger, Oleksandr Bulayenko and Giancarlo Frosio (2016), Opinion of the CEIPI on the European Commission’s Copyright Reform Proposal, with a Focus on the Introduction of Neighbouring Rights for Press Publishers in EU Law, CEIPI Research Paper No. 2016-01, available at: <https://papers.ssrn.com/sol3/papers=2921334>.

<sup>92</sup> Proposal, *supra* note 1, Art. 2(1).

<sup>93</sup> *ibid.*, Art. 3(1). Please note that, as presently drafted, the text of the article makes unclear whether “the purposes of scientific research” would apply to the “lawful access” or the “reproductions and extractions”. It seems to be implied, given the proximity, that the purpose refers to the lawful access. However, given also that different linguistic versions of the text might use different punctuation, clarification would be needed. In any event, also for this reason, we suggest dropping this purpose-specific requirement altogether.

<sup>94</sup> *ibid.*, Art. 3(3).

<sup>95</sup> *ibid.*, Art. 3(2).

According to **Article 6 of the DSM Draft Directive**—a provision common to all three new exceptions—the three-step test would apply to the new TDM exceptions. In addition, technological protection measures (TPMs) anti-circumvention provisions would also apply to all new exceptions.

Article 5(5) and the first, third and fifth subparagraphs of Article 6(4) of Directive 2001/29/EC shall apply to the exceptions and the limitation provided for under this Title.<sup>96</sup>

## Positive and Negative Impacts

As mentioned, the introduction of Article 3 of the DSM Draft Directive meets important policy goals. In particular, it is set to provide a normalised, consistent level playing field for researches across Europe to legally carry out TDM projects. The **major positive impacts** of the proposal lie in its focus on harmonisation of Member States' laws, through a **mandatory solution**. This is a welcome arrangement that promotes **harmonisation** and, therefore, the DSM. As such, a harmonised framework for TDM research will be driving **innovation** in the **DSM**, promoting EU-wide, integrated, larger research projects.

Again, harmonisation will also be supported by an expansive scope of the limitation, covering both **commercial and non-commercial uses**, and the **unenforceability of contrary contractual provisions**. This inclusion would be critical not to devoid the exception of any practical utility. As mentioned, publishers can contractually rule out mining in their licences, and transaction costs to obtain permission to mine content from multiple publishers might *de facto* make TDM projects unsustainable. Certainly, the proposal deserves praise for protecting TDM research from contractual enclosure.

The critical positive externalities that the major features of the proposed TDM exception bring about **make it certainly fit for the intended goals in general**.

However, there still remain **negative impacts** of the proposal as currently drafted that need to be assessed.

(1) Much discussion regarding this proposal does concern whether the TDM exception's beneficiaries should not be **limited to "research organisations"**. To qualify for the exception, research organisations must operate on a not-for-profit basis or by reinvesting all the profits in their scientific research, or pursuant to a public interest mission.<sup>97</sup> The Impact Assessment acknowledged that "part of the research community has expressed the concern that the concept of public interest organization could be difficult to define and apply".<sup>98</sup> According to Recital 11, a public interest mission might be "reflected through public funding or through provisions in national laws or public contracts".<sup>99</sup> The DSM Draft Directive further limits the scope of the exceptions that does not apply to research institutions controlled by commercial undertakings.<sup>100</sup> Control could be exercised "because of structural situations such as their quality of shareholders or members".<sup>101</sup> In particular, it

---

<sup>96</sup> *ibid.*, Art. 6.

<sup>97</sup> *ibid.*, Art. 2(1)(a-b).

<sup>98</sup> Commission (2016), *supra* 53, Part. 1/3, p. 109.

<sup>99</sup> Proposal, *supra* note 1, Recital 11.

<sup>100</sup> *ibid.*, Art. 2(1).

<sup>101</sup> *ibid.*, Recital 11.

does exclude research organisations providing preferential access to the results of their research to commercial entities.<sup>102</sup>

From a broader and more fundamental perspective, limiting beneficiaries would undermine a widespread assumption that the “right to read should be the right to mine”.<sup>103</sup> From a practical market-based perspective, this policy choice might **cripple opportunities for start-ups and individual researchers** in this area.<sup>104</sup> Indeed, the policy choice of excluding from the reach of the exception unaffiliated individuals and researchers—operating under the same terms as those organised in a qualifying research organisation—might fall short in terms of adequacy and proportionality. Actually, the existing UK exception for text and data analysis includes also individual researchers as beneficiaries and any person with lawful access to a work.<sup>105</sup> In any event, the Impact Assessment should have at least assessed the possibility of extending the limitation to some other defined categories of beneficiaries. For example, using automated analytical techniques in journalistic research to their fullest extent may contribute to solving some of modern media troubles (e.g., costs optimisation, “fake news” phenomenon).

In addition, the new reform package might result in a further bad trade-off for parties excluded from the application of the limitation. Apparently, they might face higher transaction costs for running TDM research as they might also have to clear—in addition to traditional copyright and *sui generis* database right—the **new neighbouring rights of press publishers** that the DSM Draft Directive would like to introduce for online uses.<sup>106</sup> Basically, copying for TDM purposes of online news publications would trigger potential liability for infringement—and thus licencing obligations—also against press publishers for all those parties excluded from the proposed exception’s application.

Limited indirect application of the new exception to private parties is given by Recital 10 of the DSM Draft Directive clarifying that “[r]esearch organisations should also benefit from the exception when they engage into **public-private partnerships**”.<sup>107</sup> Apparently, the recital refers to TDM research carried out by private businesses within the framework of a collaboration with a research organisation, unless the private entity controls the research organisation according to Article 2(1). In any event, public-private partnerships might be a limited option for start-ups as they are “time intensive and nearly impossible to handle for small teams”.<sup>108</sup>

---

<sup>102</sup> *ibid.*, Art. 2(1) and Recital 11.

<sup>103</sup> Peter Murray-Rust, Jennifer Molloy and Diane Cabell (2014), ‘Open Content Mining’, in Samuel A. Moore (ed.), *Issues in Open Research Data*, London, UK: Ubiquity Press, p. 28; Handke, Guibault, and Vallbé (2015), *supra* 72, p. 2; IFLA (2013), Statement on Text and Data Mining, p. 2, available at [https://www.ifla.org/files/assets/clm/statements/iflastatement\\_on\\_text\\_and\\_data\\_mining.pdf](https://www.ifla.org/files/assets/clm/statements/iflastatement_on_text_and_data_mining.pdf).

<sup>104</sup> See, e.g., EUA, CESAER, LERU and Liber (2017), Future-proofing European Research Excellence: A Statement from European Research Organisations on Copyright in the Digital Single Market, 10 January 2017, pp. 1-2 (strongly supporting the principle that “the right to read is the right to mine” meaning that having lawful access to content includes the right to mine that content); Bundesrat (2016), Beschluss 565/16, Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates über das Urheberrecht im digitalen Binnenmarkt, COM(2016) 593 final, Ratsdok 12254/16, § 22 (noting that the narrow scope of the EU TDM exception would put out of business many providers of data analysis in Europe) See also Martin Senftleben (2017), *EU Copyright Reform and Startups: Shedding Light on Potential Threats in the Political Black Box*, March 2017, pp. 5-9, available at: <http://www.innovatorsact.eu/wp-content/uploads/2017/03/Issues-Paper-Copyright-Directive-2.pdf>.

<sup>105</sup> See Copyright, Designs and Patents Act 1988, § 29A (UK).

<sup>106</sup> See Proposal, *supra* note 1, Art. 11(1). See also Geiger, Bulayenko and Frosio (2016), *supra* 91.

<sup>107</sup> Proposal, *supra* note 1, Recital 10.

<sup>108</sup> Allied for Startups, Policy Paper, p. 2, [http://www.innovatorsact.eu/wp-content/uploads/2017/04/AFSOnePagerCopyright\\_0417.pdf](http://www.innovatorsact.eu/wp-content/uploads/2017/04/AFSOnePagerCopyright_0417.pdf).

Last, it is worth noting that the narrow application of the limitation to research organisation does not fully provide the European DSM with the legal framework to **fill the gap with other jurisdictions** adopting opening clauses or fair use models to allow a broader number of research players to perform TDM research and promote related innovation. Actually, given the global nature of the modern economy, the Impact Assessment should have examined the impact of the proposed exception on EU's competitive advantage against other top innovative economies enabling all undertakings to carry out TDM under fair use/fair dealing models (e.g., USA, Canada, Israel).<sup>109</sup>

(2) The Impact Assessment does not provide any rationale for making **lawful access purpose-specific**. This approach disregards a large number of possible applications of TDM that might now be construed—being excluded from the reach of the exception—as within the exclusive rights of the copyright holders. In addition, this purpose-specific approach might raise subtle issues of applicability of the new limitation within research organisations enjoying lawful access to a database. For example, if a public university has lawful access to a database under a “for educational purpose” licence, would it need to pay an additional licensing fee for a “scientific research” purpose? If this is the case, would this obligation contradict the prohibition of contractually overriding the TDM exception?<sup>110</sup> Given the scope of the new limitation—and the “no-contractual-override” provision—the answer is probably not. Still, research institutions might find these possible legal uncertainties a limitation to the deployment of TDM research due to potential liability that might arise and related transaction costs that should be considered before running TDM research projects. Since the exception is already limited to research organisations, dropping restrictions to purpose-specific uses of lawfully accessed databases might avoid unwanted results.

(3) Applying the exception only to works to which the research organisations have “lawful access” would de facto subject TDM research to **private ordering**. According to the European Copyright Society, “the exception can effectively be denied to certain users by a right holder who refuses to grant ‘lawful access’ to works or who grants such access on a conditional basis only”.<sup>111</sup> In addition, subjecting TDM to lawful access will make TDM research projects harder to run by raising related costs. Possibly, publishers might price TDM into their subscription fees, if only those with lawful access can perform TDM research. Subjecting TDM research to market access does **discriminate research according to research organisations’ market power**. Only few research organisations will be able to acquire licences for all databases that are relevant for a TDM research project.<sup>112</sup> This will make comprehensive TDM projects impossible to perform for the majority of research organisations, especially those from Member States with more limited access to funding. In turn, this shall spread the gap between richer and poorer research institutions and, most likely, increase the scientific and innovation divide between developed and less developed European countries. Overall, the quality and value of TDM research will be sub-optimal as budget considerations will constrict the scope of research. Negative externalities of this policy choice on overall global research outputs become even more substantial because those organisations—mainly private—with relevant market power that might run comprehensive TDM projects are excluded from the reach of the exception, which “should [at least] be extended to anybody who has lawful access”.<sup>113</sup>

---

<sup>109</sup> See Geiger, Bulayenko, Hassler, Izyumenko, Schönherr and Seuba (2015), *supra* 70, p. 24.

<sup>110</sup> See Proposal, *supra* note 1, Art. 3(2).

<sup>111</sup> European Copyright Society (2017), *supra* 3, p. 4.

<sup>112</sup> *ibid.*

<sup>113</sup> Commission (2016), *supra* 53, Part. 1/3, p. 109.



[4] The introduction of measures to protect the security and integrity of networks and databases might allow rightholders to block access for researchers trying to conduct TDM.<sup>114</sup> However, Recital 12 spells out clearly that “those measures should not exceed what is necessary to pursue the objective of ensuring the security and integrity of the system and **should not undermine the effective application of the exception**”.<sup>115</sup> For the avoidance of doubt, the same wording should be **included in Article 3(3)** also, rather than only referring to a limitation to measures exceeding their objective. A parallel can be made between these provisions and the safeguards put in place in the context of “traffic management” by telecom operators.<sup>116</sup> According to the Telecoms Single Market Regulation, “traffic management measures” can be applied only in order to comply with EU law and public authorities bound by EU law, preserve the integrity and security of the network, and prevent or mitigate network congestion.<sup>117</sup> Apparently, the scope of the DSM Draft Directive does focus on measures to prevent congestion as Recital 12 provides that the security and integrity measures should be allowed “in view of a potentially high number of access requests to and downloads of their works or other subject-matter”.<sup>118</sup> The application of these measures should be the result of commonly agreed best practices.<sup>119</sup>

[5] It is worth noting that the application of **anti-circumvention provisions** might trample **over users’ privileged uses**. TPMs’ effects on exceptions and limitations have been highlighted by abundant literature.<sup>120</sup> TPMs might limit or prevent altogether access to works for purposes that are not restricted by authors’ rights or for uses that are actually privileged. Rightholders’ obligations to make available the means to benefit from exception and limitations do not themselves limit liability for circumvention.<sup>121</sup> As Guibault et al noted, “for even if Article 6(4) creates an obligation to provide the means to exercise a limitation, this obligation is imposed on rights owners and does not give users any authority to perform act of circumvention themselves”.<sup>122</sup> Also, inconsistent implementations across national jurisdictions of measures to guarantee application of exceptions and limitations against TPMs’ anti-circumvention provisions might effectively curtail harmonised enjoyment of the new mandatory exceptions, thus limiting DSM effectiveness.<sup>123</sup>

---

<sup>114</sup> See EUA, CESAER, LERU and Liber (2017), *supra* 104, p. 2.

<sup>115</sup> Proposal, *supra* note 1, Recital 12 (emphasis added).

<sup>116</sup> Regulation (EU) 2015/2120 of the European Parliament and of the Council of 25 November 2015 laying down measures concerning open internet access and amending Directive 2002/22/EC on universal service and users’ rights relating to electronic communications networks and services and Regulation (EU) No 531/2012 on roaming on public mobile communications networks within the Union, 2015 O.J. (L 310) 1-18, Art. 3(3).

<sup>117</sup> *ibid.*, Art. 3(3)(a-c).

<sup>118</sup> Proposal, *supra* note 1, Recital 12.

<sup>119</sup> *ibid.*, Art. 3(4).

<sup>120</sup> See, e.g., Frosio (2011), *supra* 23, pp. 99-103, 135-141 (discussing most of the relevant literature and major threats that technological protection measures pose for fair dealings, privileged and fair uses); Christophe Geiger (2008), *supra* 55; Séverine Dusollier (2003), ‘Tipping the Scale in Favour of the Right Holders: The European Anti-Circumvention Provisions’, in Eberhard Becker, Willms Buhse, Dirk Günnewig and Niels Rump (eds.), *Digital Rights Management. Technological, Economic, Legal and Political Aspects*, Berlin, Germany: Springer, pp. 462-478.

<sup>121</sup> See Directive 2001/29/EC, *supra* 8, Art. 6(4). See also Common Position No. 48/2000 of 28 September 2000 adopted by the Council, with a view to adopting a Directive of the European Parliament and of the Council on the harmonisation of certain aspects of copyright and related rights in the information society, 2000 O.J. (C 344) 1, 1 December 2000, p. 19 (noting that “Art. 6(1) protects against circumvention of all technological measures designed to prevent or restrict acts not authorized by the rightholder, regardless of whether the person performing the circumvention is a beneficiary of one of the exceptions provided for in Article 5”).

<sup>122</sup> Lucie Guibault et al (2007), *Study on the Implementation and Effect in Member States’ Laws of Directive 2001/29/EC on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society*, report prepared for the European Commission, DG Internal Market, ETD/2005/IM/D1/91.

<sup>123</sup> For example, under the French law only some of the uses permitted under exceptions and limitations are protected to some extent against application of technological protection measures preventing users from taking advantage of them. See Art. L331-31 of the French Intellectual Property Code.

## 6. CONCLUSIONS

The DSM Draft Directive does envision, *inter alia*, a number of synergic actions to promote European development by facilitating research and innovation in the upcoming DSM. The full implementations of these actions—by expanding them even further according to the suggestions included in this briefing—would be critical to European innovation and digital market unification.

In particular, the introduction of a **TDM exception is essential to unlock the potentiality of European research and unburden researchers from legal encumbrances and uncertainties** that make cross-jurisdictional research collaboration harder to promote. In general, a TDM exception will bring about much needed harmonisation in the field, therefore serving well the needs of the DSM.

Although the proposal serves well in general harmonisation and promotion of innovative practices in the DSM, several improvements would be still welcome.

1. Obviously, focus on facilitation of research, teaching and preservation of cultural heritage stands as a primary need for the promotion of the DSM. However, this reform should be an opportunity also to reflect on the future **design of an "opening clause"** to be added to the list of exempted uses to address situations that are not yet covered by existing exceptions and limitations but are justified by important public interest rationales and fundamental rights such as freedom of expression and the right to information. In particular, TDM research and innovation would profit substantially from such an opening clause.

Understood that the introduction of an opening clause would serve well innovation in the DSM, specific ameliorations of the TDM limitations currently drafted should be also carefully pondered. They reflect, *inter alia*, some of the points made above in connection with possible negative impacts of the proposed reform, to which we remand for further clarifications, if necessary.

2. For the avoidance of doubt, Recital 8 of the Draft Directive should mention that also **works** and other subject matter **not protected** by copyright or neighbouring rights **can be freely mined**.
3. The proposed reform does not define the **notion of "lawful access"**.<sup>124</sup> However, the existing exception Article 5(1) of Directive 2001/29/EC that might cover some TDM techniques involving temporary reproduction refers to "lawful use", which has been defined by the Directive and unambiguously interpreted by the CJEU.<sup>125</sup> If "lawful access" is intended to mean what **"lawful use"** means, the reform should **maintain the term already adopted in EU law** in a provision already covering some TDM techniques.

---

<sup>124</sup> See Proposal, *supra* note 1, Recital 9 and Art. 3(1).

<sup>125</sup> See Directive 2001/29/EC, *supra* 8, Recital 33 (noting that "[a] use should be considered lawful where it is authorised by the rightholder or not restricted by law"). See also CJEU, C-403/08 and C-429/08, Football Association Premier League and Others (4 October 2011), ECLI:EU:C:2011:631, § 168; CJEU, C-302/10, Order, Infopaq International (17 January 2012), ECLI:EU:C:2012:16, § 42.

4. The notion of **“normal use” of a database** might receive multiple interpretations according to the Member States in which it is applied, therefore **harmonisation** should be pursued **by including TDM** in the EU law notion of “normal use”.
5. The **TDM exception** should **not be limited to research organisations but extended to all** those **enjoying lawful access** to underlying mined materials—as the right to read should be the right to mine—especially in order not to cripple research from start-ups and independent researchers. In any event, the possibility of extending the exception to some **other defined categories of beneficiaries**, such as journalists, should be **at least** assessed.
6. In case of TDM for commercial uses carried out by some entities (e.g. other than research organizations or, possibly, individual researchers and journalists), fair remuneration could be considered provided that harm can be demonstrated on the basis of relevant empirical data.
7. Since the exception is already limited to research organisations, **dropping restrictions to purpose-specific uses** of lawfully accessed databases might avoid unwanted results.
8. In light of the increasing **research focus on the quality and verifiability**, a TDM exception should enable **storing and communication of research files** created for TDM.
9. As technological enclosure might cripple TDM research, the proposal should make clear that the exception would also be **protected from override by TPMs**, by plainly stating that TPMs cannot prevent the enjoyment of the new mandatory exception and that effective means should be made available to users to secure their removal.
10. For the avoidance of doubt, the same wording included in Recital 12 of the DSM Draft Directive should be also included in Article 3(3) by plainly stating that **security and integrity measures** should **not “undermine the effective application of the exception”**, rather than only referring to a limitation to measures exceeding their objective.
11. A provision limiting **contractual and technological override** should be **extended to any exceptions potentially covering TDM**, including for example the TDM techniques covered by Article 5(1), Directive 2001/29/EC.
12. In addition, protection against contractual and technological override should also be always extended to TDM mining materials not protected by IPRs, including those made available in a database.

Aptly, public interest and access to European knowledge in the DSM represents a critical focus of the upcoming copyright reform. The relevant institutions should not depart from this agenda in the path leading to final implementation. They should rather strengthen it as far as possible, especially by also introducing in addition to the list of specific limitations an opening clause that might serve TDM and other purposes. An opening clause might address uses that are not yet covered by existing exceptions and limitations but are justified by important public interest rationales and fundamental rights such as freedom of expression and the right to information.



## **Abstract**

This in-depth analysis, commissioned by the European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs at the request of the Committee on Legal Affairs (JURI-Committee), is a contribution to the workshop on "Text and data mining" held on 22 February 2018 in Brussels. It provides an analysis of the Commission's Proposal (which introduces in Article 3 a mandatory exception to copyright allowing to carry out text and data mining of protected works), assesses its positive and negative impacts and provides some suggestions for possible improvements. Advantages of introducing an "open clause" on top of an enumerated list of exceptions to address some of the related problems are also reviewed.

## **DISCLAIMER**

This document is addressed to the Members and staff of the European Parliament to assist them in their parliamentary work. The content of the document is the sole responsibility of its author(s) and should not be taken to represent an official position of the European Parliament.