



UvA-DARE (Digital Academic Repository)

Study on EU copyright and related rights and access to and reuse of data

Senftleben, M.R.F.

DOI

[10.2777/78973](https://doi.org/10.2777/78973)

Publication date

2022

Document Version

Final published version

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Senftleben, M. R. F. (2022). *Study on EU copyright and related rights and access to and reuse of data*. (European Commission. Research and Innovation. Independent Expert Report). Publications Office of the European Union. <https://doi.org/10.2777/78973>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Study on EU copyright and related rights and access to and reuse of data

**Independent
Expert
Report**

Written by Martin R.F. Senftleben

*Research and
Innovation*

Study on EU copyright and related rights and access to and reuse of data

European Commission
Directorate-General for Research and Innovation
Directorate A — ERA & Innovation
Unit A.4 — Open Science
Contact Alea López de San Román
Email Alea.lopez-de-san-roman@ec.europa.eu
RTD-OPEN-SCIENCE@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu

European Commission
B-1049 Brussels

Manuscript completed in March 2022

This document has been prepared for the European Commission, however it reflects the views only of the authors, and the European Commission shall not be liable for any consequence stemming from the reuse.

More information on the European Union is available on the internet (<http://europa.eu>).

PDF	ISBN 978-92-76-53632-1	doi:10.2777/78973	KI-08-22-205-EN-N
-----	------------------------	-------------------	-------------------

Luxembourg: Publications Office of the European Union, 2022

© European Union, 2022



The reuse policy of European Commission documents is implemented by Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Unless otherwise noted, the reuse of this document is authorised under a Creative Commons Attribution 4.0 International (CC-BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated. For any use or reproduction of elements that are not owned by the European Union, permission may need to be sought directly from the respective rightholders.

Study on EU copyright and related rights and access to and reuse of data

Table of Contents

EXECUTIVE SUMMARY	2
List of abbreviations	7
1 Introduction	8
1.1 Problem Statement	8
1.2 Steps of the Analysis	8
1.3 Methodology	8
2 Importance of an Adequate Copyright Framework	8
2.1 Central Dimensions of Data Access and Reuse	9
2.2 Concept of "Data"	9
2.3 Need to Reconcile Different Fundamental Rights	11
2.4 Impact of Copyright Law	13
2.5 Meeting Right Holders at Eye Level	14
3 Impact of Copyright Protection	15
3.1 Information Society Directive	16
3.1.1 Potential Illustration Requirement	17
3.1.2 Lack of Harmonisation	18
3.1.3 Requirement of Non-commercial Use	19
3.1.4 Three-Step Test Compliance	21
3.1.5 Technological Protection Measures	25
3.1.6 Contractual Restrictions	26
3.1.7 Other Exceptions and Limitations	27
3.1.8 Considerable Legal Uncertainty	28
3.2 Database Directive	29
3.2.1 Data Sharing	30
3.2.2 Scope of Exclusive Rights	33
3.2.3 Mixed Results	36
3.3 Digital Single Market Directive	36
3.3.1 No Requirement of Non-Commercial Use	38
3.3.2 Technological Protection Measures	41
3.3.3 Requirement of Lawful Access	43
3.3.4 Data Sharing	45
3.3.5 Unanswered Questions	47
3.4 Data Act Proposal	47
3.4.1 Exclusion of Sui Generis Database Protection	48
3.4.2 Raw Data Access Regimes	50
3.4.3 Lack of Research Focus	52
4. Publicly Funded Research Data	52
4.1 Copyright Data Harmonisation	53
4.1.1 Experiences in the Music Industry	54
4.1.2 Initiatives in Other Creative Industry Segments	56
4.1.3 Different Settings for Data Improvement	57
4.2 Research Data Licensing	58
4.2.1 Horizon Europe Open Access Requirements	59
4.2.2 Impact of Copyright and Sui Generis Database Protection	61
4.2.3 Impulses for Copyright Data Improvement	62
5. Conclusion	63
5.1 Imbalance	63
5.2 Recommendations for Legislative Measures	66
5.3 Recommendations for Non-legislative Measures	68

Research conducted between December 2021 and March 2022 for the European Commission, Directorate-General for Research and Innovation (DG RTD), Open Science Unit.

Martin R. F. Senftleben is Professor of Intellectual Property Law and Director of the Institute for Information Law (IViR) at the Amsterdam Law School.

EXECUTIVE SUMMARY

To safeguard freedom of expression and information, and the freedom of sciences, of researchers,¹ **it is important to improve the legal framework for scientific research in copyright, related rights and sui generis database law.** In particular, it is important to remove imbalances that pose obstacles to data access and reuse. **Article 5(3)(a) of the Information Society Directive could serve as a reference point for this legislative step.** To offer researchers a more robust and reliable legal position, it is advisable to:

- **clarify that the requirement of use as an “illustration” only concerns the teaching branch** of the use privilege and does not relate to scientific research;
- **abandon the requirement of use for a “non-commercial purpose”** and, instead, follow the approach taken in Article 3(1) of the Directive on Copyright in the Digital Single Market which, rightly understood, offers more room for public-private partnerships and more opportunities to translate research insights into products and services that can be brought to the market;
- **recalibrate the determination of lawful access.** Instead of requiring access permissions of each individual institution participating in a research project, it should be sufficient that one participating institution has lawful access;
- **clarify that,** regardless of the volume of use, **scientific research constitutes a “special case” in the sense of the three-step test** of Article 5(5) of the Information Society Directive because of the fundamental rights underpinning following from Articles 11(1) and 13 of the Charter of Fundamental Rights;
- **clarify that, in the assessment of a conflict with a normal exploitation or an unreasonable prejudice to legitimate interests of right holders** under Article 5(5), **it is necessary to take benefits into account** which right holders, such as academic publishers, derive from the work of researchers and the results of scientific research projects;
- **grant researchers the right to circumvent technological protection measures** in case right holders fail to ensure that the use privilege for scientific research remains effective when technological protection measures are applied;
- **declare Article 6(4), subparagraph 4, of the Information Society Directive inapplicable to use for the purposes of scientific research,** as already done in Article 7(2) of the Directive on Copyright in the Digital Single Market;
- **declare any contractual provision contrary to the use privilege for scientific research unenforceable,** as already stated in Article 7(1) of the Directive on Copyright in the Digital Single Market.

In contrast to the current, optional version of Article 5(3)(a) in the Information Society Directive, **this more flexible and more robust exemption of use for scientific**

¹ Articles 11(1) and 13 of the EU Charter of Fundamental Rights.

research should constitute a mandatory “shall” provision to ensure a harmonised application across Member States and comparable conditions for research teams in different countries.

The proposed more flexible and more robust exemption of research use can be expected to play a crucial role in the realisation of EU open science objectives. As the proposed broadened and strengthened version of Article 5(3)(a) would cover both – the right of making copies for research purposes (reproduction) and the right of sharing these copies (communication and making available to the public) – the provision has the potential to enable researchers to comply with open access requirements of funding schemes for scientific research, such as Horizon Europe. **With the proposed broadened and strengthened provision, copyright protection would impose less constraints on initiatives to make research data, including copyrighted material, available open access.**

To attain the described goals – an equal legal position for researchers in line with underlying fundamental rights and less barriers to open access availability of research data – **it is advisable to implement the proposed more flexible and more robust use privilege for scientific research not only in the field of copyright and related rights but also in the area of the sui generis database right.** Researchers should be able to rely on a corresponding use privileges with a congruent scope in Article 9(b) of the Database Directive.

With regard to non-legislative measures, it is important to note that **some of the aforementioned recommendations for legislative measures can also serve as an impulse for non-legislative clarifications and best practice models:**

- with regard to the overarching research exceptions in Article 5(3)(a) ISD and Article 9(b) DBD, it could be clarified that:
 - o **the requirement of use as an “illustration” only concerns the teaching branch** of the use privilege and does not relate to scientific research;
 - o regardless of the volume of use, **scientific research constitutes a “special case” in the sense of the three-step test** of Article 5(5) ISD because of the fundamental rights underpinning following from Articles 11(1) and 13 CFR;
 - o **in the assessment of a conflict with a normal exploitation or an unreasonable prejudice to legitimate interests of right holders** under Article 5(5) ISD, **it is necessary to take benefits into account** which right holders, such as academic publishers, derive from the work of researchers and the results of scientific research projects (cf. section 3.1.4);
- with regard to the specific TDM exception in Article 3 DSMD and, more specifically, the lawful access guidelines in Recital 14 DSMD, it could be clarified that:
 - o in the case of subscriptions, the persons attached to a research organisation or cultural heritage institution with the subscription are not

the only group of beneficiaries, in respect of which lawful access can be assumed (the relevant sentence in Recital 14 starts with “[f]or instance”). Rightly understood, lawful access should also be assumed with regard to researchers from other organisations or institutions in the case of joint research projects. It should thus be deemed **sufficient that one participating institution has lawful access** (cf. section 3.3.3);

- Member States should use Article 5(3)(a) ISD as a basis to complement Article 3 DSMD with a further copyright exception that permits the sharing of TDM datasets within research consortia and, for purposes such as research validation, also with the broader academic community. To ensure a harmonised approach, it seems advisable to **develop a model provision for TDM dataset sharing** that can be implemented in a uniform manner in different national contexts.

Moreover, **it seems promising to explore the intersection between** open access and related data and metadata initiatives in the academic world, and the need to improve copyright data and data management infrastructures in the creative sector. The interplay has two central aspects:

- on the one hand, an improved **copyright data infrastructure** makes it easier for researchers to obtain use permissions that are required when statutory use privileges for scientific research are inapplicable. Hence, data improvement initiatives are likely to enhance data availability for research purposes;
- on the other hand, **open access obligations in academic funding schemes** that include the sharing of datasets and the creation of corresponding, sufficiently rich, standardised and machine-actionable metadata have the potential to support data improvement strategies in the creative industries.

Given this interrelation, **it seems advisable to develop non-legislative initiatives that pave the way for the injection of copyright-related data and metadata that result from research projects into data improvement processes in the creative sector.** In exchange for valuable contributions of researchers to the improvement of copyright data, including the creation of rich metadata, right holders may be willing to offer broader support for academic initiatives that seek to ensure open access to research data that include protected works and (parts of) databases. In particular, this may be an attractive option for the creative industry if it proves to be possible to draw a boundary between open access data and metadata systems for research purposes and closed data and metadata systems that use metadata strategically as a source of information to enhance the visibility and findability of works, databases and right holders, and create new and broader licensing opportunities.

List of abbreviations

BC	Berne Convention for the Protection of Literary and Artistic Works
CFR	Charter of Fundamental Rights of the European Union (<i>OJ</i> 2000 C 364, p. 1)
CJEU	Court of Justice of the European Union
DAP	European Commission, 23 February 2022, <i>Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act)</i> , COM(2022) 68 final
DBD	Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases (<i>OJ</i> 1996 L 77, p. 20)
DSMD	Directive 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC (<i>OJ</i> 2019 L 130, p. 92)
ECLI	European Case Law Identifier
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EU	European Union
ISD	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 (<i>OJ</i> 2001 L 167, p. 10)
TDM	Text and data mining
TRIPS	Agreement on Trade-Related Aspects of Intellectual Property Rights, Annex 1C of the Marrakesh Agreement Establishing the World Trade Organization, signed on 15 April 1994, <i>ILM</i> 33 (1994), 1197
WCT	WIPO Copyright Treaty, adopted in Geneva on 20 December 1996
WIPO	World Intellectual Property Organization
WTO	World Trade Organization

1 Introduction

1.1 Problem Statement

EU legislation in the field of copyright, related rights and sui generis database rights (besides individual references to these fields of law in the following analysis, the term “copyright” will be used as an umbrella term referring to all three areas of legislation), **can have a deep impact on access to data resources for scientific research and the availability of data resulting from publicly funded research.** To establish a copyright and related rights framework that offers appropriate data access and reuse opportunities for scientific research, it is necessary to identify potential barriers and challenges that may arise from EU copyright and related rights legislation and corresponding rights management.

Against this background, **the following study will analyse the interaction between copyright and related rights law and data access and reuse for scientific research purposes.** It will **propose legislative and non-legislative measures to improve the current EU regulatory framework.** In this way, it aims to provide input for the Commission’s policy development on open science and, in particular, data access and reuse in the context of scientific research, including access to, and reuse of, publicly funded research data.

1.2 Steps of the Analysis

To attain the described objectives, **the analysis will be carried out in four steps:**

- the following **chapter 2** discusses intersections between copyright, related rights and sui generis database rights, and data access and reuse in the context of scientific research projects. It will shed light on the importance of an adequate regulatory framework for data access and reuse within copyright, related rights and sui generis database legislation;
- **chapter 3** will offer a detailed analysis of the impact which applicable provisions of EU copyright, related rights and sui generis database law may have on data access and reuse for scientific purposes. The analysis seeks to identify elements that may foster or deter data access and reuse;
- with regard to publicly funded research data, **chapter 4** provides an analysis of licensing initiatives that seek to foster access and reuse. The discussion will also address copyright data issues that may become relevant in this context;
- in **chapter 5**, the insights from the analysis will be used to develop concrete actionable recommendations for legislative and non-legislative measures and initiatives to improve the legal environment for data access and reuse, including publicly funded research data, for scientific purposes.

1.3 Methodology

The following analysis is the result of desk research, in particular based on the studying of scientific publications and commentary literature, **case law at the European and national level, legislative texts and policy documents.**

2 Importance of an Adequate Copyright Framework

To lay groundwork for the following, more detailed analysis of applicable norms in copyright, related rights and sui generis database law (chapter 3), **it seems advisable to clarify the different dimensions of data access and reuse – input and output – in which copyright issues can arise** (following section 2.1). **Moreover, the concept**

of “data” should be clarified (section 2.2) **before turning to a discussion of the primary fundamental rights framework** that provides binding guidance for the configuration of copyright norms at the lower level of secondary EU legislation (section 2.3). **After the exploration of the overarching fundamental rights framework, the impact of copyright norms can be brought to light** (section 2.4). Concluding remarks provide an overview of the insights resulting from the analysis (section 2.5).

2.1 Central Dimensions of Data Access and Reuse

As the reference to “access” and “reuse” issues in the problem statement for this study already indicates, **data use for scientific purposes may raise copyright-related questions with regard to two central dimensions:**

- **the *input* dimension concerns the data resources needed to establish the datasets that are necessary for the envisaged research.** If relevant data sources enjoy copyright, related rights or sui generis database protection, the input (or access) dimension raises the question whether statutory use permissions allow researchers to amass the required data without prior authorisation of right holders. If no statutory exemption of research use is applicable, the required data can only be obtained on the basis of rights clearance. This, however, can imply considerable transaction costs and require an additional budget for the payment of licensing fees (chapter 3);
- **the *output* dimension concerns research results, ranging from scientific publications to datasets and related metadata, that have been established for the research project or that evolve from the research work.** Obviously, these research results may attract copyright, related rights or sui generis database protection themselves. Hence, the output (or reuse) dimension also raises copyright issues. Future research teams seeking to rely on protected, pre-existing research results face the same legal issues that surround the data access and input dimension. Again, researchers must navigate between use permissions that follow from statutory use privileges for scientific research in the EU copyright acquis, and rights clearance obligations that may require considerable time investment and enhance costs substantially (chapter 4).

2.2 Concept of “Data”

Before embarking on a more detailed analysis of these issues, **it is important to clarify the concept of “data” that underlies the following examination.** Considering the focus on the legal framework that is inherent in the research questions underlying this study, **it seems appropriate to adopt the definition of “data” that can be found in the Proposal for a Data Act (“DAP”)** published on 23 February 2022.² According to Article 2(1) DAP, “data” means:

any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audio-visual recording.

The reference to “sound, visual or audio-visual recording” and, more generally, “representation” indicates that **this concept of “data” goes beyond mere raw data.**

² European Commission, 23 February 2022, *Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act)*, COM(2022) 68 final.

It includes data included in information products that enjoy copyright protection, such as publications, music recordings and films representing acts, facts or information. In addition, **the reference to “any compilation” opens the door for the inclusion of databases that present collected acts, facts or information in a structured form.**

In the light of these conceptual contours, **three different categories of data can be distinguished in the context of the following analysis of access and reuse issues:**

- **raw data:** these are data in machine-generated form. Data stemming from “smart” devices that form part of the so-called “Internet of Things” can serve as examples. Smart cars – not only self-driving cars but also commonly available cars that implement advanced safety tools – smart fridges, heating systems, GPS devices, smartphone apps, automatically generate data relating to use patterns and parameters. Current EU copyright rules usually exclude copyright protection for this type of data on the ground that the free, creative choices necessary for copyright eligibility³ are missing. With regard to potential sui generis database rights, Article 35 DAP seeks to exclude protection (for a more detailed discussion of this point, see section 3.4.1 below). Under certain conditions, however, protection for raw data may follow from the application of proprietary or quasi-proprietary legal instruments, in particular trade secret law⁴ and contractual obligations to observe access and use conditions;
- **data after substantial investment:** these are data that have been included in a database as a result of substantial investment in either the obtaining, verification or presentation of contents. The 1996 Database Directive (“DBD”)⁵ affords sui generis database protection against unauthorized acts of data extraction and re-utilization⁶ once the requirement of substantial investment is fulfilled.⁷ The exclusive rights of extraction (acts of transferring database contents)⁸ and re-utilization (acts of making available database contents)⁹ offer the database maker the opportunity to exert control over the taking of data resources that are contained in a protected database;¹⁰
- **data in literary and artistic works:** this is the case of acts, facts or information that, as such, are often not protected. Copyright protection does not extend to mere facts, ideas, styles, concepts etc.¹¹ Copyright only protects the individual expression which the author has chosen to give facts, ideas, styles, concepts etc. a specific form. However, due to the fact that relevant data are contained in a protected work of authorship – being it a literary, musical, artistic or audio-visual work reflecting free, creative choices¹² – access and reuse may be conditioned by

³ CJEU, 1 March 2012, case C-604/10, *Football Dataco/Yahoo!*, para. 38; CJEU, 1 December 2011, case C-145/10, *Painer*, para. 89; CJEU, 16 July 2009, case C-5/08, *Infopaq/DDF*, para. 45.

⁴ Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, *OJ* 2016 L 157, 1. Cf. R. Ducato/A. Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out”, *European Intellectual Property Review* 43 (2021), 322 (336).

⁵ Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996, on the Legal Protection of Databases, *OJ* 1996 L 77, 20.

⁶ For a definition of these modes of use, see Article 7(2) DBD.

⁷ The threshold for assuming a substantial investment need not be particularly high. For instance, see CJEU, 9 October 2008, case C-304/07, *Directmedia/Universität Freiburg*, para. 24, where an investment of 34.900€ had been deemed sufficient by the national judge.

⁸ Article 7(2)(a) DBD.

⁹ Article 7(2)(b) DBD.

¹⁰ Article 7(1) and (5) DBD.

¹¹ Article 2(8) BC; Article 9(2) TRIPS; Article 2 WCT. Cf. Dutch Supreme Court (*Hoge Raad*), 29 March 2013, case LJN BY8661, *Broeren/Duijsens*, *Tijdschrift voor auteurs-, media- en informatierecht* 2013, 108.

¹² CJEU, 1 March 2012, case C-604/10, *Football Dataco/Yahoo!*, para. 38; CJEU, 1 December 2011, case C-145/10, *Painer*, para. 89; CJEU, 16 July 2009, case C-5/08, *Infopaq/DDF*, para. 45.

the copyright holder's exclusive right to control the reproduction, communication to the public and making available of the work following from Articles 2 and 3 of the Information Society Directive ("ISD").¹³ Illustrative examples of relevant acts, facts or information are not only the insights that can be deduced from the contents of the work as such but also more abstract insights, such as syntactic and statistical information about language use in novels and newspapers, or physiognomy information contained in videos and photographs that may be used for the training of face recognition systems.

2.3 Need to Reconcile Different Fundamental Rights

The discussion of different forms of data already indicates that specific access and reuse questions can arise when scientific research requires the use of acts, facts or information enshrined in a database or a literary and artistic work that enjoys protection under EU database or copyright law.¹⁴ At the level of fundamental rights, **the tension between use interests of researchers** on the one hand, **and protection interests of holders of copyright and sui generis database rights** on the other, **can be described as a tension between different legal positions that have been recognized in the EU Charter of Fundamental Rights** ("CFR").

In the context of the right to property – the right to "own, use, dispose of and bequeath his or her lawfully acquired possessions"¹⁵ – Article 17(2) CFR clarifies that "[i]ntellectual property shall be protected."¹⁶ At the level of fundamental rights, **the exclusive rights granted in EU intellectual property law, including copyright and database legislation, thus fall within the scope of the right to property. This legal position, however, is not absolute.**¹⁷ It follows from Article 52(1) CFR that the right to property may be limited as long as the limitation is provided for by law and leaves the essence of the right intact. Subject to the principle of proportionality, limitations are possible if they are "necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others."¹⁸

More specifically, **the CJEU has pointed out in the context of cases concerning freedom of expression and information that a "fair balance" must be found** between "the rights and interests of authors on the one hand, and the rights of users of

¹³ 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, *OJ* 2001 L 167, 10.

¹⁴ Cf. R. Ducato/A. Strowel, "Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out", *European Intellectual Property Review* 43 (2021), 322 (323); P.B. Hugenholtz, "Artikelen 3 en 4 DSM-richtlijn: tekst- en datamining", *Tijdschrift voor auteurs-, media- en informatierecht* 2019, 167 (168-169); C. Geiger/G. Frosio/O. Bulayenko, "Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations", *International Review of Intellectual Property and Competition Law* 2018, 814 (817-819).

¹⁵ Article 17(1) CFR.

¹⁶ For a more detailed discussion of this global statement, see D.J.W. Jongsma, *Creating EU Copyright Law – Striking a Fair Balance*, Helsinki: Hanken School of Economics 2019, 163-168; J. Griffiths/L. McDonagh, 'Fundamental Rights and European IP Law – the Case of Art 17(2) of the EU Charter', in: C. Geiger (ed.), *Constructing European Intellectual Property Achievements and New Perspectives*, Cheltenham: Edward Elgar 2013, 75; C. Geiger, 'Intellectual Property Shall be Protected!? Article 17(2) of the Charter of Fundamental Rights of the European Union: A Mysterious Provision With an Unclear Scope', *European Intellectual Property Review* 31 (2009), 113.

¹⁷ CJEU, 29 July 2019, case C-469/17, *Funke Medien NRW*, para. 72.

¹⁸ Article 52(1) CFR.

protected subject-matter on the other.”¹⁹ Referring to user “rights”, the Court clarified that the legal position of users invoking freedom of expression and information was not a priori weaker than the protection status which a holder of copyright or database rights enjoys by virtue of EU law. **In *Funke Medien and Spiegel Online*, the Court confirmed that copyright exceptions and limitations serving freedom of expression and information “do themselves confer rights on the users of works or of other subject matter.”**²⁰

With regard to researchers seeking to make use of data resources that enjoy copyright or sui generis database protection, this case law is of particular relevance. **At the level of fundamental rights, a “right to research” can be derived from the guarantee of freedom of expression and, more specifically, the freedom of information** set forth in Article 11(1) CFR:

Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers.

Using the terminology of **Article 11(1) CFR**, it can be said that researchers requiring data access to arrive at new insights in the context of research projects exercise their right to “receive and impart information and ideas.” Sharing research outcomes, they exercise their freedom to “hold opinions” and their freedom of expression more generally. In addition to the recognition of freedom of expression and information in Article 11(1) CFR, **Article 13 CFR** stipulates that “scientific research shall be free of constraint.” The provision also underlines that “[a]cademic freedom shall be respected.” Taking the fundamental provisions relating to freedom of expression, information and science together, it seems safe to assume that EU legislation is under an obligation to create a favourable, enabling environment for scientific research, and to weigh this fundamental value against other objectives, such as the desire to protect the right to property, including intellectual property.²¹

In other words, there can be little doubt in the light of the described legal provisions and case law that **EU legislation must strike a proper balance between the right to property of copyright and database owners (Article 17(2) CFR), and the right to research which researchers can invoke** whose scientific projects depend on the use of protected data resources (Articles 11(1) and 13 CFR).

¹⁹ CJEU, 1 December 2011, case C-145/10, *Eva Maria Painer/Standard VerlagsGmbH*, para. 132; CJEU, 3 September 2014, case C-201/13, *Deckmyn*, para. 26; CJEU, 29 July 2019, case C-469/17, *Funke Medien NRW*, para. 67-76.

²⁰ CJEU, 29 July 2019, case C-516/17, *Spiegel Online*, para. 54; CJEU, 29 July 2019, case C-469/17, *Funke Medien NRW*, para. 70. For a more detailed discussion of this point, see T. Aplin/L. Bently, *Global Mandatory Fair Use: The Nature and Scope of the Right to Quote Copyright Works*, Cambridge: Cambridge University Press 2020, 75-84; C. Geiger/E. Izyumenko, ‘The Constitutionalization of Intellectual Property Law in the EU and the *Funke Medien*, *Pelham* and *Spiegel Online* Decisions of the CJEU: Progress, but Still Some Way to Go!’, *IIC* 51 (2020), 282 (292-298).

²¹ For a more detailed discussion of the right to research in the Charter of Fundamental Rights, see C. Geiger/B.J. Jütte, ‘The Right to Research as Guarantor for Sustainability, Innovation and Justice in EU Copyright Law’, in: T. Pihlajarinne/J. Mähönen/P. Upreti (eds.), *Rethinking the Role of Intellectual Property Rights in the Post-Pandemic World: An Integrated Framework of Sustainability, Innovation and Global Justice*, Cheltenham: Edward Elgar 2022, forthcoming; C. Geiger/B.J. Jütte, ‘Conceptualizing the Right to Research and its Implications for Copyright Law, An International and European Perspective’, *American University International Law Review* 38 (2022), forthcoming.

2.4 Impact of Copyright Law

Interestingly, **the CJEU has also explained in which way a proper balance between copyright protection and other fundamental rights and freedoms can be established.** This further guideline is of particular interest because it clarifies the impact of copyright law and underlines the necessity to devise an adequate regulatory framework for data access and reuse within copyright, related rights and sui generis database legislation.

In Pelham – a case about the creative reuse of music fragments in the context of sound sampling²² – the CJEU examined how a fair balance could be established between the property rights of copyright and related rights holders and conflicting fundamental rights of users, such as freedom of expression. **The CJEU emphasised that the required balance had to be struck within the system of exclusive rights and limitations in EU copyright law:**

[t]he mechanisms allowing those different rights and interests to be balanced are contained in Directive 2001/29 itself, in that it provides inter alia, first, in Articles 2 to 4 thereof, rightholders with exclusive rights and, second, in Article 5 thereof, for exceptions and limitations to those rights which may, or even must, be transposed by the Member States...²³

Therefore, an appropriate balance must be found *within* the existing EU framework for the protection of copyright and related rights. **It seems plausible that the CJEU will adopt the same approach in the area of sui generis database rights.** According to the CJEU, there is sufficient room within the copyright framework to safeguard the fundamental rights of users, such as the freedom of expression and the freedom of art in the case of sound sampling.

With regard to the use of protected material in the context of scientific research, **it can be derived from this judgment that the room for guaranteeing the freedom of expression, the freedom of information and the freedom of science of researchers must be found within the copyright system of exclusive rights on the one hand, and exceptions and limitations on the other.**²⁴ According to the CJEU, it is not the intention to step outside the copyright framework and override the copyright system by exempting use from the control of copyright holders on the basis of a direct reference to fundamental rights in the Charter.²⁵ Otherwise, the harmonisation

²² CJEU, 29 July 2019, case C-476/17, Pelham. For a more detailed discussion of the case, see M.R.F. Senftleben, 'Flexibility Grave – Partial Reproduction Focus and Closed System Fetishism in CJEU, Pelham', *International Review of Intellectual Property and Competition Law* 51 (2020), 751-769.

²³ CJEU, id., para. 60. Cf. CJEU, 29 July 2019, case C-469/17, Funke Medien NRW, para. 58, where the Court uses the same formula.

²⁴ Cf. ECtHR, 10 January 2013, case 36769/08, Ashby Donald/Frankrijk, para. 38, and C. Geiger/E. Izyumenko, 'Copyright on the Human Rights' Trial: Redefining the Boundaries of Exclusivity Through Freedom of Expression', *International Review of Intellectual Property and Competition Law* 2014, 316, with regard to the necessity to offer room for freedom of expression and information.

²⁵ For a discussion of the different forms of so-called "internal" and "external" balancing in the field of intellectual property, see M.R.F. Senftleben, *The Copyright/Trademark Interface – How the Expansion of Trademark Protection is Stifling Cultural Creativity*, Alphen aan den Rijn: Wolters Kluwer 2020, 500-504; S. Kulk/P. Teunissen, 'Naar een nieuw fundament – hoe het Handvest het auteursrecht hervormt (deel 1)', *Tijdschrift voor auteurs-, media- en informatierecht* 2019, 121 (126-129); T. Dreier, 'Balancing Proprietary and Public Domain Interests: Inside or Outside of Proprietary Rights?', in: R.C. Dreyfuss, D. Leenheer-Zimmerman, H. First (red.), *Expanding the Boundaries of Intellectual Property. Innovation Policy for the Knowledge Economy*, Oxford: Oxford University Press 2001, 295.

objectives underlying the European harmonisation of copyright law would be thwarted and the internal market could be disrupted.²⁶

It is therefore indispensable – despite the desirability of data access and use for scientific research from the perspective of freedom of expression, freedom of information and the freedom of science, as recognised in the Charter – **to further investigate what scope more specific legislation**, in particular EU legislation in the field of copyright, related rights and sui generis database rights, **provides to enable data use**. An adequate copyright framework is thus of central importance. The room for data access and use must be found within the system of copyright, related rights and sui generis database protection. Accordingly, **the next step in the analysis is the examination of the applicable copyright rules. As the CJEU has pointed out in *Pelham*, exceptions and limitations of protection are of particular relevance.**²⁷

2.5 Meeting Right Holders at Eye Level

At the level of fundamental rights, **holders of copyright, related rights and sui generis database rights can support their legal position by invoking the recognition of the right to property, including intellectual property, in Article 17 CFR**. With regard to the use of data in scientific research, this fundamental rights position is of particular importance because data – understood broadly as “acts, facts or information”²⁸ – may be embedded in protected databases or literary and artistic works. The possibility of obtaining access and deducing relevant data from protected source material, thus, depends on an authorisation given by the right holder or a statutory permission following from copyright or database legislation.

In the quest for access to data and freedom of use, however, researchers do not stand empty-handed. Surveying the canon of fundamental rights, **they can rely on freedom of expression and information (Article 11 CFR) and the freedom of science (Article 13 CFR) to undergird a right to research that includes the freedom to use protected data resources.** Considering the fundamental freedoms reflected in Articles 11 and 13 CFR, there can be little doubt that EU legislation is under an obligation to create a favourable, enabling environment for scientific research. **In line with the jurisprudence of the CJEU, this means that the legislator must establish a “fair balance”** between the rights and interests of the holders of copyright, related rights and database rights on the one hand, and the rights and interests of researchers on the other.

However, the CJEU added that the required balance must be found within the specific system of protection that follows from copyright, related rights and database legislation. More specifically, **the exceptions and limitations to the exclusive rights granted in copyright and database law must be employed to arrive at a proper balance. This approach taken by the Court appears problematic in the light of the described interplay of fundamental rights and freedoms.** As explained, both parties – right holders and researchers – can invoke fundamental rights to support their legal positions. In principle, right holders and researchers, thus, meet at eye level. **There is no hierarchy between fundamental rights.** One fundamental right is not weaker than the other. The right to property does not automatically prevail over freedom of expression, freedom of information and freedom of science. By contrast, they have the same status.

²⁶ CJEU, 29 July 2019, case C-476/17, *Pelham*, para. 63.

²⁷ CJEU, *id.*, para. 60.

²⁸ Article 2(1) DAP.

Against this background, **it is inconsistent to assume that protection of copyright is the rule and freedom of research is the exception.** Considering the equal status of the right to property (Article 17 CFR) and the freedom of expression, information and science (Articles 11 and 13 CFR), holders of copyright, related rights and database rights cannot expect to enjoy a legal position which, by definition, has more weight. Researchers must not be forced into a weak position by obliging them to defend data access and use activities on the basis of exceptions and limitations that may be construed restrictively. In the norm hierarchy, the fundamental rights and freedoms recognised in the Charter constitute primary sources of law. The legislation in the field of copyright, related rights and database rights, by contrast, is secondary law. The rule/exception relationship following from copyright legislation, therefore, can hardly be deemed decisive. **The CJEU acknowledged this in *Funke Medien and Spiegel Online* where it pointed out that copyright exceptions and limitations “do themselves confer rights on the users of works or of other subject matter.”**²⁹

3 Impact of Copyright Protection

As explained in the preceding chapter, data relevant to scientific research – “acts, facts or information”³⁰ – may be contained in databases or literary and artistic works enjoying protection under EU copyright and/or sui generis database law. **The status of a protected work or database offers the right holder the opportunity to control access to the data resource on the basis of the exclusive rights granted in EU law.**³¹ Apart from access questions, use of the data during the research process may require acts of reproduction and/or extraction which the right holder can prevent.³² The publication and sharing of research outcomes may also give rise to infringement questions. Insofar as research outcomes include elements of data resources that enjoy copyright or database protection, the publication and sharing of results may amount to an act of distribution, communication or making available to the public,³³ or an act of re-utilisation,³⁴ which the right holder can prohibit.³⁵ Finally, the publications, data collections or other documents following from a research project may attract copyright or sui generis database protection themselves.³⁶ Hence, access to these research results and the reuse of data embedded in these research results trigger the same protection mechanisms and the same questions of potential infringement.

In other words, **the exclusive rights granted in EU copyright, related rights and sui generis database law have a deep impact on data access, use and reuse for the purpose of scientific research.** As indicated in the preceding chapter, exceptions and limitations to copyright and database protection that address scientific research are of particular importance. As exponents of the freedom of expression, the freedom of

²⁹ CJEU, 29 July 2019, case C-516/17, *Spiegel Online*, para. 54; CJEU, 29 July 2019, case C-469/17, *Funke Medien NRW*, para. 70. For a more detailed discussion of this point, see T. Aplin/L. Bently, *Global Mandatory Fair Use: The Nature and Scope of the Right to Quote Copyright Works*, Cambridge: Cambridge University Press 2020, 75-84; C. Geiger/E. Izyumenko, ‘The Constitutionalization of Intellectual Property Law in the EU and the *Funke Medien*, *Pelham* and *Spiegel Online* Decisions of the CJEU: Progress, but Still Some Way to Go!’, *IIC* 51 (2020), 282 (292-298).

³⁰ Article 2(1) DAP.

³¹ In particular, see Articles 2, 3 and 4 ISD; Article 7(1) and (5) DBD.

³² Article 2 ISD; Article 7(1), (2)(a) and (5) DBD.

³³ Articles 3 and 4 ISD.

³⁴ Article 7(1), 2(b) and (5) DBD.

³⁵ Cf. M. Caspers/L. Guibault et al., *Future TDM – Baseline Report of Policies and Barriers of TDM in Europe*, Amsterdam: Institute for Information Law 2016, 23 and 26-27.

³⁶ For a discussion of this aspect of copyright and database protection as the traditional focus of the debate on copyright and science, see Reto M. Hilty, ‘Das Urheberrecht und der Wissenschaftler’, *Gewerblicher Rechtsschutz und Urheberrecht – International* 2006, 179-180.

information and the freedom of science of researchers, they can provide avenues for access and use that do not require the permission of right holders. In practice, **exceptions and limitations foster data access, use and reuse by exempting researchers from the obligation to obtain an individual authorisation for the intended use. In this way, exceptions and limitations also support the independence of scientific research.** The statutory use permission following from an exception or limitation prevents right holders from controlling – and potentially prohibiting – use for research purposes. Within the field of application of an exception or limitation, right holders cannot invoke their exclusive rights as veto rights to prevent use in a research project.

As the following discussion will show, **however, the exceptions and limitations in EU copyright, related rights and database law are far from offering researchers a broad, general use privilege which, as a rule, covers the use of protected data resources in the context of scientific research and shields researchers reliably from allegations of infringement.** Following the continental-European copyright tradition,³⁷ exceptions and limitations constitute precisely-circumscribed provisions that require researchers to meet several conditions before the exception or limitation can be invoked to justify use without prior authorisation of the right holder. To illustrate this point, the following analysis sheds light on relevant provisions and discusses their potential to foster data use in scientific research. The overview starts with the regulation of use “for the sole purpose [...] of scientific research”³⁸ in the Information Society Directive (following section 3.1), followed by an examination of the exemption of extractions “for the purposes of [...] scientific research” in the Database Directive (section 3.2). The specific regulation of text and data mining (“TDM”) in the Directive on Copyright in the Digital Single Market occupies centre stage in section 3.3. A discussion of the exclusion of sui generis database protection for machine-generated raw data in the proposed Data Act follows in section 3.4.

3.1 Information Society Directive

In the area of copyright and related rights, **Article 5(3)(a) ISD permits acts of reproduction³⁹ and acts of communication and making available to the public⁴⁰ without prior authorisation of the right holder with regard to:**

use for the sole purpose of illustration for teaching or scientific research, as long as the source, including the author’s name, is indicated, unless this turns out to be impossible and to the extent justified by the non-commercial purpose to be achieved;...

At first glance, this provision has a broad scope. In particular, it refers to “scientific research” in general, without reducing the scope of the permission to a specific form or mode of use. **A closer inspection of Article 5(3)(a) ISD, however, reveals at least**

³⁷ As to the different traditions (civil law and common law) in copyright law and the corresponding approaches to exceptions and limitations, see M.R.F. Senftleben, “Bridging the Differences Between Copyright’s Legal Traditions – the Emerging EC Fair Use Doctrine”, *Journal of the Copyright Society of the U.S.A.* 57 (2010), 521 (522-525). As to the impact of the different ways of regulating exceptions and limitations on use for the purpose of scientific research, see M.R.F. Senftleben/T. Margoni et al., “Ensuring the Visibility and Accessibility of European Creative Content on the World Market: The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law (JIPITEC)* 13 (2022), 67 (67-86).

³⁸ Article 5(3)(a) ISD.

³⁹ Article 2 ISD.

⁴⁰ Article 3(1) ISD (copyright: communication and making available to the public) and Article 3(2) ISD (related rights: making available to the public).

three potential obstacles: the question of use for illustration purposes (following subsection 3.1.1), the lack of harmonisation (3.1.2) and the requirement of non-commercial use (3.1.3). **In addition** to these risk factors that are inherent in the wording of the provision itself, **it is important to consider the broader regulatory context in which Article 5(3)(a) ISD is placed.** It follows from Article 5(5) ISD that the application of the exception depends on compliance with the so-called “three-step test” which prohibits conflicts with a work’s normal exploitation and unreasonable prejudices to the legitimate interests of right holders (3.1.4). Moreover, Article 6 ISD provides for the protection of technological measures which right holders may employ to prevent access to protected literary and artistic works (3.1.5). Finally, the impact of contractual stipulations must be factored into the equation. The Information Society Directive does not ban contractual terms that restrict use for scientific research and, thus, neutralise statutory permissions given on the basis of Article 5(3)(a) ISD (3.1.6). In addition, other exceptions and limitations in the Information Society Directive, only cover very specific forms of deducing data from literary and artistic works. These alternative access and use avenues do not offer a meaningful alternative to the more general provision in Article 5(3)(a) ISD (3.1.7). Surveying this spectrum of legal questions and sources of legal uncertainty, **it becomes apparent that Article 5(3)(a) ISD may fail to provide a robust, reliable basis for gathering data from literary and artistic works in the context of scientific research** (3.1.8).

3.1.1 Potential Illustration Requirement

First, it is unclear whether the word “illustration” in the provision concerns only “teaching” or both use for “teaching” and use for “scientific research”. Other language versions do not resolve this ambiguity. The French text of Article 5(3)(a) ISD, for example, refers to “*utilisation à des fins exclusives d’illustration dans le cadre de l’enseignement ou de la recherche scientifique*”. The German version uses the expression “*Nutzung ausschließlich zur Veranschaulichung im Unterricht oder für Zwecke der wissenschaftlichen Forschung*”.

In all these language versions, the scope of the use privilege remains opaque. On the one hand, Article 5(3)(a) ISD may be understood to generally permit “use for the sole purpose of [...] scientific research.”⁴¹ On the other hand, it can be argued that the provision only covers “use for the sole purpose of illustration for [...] scientific research.”⁴² **The obligations of EU Member States following from international copyright law do not preclude any of these interpretations.** Article 10(2) of the Berne Convention for the Protection of Literary and Artistic Works (“BC”) offers room for Berne Union Members⁴³ to permit the use of protected literary and artistic works “by way of illustration in publications, broadcasts or sound or visual recordings for teaching, provided such utilization is compatible with fair practice.” Considering this international provision, it does not come as a surprise that Article 5(3)(a) ISD refers to “illustration for teaching.”⁴⁴ However, it does not follow from Article 10(2) BC that the illustration

⁴¹ C. Geiger/F. Schönherr, “The Information Society Directive”, in: I. Stamatoudi/P. Torremans (eds.), *EU Copyright Law – A Commentary*, 2nd ed., Cheltenham: Edward Elgar 2021, §11.99; P.B. Hugenholtz/M.R.F. Senftleben, *Fair Use in Europe. In Search of Flexibilities*, Amsterdam: Institute for Information Law/VU Centre for Law and Governance 2011, 14-18.

⁴² This position is taken, for instance, by E. Derclaye “The Database Directive”, in: I. Stamatoudi/P. Torremans (eds.), *EU Copyright Law – A Commentary*, 2nd ed., Cheltenham: Edward Elgar 2021, §9.65; S. Bechtold, “Information Society Directive”, in: T. Dreier/P.B. Hugenholtz (ed.), *Concise European Copyright Law*, 2nd ed., Alphen aan den Rijn: Wolters Kluwer 2016, 464.

⁴³ All EU Member States are Members of the Berne Union. See the list of Member States at https://wipo.int/treaties/ShowResults?search_what=C&treaty_id=15 (last visited on 28 March 2022).

⁴⁴ For a detailed discussion of national approaches to the exemption of educational use in copyright law, see R. Xalabarder, *Study on Copyright Limitations and Exceptions for Educational Activities in*

requirement also applies to scientific research. The international rule is silent on the issue of research use.

Divergent national implementation practices show that both interpretations – Article 5(3)(a) ISD exempts use for “scientific research” globally; Article 5(3)(a) ISD only exempts use for “illustration for [...] scientific research” – have informed lawmaking in EU Member States. Without setting forth an illustration requirement, § 60c of **the German Copyright Act** (*Urheberrechtsgesetz*) generally permits use “for the purpose of non-commercial scientific research”.⁴⁵ Article L 122-5(3)(e) of **the French Intellectual Property Code** (*Code de la propriété intellectuelle*), by contrast, limits the scope of the copyright exception to use “à des fins exclusives d’illustration dans le cadre de la recherche”.

Evidently, a use permission without illustration requirement offers more room for the use of data embedded in literary and artistic works. In national systems requiring use for illustration purposes, the question arises whether the copying of protected data resources, such as texts, images, films and pieces of music, for the purpose of analysing and extracting relevant “acts, facts or information”⁴⁶ falls within the scope of the concept of “illustration” underlying the copyright exception. National systems without illustration requirement do not give rise to doubts in this respect: the copying of protected data resources is permissible in the context of scientific research, irrespective of use for illustration purposes.

Differences in scope, such as the differences arising from the transposition of the illustration requirement into national law, can have repercussions on collaborations between research teams across Member States. Whereas acts of reproduction and making available to the public may fall within the scope of the applicable national use privilege for scientific research in country A, it may be necessary to ensure a prior use authorisation from right holders in country B. To arrive at comparable data use possibilities, it may thus be necessary to take different measures in different Member States. If no agreement with right holders can be found and the research requires equal use conditions, the most restrictive national system involved may dictate the parameters of use for all research teams. **This unsatisfactory result minimises potential benefits – in terms of available data resources – that could have flown from Article 5(3)(a) ISD.**

3.1.2 Lack of Harmonisation

With regard to national differences in scope, however, the described divergence arising from the illustration requirement is only one factor in the equation. **The matrix of national differences becomes more complex when it is considered that under Article 5(3) ISD, Member States are not obliged to adopt a specific copyright limitation or exception in favour of scientific research.** The optional nature of Article 5(3)(a) ISD implies that Member States are not bound to implement a scientific use privilege in a standardised form. Countries are free to refrain from adopting a copyright exception based on Article 5(3)(a) ISD. Alternatively, they may opt for maximum implementation – in the sense of devising a national use privilege that seeks to exhaust the flexibility for exempting research use offered by the prototype laid down

North America, Europe, Caucasus, Central Asia and Israel, WIPO Document SCCR/19/8, dated 5 November 2009, Geneva: WIPO 2009, available at: [Study on Copyright Limitations and Exceptions for Educational Activities in North America, Europe, Caucasus, Central Asia and Israel \(wipo.int\)](#) (last visited on 28 March 2022).

⁴⁵ See the official English translation, available at: [Act on Copyright and Related Rights \(Urheberrechtsgesetz – UrhG\) \(gesetze-im-internet.de\)](#) (last visited on 10 April 2022).

⁴⁶ See the definition of “data” in Article 2(1) DAP.

in Article 5(3)(a) ISD. Between these two poles of regulatory responses, a country may also take a cautious approach and add further conditions for invoking the use privilege at the national level – conditions that narrow the field of application of the copyright exception in comparison with the prototype in Article 5(3)(a) ISD.

Again, the aforementioned French and German transpositions of Article 5(3)(a) ISD can serve as examples. On the one hand, § 60c of the German Copyright Act permits the reproduction, distribution and making available to the public of up to 15 percent of a protected work for a specifically delimited circle of persons for their personal and non-commercial scientific research. Full use may be made of images, pictures and other illustrations, individual articles from the same professional or scientific journal, other small-scale works and out-of-commerce works. Up to 75 percent of a protected work may be taken when the use only concerns acts of reproduction for one's own personal scientific research (without sharing the research material with others).⁴⁷ Article L 122-5(3)(e) of the French Intellectual Property Code, on the other hand, covers the reproduction, and communication to the public by any means, of extracts of protected works for the purpose of illustration in the context of scientific research not giving rise to any commercial exploitation. The use must be intended for a public that consists mainly of researchers who are directly involved in the research activity requiring the use. The extracts must not be disseminated to any third party. Moreover, the invocation of the copyright exception requires the payment of a lumpsum remuneration. Works designed for educational purposes and musical scores fall outside the scope of the use privilege from the outset.

These two examples already show clearly that the national implementation process in EU Member States can lead to use privileges for scientific research that have individual conceptual contours. The prerequisites for making reproductions and disseminating research results can differ considerably from country to country. As indicated in the preceding section, these differences can complicate data extractions on the basis of Article 5(3)(a) ISD. In research projects that require the cross-border exchange and use of data resources, the need to satisfy individual use conditions in all national copyright systems involved can reach a degree of legal complexity that makes reliance on copyright exceptions based on Article 5(3)(a) ISD unattractive, if not unfeasible.⁴⁸ **The beneficial effect of Article 5(3)(a) ISD must not be overestimated against this background. As a template for copyright exceptions and limitations in favour of scientific research, the provision appears promising. In the absence of fully harmonised national offshoots, however, the practical benefits remain limited,** in particular in the case of transnational research projects carried out in several EU Member States.

3.1.3 Requirement of Non-commercial Use

The requirement of use for a “non-commercial purpose” further enhances the legal complexity surrounding the extraction of data from literary and artistic works on the basis of Article 5(3)(a) ISD. As the discussion of further EU provisions in the following sections will show, the regulation of scientific research in EU copyright law is interspersed with the non-commercial use condition.⁴⁹ **Against this background, the criteria for determining the nature of research activities play a central role.** They function as gatekeepers that can preclude the invocation of an exception or

⁴⁷ Cf. T. Dreier, “UrhG § 60c Wissenschaftliche Forschung”, in: T. Dreier/G. Schulze (eds.), *Urheberrechtsgesetz – Kommentar*, 7th ed., Munich: C.H. Beck 2022, para. 12.

⁴⁸ C. Geiger/G. Frosio/O. Bulayenko, “Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations”, *International Review of Intellectual Property and Competition Law* 2018, 814 (826-827).

⁴⁹ In particular, see Article 9(b) DBD; Article 2(1)(a) DSMD.

limitation for scientific research from the outset. In the context of the Information Society Directive, **Recital 42 offers the following clarification with regard to the copyright exception laid down in Article 5(3)(a) ISD:**

When applying the exception or limitation for non-commercial educational and scientific research purposes, including distance learning, the non-commercial nature of the activity in question should be determined by that activity as such. The organisational structure and the means of funding of the establishment concerned are not the decisive factors in this respect.

Recital 42 mentions several parameters that, in principle, could impact the assessment of a given research activity. The organisational structure of the research institution may play a role. The means of funding of the research institution could be taken into account. In accordance with the Recital, however, these factors must not prevail. Instead, the non-commercial nature of a research activity should be determined "by that activity as such." This statement indicates that commercial motives, such as a profit orientation⁵⁰ or plans for commercial exploitation,⁵¹ bar researchers from the invocation of the copyright exception.

Nonetheless, doubts about the scope of the non-commercial use requirement remain. First, it seems unclear whether scientific research projects can still be qualified as non-commercial when they are made possible by industry funding. Recital 42 declares that the funding of the research institution is not decisive. However, it does not address commercial funding of an individual project and its impact on the assessment of the research activity "as such." **Second, Recital 42 is silent on the status of public-private partnerships.**

With regard to the first question – industry funding of scientific research – a previous national discussion of this question in the legislative debate can be found in Germany. The German legislator has taken the position that, in the light of Recital 42, the source of funding is irrelevant. Proposing new legislation to regulate the interface between copyright and science in 2017, the German government stated that research conducted at public universities and financed through private third-party funding could, in principle, fall within the scope of the copyright exception that permits scientific research use without prior authorisation of right holders.⁵² This position seems to be in line with the guideline following from Recital 42. As long as private funding does not give the research activity "as such" a commercial character, a private source, such as financial resources made available by industry sponsors, should not exclude the eligibility of researchers for invoking the copyright exception in favour of scientific research. This is different when the research no longer follows applicable rules of academic independence, or aims specifically at research results that can be exploited commercially.⁵³

Similar conclusions can be drawn with regard to public-private partnerships. As long as these partnerships focus on independent academic research seeking to enhance knowledge and science, the involvement of a private partner does not thwart the invocation of the exemption of use for scientific research. In this respect, contractual safeguards, such as a clarification in the partnership agreement that the research does

⁵⁰ Cf. Article 2(1)(a) DSMD that refers to research conducted on a "not-for-profit basis".

⁵¹ See the implementation of Article 5(3)(a) ISD in France. Article L 122-5(3)(e) of the French Intellectual Property Code makes it a condition that the use "*ne donne lieu à aucune exploitation commerciale.*"

⁵² German government (*Bundesregierung*), *Entwurf eines Gesetzes zur Angleichung des Urheberrechts an die aktuellen Erfordernisse der Wissensgesellschaft (Urheberrechts-Wissensgesellschafts-Gesetz – UrhWissG)*, Document 18/12329, dated 15 May 2017, Berlin: Deutscher Bundestag 2017.

⁵³ Cf. G. Spindler, "Text- und Datamining im neuen Urheberrecht und in der europäischen Diskussion", *Zeitschrift für Geistiges Eigentum* 2018, 273 (279-280).

not serve commercial purposes and cannot be exploited by the private partner,⁵⁴ may provide evidence that the research activity does not have a commercial character. If, however, the research is intended to generate commercially exploitable knowledge, the commercial character can no longer be denied.⁵⁵

As the discussion of appropriate parameters shows, **the guideline given in Recital 42** – requiring an assessment of the research activity “as such” – **fails to simplify the task of determining the eligibility of a research team for invoking the copyright exception** that permits the use of literary and artistic works for scientific research without prior authorisation. **Instead, the requirement of use for a non-commercial purpose in Article 5(3)(a) ISD adds further complexity.** The risk of reducing the scope of the copyright exception considerably must not be underestimated. In practice, there is a tendency of stimulating researchers to collaborate with private partners. European and national funding schemes for research may even require the involvement of private partners and make it a condition that these partners provide a part of the budget. Considering these developments, **it can be expected that the outlined questions arising from industry funding and public-private partnerships will often raise doubts** about the applicability of the copyright exception. **In consequence, the beneficial effect of Article 5(3)(a) ISD may remain limited.**

3.1.4 Three-Step Test Compliance

In addition to the legal requirements that follow from Article 5(3)(a) ISD itself, it is important to take into account that **more general, overarching conditions for the successful assertion of the use privilege for scientific research follow from Article 5(5) ISD. This latter provision contains the so-called “three-step test”** in EU copyright law that applies horizontally across all exceptions and limitations enumerated in the preceding paragraphs, including the exemption of research use in Article 5(3)(a) ISD. **The three-step test reads as follows:**

The exceptions and limitations provided for in paragraphs 1, 2, 3 and 4 shall only be applied in certain special cases which do not conflict with a normal exploitation of the work or other subject-matter and do not unreasonably prejudice the legitimate interests of the rightholder.

With its open-ended criteria, **the three-step test adds considerable complexity to the assessment of use for the purposes of scientific research.** Does a given form of research constitute a “special case” in the sense of Article 5(5) ISD? What is a “normal” exploitation of literary and artistic works? When does use for research purposes enter into “conflict” with a normal exploitation? Which competing right holder interests must be deemed “legitimate”? When does a prejudice to these legitimate interests reach an impermissible, “unreasonable” level?

In the absence of sufficient legal guidance on the right interpretation of these elastic criteria, **the three-step test can easily become a source of legal uncertainty.** Even if a form of research use complies with all requirements of Article 5(3)(a) ISD, right holders may still challenge the permissibility of the use on the ground that one of the requirements following from Article 5(5) ISD is not fulfilled. In practice, the application of the three-step test as a yardstick by judges has become a widespread practice in the EU because the provision stipulates that copyright exceptions, such as the research rule in Article 5(3)(a) ISD, “shall only be applied” when the use is compatible with the three-step test. Arguably, this determination of compliance at the level of applying the use

⁵⁴ Cf. Spindler, *id.*, 280.

⁵⁵ K. Anton, “UrhG § 60c Wissenschaftliche Forschung”, in: G. Spindler/F. Schuster (eds.), *Recht der elektronischen Medien*, 4th ed., Munich: C.H. Beck 2019, para. 4.

privilege for scientific research is the task of the judge hearing a case about copyright infringement. **To this day, however, the CJEU has not seized the opportunity of providing concrete guidelines for the right interpretation of the individual assessment criteria forming the three-step test** – despite several decisions in which the Court has relied on three-step test to support its decision.⁵⁶

Some guidance for the interpretation of the three-step test can nonetheless be found in relevant “case law.” At the international level, **the decision of a WTO Panel in the dispute settlement case *United States – Section 110(5) of the US Copyright Act* yielded a detailed discussion and interpretation of the test criteria.**⁵⁷ The WTO Panel discussed the three-step test in Article 13 of the Agreement on Trade-Related Aspects of Intellectual Property Rights (“TRIPS”). In substance, this international provision sets forth the same requirements that can be found in Article 5(5) ISD:

[WTO] Members shall confine limitations or exceptions to exclusive rights to certain special cases which do not conflict with a normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the right holder.⁵⁸

With regard to the first test (“certain special cases”), the WTO Panel distinguished between the word “certain” and the word “special”. It understood the term “certain” to mean that a copyright limitation had to be clearly defined, while there was no need “to identify explicitly each and every possible situation to which the exception could apply, provided that the scope of the exception was known and particularised.”⁵⁹ From the term “special,” the Panel derived the additional requirement that a limitation should be narrow in a quantitative as well as a qualitative sense.⁶⁰ It summarised this twofold requirement as narrowness in “scope and reach.”⁶¹ The application to Section 110(5) of the US Copyright Act shows that, pursuant to the Panel’s conception, it is for example the number of potential beneficiaries that must be sufficiently limited in order to comply with the quantitative aspect of speciality.⁶² As to the qualitative aspect, the Panel eschewed an inquiry into the legitimacy of the public policy purpose underlying the adoption of a limitation.⁶³ In the Panel’s view, the

⁵⁶ For an overview of these decisions, see M.R.F. Senftleben, “From Flexible Balancing Tool to Quasi-Constitutional Straitjacket— How the EU Cultivates the Constraining Function of the Three-Step Test”, in: J. Griffiths/T. Mylly (eds.), *Global Intellectual Property Protection and New Constitutionalism – Hedging Exclusive Rights*, Oxford: Oxford University Press 2021, 83 (83-105).

⁵⁷ World Trade Organization, 15 June 2000, *United States – Section 110(5) of the US Copyright Act*, Report of the Panel, WTO Document WT/DS160/R. For a discussion of this WTO Dispute Settlement decision, see C. Geiger/D. Gervais/M.R.F. Senftleben, “The Three-Step Test Revisited: How to Use the Test’s Flexibility in National Copyright Law”, *American University International Law Review* 29 (2014), 581 (593-597); M.R.F. Senftleben, *Copyright, Limitations and the Three-Step Test – An Analysis of the Three-Step Test in International and EC Copyright Law*, The Hague/London/New York: Kluwer Law International 2004, 134-230; M.J. Ficsor, “How Much of What? The Three-Step Test and Its Application in Two Recent WTO Dispute Settlement Cases”, *Revue Internationale du Droit d’Auteur* 192 (2002), 111; J. Oliver, “Copyright in the WTO: The Panel Decision on the Three-Step Test”, *Columbia Journal of Law and the Arts* 25 (2002), 119; D.J. Brennan, “The Three-Step Test Frenzy: Why the TRIPS Panel Decision might be considered Per Incuriam”, *Intellectual Property Quarterly* 2002, 213; J.C. Ginsburg, “Toward Supranational Copyright Law? The WTO Panel Decision and the “Three-Step Test” for Copyright Exceptions”, *Revue Internationale du Droit d’Auteur* 190 (2001), 13.

⁵⁸ Article 13 TRIPS. Explanation in brackets added by the author.

⁵⁹ Report of the WTO Panel, *ibid.*, para. 6.108.

⁶⁰ Report of the WTO Panel, *ibid.*, para. 6.109. Cf. André Lucas, “Le “triple test” de l’article 13 de l’Accord ADPIC à la lumière du rapport du Groupe spécial de l’OMC “Etats-Unis – Article 110 (5) de la loi sur le droit d’auteur””, in: Ganea/Heath/Gerhard Schrickler (eds.), *Urheberrecht Gestern – Heute – Morgen, Festschrift für Adolf Dietz zum 65. Geburtstag*, C.H. Beck: Munich 2001, 423 (430), who insists on the combination of both aspects of speciality to ensure a sufficiently rigid standard of control.

⁶¹ Report of the WTO Panel, *ibid.*, para. 6.112.

⁶² Report of the WTO Panel, *ibid.*, para. 6.127 and 6.143.

⁶³ Report of the WTO Panel, *ibid.*, para. 6.111.

qualitative aspect of speciality does not have a normative connotation. A limitation need not serve a special purpose to be qualified as a special case in the sense of Article 13 TRIPS.⁶⁴ Instead, the Panel raised conceptual issues, such as the categories of works affected by a copyright limitation and the circumstances under which it may be invoked. In this vein, the Panel lent weight to the fact that one of the use privileges following from Section 110(5) of the US Copyright Act was found limited to dramatic renditions of operas and homestyle receiving apparatus.⁶⁵

Turning to the second test (“conflict with a normal exploitation”), the WTO Panel interpreted the term “exploitation” as a reference to “the activity by which copyright owners employ the exclusive rights conferred on them to extract economic value from their rights to [musical] works.”⁶⁶ In this context, the Panel distinguished between an empirical and a normative meaning of the word “normal.” While the first connotation of the term “normal” appeared to be “of an empirical nature, i.e. what is regular, usual, typical or ordinary”, the second connotation reflected “a somewhat more normative, if not dynamic, approach, i.e., conforming to a type or standard.”⁶⁷ With regard to the empirical aspect, the WTO Panel accepted the US approach asking “whether there are areas of the market in which the copyright owner would ordinarily expect to exploit the work, but which are not available for exploitation because of [the exemption at issue].”⁶⁸ Accordingly, uses from which an owner would not ordinarily expect to receive compensation were not regarded as parts of a normal exploitation. Seeking to give meaning also to the normative aspect of the word “normal”, the Panel had recourse to the work of a study group preparing the 1967 Stockholm Conference for the Revision of the Berne Convention for the Protection of Literary and Artistic works.⁶⁹ In particular, it attached importance to the conclusion that “all forms of exploiting a work, which have, or are likely to acquire, considerable economic or practical importance, must be reserved to the authors.”⁷⁰ The Panel inferred from this formula that it was appropriate to consider, “in addition to those forms of exploitation that currently generate significant or tangible revenue, those forms of exploitation which, with a certain degree of likelihood and plausibility, could acquire considerable economic or practical importance.”⁷¹ On its merits, the normative aspect, therefore, served as a vehicle to widen the perspective. It allowed the Panel to factor into the equation both currently existing and potential future markets when determining a conflict with “a normal exploitation of the work.”

Finally, the WTO Panel’s analysis of the test of “no unreasonable prejudice to legitimate interests” remained limited to the interest in the economic value of

⁶⁴ Report of the WTO Panel, *ibid.*, para. 6.112. For literature statements supporting this view, see J.C. Ginsburg, “Toward Supranational Copyright Law? The WTO Panel Decision and the “Three-Step Test” for Copyright Exceptions”, *Revue Internationale du Droit d’Auteur* 190 (2001), 13; Sam Ricketson, *The Three-Step Test, Deemed Quantities, Libraries and Closed Exceptions*, Centre for Copyright Studies, Strawberry Hills 2002, 31. For comments insisting on the importance of a policy analysis already at this first step of the test procedure, see M.J. Ficsor, “How Much of What? The Three-Step Test and Its Application in Two Recent WTO Dispute Settlement Cases”, *Revue Internationale du Droit d’Auteur* 192 (2002), 111 (129-133 and 223-229); M.R.F. Senftleben, *Copyright, Limitations and the Three-Step Test – An Analysis of the Three-Step Test in International and EC Copyright Law*, The Hague/London/New York: Kluwer Law International 2004, 138-152, who propose to require that limitations be justified by some clear reason of public policy.

⁶⁵ Report of the WTO Panel, *ibid.*, para. 6.145, 6.146 and 6.159.

⁶⁶ Report of the WTO Panel, *ibid.*, para. 6.165.

⁶⁷ Report of the WTO Panel, *ibid.*, para. 6.166.

⁶⁸ Report of the WTO Panel, *ibid.*, para. 6.177 – 6.178.

⁶⁹ For a description of the work of the study group, see M.J. Ficsor, “How Much of What? The Three-Step Test and Its Application in Two Recent WTO Dispute Settlement Cases”, *Revue Internationale du Droit d’Auteur* 192 (2002), 111 (115-121).

⁷⁰ Report of the WTO Panel, *ibid.*, para. 6.179. The study group was composed of representatives of the Swedish government and the United International Bureaux for the Protection of Intellectual Property (BIRPI).

⁷¹ Report of the WTO Panel, *ibid.*, para. 6.180.

the exclusive rights conferred by copyright. In the absence of any objections raised by the parties, the Panel could readily qualify this interest as legitimate.⁷² With regard to the expression “not unreasonably prejudice,” the Panel noted that any copyright limitation, by definition, caused some detriment to right holders because it reduced the scope of exclusive rights. This led to the insight that in order not to erode copyright limitations altogether, “a certain amount of prejudice has to be presumed justified as ‘not unreasonable’.”⁷³ The Panel concluded that “prejudice to the legitimate interests of right holders reaches an unreasonable level if an exception or limitation causes or has the potential to cause an unreasonable loss of income to the copyright owner.”⁷⁴

As this overview of potential considerations shows, **the open-ended criteria following from the three-step test offer various starting points for right holders to challenge the compliance of a given form of research use with Article 5(5) ISD.** For instance, it may be argued that scientific research that requires large-scale use of copyrighted data resources, such as TDM,⁷⁵ cannot be qualified as a special case in the sense of the first test (“certain special cases”) because the use is not sufficiently confined from the quantitative perspective which the WTO Panel introduced. At the same time, the openness of the test criteria offers room for counterarguments. While the number of works may reach a large volume in TDM cases, the circle of beneficiaries – Article 5(3)(a) ISD only covers scientific researchers without a commercial orientation – is narrowly drawn and specific. From a qualitative, normative perspective, it may be added that use for scientific research, by definition, constitutes a special case in the sense of Article 5(5) ISD because breathing space for use in this category is indispensable to arrive at a proper balance between copyright protection and the guarantee of freedom of expression and information, and freedom of sciences, in Articles 11 and 13 CFR.⁷⁶

Another recurring theme in the three-step test debate is the argument that use privileges for scientific research cause a conflict with the normal exploitation of works that are specifically made for research use, such as academic publications. The power of persuasion of this argument must not be underestimated. The national implementation of Article 5(3)(a) ISD in Article L 122-5(3)(e) of the French Intellectual Property Code (*Code de la propriété intellectuelle*), for instance, explicitly excludes works made for educational purposes (“*sous réserve des oeuvres conçues à des fins pédagogiques*”) from the scope of the use privilege. Again, counterarguments are conceivable. In the case of academic publications, it must be considered that publishers of academic books and journals benefit from the work of researchers in scientific projects to quite some extent. The use of existing academic publications as data sources for the purpose of finding new knowledge – leading to new academic publications – appears as a good investment in the continuous evolution of new publishable material. The cyclic process in which new knowledge arises from the scientific analysis of pre-existing academic sources⁷⁷ begs the question whether academic publishers have legitimate

⁷² Report of the WTO Panel, *ibid.*, para. 6.226.

⁷³ Report of the WTO Panel, *ibid.*, para. 6.229.

⁷⁴ Report of the WTO Panel, *ibid.*, para. 6.229.

⁷⁵ Cf. M.R.F. Senftleben/T. Margoni et al., “Ensuring the Visibility and Accessibility of European Creative Content on the World Market – The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 13 (2022), 67 (70-71).

⁷⁶ For a more detailed discussion of the fundamental rights impact on the analysis, see C. Geiger/D. Gervais/M.R.F. Senftleben, “The Three-Step Test Revisited: How to Use the Test’s Flexibility in National Copyright Law”, *American University International Law Review* 29 (2014), 581 (581-626); C. Geiger/J. Griffiths/R.M. Hilty, “Declaration on a Balanced Interpretation of the “Three-Step Test” in Copyright Law”, *International Review of Intellectual Property and Competition Law* 39 (2008), 707; C. Geiger, “‘Constitutionalising’ Intellectual Property Law? The Influence of Fundamental Rights on Intellectual Property in the European Union”, *International Review of Intellectual Property and Competition Law* 37 (2006), 371.

⁷⁷ For a discussion of the role of copyright law in enabling cyclic innovation in the literary and artistic sector, see M.R.F. Senftleben, *The Copyright/Trademark Interface – How the Expansion of*

interests in the sense of the three-step test that can be asserted against use privileges for scientific research. Following this line of reasoning, arguments based on an erosion of the market for academic publications seem unfounded.

These examples show that various non-compliance arguments may be developed to declare research use based on Article 5(3)(a) ISD impermissible. The three-step test in Article 5(5) ISD is thus a source of legal uncertainty that can have a corrosive effect on research use. Even if non-compliance arguments can finally be rebutted, **the three-step test offers right holders a broad arsenal of ammunition to cast doubt upon use privileges for scientific research and destabilise the legal position which Article 5(3)(a) ISD seeks to bestow upon researchers.**⁷⁸

3.1.5 Technological Protection Measures

Besides the legal tools following from Article 5(5) ISD, **right holders can also rely on technological measures to prevent or restrict research use which they have not specifically authorised.** With access control and the application of protection processes, such as encryption, distortion or other transformations of works, they may control the use of works and achieve a level of protection that leaves hardly any room for scientific use on the basis of a copyright exception, such as Article 5(3)(a) ISD.⁷⁹ A password-protected login to a website with data resources that are required for research purposes can serve as an example of a technical protection measure – an electronic “fence” – that can restrict access and use possibilities for researchers substantially. Despite this risk of an impoverishment of data resources, **Article 6(1) ISD explicitly protects copyright holders against the circumvention of effective technological measures:**

Member States shall provide adequate legal protection against the circumvention of any effective technological measures, which the person concerned carries out in the knowledge, or with reasonable grounds to know, that he or she is pursuing that objective.

Article 6(2) ISD further strengthens the position of right holders by offering additional protection against the provision of equipment or services that can be employed to circumvent technological measures.

Admittedly, the protection of technological measures is not intended to undermine the exemption of research use on the basis of Article 5(3)(a) ISD. Article 6(4) ISD provides in its first subparagraph that Member States shall take appropriate measures to ensure that right holders make available to the beneficiary of a research exemption based on Article 5(3)(a) ISD the means of benefiting from that copyright exception, where the researcher has legal access to the protected material at issue. **It is unclear, however, to which extent this safeguard clause shields**

Trademark Protection Is Stifling Cultural Creativity, The Hague/London/New York: Kluwer Law International 2020, 26-35.

⁷⁸ As to the recognition of the user “rights” status following from copyright exceptions that serve freedom of expression and information, see CJEU, 29 July 2019, case C-516/17, Spiegel Online, para. 54; CJEU, 29 July 2019, case C-469/17, Funke Medien NRW, para. 70. Cf. section 2.2.

⁷⁹ C. Geiger/G. Frosio/O. Bulayenko, “Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations”, *International Review of Intellectual Property and Competition Law* 2018, 814 (826).

researchers effectively from the erosion of applicable statutory use permissions in national law.⁸⁰

Even more importantly, however, the safeguard clause for certain use privileges, including the exemption of research use by virtue of Article 5(3)(a) ISD, is no longer applicable when access to protected works is offered on demand. This follows from Article 6(4), subparagraph 4, ISD:

The provisions of the first and second subparagraphs shall not apply to works or other subject-matter made available to the public on agreed contractual terms in such a way that members of the public may access them from a place and at a time individually chosen by them.

In practice, this means that researchers can no longer insist on access to websites with required data sources, such as large databases belonging to academic publishers, online music stores and audiovisual services, when the right holder makes protected content available on demand behind a technical fence and offers access on the condition that the researcher enter into a contractual agreement first. As Professor Reto Hilty has pointed out with regard to the substantial restriction of access to data resources that follows from this rule:

In the light of this norm, the discussion about exceptions and limitations in general – and those for science in particular [...] – simply becomes obsolete as soon as information offers are only made available online.⁸¹

In other words: right holders may use the protection of technological protection measures following from Article 6(1) and (4), subparagraph 4, ISD to prevent access to literary and artistic works that constitute important data sources for scientific research. **Therefore, the protection of technological measures in the Information Society Directive has the potential to pose major obstacles to data access and data use that is required for research.**

3.1.6 Contractual Restrictions

As the discussion of technological protection measures already indicated, **the Information Society Directive allows contractual stipulations to prevail over statutory exemptions of research use**, such as use privileges based on Article 5(3)(a) ISD. In the context of protection against the circumvention of technological measures, **Article 6(4), subparagraph 4, ISD reflects this policy decision.** As explained, Article 6(4), subparagraph 4, ISD explicitly lets contractual terms reign supreme. Hence, it cannot be ruled out that researchers must accept contractual stipulations that are contrary to copyright exceptions in their favour.⁸² In particular, this risk arises when

⁸⁰ Cf. T. Margoni/M. Kretschmer, "A Deeper Look Into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology", *CREATE Working Paper 2021/7*, Glasgow: CREATE Centre 2021, 26; C. Geiger/F. Schönherr, "The Information Society Directive", in: I. Stamatoudi/P. Torremans (eds.), *EU Copyright Law – A Commentary*, 2nd ed., Cheltenham: Edward Elgar 2021, §11.109.

⁸¹ R.M. Hilty, "Das Urheberrecht und der Wissenschaftler", *Gewerblicher Rechtsschutz und Urheberrecht – International* 2006, 179 (187): "Im Lichte dieser Norm wird die Diskussion über Schrankenbestimmungen im Allgemeinen - und jene für die Wissenschaft im Besonderen [...] - schlicht obsolet, sobald Informationsangebote nur noch online zur Verfügung gestellt werden." (English translation by the author of this report).

⁸² Cf. Hilty, id., 186-187; C. Geiger/F. Schönherr, "The Information Society Directive", in: I. Stamatoudi/P. Torremans (eds.), *EU Copyright Law – A Commentary*, 2nd ed., Cheltenham: Edward Elgar 2021, §11.112; T. Dreier/M.R.F. Senftleben, "Das Verhältnis des Urheberrechts zum Vertragsrecht – Grenzen des Vertragsrechts durch Intellectual Property Law", in: M. Lejeune (ed.),

access to a protected data source, such as a specific work repertoire, is indispensable for the envisaged research project and the research team has no acceptable alternatives.

3.1.7 Other Exceptions and Limitations

Besides the specific rule in Article 5(3)(a) ISD, the list of permissible exceptions and limitations in the Information Society Directive contains several other provisions that can become relevant in the context of scientific research. Article 5(3)(d) ISD, for instance, offers room for the exemption of “quotations for purposes such as criticism or review” – a quotation right which researchers can invoke when establishing scientific publications. These publications, in turn, constitute sources of data that may become relevant to future research (cf. the data categories distinguished in section 2.2). As to legal prerequisites for the successful assertion of the right of quotation, the CJEU stated in *Pelham* that for a quotation to be justified, it is necessary that a work be used “for the purposes of illustrating an assertion, of defending an opinion or of allowing an intellectual comparison between that work and the assertions of that user.”⁸³ A researcher invoking the right of quotation must have the intention of “entering into ‘dialogue’ with that work.”⁸⁴ Apart from this support for the academic discourse leading to new publications and data sources, Article 5(3)(d) ISD does not seem to be particularly relevant to the access to, and reuse of, research data.

Another provision that may enter the picture in research cases is **the mandatory exemption of temporary copying in Article 5(1) ISD:**

[t]emporary acts of reproduction referred to in Article 2, which are transient or incidental [and] an integral and essential part of a technological process and whose sole purpose is to enable:

- (a) a transmission in a network between third parties by an intermediary, or
- (b) a lawful use

of a work or other subject-matter to be made, and which have no independent economic significance, shall be exempted from the reproduction right provided for in Article 2.

Evidently, **this is quite a complex provision with five central requirements that must be satisfied cumulatively in order to benefit from the use privilege.**⁸⁵ As regards the first condition, the existence of a “temporary” reproduction can be assumed, for example, when the copies are immediately deleted or replaced automatically.⁸⁶ A reproduction can be deemed “transient” when the conservation period of copies is limited to the time necessary for the technical process of making the reproduction and the copies are automatically erased after completion of the process.⁸⁷ A reproduction is “incidental” where it is not self-contained with respect to the technical process of which it forms part. Thus, copies resulting from temporary reproductions should have no purpose that is

Der E-Commerce-Vertrag nach amerikanischem Recht, Köln: Otto Schmidt 2001, 81 (81-120). For an in-depth analysis, see L.M.C.R. Guibault, *Copyright Limitations and Contracts – An Analysis of the Contractual Overridability of Limitations on Copyright*, The Hague/London/Boston: Kluwer Law International 2002.

⁸³ CJEU, 29 July 2019, case C-476/17, *Pelham*, para. 71.

⁸⁴ CJEU, *id.*, para. 71.

⁸⁵ CJEU, 16 July 2009, case C-5/08, *Infopaq*, para. 55.

⁸⁶ CJEU, 5 June 2014, case C-360/13, *Meltwater*, para. 26.

⁸⁷ CJEU, *id.*, para. 40.

separate from the one for which they have been made in the framework of the technical process.⁸⁸

These conceptual contours indicate clearly that **Article 5(1) ISD only offers limited possibilities in the context of scientific research.**⁸⁹ As copies based on Article 5(1) ISD cannot be retained for a longer period, the provision does not permit the creation of source data repositories that could be used for replicability studies to validate research results. Moreover, the transient nature of the copies excludes reuse from the outset.

Nonetheless, Article 5(1) ISD may play a role in the analysis of online data that can be analysed directly and processed in the format in which they are available on webpages. For this form of web scraping and computational analysis, the requirements of a temporary and transient nature need not constitute insurmountable hurdles. The invocation of the use privilege in connection with new research technologies, such as TDM, also seems in line with the general objectives of the provision.⁹⁰ The CJEU has recognised that, in order to protect the effectiveness of the copyright exception and safeguard its purpose, Article 5(1) ISD must be understood to allow the development and operation of new technologies and ensure a fair balance between the rights and interests of right holders and those of users.⁹¹ **It seems consistent to assume that, as long as the individual requirements of the provision are fulfilled and the inherent use restrictions do not compromise academic standards of reliability and replicability, researchers can belong to the circle of users who can benefit from Article 5(1) ISD, for instance, in the context of TDM.**

3.1.8 Considerable Legal Uncertainty

Surveying the conditions and requirements that accompany the exemption of use for the purpose of scientific research in Article 5(3)(a) ISD, the conclusion seems inescapable that researchers are confronted with a complex legal framework that can easily lead to considerable legal uncertainty. It is unclear whether the use privilege following from Article 5(3)(a) ISD is confined to use for mere illustrative purposes. The status of collaborations with partners from the private sector remains opaque because of the exclusion of research use with a commercial orientation. Moreover, researchers may have to justify the intended use as a special case in the sense of the three-step test laid down in Article 5(5) ISD. They may also have to rebut allegations that research use has a corrosive effect on the normal exploitation of works serving as data sources, or prejudices legitimate interests of right holders in some other unreasonable way.

Apart from these legal issues, researchers may be exposed to restrictions of use privileges following from Article 5(3)(a) ISD that arise from the application of technological protection measures that serve as electronic fences preventing access

⁸⁸ CJEU, *id.*, para. 43.

⁸⁹ Cf. C. Geiger/G. Frosio/O. Bulayenko, "Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations", *International Review of Intellectual Property and Competition Law* 2018, 814 (821-822); R.M. Hilty/H. Richter, "Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules – Part B: Exceptions and Limitations – Art. 3 Text and Data Mining", *Max Planck Institute for Innovation and Competition Research Paper Series* 2017-02, 2.

⁹⁰ Cf. T. Margoni/M. Kretschmer, "A Deeper Look Into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology", *CREATE Working Paper* 2021/7, Glasgow: CREATE Centre 2021, 18-19.

⁹¹ CJEU, 4 October 2011, joined cases C-403/08 and C-429/08, Football Association Premier League, para. 163-164; CJEU, 5 June 2014, case C-360/13, Meltwater, para. 24.

and use for research purposes. **Restrictions may also follow from contractual terms** that exclude forms of use necessary for research.

In addition to these barriers that follow from the harmonised rules in EU copyright law, it must not be overlooked that **the research exemption in Article 5(3)(a) ISD is an optional “may” provision.** In certain Member States, the use privileges may be sought in vain. In others, the national implementation may have led to more nuanced domestic rules, adding further conditions and legal requirements. In the case of research consortia with partners in several Member States, the scope of use privileges that are available at the national level may thus differ from one research partner to the other. For acts of copying or data sharing that involve protected works and must be carried out by all research partners in a consortium, the group as a whole, thus, can only rely on the smallest common denominator that falls within the scope of the most restrictive national rule following from Article 5(3)(a) ISD.

The regulatory framework in the Information Society Directive must therefore be deemed problematic. **Instead of being conducive to research, the broad spectrum of legal and practical prerequisites and restrictions can easily frustrate the successful invocation of national exponents of Article 5(3)(a) ISD.**

3.2 Database Directive

If data relevant to a research project are incorporated in a protected database, researchers must observe the specific provisions of the Database Directive. This compliance obligation is twofold. **On the one hand, Article 3(1) DBD clarifies that databases enjoy copyright protection when, by reason of the selection or arrangement of their contents, they constitute the author’s own intellectual creation.**⁹² This copyright aspect of database protection does not lead to legal requirements that differ substantially from those discussed in the preceding section. **Article 5(2)(b) DBD contains an exemption of use for scientific research** that corresponds with the copyright exception laid down in Article 5(3)(a) ISD.

In addition to copyright protection, however, Article 7(1) DBD provides for a sui generis database right that arises from “substantial investment in either the obtaining, verification or presentation of the contents” of the database.⁹³ With regard to this sui generis database right, **Article 9(b) DBD offers room for Member States to stipulate that “lawful users of a database which is made available to the public in whatever manner” may, without prior authorisation of the right holder, extract a substantial part of the database content**

for the purposes of illustration for teaching or scientific research, as long as the source is indicated and to the extent justified by the non-commercial purpose to be achieved,...

Evidently, **this provision poses several difficulties that also arise in the context of copyright protection.** As Article 5(3)(a) ISD, Article 9(b) DBD raises the question whether the illustration requirement only concerns “teaching” or covers both “teaching” and “scientific research” (cf. section 3.1.1 above). Moreover, Article 9(b) DBD is an optional “may” provision. Member States are thus free to refrain from implementing the

⁹² With regard to the applicable test of free, creative choices in content selection and arrangement, see CJEU, 1 March 2012, case C-604/10, *Football Dataco/Yahoo!*, para. 38.

⁹³ As already indicated above, the threshold for assuming a substantial investment need not be particularly high. For instance, see CJEU, 9 October 2008, case C-304/07, *Directmedia/Universität Freiburg*, para. 24, where an investment of 34.900€ had been deemed sufficient by the national judge.

use privilege for research. They may also opt for a more restrictive implementation and add further conditions. Recital 51 DBD leaves no doubt that Member States may limit the use permission “to certain categories of teaching or scientific research institution” (cf. 3.1.2). Finally, the requirement of “non-commercial” research reappears in Article 9(b) DBD and gives rise to the described discussion about private funding of scientific research and public-private partnerships (cf. 3.1.3).

Further parallels with the regulatory approach underlying Article 5(3)(a) ISD come the fore when it is considered that Article 9(b) DBD can only be invoked by “lawful users.” In this respect, Recital 34 DBD provides the following guidance:

Whereas, nevertheless, once the rightholder has chosen to make available a copy of the database to a user, whether by an on-line service or by other means of distribution, that lawful user must be able to access and use the database for the purposes and in the way set out in the agreement with the rightholder, even if such access and use necessitate performance of otherwise restricted acts.

It follows from this explanation that **right holders can shape the modalities of use in the contractual terms that accompany the making available or distribution of the database.** Hence, the issue of a potential contractual restriction of use for research purposes reappears (cf. 3.1.6). **It is also conceivable that right holders employ technological measures to control access to the database.** The circumvention of these technological measures against the will of the right holder is then likely to erode the status of “lawful user” which is a precondition for invoking the exception in favour of scientific research (cf. 3.1.5).

In addition to these recurring themes that have already been explored in the copyright analysis, **Article 9(b) DBD raises several issues that impact research use of data.** In particular, it must not be overlooked that the use permission following from Article 9(b) DBD covers less exclusive rights than the copyright exception in Article 5(3)(a) ISD. An exemption from the obligation to seek the right holder’s authorisation for acts of making extracted data available within a broader research group or to the general public is sought in vain (following subsection 3.2.1). As a potential counterbalance, CJEU jurisprudence may nonetheless provide breathing space for use in the context of scientific research that does not depend on the specific exception laid down in Article 9(b) DBD (3.2.2). **The analysis of sui generis database rights, thus, yields mixed results. While the specific research exemption in this regulatory framework does not offer much flexibility, case-law developments point in the direction of growing support for the use of data resources in the context of scientific research** (3.2.3).

3.2.1 Data Sharing

As explained in section 3.1, **the exemption of scientific research use laid down in Article 5(3)(a) ISD covers not only the right of reproduction⁹⁴ but also the right of communication to the public and the right of making available to the public.⁹⁵** In other words, EU copyright law offers researchers the opportunity to copy literary and artistic works for the purpose of gathering data. It also offers room to share data resources and research results within the group of researchers belonging to a research consortium and the broader public, even if these data resources or research results include protected elements of literary and artistic works.

⁹⁴ Article 2 ISD.

⁹⁵ Article 3 ISD.

The exemption of scientific research use following from Article 9(b) DBD, by contrast, only covers the right of extraction. This exclusive right of the database maker is defined in Article 7(2)(a) DBD as follows:

“extraction” shall mean the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form;...

From the outset, the use permission following from Article 9(b) DBD is thus incomplete in comparison with the scope of Article 5(3)(a) ISD. As the definition of “extraction” shows, this type of database use corresponds with acts of reproduction in a copyright context. Article 9(b) DBD, however, does not cover acts of re-utilising data extracted from a protection database:

“re-utilization” shall mean any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission.

Hence, EU sui generis database law does not provide for a use permission that corresponds with the limitation of the right of communication to the public and making available to the public in copyright law. As Professor Estelle Derclaye has pointed out, **this asymmetry is likely to pose particular difficulties in the context of research projects:**

The teaching and research exceptions in the two chapters [of the Database Directive] are identical except on one crucial point: in the sui generis right chapter, it is limited to extraction while in the copyright chapter (Art. 6(3) [DBD]) the lawful user can “use” a substantial part of the database, so the term is much broader and includes all restricted acts. The corresponding exception in the sui generis right chapter is therefore far more restricted and in effect quasi unusable since to teach and research one almost always has to communicate to the public.⁹⁶

In particular, the problem of a “quasi unusable”⁹⁷ exception to the sui generis database right can arise in two contexts. On the one hand, the lack of an entitlement to make protected elements of a database available to the public can lead to a situation where researchers in a larger consortium are inhibited from sharing data resources (extracted from a protected database) with colleagues. **On the other hand,** the missing limitation of the right of making available to the public will prevent researchers from sharing research results with the broader academic community – or the public at large – if these research results contain protected elements of a database.

As to the first problem scenario (no sharing of data resources within a research consortium), it must be considered that the concept of “public” in EU copyright and database law does not require a particularly large group of persons. In the context of broadcasting and cable retransmission, the CJEU has held that the term “public” refers to an “indeterminate number” of persons.⁹⁸ As a general rule, the existence of a public can be assumed when “a fairly large number of persons are involved.”⁹⁹ This CJEU jurisprudence, however, only describes a clear-cut case in which the existence of a public can readily be assumed. The minimum threshold for assuming a relevant public is set at a substantially lower level in EU Member State law. In national

⁹⁶ E. Derclaye “The Database Directive”, in: I. Stamatoudi/P. Torremans (eds.), *EU Copyright Law – A Commentary*, 2nd ed., Cheltenham: Edward Elgar 2021, §9.65 (clarifications in brackets added by the author).

⁹⁷ Derclaye, id., §9.65.

⁹⁸ CJEU, 7 December 2006, case C-306/05, SGAE/Rafael Hoteles, para. 37.

⁹⁹ CJEU, id., para. 38.

law and practice, it is usually sufficient that protected material is made available to a circle of persons that is larger than the immediate social environment – larger than the limited group of persons with direct personal connections, such as family members and close friends.¹⁰⁰ The CJEU came closer to this minimum threshold in *SCF/Marco del Corso* where it described the public as a circle of persons “not restricted to specific individuals belonging to a private group”¹⁰¹ and clarified that “the concept of public encompasses a certain de minimis threshold, which excludes from the concept groups of persons which are too small, or insignificant.”¹⁰² In the concrete case, these considerations of the Court concerned the circle of persons present in a private dental practice.¹⁰³

In the light of this case law and the national traditions in EU Member States, it cannot be ruled out that the circle of researchers belonging to a broader research consortium, such as a group of researchers consisting of several teams in different EU Member States, constitutes a relevant public in the sense of copyright and sui generis database law. Accordingly, the sharing of protected database contents within this circle of researchers amounts to an act of making available to the public. As the research exemption in Article 9(b) DBD does not cover the re-utilisation – the making available to the public¹⁰⁴ – of protected database contents, this use falls outside the scope of the use privilege and requires an authorisation for each protected database element. Quite clearly, the transaction costs and licensing budget required for this rights clearance task can easily frustrate the sharing of protected data resources within a research group. As a result, the harmonisation of different datasets, such as the harmonisation of datasets stemming from teams in several Member States, becomes difficult, if not impossible.

The same can be said about initiatives to share data resources with a more general public, such as the academic community with an interest in the research (the second problem scenario mentioned above). In this case, the missing entitlement to share database material can pose particular problems from the perspective of academic integrity. If research results and underlying data resources cannot be made available because they contain protected database elements, it is hardly possible to check the replicability of the scientific analysis and verify research results.

The practical impact of the missing entitlement to share protected database elements for the purposes of scientific research must not be underestimated in the light of the broad database concept underlying the Database Directive. Any collection of independent works, data or other materials is eligible for sui generis database protection when it is arranged in a systematic or methodical way and individually accessible by electronic or other means.¹⁰⁵ Moreover, sui generis database protection requires a qualitatively and/or quantitatively substantial investment in the obtaining, verification or presentation of database contents.¹⁰⁶ The threshold for assuming a substantial investment, however, need not be particularly high. In *Directmedia/Universität Freiburg*, for instance, the CJEU dealt with a case where the national judge had deemed an investment of 34.900€ sufficient to assume sui generis database protection.¹⁰⁷

Considering the low threshold for asserting sui generis database protection, the lack of an entitlement to share protected data resources can play a role in

¹⁰⁰ I. Stamatoudi/P. Torremans/S. Karapapa, “The Information Society Directive”, in: I. Stamatoudi/P. Torremans (eds.), *EU Copyright Law – A Commentary*, 2nd ed., Cheltenham: Edward Elgar 2021, §11.35. For example, see § 15(3) of the German Copyright Act (*Urheberrechtsgesetz*); Article L 122-5(1) of the French Intellectual Property Code (*Code de la propriété intellectuelle*); Article 12(4) of the Dutch Copyright Act (*Auteurswet*).

¹⁰¹ CJEU, 15 March 2012, case C-135/10, *SCF/Marco del Corso*, para. 85.

¹⁰² CJEU, *id.*, para. 86.

¹⁰³ CJEU, *id.*, para. 30.

¹⁰⁴ Cf. the definition in Article 7(2)(b) DBD.

¹⁰⁵ Article 1(2) DBD.

¹⁰⁶ Article 7(1) DBD.

¹⁰⁷ CJEU, 9 October 2008, case C-304/07, *Directmedia/Universität Freiburg*, para. 24.

various research contexts. The problem is not limited to big data collections that are exploited commercially, such as specific image databases or repositories with academic publications. By contrast, use restrictions can arise from the sui generis database right with regard to much less comprehensive databases, such as open website contents that is structured in a systematic and methodical way. Online collections of addresses, hyperlinks, images, blog posts etc. may attract sui generis database protection. In national law, protection has been confirmed, for instance, with regard to a website that referred to various web pages by means of hyperlinks. Protection has also been assumed in the case of an indexed overview of vacancy advertisements.¹⁰⁸ Therefore, it cannot be excluded that the scraping of freely accessible online resources includes data collections that must be qualified as protected databases in the sense of the Database Directive. In other words: even limited web scraping activities can raise the question whether the research exemption in Article 9(b) DBD allows for the sharing of data resources that have been obtained in this way.

It is also important to note that a different arrangement and structuring of collected database elements does not change the equation. The making available of these data resources within a broader research group, the academic community, or the public at large, can still amount to an infringement of the reutilisation right of the database owner. It has already been established in CJEU case law that a different arrangement or organisation of elements of a database is not sufficient to rebut infringement arguments. The Court pointed out that:

bearing in mind the technical possibilities of reorganisation which are possible with electronic databases, the fact that all or part of the contents of a database protected by the sui generis right is found in a modified form in another database does not, as such, preclude a finding that there has been extraction.¹⁰⁹

This statement of the Court concerns the database owner's right of extraction. However, it is also relevant to the assessment of an encroachment upon the right of making available to the public – "re-utilization" in sui generis database terminology.¹¹⁰ A different arrangement of database elements in data collections for scientific research does not exclude a finding that a substantial part of a protected database has been taken. The moment substantial database elements – contained in a data collection for research – are shared among members of a relevant public, such as a larger group of researchers belonging to a research consortium, the database owner's right of re-utilisation may be infringed. Article 9(b) DBD does not immunise researchers against this form of infringement. To avoid lawsuits, it is thus necessary to obtain licenses for the sharing of protected database contents. In research projects requiring input from various data sources, such as a large number of websites, the transaction costs and licensing budget necessary for rights clearance can put an end to the research project as a whole.

3.2.2 Scope of Exclusive Rights

Despite the limited scope of the research exemption in Article 9(b) DBD, case-law developments in the field of the sui generis database right may still offer some support for the sharing of protected database elements without prior

¹⁰⁸ Cf. J.H. Spoor/D.W.F. Verkade/D.J.G. Visser, *Auteursrecht, portretrecht, naburige rechten en databankenrecht*, 4th ed., Deventer: Wolters Kluwer 2019, 792-797; P.B. Hugenholtz, "Database Directive", in: T. Dreier/P.B. Hugenholtz (ed.), *Concise European Copyright Law*, 2nd ed., Alphen aan den Rijn: Wolters Kluwer 2016, 390.

¹⁰⁹ CJEU, 5 March 2009, case C-545/07, *Apis/Lakorda*, para. 48.

¹¹⁰ Article 7(2)(b) DBD.

authorisation of the right holder. The exclusive rights granted in sui generis database law – the rights of extraction and re-utilisation¹¹¹ – put the right holder in a position which is in many ways comparable to the position following from the grant of the right of reproduction and the right of communication to the public in copyright law.¹¹² In particular, the exclusive rights of database makers encompass the permanent or temporary transfer and making available to the public of all or part of the contents of a protected database.¹¹³ Acts of extraction and re-utilisation amount to infringement when elements are taken that constitute a substantial part of a database from a quantitative or qualitative perspective.¹¹⁴ The qualitative assessment requires an assessment of the investment which the right holder had to make in order to create that part of the database.¹¹⁵

However, **the taking of database contents need not always fall within the scope of sui generis database protection.**¹¹⁶ According to recent case law of the CJEU, the unauthorised use of substantial parts of a database only constitutes an infringement if the additional condition is met that the maker of the database is deprived of income that must be deemed necessary to cover the costs of his investment:

It follows that such a transfer of the substantial contents of the databases concerned and such a making available of those data to the public, without the consent of the person who created them, are, respectively, measures of extraction and re-utilisation of those databases, prohibited by Article 7(1) of Directive 96/9, provided that they have the effect of depriving that person of income intended to enable him or her to redeem the cost of that investment.¹¹⁷

Hence, **the extraction or re-utilisation of substantial parts must entail the risk that the investment in the creation of the database may not be recouped.**¹¹⁸ Arguably, this gatekeeper criterion – requiring the frustration of amortisation options – offers an alternative basis for the exemption of scientific research from the control of database owners. It can serve as an alternative because it prevents a finding of infringement from the outset. If no infringement can be found because the research use does not inhibit the right holder from redeeming his investment, there is no need for researchers to invoke the research exemption in Article 9(b) DBD. A priori, the use in a research setting has no relevance in the sense of sui generis database law.

More specifically, it is conceivable in the light of the described CJEU case law to take the special purpose of scientific research into account when determining the scope of database protection and deciding on a potential prohibition of data sharing. After all, scientific research concerns use that does not aim at substituting demand for the primary database. This starting point – the absence of parasitic

¹¹¹ Article 7(2) DBD.

¹¹² Articles 2 and 3 ISD.

¹¹³ Article 7(2) DBD.

¹¹⁴ Article 7(1) DBD. Cf. P.B. Hugenholtz, "Database Directive", in: T. Dreier/P.B. Hugenholtz (ed.), *Concise European Copyright Law*, 2nd ed., Alphen aan den Rijn: Wolters Kluwer 2016, 405-406.

¹¹⁵ CJEU, 5 March 2009, case C-545/07, *Apis/Lakorda*, para. 68.

¹¹⁶ Cf. R. Ducato/A. Strowel, "Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out", *European Intellectual Property Review* 43 (2021), 322 (335), who develop a further line of argument against the invocation of the sui generis database right based on the consideration that research use may remain confined to a mere act of "consultation".

¹¹⁷ CJEU, 3 June 2021, case C-762/19, *CV-Online Latvia/Melons*, para. 37. Cf. M. van Eechoud, "Database Rights in the EU's Data Strategy: A Question of Sport?", in: M.R.F. Senftleben/J. Poort et al. (eds.), *Intellectual Property and Sports – Essays in Honour of Bernt Hugenholtz*, The Hague/London/New York: Kluwer Law International 2021, 251 (257-258); E. Derclaye/M. Husovec, "Sui Generis Database Protection 2.0: Judicial and Legislative Reforms", *European Intellectual Property Review* 44 (2022), forthcoming, available at: <https://ssrn.com/abstract=3964943> (last visited on 14 April 2022).

¹¹⁸ CJEU, id., para. 46.

competition¹¹⁹ – offers the possibility to argue that use in the context of scientific research, a priori, does not constitute infringing use within the meaning of the Database Directive. Even the making available of database contents to the public does not aim at providing alternative access to the database and undermining the market for the original collection of data.¹²⁰ Instead, database elements are shared within a group of researchers to jointly carry out the research project. Database contents may also be shared within the academic community to allow for replicability studies that confirm the validity of research results. Insofar as final research publications contain traces of protected databases, these data are also not shared to erode the primary market for the database. Instead, the publication serves the purpose of informing the public about the outcome of the research project. Because of the specific research context – the creation of new knowledge; not the erosion of the market for the database – researchers may thus escape the verdict of infringement. As explained, the CJEU requires right holders to demonstrate that the unauthorised use creates the risk of frustrating amortisation options.

In the absence of a taking of substantial parts, use in the context of scientific research may still amount to infringement if the research requires the repeated and systematic extraction or re-utilisation of insubstantial parts of a database.¹²¹ This latter form of infringement, however, can only be assumed when the use conflicts with a normal exploitation of the database or prejudices the legitimate interests of the database maker in an unreasonable manner.¹²² Thus, the regulatory framework itself sets forth the infringement condition that the unauthorised use erode the market for the primary database and deprive the database owner of opportunities to redeem his investment. As explained above and in the context of the copyright analysis (section 3.1.4), research use is unlikely to cause this type of detriment. It is unlikely to conflict with a normal exploitation or culminate in an unreasonable prejudice.¹²³

However, it would also be an exaggeration to present the unwritten case-law requirement of use that frustrates amortisation opportunities as an ultimate remedy for the lack of an exception to the sui generis database right that covers the making available of data resources in the context of scientific research (cf. the discussion of the scope of Article 9(b) DBD in the preceding section). While the infringement requirement following from CJEU jurisprudence, as explained, may avert a finding of infringement in several cases, this solution seems less reliable and robust than a statutory use permission in favour of researchers that encompasses not only acts of copying but also acts of sharing database contents with researchers in a larger project, the academic community or the public at large. **The case-law approach asking whether research use deprives the database owner of income necessary to recoup his investment has obvious limits.** Whenever a right holder can establish that use of a database in a scientific context constitutes an important segment of the database commercialisation strategy, the counterargument of scientific use not curtailing amortisation options will fail. **The described case-law solution, thus, is ineffective in respect of databases that are specifically made for academic use,** such as collections of scientific articles, and databases that generate substantial income because they are often used in scientific research.

¹¹⁹ Cf. CJEU, 19 December 2013, case C-202/12, Innoweb/Wegener, para. 48-53.

¹²⁰ As to the criterion of alternative access that makes the consultation of the original database obsolete, see CJEU, id., para. 48-53.

¹²¹ Article 7(5) DBD.

¹²² Article 7(5) DBD.

¹²³ R. Ducato/A. Strowel, "Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out", *European Intellectual Property Review* 43 (2021), 322 (336); C. Geiger/G. Frosio/O. Bulayenko, "Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations", *International Review of Intellectual Property and Competition Law* 2018, 814 (825).

3.2.3 Mixed Results

On balance, the analysis of sui generis database law yields mixed results. On the one hand, the EU acquis in this area contains several elements that have already been identified as sources of legal uncertainty in the copyright analysis (section 3.1). As Article 5(3)(a) ISD, Article 9(b) DBD raises the question of a potential illustration requirement. It may restrict collaborations with the private sector by requiring use for a non-commercial purpose. In comparison with Article 5(3)(a) ISD, Article 9(b) DBD even imposes further constraints on research use because it only exempts acts of copying from the control of database owners. The right of making available to the public – which may be affected when protected database elements are shared with members of a larger research consortium, the academic community or the public at large – falls outside the reach of the use privilege from the outset.

On the other hand, developments in CJEU jurisprudence point in the direction of a general immunisation of research use against allegations of sui generis database infringement. For a database owner to have success with an infringement claim against researchers, he must establish that use for the purpose of scientific research deprives him of income intended to enable him to redeem the investment made in the database. Arguably, the right holder will have difficulty to produce relevant evidence unless the database at issue has specifically been created for academic use or is often used for research purposes. The income-related infringement criterion stemming from CJEU case law, thus, may enhance the freedom of researchers to use protected database contents without any need to rely on the exemption of scientific research in Article 9(b) DBD. Applying the gatekeeper criterion that requires a loss of substantial income opportunities, judges may be able to deny prima facie sui generis database infringement in many cases.

3.3 Digital Single Market Directive

With the adoption of the Directive on Copyright and Related Rights in the Digital Single Market (“DSMD” or “DSM Directive”),¹²⁴ the EU legislator has added an important new use privilege to the canon of copyright provisions that seek to remove barriers to scientific research use. From the outset, the legislative initiative aimed at the elimination of legal uncertainty to encourage the use of TDM tools:

The need to better reflect technological advances and avoid uneven situations in the single market is also clear with text-and-data mining (TDM), through which vast amounts of digital content are read and analysed by machines in the context of science and research. The lack of a clear EU provision on TDM for scientific research purposes creates uncertainties in the research community. This harms the EU’s competitiveness and scientific leadership at a time when research and innovation (R&I) activities within the EU must increasingly take place through cross-border and cross-discipline collaboration and on a larger scale, in response to the major societal challenges that R&I addresses.¹²⁵

¹²⁴ Directive 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC, OJ 2019 L 130, 92.

¹²⁵ European Commission, 9 December 2015, *Towards a Modern, More European Copyright Framework*, Document COM(2015) 626 final, 7. For an analysis of the impact of the copyright framework on TDM research, see C. Handke/L. Guibault/J.-J. Vallbé, “Copyright’s Impact on Data Mining in Academic Research”, *Managerial and Decision Economics* 42 (2021), 1999 (1999-2016).

Accordingly, **Articles 3 and 4 DSMD set forth two exceptions to copyright, related rights and database protection that can play an important role in the context of research projects** that require the extraction of data from protected literary and artistic works and/or databases. Addressing scientific research directly, **Article 3(1) DSMD sets forth an obligation for Member States to exempt from copyright, related rights and sui generis database protection:**

reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access.

In several respects, this new use privilege overcomes obstacles to research use that have been identified in the analysis of the Information Society Directive and the Database Directive in the preceding sections. In contrast to the optional limitations of protection following from Article 5(3)(a) ISD and Article 9(b) DBD, **the new TDM provision in Article 3(1) DSMD is a mandatory “shall” provision.** Hence, all Member States are bound to implement this use privilege in their domestic legislation. This minimises the risk of non-implementation in individual countries and divergent national approaches (cf. subsection 3.1.2). Moreover, Article 2(2) DSMD contains a harmonised TDM definition:

“text and data mining” means any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations.

The existence of a harmonised TDM concept at EU level paves the way for the uniform application of the use privilege across Member States. **The DSM Directive also takes an important step in eliminating legal uncertainty that may arise from contractual stipulations seeking to restrict use for research purposes.** Article 7(1) DSMD declares any contractual provision unenforceable that runs counter to the TDM provision in Article 3 DSMD. Researchers are thus no longer exposed to the risk of use restrictions that may arise from contractual terms (cf. 3.1.6).

In several other respects, however, Article 3 DSMD raises complex questions instead of creating legal certainty. First, the involvement of private partners with a commercial orientation can lead to questions about the applicability of the TDM exemption that requires use for the purpose of scientific research (following subsection 3.3.1). **Second,** the regulatory framework surrounding Article 3 DSMD does not alleviate the problem of access restrictions that can arise from the use of technological protection measures (3.3.2). **Third,** Article 3(1) DSMD sets forth a lawful access requirement that can lead to an impoverishment of the data reservoir that is available in research consortia including partners from different institutions (3.3.3). **Fourth,** the mandatory exemption of scientific TDM research only covers acts of copying. As in the case of the sui generis database right (cf. 3.2.1), the absence of a permission to make protected source material available to a group of researchers or the broader academic community can make it difficult, if not impossible, to share and harmonise datasets within a research consortium and allow replicability studies to verify research results (3.3.4). **In sum, the new TDM provision for scientific research is an important addition to the optional research exemptions in Article 5(3)(a) ISD and Article 9(b) DBD.** In an attempt to establish a detailed regulation of scientific TDM research, **however, EU legislation has added several complicating factors to the equation that reduce the legal certainty for larger research collaborations with teams in several Member States and mixed consortia with cultural heritage institutions and commercial entities** (3.3.5).

3.3.1 No Requirement of Non-Commercial Use

In contrast to the provisions in copyright and sui generis database law, **Article 3(1) DSMD does not contain the requirement that the exempted use must be for a non-commercial purpose.**¹²⁶ **Nonetheless, a confinement to non-commercial TDM activities follows from the definition of the beneficiaries of the copyright exception.** Only research organisations and cultural heritage institutions can invoke Article 3 DSMD. According to Article 2(1) DSMD, **“research organisation” means**

a university, including its libraries, a research institute or any other entity, the primary goal of which is to conduct scientific research or to carry out educational activities involving also the conduct of scientific research:

(a) on a not-for-profit basis or by reinvesting all the profits in its scientific research; or

(b) pursuant to a public interest mission recognised by a Member State;

in such a way that the access to the results generated by such scientific research cannot be enjoyed on a preferential basis by an undertaking that exercises a decisive influence upon such organisation;...

While Recital 42 ISD excludes the organisational structure as a decisive factor for determining the non-commercial nature of research in the context of Article 5(3)(a) ISD (cf. section 3.1.3), the public interest mission or non-profit character of the organisation – in the sense of excluding any decisive influence of an undertaking – is a precondition for qualifying a research institution as a “research organisation” that may rely on Article 3 DSMD. Other organisational aspects are irrelevant to the assessment. Recital 11 DSMD clarifies that, if the requirement of a public interest mission or non-profit nature is fulfilled, the diversity of “legal forms and structures” of research institutions across EU Member States does not matter. For instance, it is not decisive whether a research institution is a body of public or private law.¹²⁷

This configuration of the concept of “research organisation” indicates that, **as long as the public interest or non-profit orientation of the research institution remains intact, Article 3 DSMD may be invoked for the purposes of scientific research even if industry funding has been secured for the underlying research project.** The explicit reference to “scientific research” does not militate against this conclusion. As already explained in section 3.1.3, the source of funding does not pose an obstacle if it does not alter the scientific setting. In the context of Article 3 DSMD, the yardstick for assessing a potential incompatibility is the requirement of academic independence that can be deduced from the definition in Article 2(1) DSMD. **The moment an undertaking obtains a “decisive influence upon such organisation,” a research institution loses the status of a “research organisation” that can benefit from the TDM privilege.** Once again, however, it must be emphasised that in contrast to Article 5(3)(a) ISD and Article 9(b) DBD, Article 3 DSMD does not set forth a requirement of research for a “non-commercial purpose.” This deliberate departure from previous research provisions in the EU copyright and database acquis broadens the scope of the use privilege substantially. Rightly understood, Article 3 DSMD remains applicable even if a research project aims at developing knowledge that can be translated into practical tools and brought to the market in the end. As Professor Eleonora Rosati has stated, the absence of a restriction to use for non-commercial purposes implies that:

¹²⁶ Article 5(3)(a) ISD; Article 9(b) DBD.

¹²⁷ Cf. E. Rosati, *Copyright in the Digital Single Market*, Oxford: Oxford University Press 2021, 42.

[t]he end goal of the scientific research activity at issue, for example whether it is ultimately aimed at generating a profit or is used for profit-making purposes, is not relevant per se.¹²⁸

The departure from earlier research provisions in Article 5(3)(a) ISD and Article 9(b) DBD also concerns the regulation of public-private partnerships. Recital 11 DSMD points out in this regard that current EU research policy encourages universities and research institutes to collaborate with the private sector. Against this background, **research organisations should also benefit from the exemption of TDM in Article 3 DSMD when their research activities are carried out in the framework of public-private partnerships:**

While research organisations and cultural heritage institutions should continue to be the beneficiaries of that exception, they should also be able to rely on their private partners for carrying out text and data mining, including by using their technological tools.¹²⁹

This statement in Recital 11 DSMD offers clarity with regard to the status of the research organisation: the fact that TDM support is offered by a private partner does not inhibit the research team from relying on Article 3 DSMD. **The practical implementation of the TDM process may be “outsourced” to a private undertaking in the research consortium.¹³⁰ While this clarification is good news for the research organisation itself, the status of the private partner remains unclear.** Which benefits can a private partner derive from the research project without thwarting the invocation of Article 3 DSMD? Can it be attractive for a private partner to join a scientific research consortium that must keep its TDM research within the conceptual contours of Article 3 DSMD?

The answer to these questions seems to lie, again, in the fact that Article 3 DSMD does not require that the research serve a non-commercial purpose. As explained, the ultimate goal of the research may be a practical implementation of research outcomes. The research can culminate in the development of tools that can be exploited commercially. As long as this practical, marketable offspring does not erode the independent, scientific setting of the research¹³¹ and the non-profit nature of the research organisation,¹³² Article 3 DSMD remains applicable.

The absence of a non-commercial use requirement in Article 3 DSMD is thus an important recalibration of the approach to use privileges for scientific research. In line with the objective to foster collaborations with the private sector,¹³³ the legislative decision to abandon this requirement offers additional room for public-private research partnerships and the translation of research results into practical tools that can be brought to the market in the form of new products and services.

¹²⁸ Rosati, id., 43; J. Griffiths/T. Synodinou/R. Xalabarder, *Comment of the European Copyright Society Addressing Selected Aspects of the Implementation of Articles 3 to 7 of Directive (EU) 2019/790 on Copyright in the Digital Single Market*, ECS 2022, 14-15, available at: <https://europeancopyrightsociety.org/> (last visited: 7 June 2022). However, see also the more nuanced positions taken by I. Stamatoudi/P. Torremans, “The Digital Single Market Directive”, in: I. Stamatoudi/P. Torremans (eds.), *EU Copyright Law – A Commentary*, 2nd ed., Cheltenham: Edward Elgar 2021, §17.94; T. Margoni/M. Kretschmer, “A Deeper Look Into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology”, *CREATE Working Paper 2021/7*, Glasgow: CREATE Centre 2021, 24-25.

¹²⁹ Recital 11 DSMD.

¹³⁰ R. Ducato/A. Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out”, *European Intellectual Property Review* 43 (2021), 322 (324).

¹³¹ See the requirement of use “for the purposes of scientific research” in Article 3 DSMD.

¹³² See the definition of “research organisation” in Article 2(1) DSMD.

¹³³ Recital 11 DSMD.

It is important to note that, in addition to the exemption of scientific TDM research in Article 3 DSMD, Article 4(1) DSMD contains a general exemption of TDM that is not limited to scientific research. Under this additional TDM rule, anyone, including private partners in a research consortium with academic institutions, may make copies of works or databases for the purposes of TDM and retain them as long as necessary for the TDM project.¹³⁴ However, right holders may prevent TDM on the basis of Article 4(3) DSMD by expressly reserving the use of their works and other materials in an appropriate manner, for instance, by employing machine-readable means.¹³⁵

In cases where a private member in a broader research consortium, or the research consortium as a whole, cannot rely on the scientific research exemption in Article 3 DSMD, **Article 4 DSMD can thus offer an alternative avenue to carry out TDM for research purposes without obtaining a prior authorisation from each individual right holder. With the rights reservation option in Article 4(3) DSMD, however, a complicating factor enters the picture that must not be underestimated.** In the case of content made available online, the proviso that right holders can exclude TDM via a machine-readable reservation of rights means that the researcher must take into account metadata, such as robots.txt files, but also the terms and conditions of a website or online service, in order to assess whether TDM is permitted with regard to a particular work or database.¹³⁶ In principle, technical safeguards, such as robots.txt files, constitute technical indications which crawler software and other TDM tools can read and “understand” easily. But the question is whether TDM tools are also capable of identifying and “interpreting” electronic caveats laid down in general terms and conditions of a website.¹³⁷ If an automated, machine-based processing of relevant terms and conditions is not possible, the rights reservation option is likely to render Article 4 DSMD de facto mute as an alternative basis for TDM research. TDM normally requires the availability of vast amounts of content and data in order to achieve reliable results that may finally lead to new scientific and technological advancements.¹³⁸ **The moment researchers are obliged to check use conditions and obtain permissions at the level of individual works or databases, the burden of rights clearance can easily put an end to the research project as a whole.**¹³⁹ From this perspective, **Article 4 DSMD does not offer a meaningful alternative to the scientific research rule in Article 3 DSMD.**

¹³⁴ As to the entitlement to retain TMD copies, see Article 4(2) DSMD.

¹³⁵ For a discussion of the additional room for copyright exceptions that can follow from a rights reservation option, see M.R.F. Senftleben, “How to Overcome the Normal Exploitation Obstacle: Opt-Out Formalities, Embargo Periods, and the International Three-Step Test”, *Berkeley Technology Law Journal Commentaries* 1, No. 1 (2014), 1 (12-16).

¹³⁶ See Recital 18 DSMD which clarifies that “[i]n the case of content that has been made publicly available online, it should only be considered appropriate to reserve those rights by the use of machine-readable means, including metadata and terms and conditions of a website or a service.” Cf. P.B. Hugenholtz, “Artikelen 3 en 4 DSM-richtlijn: tekst- en datamining”, *Tijdschrift voor auteurs-, media- en informatierecht* 2019, 167 (170).

¹³⁷ For a confirmation of the legal relevance of conditions made in this form, see Recital 18 DSMD.

¹³⁸ Cf. M.R.F. Senftleben/T. Margoni et al., “Ensuring the Visibility and Accessibility of European Creative Content on the World Market – The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 13 (2022), 67, para. 7.

¹³⁹ Cf. R.M. Hilty/H. Richter, “Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules – Part B: Exceptions and Limitations – Art. 3 Text and Data Mining”, *Max Planck Institute for Innovation and Competition Research Paper Series* 2017-02, 1.

3.3.2 Technological Protection Measures

As explained, **the absence of a requirement of non-commercial use enhances the flexibility of the scientific TDM rule in Article 3 DSMD. A further step towards a more flexible approach comes to the fore when exploring the rules concerning technological protection measures.**

On the one hand, Article 7(2) DSMD confirms the approach that has been introduced in the Information Society Directive. With the incorporation of Article 6(4) ISD by reference, Article 7(2) DSMD leaves little doubt that, **in principle, the protection of technological measures remains intact.** As discussed in section 3.1.5, Article 6(4) ISD is a safeguard clause for copyright exceptions that has been added to counterbalance the protection of technological measures following from Article 6(1) ISD (“[n]otwithstanding the legal protection provided for in paragraph 1...”).

In this regulatory framework known from the Information Society Directive, **Article 7(2) DSMD adds the TDM exemptions** laid down in Articles 3 and 4 DSMD **to the portfolio of copyright exceptions that, based on voluntary arrangements or Member State intervention,¹⁴⁰ ought to prevail over technological protection measures.** If no voluntary solutions evolve in the marketplace, Member States are bound to take “appropriate measures” to ensure that right holders make available to research organisations¹⁴¹ the means of benefiting from the TDM exemption in Article 3 DSMD “to the extent necessary to benefit from that exception or limitation and where [the research organisation] has legal access to the protected work or subject-matter concerned.”¹⁴² With regard to this aspect of the reference to Article 6(4) ISD made in Article 7(2) DSMD, the critical assessment made above (section 3.1.5) applies *mutatis mutandis*. **The safeguard clause of Article 6(4) ISD does not offer a sufficiently robust basis to prevent technological protection measures from impeding research activities.¹⁴³**

On the other hand, it must not be overlooked that Article 7(2) DSMD only incorporates selected subparagraphs of Article 6(4) ISD. Article 7(2) DSMD only refers to subparagraphs 1, 3 and 5. Hence, the problematic subparagraph 4 of Article 6(4) ISD (cf. section 3.1.5) does not apply in the context of TDM research. In practice, **this means that the TDM rule for scientific research (Article 3 DSMD) can be invoked** – and researchers should be able to rely on Member State measures to ensure the effectiveness of the TDM exemption – **even if protected works are made available on demand on the basis of contractual agreements.¹⁴⁴ The only remaining obstacle is the requirement of “lawful access”¹⁴⁵ or “legal access”¹⁴⁶** that can be found in Article 3 DSMD and the first subparagraph of Article 6(4) ISD. The legal framework for TDM does not entitle researchers to exact access to protected source material from right holders. Instead, it is a prerequisite for invoking the copyright exception that the research team ensure lawful access.¹⁴⁷ Once access has been obtained in a lawful way, however, the right holder can no longer prohibit TDM activities. As already pointed out, Article 7(1) DSMD renders any contractual stipulation seeking to

¹⁴⁰ Article 6(4), first subparagraph, ISD.

¹⁴¹ Article 2(1) DSMD.

¹⁴² Article 6(4), first subparagraph, ISD, as included in the regulation of TDM by Article 7(2) DSMD.

¹⁴³ Cf. R.M. Hilty, “Das Urheberrecht und der Wissenschaftler”, *Gewerblicher Rechtsschutz und Urheberrecht – International* 2006, 179 (186-187).

¹⁴⁴ Cf. E. Rosati, *Copyright in the Digital Single Market*, Oxford: Oxford University Press 2021, 57.

¹⁴⁵ Article 3 DSMD.

¹⁴⁶ Article 6(4), first subparagraph, ISD, as incorporated by virtue of the reference made in Article 7(2) DSMD.

¹⁴⁷ For a critique of this approach and an alternative regulatory proposal, see R.M. Hilty/H. Richter, “Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules – Part B: Exceptions and Limitations – Art. 3 Text and Data Mining”, *Max Planck Institute for Innovation and Competition Research Paper Series* 2017-02, 7-8.

undermine the freedom of TDM research unenforceable. The fact that Article 7(2) DSMD avoids a reference to subparagraph 4 of Article 6(4) ISD confirms this regulatory design.

Nonetheless, it seems unjustified to jump to the conclusion that the approach chosen in Article 7(2) DSMD is a major improvement in comparison with the problematic legal framework established in Article 6(4) ISD. **For at least two reasons, the protection of technological measures which right holders employ to protect copyrighted source material remains a stumbling block for scientific research and, in particular, a stumbling block for TDM-based research that requires an analysis of a large number of sources.**

First, it is doubtful that Member States are capable of devising satisfactory solutions on the basis of the first subparagraph of Article 6(4) ISD **to effectively remove research obstacles posed by technological measures that are used to protect copyrighted source material.**¹⁴⁸ While the first subparagraph of Article 6(4) ISD states that Member States should take appropriate measures to ensure that right holders pull down their technological fences and make protected works available to researchers, it is an open question how this result can be achieved in an efficient way in practice – without creating administrative burdens and transaction costs that make it unattractive for researchers to rely on this mechanism to enable TDM research.

Second, it must not be overlooked that the new rules for scientific TDM offer right holders an additional opportunity to introduce technological measures that can pose obstacles to research use. In line with Article 3(3) DSMD, right holders are allowed to:

apply measures to ensure the security and integrity of the networks and databases where the works or other subject matter are hosted. Such measures shall not go beyond what is necessary to achieve that objective.

Providing concrete examples, Recital 16 DSMD explains that the measures mentioned in Article 3(3) DSMD could be used to ensure that only persons having lawful access to data resources, such as work repositories and databases, can access them. The Recital also clarifies that measures could include an access verification through IP address validation or user authentication. As innocent as these measures may appear when considering that the scientific TDM rule in Article 3(1) DSMD requires “lawful access” anyway, it is noteworthy that Recital 16 DSMD itself foresees the danger of excessive and overly restrictive use of technological verification systems by right holders:

Those measures should remain proportionate to the risks involved, and 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.

¹⁴⁸ Cf. J. Griffiths/T. Synodinou/R. Xalabarder, *Comment of the European Copyright Society Addressing Selected Aspects of the Implementation of Articles 3 to 7 of Directive (EU) 2019/790 on Copyright in the Digital Single Market*, ECS 2022, 12-14, available at: <https://europeancopyrightsociety.org/> (last visited: 7 June 2022); R. Ducato/A. Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out”, *European Intellectual Property Review* 43 (2021), 322 (329); T. Margoni/M. Kretschmer, “A Deeper Look Into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology”, *CREATE Working Paper 2021/7*, Glasgow: CREATE Centre 2021, 26; C. Geiger/F. Schönherr, “The Information Society Directive”, in: I. Stamatoudi/P. Torremans (eds.), *EU Copyright Law – A Commentary*, 2nd ed., Cheltenham: Edward Elgar 2021, §11.109; C. Geiger/G. Frosio/O. Bulayenko, “Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations”, *International Review of Intellectual Property and Competition Law* 2018, 814 (837-838).

Evidently, it cannot be ruled out that right holders use the option of technological access control on the basis of Article 3(3) DSMD as a backdoor to impose limits on TDM research which, because of the ban on contractual restrictions following from Article 7(1) DSMD, cannot be introduced via contractual stipulations. Recital 16 DSMD openly addresses right holder concerns about “a potentially high number of access requests to, and downloads of, their works or other subject matter.” Without much imagination, it is conceivable in the light of this line of reasoning that right holders and researchers will not always agree on the intensity of data access and use that is appropriate and justifiable in the context of TDM.

Hence, **Article 3(3) DSMD may take away what has been given in Article 7(1) DSMD** (TDM exemption prevails over contractual restrictions) **and Article 7(2) DSMD** (TDM exemption survives the application of technological protection measures). With the option of implementing access verification systems to prevent excessive access requests and downloads, Article 3(3) DSMD adds a new complicating factor.¹⁴⁹ **Like the grant of protection for technological measures in Article 6 ISD** (cf. section 3.1.5), **it provides a statutory basis for the use of technological measures to police access to data resources and the volume of downloads. This can easily lead to practical hurdles** – in the form of unjustified access restrictions – **that may deprive researchers of the benefits that could accrue from the exemption of scientific TDM research in Article 3(1) DSMD.**

3.3.3 Requirement of Lawful Access

For research organisations to benefit from the exemption of scientific TDM research in Article 3(1) DSMD, they must ensure “lawful access” to the data resources they intend to mine. As already indicated in the preceding section, **this requirement is not new.**¹⁵⁰ With regard to the copyright exception for scientific research laid down in Article 5(3)(a) ISD, Article 6(4) ISD makes it a condition that researchers have “legal access” to benefit from measures that remove use barriers following from the application of technological protection measures. Similarly, the scientific research rule in Article 9(b) DBD can only be invoked when researchers have the status of “lawful users.” EU legislation in the field of copyright, related rights and sui generis database rights, thus, hesitates to grant researchers primary access rights. **The exemption of use for scientific research does not automatically imply that researchers can demand access to protected source material.** The moment a rights holder builds a technological or contractual fence around protected source material, researchers must first comply with the access conditions set by the right holder before they can rely on the use privilege for scientific use.¹⁵¹

¹⁴⁹ R. Ducato/A. Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out”, *European Intellectual Property Review* 43 (2021), 322 (329-330); C. Geiger/G. Frosio/O. Bulayenko, “Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations”, *International Review of Intellectual Property and Competition Law* 2018, 814 (836-837). For an analysis underlining the necessity of technical control options, see R.M. Hilty/H. Richter, “Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules – Part B: Exceptions and Limitations – Art. 3 Text and Data Mining”, *Max Planck Institute for Innovation and Competition Research Paper Series* 2017-02, 9.

¹⁵⁰ Cf. CJEU, 4 October 2011, cases C-403/08 and C-429/08, *Football Association Premier League and Others*, para. 168; CJEU, 17 January 2012, C-302/10, *Infopaq*, para. 42.

¹⁵¹ For a critique of this approach, see R.M. Hilty, “Das Urheberrecht und der Wissenschaftler”, *Gewerblicher Rechtsschutz und Urheberrecht – International* 2006, 179 (186-187); P.B. Hugenholtz, “Artikelen 3 en 4 DSM-richtlijn: tekst- en datamining”, *Tijdschrift voor auteurs-, media- en informatierecht* 2019, 167 (170); T. Margoni/M. Kretschmer, “A Deeper Look Into the

In principle, this element of the copyright acquis has been maintained in the DSM Directive. Article 7(1) DSMD adds an important nuance by preventing right holders from eroding the exemption of scientific TDM research with contractual stipulations. Apart from this clarification, however, the requirement of lawful access – in accordance with the access conditions set by the right holder – remains intact. As pointed out, the lawful access requirement made its way into Article 3(1) DSMD. **Recital 14 DSMD confirms the dependency of researchers on the access framework shaped by the right holder**, including subscription fees¹⁵² and volume restrictions that concern the number of persons covered by subscriptions:

Lawful access should be understood as covering access to content based on an open access policy or through contractual arrangements between rightholders and research organisations or cultural heritage institutions, such as subscriptions, or through other lawful means. For instance, in the case of subscriptions taken by research organisations or cultural heritage institutions, the persons attached thereto and covered by those subscriptions should be deemed to have lawful access.

In practice, this configuration of the lawful access requirement can pose substantial difficulties in research collaborations with partners from different institutions. Recital 14 DSMD reflects the current practice of access entitlements that are confined to individual research organisations and cultural heritage institutions and their respective staff members (“persons attached thereto”). The lawful access test is not aligned with access needs within an individual research consortium with members from different institutions. Recital 14 DSMD does not state that all members within a research project are deemed to have lawful access if at least one participating institution has a subscription. Instead, each individual participating institution must ensure lawful access by concluding access deals with right holders to ensure that the TDM exemption of Article 3 DSMD applies across all consortium members.

This configuration of the lawful access requirement can pose substantial difficulties. It multiplies access clearance obligations. Moreover, it imposes an obligation on researchers in consortia with partners from different institutions to check whether all members participating in TDM research have sufficient access credentials. As existing access contracts can differ from one participating institution to the other, the verification of congruent access entitlements can be a matter of quite some complexity.

If parallel access contracts are missing, **the lawful access requirement can impede TDM research that requires the involvement of several partners in a consortium.** A cultural heritage institution participating in a research project and possessing valuable access credentials covering a broad range of required data resources, for instance, will be unable to make these data resources available to TDM specialists of a consortium partner unless this other consortium partner has the same access entitlement. Similar problems can arise when a research project aims at exploring data from different cultural backgrounds and language groups. If 10 research organisations in a consortium seeking to explore 10 different cultural backgrounds are combining their forces to arrive at cross-cultural insights, the consortium partners are unlikely to have congruent access entitlements. An Estonian consortium partner, for instance, seems unlikely to have concluded access agreements covering Portuguese literature and vice versa. The Dutch consortium is unlikely to have access credentials for Estonian or Portuguese sources etc.

EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology”, *CREATE Working Paper 2021/7*, Glasgow: CREATE Centre 2021, 26-27.

¹⁵² R. Ducato/A. Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out”, *European Intellectual Property Review* 43 (2021), 322 (330); C. Geiger/G. Frosio/O. Bulayenko, “Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations”, *International Review of Intellectual Property and Competition Law* 2018, 814 (836-837).

Accordingly, the requirement of lawful access will prevent the 10 consortium partners from pooling their data resources and conducting the TDM research centrally. Instead, each individual consortium partner will have to carry out TDM with regard to its own individual data portfolio. Moreover, the consortium partners will have to ensure that their respective country portfolios have a comparable scope to arrive at consistent research results.

To alleviate the burden of lawful access credentials for each individual research institution, Recital 14 DSMD only confirms that “[l]awful access should also cover access to content that is freely available online.” While this is a welcome clarification, this sentence in **Recital 14 DSMD only safeguards a bare minimum of access entitlements. It would be a worrisome impoverishment of data access if researchers even had to clear access rights for each individual work or database that is freely made available online.**

3.3.4 Data Sharing

The described obstacles posed by the lawful access requirement are further increased by the fact that the scientific TDM rule in Article 3 DSMD only covers acts of copying: the copyright branch of the provision only concerns the right of reproduction; the sui generis database branch only affects the right of extraction. Researchers with the requisite access entitlement may thus make copies of works and databases to establish the corpus of data resources that is necessary for the intended TDM analysis. However, **Article 3 DSMD does not permit the researchers to share this corpus of protected works and databases with other researchers if this sharing amounts to an act of making works available to the public¹⁵³ and/or an act of re-utilising protected database contents.¹⁵⁴**

The difficulties arising from the lack of an entitlement to share data resources have already been described in the context of sui generis database rights. As explained in section 3.2.1, it cannot be ruled out that the circle of researchers belonging to a broader research consortium, such as a group of researchers consisting of several teams in different EU Member States, constitutes a relevant public in the sense of copyright and sui generis database law. Accordingly, **the sharing of TDM copies of protected works and database contents within this circle of researchers amounts to an act of making available to the public.** As the TDM exemption does not cover the making available of protected works or databases, this use falls outside the scope of Article 3 DSMD and requires an authorisation for each individual work and protected database element. Quite clearly, **the transaction costs and licensing budget required for this rights clearance task can easily frustrate the sharing of TDM copies within a research group.** As a result, **the harmonisation of different datasets for a TDM analysis**, such as the harmonisation of datasets stemming from teams in several Member States, **becomes difficult, if not impossible.**

The same can be said about initiatives to share TDM copies with a more general public, such as the academic community with an interest in the research. In this case, the missing entitlement to share TDM source material can pose particular problems from the perspective of academic integrity. **If TDM corpora cannot be made**

¹⁵³ Article 3 ISD.

¹⁵⁴ Article 7(2)(b) DBD. Cf. T. Margoni/M. Kretschmer, “A Deeper Look Into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology”, *CREATE Working Paper 2021/7*, Glasgow: CREATE Centre 2021, 25.

available, it is hardly possible to check the replicability of the TDM analysis and verify research results.¹⁵⁵

Transposing Article 3 DSMD into national law, the dilemmas arising from the missing entitlement to share TDM copies have not gone unnoticed. Seeking to offer a solution, **the German legislator**, for instance, supplemented the national counterpart of Article 3(1) DSMD with the following rule addressing the sharing of TDM copies:

Those authorised in accordance with subsections (2) and (3) [research organisations, cultural heritage institutions and individual researchers] and pursuing non-commercial purposes may make reproductions made pursuant to subsection (1) available to the following persons:

1. a specifically delimited circle of persons for their joint scientific research and
2. individual third persons for the purpose of monitoring the quality of the scientific research.

The making available to the public must be terminated as soon as the joint scientific research or the monitoring of the quality of the scientific research has been concluded.¹⁵⁶

This German example shows that EU copyright law offers flexibility to permit the sharing of TDM copies to some extent. However, it is important to point out that this solution is far from offering researchers in international consortia a reliable legal framework. First, the legal basis for the German solution is Article 5(3)(a) ISD.¹⁵⁷ Accordingly, the requirement of use for “non-commercial purposes” re-enters the picture, whereas this requirement has been abandoned in the context of the TDM regulation in Article 3 DSMD (cf. section 3.3.1). The sharing entitlement following from the German provision is thus incongruent – and potentially more limited – than the reproduction entitlement following from Article 3 DSMD. In consequence, it will be impossible to share the TDM corpus if the research does not fulfil the requirement of use for a non-commercial purpose. Second, reliance on Article 5(3)(a) ISD as a legal basis for the German provision means that this solution rests on an optional copyright exception in the EU acquis that may be unavailable in other Member States (cf. section 3.1.2). In the case of research consortia with partners in different Member States, the permission given by the German legislator may thus be insufficient to allow the sharing of TDM copies with researchers in other Member States where no comparable rule exists. Third, Article 5(3)(a) ISD – as a legal basis for the German provision – only covers the right of making available to the public in the field of copyright and related rights.¹⁵⁸ The corresponding provision in the field of sui generis database rights – Article 9(b) DBD – does not encompass acts of making

¹⁵⁵ R. Ducato/A. Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out”, *European Intellectual Property Review* 43 (2021), 322 (327); C. Geiger/G. Frozio/O. Bulayenko, “Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations”, *International Review of Intellectual Property and Competition Law* 2018, 814 (819).

¹⁵⁶ § 60d(4) of the German Copyright Act (*Urheberrechtsgesetz*), official English translation, available at: https://www.gesetze-im-internet.de/englisch_urhg/ (last visited on 10 April 2022). The clarification in brackets has been added by the author.

¹⁵⁷ M. Stieper, “Das Verhältnis der verpflichtenden Schranken der DSM-RL zu den optionalen Schranken der InfoSoc-RL”, *Gewerblicher Rechtsschutz und Urheberrecht* 2020, 1 (4); T. Dreier, “UrhG § 60d Text und Data Mining für Zwecke der wissenschaftlichen Forschung”, in: T. Dreier/G. Schulze (eds.), *Urheberrechtsgesetz – Kommentar*, 7th ed., Munich: C.H. Beck 2022, para. 2 and 10.

¹⁵⁸ Article 3 ISD.

available to the public.¹⁵⁹ The sharing entitlement following from the German provision remains limited to parts of a TDM corpus that enjoy copyright protection (cf. section 3.2.1). The sharing of protected database elements falls outside the scope of the copyright exception¹⁶⁰ and can only be deemed permissible on the basis of an interpretation of the sui generis database right that leaves use for research purposes unaffected from the outset (section 3.2.2).

3.3.5 Unanswered Questions

In sum, Article 3 DSMD offers an important addition to the general exemptions of scientific research in Article 5(3)(a) ISD and Article 9(b) DBD. With the introduction of a mandatory, harmonised rule for scientific TDM, the provision establishes a legal regime that can be applied universally across Member States. In connection with the ban on contractual restrictions in Article 7(1) DSMD, Article 3 DSMD overcomes problems that may arise from contractual terms seeking to exclude TDM activities. Abandoning the traditional requirement of non-commercial use, Article 3 DSMD also offers room for public-private partnerships. It enhances opportunities for the translation of research results into practical solutions that can be brought to the market.

A closer look at the configuration of Article 3 DSMD, however, also reveals several new problem areas that can lead to legal uncertainty. Instead of removing obstacles that can arise from the use of technological protection measures, Article 3(3) DSMD adds a new legal basis for the application of technological measures to police access requests and download volumes. With its focus on individual access permissions of research organisations and their staff members, the requirement of lawful access in Article 3(1) DSMD can easily pose obstacles to the sharing of data resources within a research consortium that consists of partners from different institutions. Similarly, the fact that Article 3(1) DSMD does not cover the right of making TDM copies available to the public impedes the sharing of data resources among members of larger consortia who constitute a relevant public in the sense of copyright law. This, in turn, may frustrate dataset harmonisation and replicability studies seeking to verify TDM results.

3.4 Data Act Proposal

As the analysis in the preceding sections has shown, the EU acquis in the field of copyright, related rights and sui generis database protection contains several rules that regulate the use of data stemming from literary and artistic works and protected databases. In the area of machine-generated raw data (see the different data categories described in section 2.2), a specific regime for research use has not evolved yet. Established CJEU case law indicates that this type of data fall outside the province of sui generis database protection.¹⁶¹ Confirming this approach, **the proposed Data Act seeks to keep machine-generated raw data outside the scope of the sui generis database right.**

From the perspective of scientific research, this is a positive development (following section 3.4.1). The clarification that raw data do not attract sui generis database protection prevents data holders from asserting an intellectual property right

¹⁵⁹ See the definition of "re-utilization" in Article 7(2)(b) DBD and the missing reference to reutilisation in Article 9(b) DBD.

¹⁶⁰ T. Dreier, "UrhG § 60d Text und Data Mining für Zwecke der wissenschaftlichen Forschung", in: T. Dreier/G. Schulze (eds.), *Urheberrechtsgesetz – Kommentar*, 7th ed., Munich: C.H. Beck 2022, para. 10.

¹⁶¹ CJEU, 9 November 2004, case C-203/02, British Horseracing Board/William Hill, para. 31-32.

against use in scientific research. It shields researchers from allegations of sui generis database infringement. **For a researcher to be able to use machine-generated raw data in the first instance, however, it is crucial to ensure data access. In this regard, the proposed Data Act contains two access avenues that take the data rights of users and public sector bodies as starting points (3.4.2). However,** a closer look at these access options reveals that **they may be unsatisfactory.** Access for research purposes is based on a secondary, indirect entitlement of researchers that depends on the way in which users and public sector bodies exercise their primary rights. If research institutions are public sector bodies themselves, the access right is still confined to specific emergency situations (3.4.3).

3.4.1 Exclusion of Sui Generis Database Protection

As already indicated (cf. section 3.2.2), **the legal framework for sui generis database protection contains a relatively broad concept of “database.”** In principle, a collection of independent works, data or other materials is eligible for sui generis database protection when it is arranged in a systematic or methodical way and individually accessible by electronic or other means.¹⁶² **Moreover, sui generis database protection requires a qualitatively and/or quantitatively substantial investment in the obtaining, verification or presentation of database contents.**¹⁶³

Given the elasticity of these protection requirements, it can be difficult to draw a clear boundary line between data collections that enjoy sui generis database protection and others that remain outside the scope of the sui generis protection regime. In particular, a discussion about the limits of the database concept has arisen **with regard to machine-generated raw data,** such as data stemming from “smart” devices that routinely and continuously amass data (cf. section 2.2).¹⁶⁴

In respect of this latter category of data, the CJEU has provided important guidelines in *British Horseracing Board/William Hill*. Discussing sui generis database protection for horse racing data, including the names of horses in a race, the date, time and name of the race and the name of the racecourse,¹⁶⁵ the Court clarified that for investment in the obtaining of database contents to be relevant to the assessment of eligibility for sui generis database protection, this investment had to relate to:

the resources used to seek out existing independent materials and collect them in the database, and not to the resources used for the creation as such of independent materials. The purpose of the protection by the sui generis right provided for by the directive is to promote the establishment of storage and

¹⁶² Article 1(2) DBD.

¹⁶³ Article 7(1) DBD.

¹⁶⁴ European Commission, 23 February 2022, *Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act)*, COM(2022) 68 final, Explanatory Memorandum, 4. As to the discussion on rights for data producers in this area, see J. Drexler et al., “Max Planck Institute for Innovation and Competition – Position Statement of 26 April 2017 on the European Commission’s ‘Public Consultation on Building the European Data Economy’”, *Max Planck Institute for Innovation & Competition Research Paper Series* 17-08, Munich: Max Planck Institute for Innovation and Competition 2017, available at: www.ssrn.com/abstract=2959924; A. Gartner/K. Brimsted, “Let’s Talk About Data Ownership”, *European Intellectual Property Review* 2017, 461 (464); W. Kerber, “A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis”, *Gewerblicher Rechtsschutz und Urheberrecht – International* 2016, 989; H. Zech, “A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data”, *Journal of Intellectual Property Law and Practice* 2016, 460.

¹⁶⁵ CJEU, 9 November 2004, case C-203/02, *British Horseracing Board/William Hill*, para. 19.

processing systems for existing information and not the creation of materials capable of being collected subsequently in a database.¹⁶⁶

Hence, the mere creation of data that may subsequently serve as source materials for a database is not sufficient to attract sui generis database rights.¹⁶⁷ Applying this insight to machine-generated data, it can be said that **the mere putting into operation of an automated process that constantly creates data is not sufficient to acquire sui generis database rights.**¹⁶⁸ Seeking to further clarify this point – in particular in light of the risk of an “accidental or unintended application of the sui generis right to databases containing machine-generated data”¹⁶⁹ – **Article 35 DAP states that:**

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

According to the definition in Article 2(2) DAP, “product” refers in this context to tangible, movable items that may be incorporated in an immovable item, and that obtain, generate or collect data concerning its use or environment. The product must be able to communicate the data via a publicly available electronic communications service. Its primary function must be different from the function of storing and processing data. In line with Article 2(3) DAP, a “related service” means a digital service, including software, which interacts with a product in such a way that its absence would prevent the product from performing one of its functions. Quite clearly, the exclusion of sui generis database protection in **Article 35 DAP, thus, focuses on machine-generated raw data.**

While the provision is intended to create further clarity, the specific regulatory design of Article 35 DAP may still give rise to doubts. The reason for this lies in the combination of a statement excluding sui generis database protection with an explanation of the underlying intention of the legislator. The second part of the provision (after the comma) excludes the invocation of sui generis database protection for machine-generated raw data. The first part, however, indicates that this is necessary to safeguard the rights of data access and data sharing laid down in Articles 4 and 5 DAP. Therefore, the question arises whether the exclusion statement in the second part of Article 35 DAP must be understood in an absolute sense. Is sui generis database protection for machine-generated raw data excluded categorically? Regardless of whether this question arises in the context of Articles 4 and 5 DAP?

Or does Article 35 DAP only seek to safeguard the access and use rights granted in Articles 4 and 5 DAP? Does sui generis database protection remain possible as long as it does not interfere with the access and use rights following from Articles 4 and 5 DAP? Evidently, this more restrictive interpretation of the exclusion statement in Article 35 DAP could lead to the result that sui generis database protection is only excluded to the extent to which this is necessary to safeguard the rights in Articles 4 and 5 DAP. **To find the right interpretation, recourse may be had to Recital 84 DAP:**

¹⁶⁶ CJEU, id., para. 31.

¹⁶⁷ CJEU, id., para. 32.

¹⁶⁸ Cf. R. Fisher/J. Chicot et al., *Study in Support of the Evaluation of Directive 96/9/EC on the Legal Protection of Databases*, Brussels: European Commission 2018, 111-115; M. Leistner, “Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform”, in: S. Lohsse/R. Schulze/D. Staudenmayer (eds.), *Trading Data in the Digital Economy: Legal Concepts and Tools*, Baden-Baden: Nomos 2017, 27 (28).

¹⁶⁹ European Commission, Explanatory Memorandum, id., 9.

In order to eliminate the risk that holders of data in databases obtained or generated by means of physical components, such as sensors, of a connected product and a related service claim the sui generis right under Article 7 of Directive 96/9/EC where such databases do not qualify for the sui generis right, and in so doing hinder the effective exercise of the right of users to access and use data and the right to share data with third parties under this Regulation, this Regulation should clarify that the sui generis right does not apply to such databases as the requirements for protection would not be fulfilled.

This Recital offers important insights. The last words “as the requirements for protection would not be fulfilled” reflect the assumption that, from the outset, the machine-generated raw data addressed in Article 35 DAP are not eligible for sui generis database protection. This premise is in line with relevant CJEU case law, such as the decision in *British Horseracing Board/William Hill*. **Article 35 DAP is thus of a declaratory nature.** It points out the absence of protection that may follow from the correct application of the protection requirements in the Database Directive anyway. **Viewed from this perspective, the words “where such databases do not qualify for the sui generis database right” in Recital 84 must be understood as a confirmatory statement** (“a priori, databases containing machine-generated raw data do not enjoy protection”). **It would be wrong to read this element of the Recital as an indication that there might be other cases where such databases, in fact, do qualify for the sui generis database right.**¹⁷⁰ **Nonetheless, some legal uncertainty remains.** As Professor Estelle Derclaye and Dr. Martin Husovec have pointed out:

The sui generis database protection undoubtedly can apply to *some* scenarios where IoT devices collect data about the use of the products, as defined in Article 2 [of the proposed Data Act]. If the installation of sensors can be viewed as an investment in *obtaining* the data in a given context, it cannot be ruled out that it constitutes a separable relevant investment in the CJEU’s test.¹⁷¹

The basic assumption that machine-generated raw data, by definition, are not eligible for sui generis database protection may thus prove to be less reliable than expected in the framework of the proposed Data Act.

3.4.2 Raw Data Access Regimes

An effective exclusion of sui generis database protection – following from CJEU jurisprudence and confirmed in Article 35 DAP – **would eliminate one obstacle to the use of data in the context of scientific research:** the use of machine-generated raw data would not amount to an infringement of sui generis database rights – at least as long as Member States refrain from introducing national protection regimes in the absence of harmonised sui generis database protection.¹⁷² **This clarification, however, does not automatically offer researchers access to machine-generated raw data.** In the absence of sui generis database rights, holders of machine-generated raw data

¹⁷⁰ Cf. European Commission, Explanatory Memorandum, id., 12.

¹⁷¹ E. Derclaye/M. Husovec, “Why the Sui Generis Database Clause in the Data Act is Counter-productive and How to Improve It?”, available at: <https://ssrn.com/abstract=4052390> (last visited on 14 April 2022), 2; I. Graef/M. Husovec, “Seven Things to Improve in the Data Act”, available at: <https://ssrn.com/abstract=4051793> (last visited on 14 April 2022), 4. Cf. R. Fisher/J. Chicot et al., *Study in Support of the Evaluation of Directive 96/9/EC on the Legal Protection of Databases*, Brussels: European Commission 2018, 30-31.

¹⁷² Cf. Derclaye/Husovec, id., 2-3, who recommend to exclude national Member State approaches by making it clear that “the databases can never be protected by the sui generis right nor can Member States enact national rights similar to the database sui generis right to protect them.”

collections may still employ trade secret protection,¹⁷³ technological protection measures and contractual obligations to block access to raw data resources.

Therefore, it is of particular importance to clarify whether the proposed Data Act ensures data access for the purposes of scientific research. In this regard, Article 1 DAP reveals that the supply of data for scientific research is not the primary purpose of the proposed new legislation. The Data Act seeks to make machine-generated data available to users, trade and business persons and, in cases of exceptional need, to public sector bodies.¹⁷⁴ Given this starting point, researchers can only obtain access by tapping into the data stream to users, or invoking the provisions relating to public bodies.

As to the first option – benefitting from the data stream to users – **Article 5(1) DAP makes it clear that users may request the sharing of data with a third party** “without undue delay, free of charge to the user, of the same quality as is available to the data holder and, where applicable, continuously and in real-time.” Recital 29 adds the further clarification that “[a] third party to whom data is made available may be an enterprise, a research organisation or a not-for-profit organisation.” EU legislation, thus, explicitly contemplates the possibility of users sharing data with researchers in the context of research projects.

The second option – data access for public bodies – rests on Article 14(1) DAP which entitles public sector bodies to request data in situations of exceptional need. Recital 56 points out that “[r]esearch-performing organisations and research-funding organisations could also be organised as public sector bodies or bodies governed by public law.” Hence, researchers in public sector research organisations may be able to rely on Article 14(1) DAP and obtain direct access via this avenue. **In addition, Article 21(1) DAP allows public bodies that receive data on the basis of Article 14(1) DAP to share these data “with individuals or organisations in view of carrying out scientific research** or analytics compatible with the purpose for which the data was requested.” In cases where a research institution is not organised as a public body itself, indirect data access may thus follow from a collaboration with an eligible public sector body.

Articles 14(1) and 21(1) DAP, however, only cover situations of exceptional need, in particular public emergency situations and scenarios where a public sector body depends on the data (which are not available on the market) to fulfil a specific public interest task.¹⁷⁵ **From the outset, the scope of this data access option is thus quite narrow.** In addition, Article 21(2) sets forth a non-profit requirement. Recital 68 explains this requirement as follows:

Individuals conducting research or research organisations with whom these data may be shared should act either on a not-for-profit basis or in the context of a public-interest mission recognised by the State. Organisations upon which commercial undertakings have a decisive influence allowing such undertakings to exercise control because of structural situations, which could result in preferential

¹⁷³ As to available protection instruments in the field of trade secrets, see Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, OJ 2016 L 157, 1. Cf. T. Aplin, “Right to Property and Trade Secrets”, in: C. Geiger (ed.), *Research Handbook on Human Rights and Intellectual Property*, Cheltenham: Edward Elgar 2015, 421-437; L. Specht-Riemenschneider, *Studie zur Regulierung eines privilegierten Zugangs zu Daten für Wissenschaft und Forschung durch die regulatorische Verankerung von Forschungsklauseln in den Sektoren Gesundheit, Online-Wirtschaft, Energie und Mobilität*, Bonn 2021, 21-22.

¹⁷⁴ Articles 1 and 2(7) DAP.

¹⁷⁵ Article 15 DAP; Recitals 56 and 62 DAP.

access to the results of the research, should not be considered research organisations for the purposes of this Regulation.

For regular scientific research – in the sense of research projects initiated independently by the academic community itself – **the public sector avenue** (Articles 14(1) and 21(1) DAP) **has little to offer in the light of the described preconditions for data access.** In the context of Article 21(1) DAP, some flexibility for broader use of data may potentially follow from use of the formula “carrying out scientific research or analytics compatible with the purpose for which the data was requested” in the provision. Arguably, research “compatible” with the purpose need not focus exclusively on remedying the emergency situation. The title of Article 21 DAP, however, leaves no doubt that the research must be conducted “in the context of exceptional needs”. Hence, the provision does not allow researchers to develop the research questions themselves because the research design must be aligned with the exceptional circumstances justifying the data request. Moreover, the emergency situation or other situation of exceptional need can hardly be foreseen. **Instead of being self-determined and following an autonomous research agenda, the research reacts to difficult circumstances that have arisen.**

3.4.3 Lack of Research Focus

In sum, the raw data rules following from the proposed Data Act can hardly be described as a particularly research-friendly regime. Robust access and use guarantees for scientific research are missing. The facilitation of research does not lie at the heart of the proposed new legislation. The norms that address data use for research purposes appear as accessory rules that add a research perspective to the primary access provisions for users and public sector bodies. In the proposed regulatory matrix, **the most promising access instrument for research teams is the option of obtaining data as a result of sharing requests which users make in favour of a research organisation as a third party. However, it is an open question whether, in practice, this access avenue – which depends on user collaboration – offers sufficiently broad data access to compile representative data samples that allow scientifically sound analytical work.**

4. Publicly Funded Research Data

The foregoing discussion of EU legislation in the field of copyright, related rights and sui generis database rights **has shown that the use of protected works and databases as sources for research data** (see the different data categories described in section 2.2) **raises complex legal questions.** Existing exemptions of research use require compliance with specific legal requirements (sections 3.1 and 3.2). Outside the mandatory TDM rules in the DSM Directive (section 3.3), these legal requirements may also differ from Member State to Member State.

Against this background, research teams – in particular consortia with partners in different countries – **will often have to consider alternatives, in particular licensing agreements,** to fill gaps in the use permissions following from the statutory exemptions of research use. A licensing approach avoids the legal uncertainty arising from the complexity of the legal prerequisites for invoking the statutory use privileges for scientific research.

As licensing is an important alternative, it is of particular importance that sufficient licensing opportunities are available to cover the broad spectrum of data resources that may be necessary for a research project. Moreover, it is

essential that licensing terms, including licensing fees, are favourable to academic research. In this latter regard, open access licenses can be particularly attractive. They may offer clarity about permitted modes of use without requiring a large budget for rights clearance and remuneration payments.

Evidently, the academic community itself can contribute to an improvement of the licensing landscape for scientific research. Making research data easily accessible for reuse, research teams can create a data reservoir that supports future research endeavours by alleviating the burden of rights clearance and data acquisition. **As already pointed out in section 2.1, two dimensions can be distinguished in the analysis of licensing initiatives:**

- **the input dimension concerns the datasets** (and related source material, such as protected works and databases) **that have served as a basis for the scientific analysis.** Access and reuse possibilities are particularly important in this respect to allow:
 - replicability studies that give the academic community the opportunity to verify research results and confirm conclusions and recommendations that have been based on the data;
 - future research teams to build on the data resources that have already been amassed in the context of previous research.
- **the output dimension concerns research results, such as publications and data collections, that are created in the framework of a research project.** In the research innovation cycle,¹⁷⁶ these research products constitute themselves data sources that can be put to good use in future research projects.

Considering the outlined benefits, it is of particular interest to take stock and analyse licensing initiatives that seek to foster access to, and reuse of, publicly funded research data. Before embarking on this analysis in section 4.2, however, **the following section 4.1 places the discussion in the broader context of data harmonisation initiatives** that seek to enhance licensing and content recommendation opportunities for the European creative industry (following section 4.1). In the concluding section 4.3, the issue of copyright data improvement will be revisited in the light of the insights from the discussion of open access licensing requirements in academic funding schemes, such as Horizon Europe.

4.1 Copyright Data Harmonisation

In the European Strategy for Data, the European Commission highlighted the EU's ambition "to acquire a leading role in the data economy."¹⁷⁷ At the same time, the Commission conceded that the EU would have to "increase its pools of quality data available for use and re-use."¹⁷⁸ In the creative industries,¹⁷⁹ this need for enhanced data

¹⁷⁶ For a discussion of the role of copyright law in enabling cyclic innovation in the literary and artistic sector, see M.R.F. Senftleben, *The Copyright/Trademark Interface – How the Expansion of Trademark Protection Is Stifling Cultural Creativity*, The Hague/London/New York: Kluwer Law International 2020, 26-35.

¹⁷⁷ European Commission, 19 February 2020, "A European Strategy for Data", Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Document COM(2020) 66 final, 1, available at: <https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy>.

¹⁷⁸ European Commission, id., 1.

¹⁷⁹ As to the contours of the cultural and creative sectors, see the definition provided in Article 2(1) of the Regulation 1295/2013 of 11 December 2013 establishing the Creative Europe Programme (2014 to 2020), available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013R1295> : "all sectors whose activities are based on cultural

quality and interoperability is particularly strong. **Without data improvement, unprecedented opportunities for monetising the wide variety of creative content in EU Member States and making this content available for new technologies, such as AI systems, will most probably be lost.**¹⁸⁰

Appropriate copyright data management and licensing infrastructures, however, are not only desirable to offer the creative industries the opportunity of exploiting the promising new market for AI training data. **More generally, improved copyright data, including accurate metadata with information on right holders, can be expected to enhance licensing opportunities** by increasing the visibility of content and lowering the transaction costs for the conclusion of agreements.¹⁸¹ **The improvement of copyright data also plays an important role in the field of automated content recommender systems.** Without appropriate metadata that enhance the visibility of European content, the lack of niche repertoire recommendations may be due to inaccurate or missing data rather than being the result of a discriminatory mainstream orientation of content recommender systems.¹⁸²

A scenario with the described beneficial developments in the area of repertoire licensing and content recommender systems, however, will only arise if the considerable problems in the field of copyright data creation and management can be overcome. **To better illustrate data obstacles in European creative industries, the situation in the music sector can serve as a starting point** (following section 4.1.1). Experiences in other creative industry sectors show that even bigger problems are lying ahead, for instance, in the field of visual arts (section 4.1.2).

4.1.1 Experiences in the Music Industry

The music segment of the creative industry offers several well-known examples of data infrastructures, such as the Common Information System (“**CIS**”) of the International Confederation of Societies of Authors and Composers (“**CISAC**”). With its different nodes in several regions of the world, the CIS-Net system and accompanying standards constitute a global tool seeking to facilitate music licensing and the distribution of revenues.¹⁸³ **In terms of data standardisation,** the International Standard Work Code (“**ISWC**”) of the music publishing industry,¹⁸⁴ the International Standard Recording

values and/or artistic and other creative expressions, whether those activities are market- or non-market-oriented, whatever the type of structure that carries them out, and irrespective of how that structure is financed. Those activities include the development, the creation, the production, the dissemination and the preservation of goods and services which embody cultural, artistic or other creative expressions, as well as related functions such as education or management. The cultural and creative sectors include inter alia architecture, archives, libraries and museums, artistic crafts, audiovisual (including film, television, video games and multimedia), tangible and intangible cultural heritage, design, festivals, music, literature, performing arts, publishing, radio and visual arts.”

¹⁸⁰ M.R.F. Senftleben/T. Margoni et al., “Ensuring the Visibility and Accessibility of European Creative Content on the World Market: The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 13 (2022), 67 (70-74).

¹⁸¹ Senftleben/Margoni et al., id., 79-81.

¹⁸² Senftleben/Margoni et al., id., 73-74.

¹⁸³ See <https://www.cisac.org/What-We-Do/Information-Services/CIS-Net>.

¹⁸⁴ ISWC has been developed by CISAC, in collaboration with ISO, as “a unique, permanent, and internally recognized reference number for the identification of musical works”. As an example of a further unique identifier system, see also GRiD (Global Release Identifier) which has been developed by IFPI. Cf. A. Katz, “The Potential Demise of Another Natural Monopoly: New Technologies and the Administration of Performing Rights”, *Journal of Competition Law and Economics* 1 (2005), 276.

Code (“ISRC”) of the recording industry, the Interested Party Information (“IPI”) number, and the International Standard Name Identifier (“ISNI”) **offer prime examples of existing initiatives** to enable the exchange of accurate data related to the identification of repertoire or related to the mitigation of ex post transaction costs that arise in relation to the operation of licensing agreements.

At the same time, these examples reveal data deficiencies and interoperability problems arising from different sets of metadata and different approaches to data identification and verification. **To this day, initiatives to harmonize ISWC and ISRC metadata and incorporate them into a single, comprehensive database have failed.** In the EU, former Commissioner Neelie Kroes launched a working group to stimulate the establishment of a Global Repertoire Database in 2008. While the working group participants, including producers, collecting societies and distribution platforms, arrived at recommendations on the way forward,¹⁸⁵ the project was abandoned in 2014.¹⁸⁶ Other unsuccessful attempts include the International Music Joint Venture in 2000, which was formed by several collecting societies in Europe and North America, and a project initiated by the World Intellectual Property Organization (“WIPO”) aiming at the establishment of a common rights database in 2011.¹⁸⁷

Despite discouraging past experiences, the quest for appropriate data improvement strategies continues. In the US, a new initiative to form a comprehensive database followed from the 2018 Music Modernization Act (“MMA”).¹⁸⁸ In Title I, the MMA established the Mechanical Licensing Collective (“MLC”) as a one-stop shop for obtaining music licenses.¹⁸⁹ The new licensing hub became operational on 1 January 2021 and has offered a US-wide platform for licence administration, enforcement and royalty processing since that time.¹⁹⁰ New initiatives have also been taken in Europe.¹⁹¹ The Technical Online Working Group Europe (“TOWGE”) brings together a large group of European collecting societies, music publishers and rights agencies developing a digital royalty processing system. TOWGE is based on a small group of direct licensors reporting back to local societies.¹⁹² An initiative with similar objectives has been taken by the Finnish collecting society Teosto. A collaboration between Teosto and the start-up company Mind Your Rights has led to the “Concertify” platform seeking to provide—on top of existing industry structures—an efficient and transparent cross-

¹⁸⁵ Cf. M. Isherwood, “Global Repertoire Database”, presented at: World Intellectual Property Organization, *Enabling Creativity in the Digital Environment: Copyright Documentation and Infrastructure*, WIPO Meeting wipo_cr_doc_ge_11, 13-14 October 2011, Geneva: WIPO 2011, available at: https://www.wipo.int/meetings/en/2011/wipo_cr_doc_ge_11/prov_program.html.

¹⁸⁶ Cf. P. Resnikoff, “Global Repertoire Database Declared a Global Failure”, *Digital Music News*, 10 July 2014, available at: <https://www.digitalmusicnews.com/2014/07/10/global-repertoire-database-declared-global-failure/>; S.F. Schwemer, *Licensing and Access to Content in the European Union. In Licensing and Access to Content in the European Union: Regulation between Copyright and Competition Law*, Cambridge: Cambridge University Press 2019, 68-73.

¹⁸⁷ Schwemer, id., 69-70.

¹⁸⁸ House Report 1551, Pub. L. 115-264.

¹⁸⁹ Cf. F. Lyons/H. Sun et al., *Music 2025 – The Music Data Dilemma: Issues Facing the Music Industry in Improving Data Management*, Newport: UK Intellectual Property Office 2019, available at: <https://www.gov.uk/government/publications/music-2025-the-music-data-dilemma>, 34.

¹⁹⁰ See <https://www.themlc.com/press/mechanical-licensing-collective-begins-full-operations-envisioned-music-modernization-act>. As to the underlying planning and preparations, see U.S. Copyright Office Library of Congress, MLC Comments in Reply to the Designation Proposal of the American Music Licensing Collective, Inc., Docket No. 2018-11, 21, available at: https://bw-98d8a23fd60826a2a474c5b4f5811707-bwcore.s3.amazonaws.com/photos/Proposed_MLC_-_Reply_Comments.pdf.

¹⁹¹ For a discussion of further data integration and harmonisation opportunities in the EU, see Norbert Gronau and Martin Schaefer, “Why Metadata Matters for the Future of Copyright”, *European Intellectual Property Review* 43 (2021), 488-494; Martin Schaefer, “Why Metadata Matter for the Future of Copyright”, *Kluwer Copyright Blog*, 27 November 2020, available at: <http://copyrightblog.kluweriplaw.com/2020/11/27/why-metadata-matter-for-the-future-of-copyright/>.

¹⁹² See <https://www.digitalmusicnews.com/2019/07/26/towge-digital-royalty-group/>.

border copyright licensing system. Concertify allows artists, copyright holders, including collecting societies, music publishers and event organisers to interact directly by using modules, such as a module for setlist reporting.¹⁹³ With the support of the Slovak Art Council, a collaboration between the collecting society SOZA and various stakeholders has led to the creation of a prototype for a comprehensive data and metadata database of the Slovak music repertoire. The consortium also created the prototype of a “Listen Local” recommender system that meets the requirements of the trustworthy AI recommendations of the High-Level Working Group on AI.¹⁹⁴ The accompanying feasibility study highlighted and quantified the problems that arise from incomplete copyright data in existing databases and commercial AI-solutions. For example, it demonstrated that at least 15% of Slovak, Estonian, Hungarian and Dutch works are unlikely to be ever exploited due to data problems.¹⁹⁵

4.1.2 Initiatives in Other Creative Industry Segments

Other sectors of the creative industry are facing similar data problems and have embarked on initiatives for data improvement, harmonisation and combination as well. In the field of book publishing, industry initiatives, such as the establishment of different e-book platforms and catalogues, play an important role. Flickr and Google Images offer a search option for material covered by creative commons licences.¹⁹⁶ Another example is the Entertainment Identifier Registry (“EIDR”), which is a universal unique identifier system for movie and television assets based on DOI technology.¹⁹⁷

As to standardisation, the International Standard Book Number (“ISBN”), the International Standard Serial Number (“ISSN”) for journals, the International Standard Music Number (“ISMN”) for notated music, and the International Standard Audiovisual Number (“ISAN”) for audiovisual works can serve as examples. Moreover, the standardisation work of the international EDItEUR group – leading to the “ONIX” family of standards¹⁹⁸ – is important in the field of books, e-books and serials.¹⁹⁹ With regard to the digital environment, the International DOI Foundation provides the aforementioned Digital Object Identifier (“DOI”) services and registration: a technical and social infrastructure for the registration and use of persistent interoperable identifiers for use on digital networks, including identifiers for literary and artistic works.²⁰⁰

In the area of visual arts, CISAC’s Visual Arts Council has extended its initial work on the right of resale and established an online licensing hub²⁰¹ under the umbrella of the International Council of Creators of Graphic, Plastic and Photographic Arts (“CIAGP”).²⁰² OnLineArt (“OLA”) is a one-stop shop for obtaining licenses for worldwide online use of works of visual art currently encompassing works of 60,000 artists.²⁰³ While existing initiatives in the visual arts sector – in particular museums and other cultural heritage institutions digitising works in their holdings – have substantially extended the data coverage of works of fine art, the situation in the field of photography and illustrations is

¹⁹³ See <https://www.mindyourrights.fi/>.

¹⁹⁴ See <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

¹⁹⁵ Daniel Antal, *Feasibility Study On Promoting Slovak Music in Slovakia and Abroad*, The Hague: Reprex 2020, available at: https://reprex.nl/publication/listen_local_2020/.

¹⁹⁶ See <https://www.theverge.com/2020/8/31/21408305/google-images-photo-licensing-search-results> (Google Images) and <https://www.flickr.com/creativecommons/> (Flickr).

¹⁹⁷ See <https://www.eidr.org/>.

¹⁹⁸ See <https://www.editeur.org/8/ONIX/>.

¹⁹⁹ See <https://www.editeur.org/2/About/#Intro>.

²⁰⁰ See <https://www.doi.org/>.

²⁰¹ See <https://www.cisac.org/What-We-Do/Creators-Relations/CIAGP>.

²⁰² See <http://www.ciagp.org/>.

²⁰³ See <https://onlineart.info/>.

much less transparent.²⁰⁴ Major visual arts libraries, such as Getty Images, may consistently use data management tools. The costs of properly documenting individual works, however, may be prohibitively high for smaller providers of photography and illustrations in the light of the low average value of individual works.²⁰⁵ In comparison with the status quo reached in the field of music, the process of harmonising, attaching and bundling metadata still seems in its infancy in the area of visual arts.

4.1.3 Different Settings for Data Improvement

The described experiences with existing data infrastructures in different branches of the creative industry, and current initiatives to arrive at harmonised and more comprehensive metadata, shed light on different settings for data improvement that would also be beneficial for researchers seeking to deduce data from literary and artistic works and related databases. As the discussion has shown, **the initiative to harmonise, combine and enhance the coverage of work-related data may come from different actors in the public and private sphere, and employ different tools of public and private law:**

- **legislation:** the MLC, for instance, is the result of US legislation that explicitly mandates the establishment of a nationwide licensing hub for mechanical music rights. In the EU, Article 17(4)(b) DSMD, indirectly, may have similar effects if the new obligations to license user-uploaded content and exchange work-related data for content moderation purposes leads to shared data standards and content identification libraries.²⁰⁶ In addition, the 2014 Directive on Collective Management of Copyright and Related Rights²⁰⁷ incentivises collecting societies to cooperate in licensing hubs for multi-territorial licensing of online rights in musical works and adopt voluntary industry standards to improve efficiency in the exchange of data;
- **public institutions:** impulses for the further development of the data infrastructure may also arise from non-legislative initiatives taken by national, European or international public bodies. The 2008 group of specialists working on a Global Repertoire Database, for instance, came together under the auspices of former Commissioner Neelie Kroes. WIPO initiated the aforementioned 2011 project for the establishment of a common rights database and has embarked on surveys on voluntary registration systems for copyright and related rights in 2005, 2010 and 2021;²⁰⁸

²⁰⁴ For a closer analysis of the particular situation and dynamics in the visual arts sector, see the report by Tristan Azzi and Yves El Hage, *Les métadonnées liées aux images fixes*, Paris: CSPLA 2021. Cf. also G. Pessach, "Museums, Digitization and Copyright Law: Taking Stock and Looking Ahead", *Journal of International Media and Entertainment Law* 1 (2007), 253.

²⁰⁵ Cf. R.A. Posner, "Transaction Costs and Antitrust Concerns in the Licensing of Intellectual Property", *John Marshall Review of Intellectual Property Law* 4 (2005), 325.

²⁰⁶ For a more detailed discussion of this point, see M.R.F. Senftleben/T. Margoni et al., "Ensuring the Visibility and Accessibility of European Creative Content on the World Market: The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive", *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 13 (2022), 67 (81-85). Cf. J. Reda/P. Keller, "A Proposal to Leverage Article 17 to Build a Public Repository of Public Domain and Openly Licensed Works", Kluwer Copyright Blog, 23 September 2021, available at: <http://copyrightblog.kluweriplaw.com/2021/09/23/a-proposal-to-leverage-article-17-to-build-a-public-repository-of-public-domain-and-openly-licensed-works/>.

²⁰⁷ Directive 2014/26/EU of the European Parliament and of the Council of 26 February 2014 on collective management of copyright and related rights and multi-territorial licensing of rights in musical works for online use in the internal market, *OJ* 2014 L 84, 72.

²⁰⁸ WIPO Survey of National Legislation on Voluntary Registration Systems for Copyright and Related Rights, prepared by the Secretariat, SCCR/13/2, November 9, 2015, available at: https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=52829; WIPO Second Survey on

- **private entities**: the initiatives that have led to TOWGE, the Concertify platform and SOZA's Listen Local platform show that private entities, in particular collecting societies, may play a decisive role in the further harmonisation and combination of copyright-related data. In addition, individual companies, such as Apple and Spotify, may obtain a market position that allows them to bring together an unprecedented volume of data and establish de facto data standards with a major impact on the sector. External technology start-ups also invest heavily in solutions based on blockchain or related technologies.²⁰⁹

For the analysis of data issues in the field of copyright and related rights, it is important to bear these different settings in mind. **To arrive at a substantial improvement of the data infrastructure, including the data infrastructure that is available for scientific research, it may be necessary to combine public and private initiatives and seek to offer incentives at different levels.**

4.2 Research Data Licensing

The distinction between different tools stemming from the public and private sphere in the preceding section offers a useful matrix for the discussion of data improvement and licensing initiatives that concern data resources resulting from publicly funded research. **As a prototype of open access initiatives, the licensing requirements relating to research data under Horizon Europe can serve as a reference point for the analysis.** As the discussion of open access licensing requirements will show (following section 4.2.1), the Horizon Europe rules can be qualified as a non-legislative initiative taken by a body of public law.

An assessment of the Horizon Europe provisions in the light of applicable copyright and database rules, however, reveals certain weaknesses of this instrument. **As a non-legislative measure, the Horizon Europe open access rules are subordinated to applicable copyright norms.** In consequence, gaps in statutory use privileges for scientific research, as discussed in the preceding chapter 3, can thwart the realisation of the objective to make research data more broadly available for the validation of research results and data reuse (section 4.2.2).

To arrive at more satisfactory solutions, it seems advisable to broaden the room in copyright and sui generis database law for open access initiatives in the academic sector that involve the digitisation of protected content and the creation of accurate metadata. As a countermove, data and metadata collections that come into existence as a result of research projects could be used more broadly to contribute to the improvement of the overall copyright data infrastructure with the ultimate goal of enhancing licensing and content recommendation opportunities for the creative industries (section 4.2.3). This "cross-fertilisation" of the copyright data infrastructure, however, should not become an end in itself. The enhancement of the data infrastructure for research must remain the primary objective of open access initiatives and collaborations with the creative industries should contribute to this ultimate goal.

Voluntary Registration and Deposit Systems (2010), available at: [WIPO Second Survey on Voluntary Registration and Deposit Systems](#); and WIPO Survey on Voluntary Copyright Registration Systems: Final Report, prepared by S. van Gompel/S. Massalina, Amsterdam, 23 April 2021, available at: https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_crr_ge_2_21/wipo_crr_ge_2_21_report.pdf.

²⁰⁹ Cf. B. Balász/D. Gervais/J.P. Quintais, "Blockchain and Smart Contracts: The Missing Link in Copyright Licensing?", *International Journal of Law and Information Technology* 26 (2018), 311 (311-336).

4.2.1 Horizon Europe Open Access Requirements

The Horizon Europe programme attaches particular importance to the open access distribution of research results, including research data. The Programme Guide explains in this respect:

Open access to generated research data is required under the premise “as open as possible as closed as necessary”, meaning that there can be exceptions to this. Data management plans are mandatory for all projects generating or reusing data and should be aligned with the [dissemination and exploitation] plan. Additionally, we recommend that you provide open access to research outputs beyond publications and data (e.g. software tools, models, apps, etc) and share them as early and openly as possible providing guidance for potentially interested users.²¹⁰

To foster open science practices, **it is mandatory under Horizon Europe to manage research data in line with the FAIR principles** of “Findability”, “Accessibility”, “Interoperability” and “Reusability.” This requires in particular the generation and use of data in accordance with a data management plan and, to the largest extent possible, the grant of open access. The underlying open science concept encompasses metadata that provide information on research outputs, tools and instruments. Access to results, including research data, is deemed essential to validate research outcomes, in particular conclusions drawn in scientific publications. Horizon Europe also strives for open access to research data in order to enhance opportunities for data reuse in future research projects.²¹¹

Within the specific communication, dissemination, open science and visibility rules set forth in Annex 5 to the Horizon Europe Model Grant Agreement, **the research data management obligations, including open access, rest on Article 17.**²¹² In accordance with this provision, **grant beneficiaries must manage research data generated in the framework of their research projects responsibly, in line with the FAIR principles and by taking all of the following actions:**

- **establish a data management plan** (“DMP”) (and regularly update it);
- as soon as possible and within the deadlines set out in the DMP, **deposit the data in a trusted repository**; if required in the call conditions, this repository must be federated in the EOSC in compliance with EOSC requirements;
- as soon as possible and within the deadlines set out in the DMP, **ensure open access – via the repository – to the deposited data, under** the latest available version of the Creative Commons Attribution International Public License (**CC BY**) or Creative Commons Public Domain Dedication (**CC 0**) **or a licence with equivalent rights, following the principle “as open as possible as closed as necessary”**, unless providing open access would in particular:
 - be against the beneficiary’s legitimate interests, including regarding commercial exploitation, or;
 - be contrary to any other constraints, in particular the EU competitive interests or the beneficiary’s obligations under this Agreement; if open access is not provided (to some or all data), this must be justified in the DMP.

²¹⁰ European Commission, *Horizon Europe (HORIZON) – Programme Guide*, Version 1.5, dated 1 February 2022, Brussels: European Commission 2022, 33-34 (explanation in brackets added by the author).

²¹¹ Programme Guide, id., 39.

²¹² European Commission, *Horizon Europe (HORIZON) Euratom Research and Training Programme (EURATOM) – General Model Grant Agreement EIC Accelerator Contract*, Version 1.1, dated 15 December 2021, Brussels: European Commission, 110.

- **provide information** via the repository about **any research output or any other tools and instruments needed to re-use or validate the data.**²¹³

Importantly, **Article 17 also contains rules on metadata.** To the largest extent possible in the light of competing legitimate interests or other constraints, metadata relating to deposited research data must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent. It follows from the FAIR principles that the metadata must be machine-actionable and follow a standardised format.²¹⁴ In line with Annex 5, they should, as a minimum, provide information on:

- datasets (description, date of deposit, author(s), venue and embargo);
- Horizon Europe or Euratom funding;
- grant project name, acronym and number;
- licensing terms;
- persistent identifiers for the dataset, the authors involved in the action, and, if possible, for their organisations and the grant.²¹⁵

For the dataset, digital object identifiers (DOIs) may be used as persistent identifiers; authors involved in the research should be identifiable via ORCIDs or ResearcherIDs. Ideally, persistent identifiers also indicate their research organisations (via ROR IDs) and the relevant grant (using grant DOIs).²¹⁶ Where applicable, the metadata must also include persistent identifiers for related publications and other research outputs.²¹⁷

Article 17 also includes provisions with regard to the validation of scientific publications. In particular, it requires that:

[w]here the call conditions impose additional obligations regarding the validation of scientific publications, the beneficiaries must provide (digital or physical) access to data or other results needed for validation of the conclusions of scientific publications, to the extent that their legitimate interests or constraints are safeguarded (and unless they already provided the (open) access at publication).

In the Horizon Europe Work Programme 2021-2022, the additional obligation regarding the validation of scientific publications applies to all calls.

With regard to the described obligations, **the Annotated Grant Agreement** clarifies that the requirements for research data management apply only to data that are generated in the course of the funded research project.²¹⁸ Data must be deposited as soon as possible after they have been produced or generated, and the necessary quality checks have been completed. To the extent technically feasible, the underlying concept of "data" includes raw data, especially if access to raw data is crucial to enable the reanalysis or reproducibility of research results and the reuse of data. The deposition of data in a trusted repository should take place in a way that provides value and context. More specifically, this means that metadata should be sufficiently rich to render the data findable and offer sufficient background information.²¹⁹

As to constraints that may make it impossible to offer open access to research data, the Annotated Grant Agreement explicitly refers to "data protection rules, privacy,

²¹³ Model Grant Agreement, id., 111 (emphases added by the author).

²¹⁴ European Commission, *EU Grants AGA – Annotated Model Grant Agreement EU Funding Programmes 2021-2027*, pre-draft, dated 30 November 2021, Brussels: European Commission 2021, 159.

²¹⁵ Model Grant Agreement, id., 111-112.

²¹⁶ Annotated Grant Agreement, id., 159.

²¹⁷ Model Grant Agreement, id., 111-112.

²¹⁸ Annotated Grant Agreement, id., 157.

²¹⁹ Annotated Grant Agreement, id., 158.

confidentiality, trade secrets, Union competitive interests, security rules, intellectual property rights.”²²⁰ A Creative Commons Public Domain Mark or equivalent should be applied to raw research data “unless the data meet the requirements to be protected by copyright/database right.”²²¹ In cases where datasets remain closed because of constraints but there are no compelling reasons that the related metadata should not be findable and accessible, at least the metadata should be published open access.²²²

4.2.2 Impact of Copyright and Sui Generis Database Protection

As the overview of Horizon Europe obligations indicates, copyright and database protection pose obstacles to the attainment of open science goals. The moment data resources used as a basis for research (input dimension) or data resources evolving from research (output dimension) contain protected third party elements, the intellectual property status constitutes a constraint that prevents researchers in Horizon projects from making research data available open access. Again (cf. section 2.4), it becomes apparent that the realisation of research-related policy goals depends on sufficient flexibility in copyright, related rights and sui generis database law.

Hence, it is necessary to revisit the scientific research rules that have been discussed in chapter 3. With regard to the policy goal of ensuring open access, particular problems arise from the fact that the database rule in Article 9(b) DBD and the TDM rule in Article 3(1) DSMD do not exempt researchers from the obligation to obtain the authorisation of right holders before sharing research data that contain protected works or database elements (cf. sections 3.2.1 and 3.3.4). In practice, this means that a harmonised and standardised dataset that has been prepared for a TDM analysis cannot be made available open access when it includes protected material, such as images and photographs, sound or music fragments, snippets of films or videos, excerpts from books, articles or newspapers etc.²²³ Similarly, a dataset that has been prepared for research use in a Horizon project is beyond reach for open access if it includes protected parts of third party source databases. It does not matter in this context that the preparation of the research dataset may have led to a reorganisation of database elements.²²⁴

Admittedly, copyright and sui generis database law are not completely silent on the need to store research data for scientific purposes. Article 3(2) DSMD states that:

[c]opies of works or other subject matter made in compliance with paragraph 1 shall be stored with an appropriate level of security and may be retained for the purposes of scientific research, including for the verification of research results.

Therefore, the DSM Directive seeks to offer a solution that safeguards the academic and public interest in the validation of research results. This, however, does not include open access. Recital 15 DSMD underlines that the TDM copies should be stored in a secure environment. After discussions with relevant stakeholders, Member States may introduce specific national arrangements for retaining the copies. They may also appoint trusted bodies for the purpose of data storing. To avoid overbroad inroads into the TDM

²²⁰ Annotated Grant Agreement, id., 158.

²²¹ Annotated Grant Agreement, id., 158.

²²² Annotated Grant Agreement, id., 159.

²²³ CJEU case law indicates in this regard that the threshold for assuming copyright protection is low. In the case of texts, for instance, a fragment consisting of 11 words may already enjoy the status of a protected literary work. See CJEU, 16 July 2009, case C-5/08, Infopaq/DDF, para. 47. With regard to visual material, such as portrait photographs, a similar, elastic approach follows from CJEU, 1 December 2011, case C-145/10, Painer, para. 91.

²²⁴ CJEU, 5 March 2009, case C-545/07, Apis/Lakorda, para. 48. Cf. the discussion in section 3.2.1.

exemption, national arrangements should be “proportionate and limited to what is needed for retaining the copies in a safe manner and preventing unauthorised use.”²²⁵

With regard to other uses for the purpose of scientific research, such as scientific peer review and joint research, Recital 15 DSMD relies on the freedom of Member States to introduce further copyright exceptions by virtue of Article 5(3)(a) ISD. As explained in sections 3.1 and 3.3.4, however, Article 5(3)(a) ISD can hardly be expected to pave the way for EU-wide open access solutions in the area of research data. Without repeating all inherent limits of the provision and related sources of legal uncertainty, it suffices to recall that Article 5(3)(a) ISD only permits acts of making available to the public with regard to works enjoying copyright protection. Sui generis databases – and datasets including protected elements stemming from sui generis databases – fall outside the scope of the provision (section 3.2.1). Moreover, Article 5(3)(a) ISD is an optional copyright exception which Member States have implemented in different ways in their national law (section 3.1.2). This makes it unlikely that a satisfactory European open access regime could emerge from the application of the individual national exponents of the exception prototype in Article 5(3)(a) ISD.

4.2.3 Impulses for Copyright Data Improvement

Considering the described shortcomings of the copyright data infrastructure and the difficulty of implementing data improvement initiatives (section 4.1), the open access restrictions following from copyright and sui generis database protection can be placed in a broader policy context. As explained in section 4.2.1, open access requirements in the academic world, such as Horizon Europe obligations, concern the deposition of research data in trusted repositories. The data deposit must be accompanied by the provision of metadata in machine-actionable and structured form.²²⁶ In addition to other information, metadata must include descriptions of the deposited datasets.²²⁷

Quite clearly, this type of data depositions, accompanied by rich metadata in the described format, could have beneficial effects on the overall copyright data infrastructure. **Open access deposits of research data that contain works and databases, and provide accurate metadata, could offer important impulses for a more general improvement of copyright data.** The visual arts sector can serve as an example. As explained in section 4.1.2, initiatives for the creation of digital work repositories, including metadata, often come from museums and other cultural heritage institutions that embark on the digitisation of works in their holdings. This has already extended the data coverage of works of fine art, while the situation in the field of photography and illustrations remains unsatisfactory.²²⁸

Considering the data improvement dynamic triggered by cultural heritage projects, it is of particular interest that the exemption of TDM research in Article 3(1) DSMD concerns not only research organisations but also cultural heritage institutions. EU legislation itself foresees fruitful collaborations on the basis of the new rules for scientific TDM research. **Collaborations between researchers and cultural heritage specialists could**

²²⁵ Recital 15 DSMD.

²²⁶ European Commission, *EU Grants AGA – Annotated Model Grant Agreement EU Funding Programmes 2021-2027*, pre-draft, dated 30 November 2021, Brussels: European Commission 2021, 159.

²²⁷ European Commission, *Horizon Europe (HORIZON) Euratom Research and Training Programme (EURATOM) – General Model Grant Agreement EIC Accelerator Contract*, Version 1.1, dated 15 December 2021, Brussels: European Commission, 111-112.

²²⁸ For a closer analysis of the particular situation and dynamics in the visual arts sector, see the report by Tristan Azzi and Yves El Hage, *Les métadonnées liées aux images fixes*, Paris: CSPLA 2021. Cf. also G. Pessach, “Museums, Digitization and Copyright Law: Taking Stock and Looking Ahead”, *Journal of International Media and Entertainment Law* 1 (2007), 253.

indeed amplify existing data improvement tendencies. Joint initiatives could become a propelling force for the creation of data and metadata. In line with the FAIR principles, these initiatives at the intersection of research and cultural heritage preservation can be expected to increase the findability, accessibility, interoperability and reusability of works.

The beneficial effects of collaborative research/cultural heritage projects in the area of TDM, however, need not remain limited to the open science agenda. For the creative industry sector as a whole, TDM collaborations culminating in the creation of data and metadata in standardised, machine-actionable formats seem to have the potential to substantially enhance:

- licensing opportunities;
- the accuracy and completeness of repartitioning schemes of collecting societies; and
- the visibility of European works in automated recommender systems.²²⁹

For these positive side-effects of data and metadata creation to become a reality, however, it seems indispensable to remove barriers that prevent open access. As pointed out in the preceding section, such barriers can follow from the current copyright and database framework. **The overly cautious approach that has been adopted with regard to the sharing of research data that include protected works and database elements (sections 3.2.1 and 3.3.4) appears short-sighted.** In the light of substantial benefits that may accrue from the improvement of copyright data, a more flexible regime that broadens open access to datasets and metadata covering protected works and databases seems to be the preferable option.

In this regard, it need not be feared that open access libraries of works and (parts of) databases resulting from research projects substitute demand for primary information products and erode the licensing market instead of fostering licensing deals. After all, it must not be overlooked that copyright provisions allowing for the making available of works to the public, such as Article 5(3)(a) ISD, have a focus on use for scientific research. Hence, it is possible to draw a line between open access availability to achieve open science goals and enable future research, and the use of work libraries in closed formats for commercial purposes. **As long as accurate metadata are available that offer the information necessary to find content and right holders, it is conceivable to arrive at a beneficial co-existence of open access arrangements for scientific research on the one hand, and commercial exploitation platforms that lead to licensing agreements on the other.**

5. Conclusion

5.1 Imbalance

From a fundamental rights perspective, researchers and holders of copyright, related rights or sui generis database rights meet at eye level. The recognition of the right to property, including intellectual property, in Article 17(2) CFR is not stronger or more important than the freedom of expression and information, and the freedom of sciences, guaranteed in Articles 11(1) and 13 CFR. Instead, these fundamental rights and freedoms have an equal rank in the norm hierarchy (section 2.3). This concerns an equal rank at the highest level of the legal order: the Charter of Fundamental Rights constitutes primary EU law.

²²⁹ Cf. M.R.F. Senftleben/T. Margoni et al., "Ensuring the Visibility and Accessibility of European Creative Content on the World Market: The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive", *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* (JIPITEC) 13 (2022), 67 (71-74).

Nonetheless, the equal status of researchers and research policies, such as the open science agenda, is hardly discernible at the level of secondary copyright law. Under the EU Directives regulating copyright, related rights and sui generis database protection, right holders enjoy broad, robust exclusive rights of reproduction,²³⁰ communication and making available to the public,²³¹ extraction²³² and re-utilisation.²³³ Given the flexible, elastic circumscription of these exclusive rights – covering “any”²³⁴ form of use falling within the relevant category, or direct and indirect use “by any means and in any form”²³⁵ – these legal positions are likely to absorb whatever new form of use emerges. TDM, for instance, gives rise to the question whether the automated, computational analysis of protected works constitutes a relevant form of use “as a work” at all.²³⁶ In the EU, however, this new form of use has been subordinated to the control of copyright and database right holders to a large extent.²³⁷ **It seems that EU legislation, almost intuitively, opts for the grant of exclusive rights which right holders can invoke as veto rights against use without prior authorisation.**

As a result, **researchers seeking access to protected works and databases to identify relevant data sources and compile appropriate datasets for their scientific research** (see the distinction between different categories of data in section 2.2) **face a discouraging phalanx of legal conditions and requirements which must be fulfilled before they can benefit from a statutory use privilege for scientific use. The foregoing analysis** (chapter 3) **has shed light on:**

²³⁰ Article 2 ISD.

²³¹ Article 3 ISD.

²³² Article 7(1), (2)(a) and (5) DBD.

²³³ Article 7(1), (2)(b) and (5) DBD.

²³⁴ Article 3(1) ISD.

²³⁵ Article 2 ISD.

²³⁶ Cf. R. Ducato/A. Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out”, *European Intellectual Property Review* 43 (2021), 322 (334); R. Ducato/A. Strowel, “Limitations to Text and Data Mining and Consumer Empowerment: Making the Case for a Right to ‘Machine Legibility’”, *International Review of Intellectual Property and Competition Law* 50 (2019), 649 (649-684); P.B. Hugenholtz, “Artikelen 3 en 4 DSM-richtlijn: tekst- en datamining”, *Tijdschrift voor auteurs-, media- en informatierecht* 2019, 167 (169); R.M. Hilty/H. Richter, “Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules – Part B: Exceptions and Limitations – Art. 3 Text and Data Mining”, *Max Planck Institute for Innovation and Competition Research Paper Series* 2017-02, 2-5; European Copyright Society, *General Opinion on the EU Copyright Reform Package*, ECS 2017, 5-6, available at: <https://europeancopyrightsociety.org/> (last visited: 7 June 2022). For a nuanced approach asking whether the author can be expected to have taken into account “de-intellectualized” use, see M. Borghi/S. Karapapa, “Non-display Uses of Copyright Works: Google Books and Beyond”, *Queen Mary Journal of Intellectual Property* 1 (2011), 21 (45-52). For a national approach that follows this premise and places TDM outside the control of copyright holders, see T. Ueno, “The Flexible Copyright Exception for ‘Non-Enjoyment’ Purposes Recent Amendment in Japan and Its Implication”, *Gewerblicher Rechtsschutz und Urheberrecht International* 70 (2021), 145 (145-152). Cf. also M. W. Carroll, “Copyright and the Progress of Science: Why Text and Data Mining Is Lawful”, *U.C. Davis Law Review* 53 (2019), 893.

²³⁷ For a critique of this development in the light of more flexible TDM regimes in other regions, see M.R.F. Senftleben/T. Margoni et al., “Ensuring the Visibility and Accessibility of European Creative Content on the World Market: The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law (JIPITEC)* 13 (2022), 67 (72-73); C. Geiger, “The Missing Goal-Scorers in the Artificial Intelligence Team: Of Big Data, the Right to Research and the Failed Text-and-Data Mining Limitations in the CSDM Directive”, in: M.R.F. Senftleben, J. Poort et al. (eds.), *Intellectual Property and Sports – Essays in Honour of Berndt Hugenholtz*, The Hague/London/New York: Kluwer Law International 2021, 383-394; C. Handke/L. Guibault/J.-J. Vallbé, “Is Europe Falling Behind in Data Mining? Copyright’s Impact on Data Mining in Academic Research”, in: B. Schmidt/M. Dobрева (eds.), *New Avenues for Electronic Publishing in the Age of Infinite Collections and Citizen Science: Scale, Openness and Trust - Proceedings of the 19th International Conference on Electronic Publishing, IOS 2015*, 120-130.

- the potential restriction of use privileges to use for mere illustrative purposes (Article 5(3)(a) ISD; Article 9(b) DBD);
- the exclusion of collaborations with partners from the private sector that may have a commercial orientation (Article 5(3)(a) ISD; Article 9(b) DBD);
- the restriction of use privileges to acts of reproduction and extraction (not covering communication and making available to the public), preventing the sharing of research data with members of a larger research consortium, the academic community and the public at large (Article 9(b) DBD; Article 3 DSMD);
- difficulties for transnational research consortia to establish and use shared datasets. These difficulties arise from the fact that each participating institution must demonstrate individual lawful access credentials (Article 3(1) DSMD);
- the necessity to justify research use as a special case in the sense of the three-step test (Article 5(5) ISD; Article 7(2) DSMD);
- the necessity to rebut allegations that research use has a corrosive effect on the normal exploitation of works and databases serving as data sources (Article 5(5) ISD; Article 7(2) DSMD);
- the necessity to rebut allegations that research use unreasonably prejudices legitimate interests of right holders (Article 5(5) ISD; Article 7(2) DSMD);
- the restriction of use privileges that follows from technological protection measures that serve as electronic fences preventing access and use for research purposes (Article 6(1) and (4), subparagraph 4, ISD; Article 3(3) DSMD);
- the restriction of use privileges that follows from contractual terms that exclude forms of use necessary for research purposes (Articles 2, 3 and 6(4), subparagraph 4, ISD; Articles 7(1) and (5) DBD).

In addition to these barriers that follow from the harmonised rules in EU copyright, related rights and sui generis database law, it must not be overlooked that the research exemptions in Article 5(3)(a) ISD and Article 9(b) DBD are optional “may” provisions. In certain Member States, these use privileges may be sought in vain. In others, the national implementation may have led to more nuanced domestic rules, adding further conditions and requirements. In the case of research consortia with partners in several Member States, the scope of statutory use privileges that are available at the national level may, as a result, differ from one research partner to the other. For acts of copying or data sharing that must be carried out by all research partners, the group as a whole, thus, can only rely on the smallest common denominator that falls within the scope of the most restrictive national rule.

Considering this configuration of use privileges for scientific research, the conclusion is inescapable that there is an imbalance in the EU acquis. Whereas EU copyright, related rights and sui generis database law offers right holders broad, robust exclusive rights, researchers must content themselves with narrow, unreliable research exemptions. The legal uncertainty arising from the obligation to comply with the described conditions and requirements can easily prevent researchers from invoking use privileges from the outset. **The frustration of research use because of legal uncertainty is a serious flaw of the legal framework.** At the level of fundamental rights, it has been recognised that legal uncertainty must be factored into the equation. In particular, legal risks surrounding a copyright exception can have a deterrent effect on use that is desirable from the perspective of freedom of expression and information. The legal uncertainty arising from the complexity of use conditions can render a given norm unconstitutional when it leads to a situation where beneficiaries are likely to refrain from asserting the use privilege. As a result, the intended reconciliation of intellectual property protection with competing fundamental freedoms no longer takes place and the copyright framework becomes imbalanced.²³⁸

²³⁸ Cf. German Federal Constitutional Court, 31 May 2016, case 1 BvR 1585/13, ‘Sampling’, para. 100. English translation available at:

With regard to the use of protected works and databases as data sources for scientific research, it must finally be added that licensing will not always constitute a promising alternative avenue to arrive at the use permissions required for research projects. As also explained in the analysis (section 4.1), the copyright data infrastructure that could support rights clearance is incomplete and fragmented in the EU. Comprehensive collections of data and metadata relating to literary and artistic creations have not been established yet. Instead, missing and inaccurate data and metadata can make it impossible for researchers to find relevant work repertoires and right holders. At least for projects requiring large-scale use permissions, licensing is unlikely to solve problems that arise from overly complex and nuanced statutory use privileges.

Given these difficulties in the area of data resources enjoying copyright, related rights or sui generis database protection, **it is a positive development that the proposed Data Act seeks to keep machine-generated raw data** (see the distinction of different data categories in section 2.2) **outside the scope of sui generis database protection.** While Article 35 DAP seems to be intended to state the principle of non-protection with sufficient clarity, the provision may nonetheless give rise to the question whether the protection ban is applicable beyond the data access and sharing rights set forth in the proposed Data Act itself (section 3.4).

5.2 Recommendations for Legislative Measures

To improve the legal framework for scientific research and remove the described imbalance that poses obstacles to data access and reuse in copyright, related rights and sui generis database law, **several legislative measures can be considered.**

In particular, it seems important to allow researchers – in line with the equal status of underlying fundamental freedoms²³⁹ – to meet right holders at eye level. **To offer a legal position in copyright law that is comparable with the position following from the grant of broad exclusive rights, it is advisable to introduce a robust, flexible use privilege. Article 5(3)(a) ISD could serve as a reference point for this legislative step.** In its current form, Article 5(3)(a) ISD permits acts of reproduction²⁴⁰ and acts of communication and making available to the public²⁴¹ without prior authorisation of the right holder with regard to:

use for the sole purpose of illustration for teaching or scientific research, as long as the source, including the author's name, is indicated, unless this turns out to be impossible and to the extent justified by the non-commercial purpose to be achieved;...

To offer researchers a more robust and reliable legal position, it is advisable to:

- **clarify that the requirement of use as an "illustration" only concerns the teaching branch** of the use privilege and does not relate to scientific research;
- **abandon the requirement of use for a "non-commercial purpose"** and, instead, follow the approach taken in Article 3(1) DSMD which, rightly understood, offers more room for public-private partnerships and more opportunities to

https://www.bundesverfassungsgericht.de/e/rs20160531_1bvr158513en.html (last visited on 14 April 2022).

²³⁹ As explained, Articles 11 and 13 CFR are not in any way weaker than the right to (intellectual) property in Article 17(2) CFR.

²⁴⁰ Article 2 ISD.

²⁴¹ Article 3(1) ISD (copyright: communication and making available to the public) and Article 3(2) ISD (related rights: making available to the public).

- translate research insights into products and services that can be brought to the market (cf. section 3.3.1);
- **recalibrate the determination of lawful access.** Instead of requiring access permissions of each individual institution participating in a research project, it should be sufficient that one participating institution has lawful access (cf. section 3.3.3);
 - **clarify that**, regardless of the volume of use, **scientific research constitutes a “special case” in the sense of the three-step test** of Article 5(5) ISD because of the fundamental rights underpinning following from Articles 11(1) and 13 CFR;
 - **clarify that, in the assessment of a conflict with a normal exploitation or an unreasonable prejudice to legitimate interests of right holders** under Article 5(5) ISD, **it is necessary to take benefits into account** which right holders, such as academic publishers, derive from the work of researchers and the results of scientific research projects (cf. section 3.1.4);
 - **grant researchers the right to circumvent technological protection measures** in case right holders fail to ensure that the use privilege for scientific research remains effective when technological protection measures are applied (cf. section 3.1.5);
 - **declare Article 6(4), subparagraph 4, ISD inapplicable to use for the purposes of scientific research**, as already done in Article 7(2) DSMD (cf. section 3.3.2);
 - **declare any contractual provision contrary to the use privilege for scientific research unenforceable**, as already stated in Article 7(1) DSMD.

In contrast to the current, optional version of Article 5(3)(a) ISD, **this more flexible and more robust exemption of use for scientific research should constitute a mandatory “shall” provision** to ensure a harmonised application across Member States and comparable conditions for research teams in different countries.

The proposed more flexible and more robust exemption of research use can also play a crucial role in the realisation of EU open science objectives. As the proposed broadened and strengthened version of Article 5(3)(a) ISD would cover both – the right of making copies for research purposes (reproduction)²⁴² and the right of sharing these copies (communication and making available to the public)²⁴³ – the provision has the potential to enable researchers to comply with open access requirements of funding schemes for scientific research, such as Horizon Europe (cf. section 4.2.1). With the proposed broadened and strengthened provision, copyright protection would impose less constraints on initiatives to make research data, including copyrighted material, available open access.

To attain the described goals – an equal legal position for researchers in line with underlying fundamental rights and less barriers to open access availability of research data – **it is advisable to implement the proposed more flexible and more robust use privilege for scientific research not only in the field of copyright and related rights** (Information Society Directive) **but also in the area of the sui generis database right** (Database Directive). Researchers should be able to rely on corresponding use privileges with a congruent scope, laid down in Article 5(3)(a) ISD and Article 9(b) DBD (cf. sections 3.2.1 and 4.2.2).²⁴⁴ Apart from the amendments listed above and the switch to a mandatory “shall” provision, this step in the direction of corresponding use privileges implies that the research exemption in sui generis database

²⁴² Article 2 ISD.

²⁴³ Article 3 ISD.

²⁴⁴ For a more detailed discussion of benefits accruing from a system of corresponding use privileges in different fields of intellectual property, see M.R.F. Senftleben, “Overprotection and Protection Overlaps in Intellectual Property Law – the Need for Horizontal Fair Use Defences”, in: A. Kur/V. Mizaras (eds.), *The Structure of Intellectual Property Law: Can One Size Fit All?*, Cheltenham: Edward Elgar 2011, 136 (136-181).

law would also have to cover both the right of making copies for research purposes (extraction)²⁴⁵ and the right to share these research copies (re-utilisation).²⁴⁶

Finally, it is important to add that, as long as more specific provisions do not in any way limit the scope and applicability of the proposed overarching research exemption in Article 5(3)(a) ISD and Article 9(b) DBD,²⁴⁷ more specific use privileges for scientific research, such as the specific TDM rules in Article 3 DSMD, fulfil an important and positive function. Clarifying modalities of permitted use in specific research circumstances and with regard to specific research techniques, they can be expected to enhance legal certainty and provide confidence in the possibility to carry out the exempted form of use without prior authorisation and transaction costs for rights clearance. Whenever this makes sense from the perspective of open science principles, these more specific use privileges – in line with the overarching exceptions in Article 5(3)(a) ISD and Article 9(b) DBD – should cover reproduction/extraction rights as well as communication and making available to the public/re-utilisation rights. As discussed, it is a shortcoming of the new TDM rules in Article 3 DSMD that this more specific use privilege does not cover the right of making available to the public. In the absence of a permission to share data resources used for TDM research, Article 3 DSMD can hardly be expected to contribute to the realisation of open access goals (cf. sections 3.4.4 and 4.2.2).

5.3 Recommendations for Non-legislative Measures

Some of the aforementioned recommendations for legislative measures can also serve as an impulse for non-legislative clarifications and best practice models. From the outset, however, it must be underlined that, in comparison with legislative amendments, non-legislative initiatives, inevitably, give researchers a **less robust and reliable legal position because the proposed interpretation and application of the law is not prescribed in the legislation itself.** Bearing this “minus” in terms of legally binding obligations in mind, it is conceivable to develop **clarifications and best practice standards** with regard to the following points:

- with regard to the overarching research exceptions in Article 5(3)(a) ISD and Article 9(b) DBD, it could be clarified that:
 - o **the requirement of use as an “illustration” only concerns the teaching branch** of the use privilege and does not relate to scientific research;
 - o regardless of the volume of use, **scientific research constitutes a “special case” in the sense of the three-step test** of Article 5(5) ISD because of the fundamental rights underpinning following from Articles 11(1) and 13 CFR;
 - o **in the assessment of a conflict with a normal exploitation or an unreasonable prejudice to legitimate interests of right holders** under Article 5(5) ISD, **it is necessary to take benefits into account** which right holders, such as academic publishers, derive from the work of researchers and the results of scientific research projects (cf. section 3.1.4);
- with regard to the specific TDM exception in Article 3 DSMD and, more specifically, the lawful access guidelines in Recital 14 DSMD, it could be clarified that:

²⁴⁵ Article 7(1), (2)(a) and (5) DBD.

²⁴⁶ Article 7(1), (2)(b) and (5) DBD.

²⁴⁷ Cf. in that regard the safeguard clause in Article 25 DSMD. Cf. M. Stieper, “Das Verhältnis der verpflichtenden Schranken der DSM-RL zu den optionalen Schranken der InfoSoc-RL”, *Gewerblicher Rechtsschutz und Urheberrecht* 2020, 1 (3-4).

- in the case of subscriptions, the persons attached to a research organisation or cultural heritage institution with the subscription are not the only group of beneficiaries, in respect of which lawful access can be assumed (the relevant sentence in Recital 14 starts with “[f]or instance”). Rightly understood, lawful access should also be assumed with regard to researchers from other organisations or institutions in the case of joint research projects. It should thus be deemed **sufficient that one participating institution has lawful access** (cf. section 3.3.3);
- Member States should use Article 5(3)(a) ISD as a basis to complement Article 3 DSMD with a further copyright exception that permits the sharing of TDM datasets within research consortia and, for purposes such as research validation, also with the broader academic community.²⁴⁸ To ensure a harmonised approach, it seems advisable to **develop a model provision for TDM dataset sharing** that can be implemented in a uniform manner in different national contexts.

In addition to these non-legislative measures that mirror several legislative initiatives that have been proposed above, **the foregoing analysis has shed light on the intersection between open access and related data and metadata initiatives in the academic world, and the need to improve copyright data and data management infrastructures in the creative sector.** The interplay has two central aspects:

- on the one hand, an improved copyright data infrastructure makes it easier for researchers to obtain use permissions that are required when statutory use privileges for scientific research are inapplicable. Hence, data improvement initiatives are likely to enhance data availability for research purposes;
- on the other hand, open access obligations in academic funding schemes that include the sharing of datasets and the creation of corresponding, sufficiently rich, standardised and machine-actionable metadata have the potential to support data improvement strategies in the creative industries.

Given this interrelation, **it seems advisable to develop non-legislative initiatives that pave the way for the injection of copyright-related data and metadata that result from research projects into data improvement processes in the creative sector** (cf. section 4.2.1).²⁴⁹

Considering the mutual interest in an improvement of copyright data and metadata (for research purposes in the academic community; for enhanced licensing and content recommendation opportunities in the creative industries), **there may also be room for non-legislative measures that explore a *quid pro quo***: in exchange for valuable contributions of researchers to the improvement of copyright data, including the creation of rich metadata, right holders may be willing to offer broader support for academic initiatives that seek to ensure open access to research data that include protected works and (parts of) databases. In particular, this may be an attractive option for the creative industry if it proves to be possible to draw a boundary between open access data and metadata systems for research purposes and closed data and metadata systems that use metadata strategically as a source of information to enhance the visibility and findability of works, databases and right holders, and create new and broader licensing opportunities (cf. section 4.2.3).

²⁴⁸ R. Ducato/A. Strowel, “Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out”, *European Intellectual Property Review* 43 (2021), 322 (332).

²⁴⁹ In this regard, new insights can also be expected from the Study on Copyright and New Technologies, SMART 2019/0038.

Getting in touch with the EU

IN PERSON

All over the European Union there are hundreds of Europe Direct information centres.

You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

ON THE PHONE OR BY EMAIL

Europe Direct is a service that answers your questions about the European Union.

You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by email via: https://europa.eu/european-union/contact_en

Finding information about the EU

ONLINE

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU PUBLICATIONS

You can download or order free and priced EU publications from:

<https://op.europa.eu/en/publications>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU LAW AND RELATED DOCUMENTS

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: <http://eur-lex.europa.eu>

OPEN DATA FROM THE EU

The EU Open Data Portal (<http://data.europa.eu/euodp/en>) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

EU legislation in the field of copyright, related rights and sui generis database rights can have a deep impact on access to data resources for scientific research and the availability of data resulting from publicly funded research. To establish a copyright and related rights framework that offers appropriate data access and reuse opportunities for scientific research, it is necessary to identify potential barriers and challenges that may arise from EU copyright and related rights legislation and corresponding rights management. This study analyses the interaction between copyright and related rights law and data access and reuse for scientific research purposes. It proposes legislative and non-legislative measures to improve the current EU regulatory framework.

Studies and reports



Publications Office
of the European Union