

REGULATING TECHNOLOGY THROUGH COPYRIGHT LAW: THE WAY FORWARD FOR BUILDING A DIGITAL SINGLE MARKET?

JULIEN CABAY

*PROFESSOR AND DIRECTOR OF JURISLAB AT UNIVERSITÉ LIBRE DE BRUXELLES
ASSOCIATE PROFESSOR AT UNIVERSITÉ DE LIÈGE*

and

ELEONORA ROSATI

*PROFESSOR OF INTELLECTUAL PROPERTY LAW
DIRECTOR OF THE INSTITUTE FOR INTELLECTUAL PROPERTY AND MARKET LAW (IFIM),
STOCKHOLM UNIVERSITY*

I. INTRODUCTION: THE LAWS OF EVERYTHING IN THE DIGITAL WORLD

“Data” is everywhere. It might even be everything, all our surroundings being more and more subject to data reductionism. The EU legislator is coping with this tendency through the development of some sort of “Data Law,” consisting of a growing set of laws dealing with “data.” Though personal data has been long addressed through Directives,¹ it seems that the adoption of the GDPR² actually paved the way for an extensive legislative activity in all fields of data regulation³. And when one considers

1. See in particular Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, *OJ L 281* of 23 November 1995, pp. 31-50; Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), *OJ L 201* of 31 July 2002, pp. 37-47.

2. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), *OJ L 119 of 4 May 2016*, pp. 1-88.

3. One can cite amongst many the: Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, *OJ L 303* of 28 November 2018, pp. 59-68; Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public

the definitions of “data” in these various laws (“data”,⁴ “personal data”,⁵ “non-personal data”,⁶ “documents”,⁷ *etc.*), it is hard to see what would fail to qualify as such pursuant to at least one of those legal instruments. As “data” is everywhere, so is Data Law, more and more.

Besides, in recent times, another legal instrument proved extremely pervasive in the EU, especially at the instigation of the Court of justice (CJEU). Indeed, as soon as it was given the same value as the Treaties,⁸ the Charter of Fundamental Rights has been relied upon by the CJEU in many cases. Its scope is broad, as it applies to the institutions and bodies of the Union, as well as to the Member States “when they are implementing Union Law.”⁹ As the CJEU stated in *Fransson*: “[...] situations cannot exist which are covered in that way by European Union Law without those fundamental rights being applicable. The applicability of European Union Law entails applicability of the fundamental rights guaranteed by the Charter.”¹⁰ Hence, the more active the EU legislator, the more expanding the realms of the Charter.

sector information (recast), *OJ L* 172 of 26 June 2019, pp. 56-83; Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), *OJ L* 152 of 3 June 2022, pp. 1-44; Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), *OJ L* 277 of 27 October 2022, pp. 1-102; Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), *OJ L* 265 of 12 October 2022, pp. 1-66. See also the 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), Brussels, 23 February 2022, COM(2022) 68 final; Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Brussels, 21 April 2021, COM(2021) 206 final.

4. Art. 2(1) Data Act Proposal: “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.”

5. Art. 4(1) GDPR: “personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.”

6. Art. 3(1) Regulation (EU) 2018/1807 on the free flow of non-personal data: “data’ means data other than personal data as defined in point (1) of Article 4 of Regulation (EU) 2016/679.”

7. Art. 2(6) Regulation (EU) 2022/868 on European Data Governance: “document’ means: (a) any content whatever its medium (paper or electronic form or as a sound, visual or audiovisual recording); or (b) any part of such content.”

8. See Art. 6(1) Treaty on the European Union.

9. Art. 51(1) Charter of Fundamental Rights.

10. CJEU, 26 February 2013, *Åkerberg Fransson*, Case C-617/10, EU:C:2013:105, at [21].

In relation to digital issues in the EU, those two general laws (Data Law, Fundamental Rights) operate as the “Laws of everything” at the micro- and macro-levels.

Making the parallel with Physics, one can easily figure data as the atoms of digital structures. In the legal digital world, Data Law would be similar to Quantum Mechanics in the physical world, governing the infinitesimally small.

In the same vein, though less obvious, Fundamental rights might appear the functional equivalent to the General Relativity. Indeed, where Einstein’s theory describes the physical laws governing the Universe, Fundamental Rights actually define a general framework that any situation governed by EU Law must fit. This is even the more true in the digital world, on the Internet in particular, given the relevance of several rights protected by the Charter, such as: Privacy (Art. 7), the Protection of personal data (Art. 8), the Freedom of expression and information (Art. 10), the Freedom to conduct business (Art. 16) and the Protection of intellectual property (Art. 17(2)). One can even wonder whether in the EU, it is conceivable that a digital issue would not be concerned by any of those fundamental rights. In particular if one does consider that similar to the universe in Einstein’s theory, the Charter expands.¹¹

Arguably, the parallel between Natural Laws and Human Laws comes short when one thinks of the possibility to reconcile the micro- and macro-levels. Whereas reconciliation of general relativity with the laws of quantum physics remains a problem nowadays, in the field of EU Law it seems that Copyright might bridge the gap between the two levels.

Indeed, a work will be protected as soon as it meets the two conditions set by the CJEU: “First, it entails an original subject matter which is the author’s own intellectual creation and, second, it requires the expression of that creation.”¹² Though the CJEU has made it clear that the originality criterion is to be assessed rigorously,¹³ this is not always the case at the level of national courts.¹⁴ Then in practice, as the enjoyment of copyright

11. See in particular the debate and evolution of the case law on the horizontal effect of the Charter, E. FRANTZIOU, “(Most of) the Charter of Fundamental Rights is Horizontally Applicable”, *European Constitutional Law Review*, 2019, Vol. 15, No. 2, pp. 306-323.

12. See *e.g.* CJEU, 11 June 2020, *Brompton*, Case C-833/18, EU:C:2020:461, at [22].

13. See in particular CJEU, 12 September 2019, *Cofemel*, Case C-683/17, EU:C:2019:721, at [51]; CJEU, 11 June 2020, *Brompton*, Case C-833/18, EU:C:2020:461, at [32].

14. See in particular the Opinion of Advocate General SZPUNAR in CJEU, 12 September 2019, *Cofemel*, Case C-683/17, EU:C:2019:363, at [54]; in CJEU, 29 July 2019, *Funke Medien*, Case C-469/17, EU:C:2018:870, at [20-22].

cannot be subject to any formality,¹⁵ it appears that much of the “data” that qualifies as a “work” could actually qualify for protection. Accordingly, in this contribution we use the term “data” to designate objects that fall within the scope of one of the data laws referred above and, at the same time, qualify for copyright protection, including database protection.

The condition that the work must be “expressed” does not bar the applicability of Copyright Law to the uses of data. To the contrary, it stems from the *Infopaq* case that provided the elements reproduced are original, “[a]n act occurring during a data capture process, which consists of storing an extract of a protected work comprising 11 words [...] is such as to come within the concept of reproduction [...]”.¹⁶ The CJEU even stated that, still provided the originality criterion is met, “[...] the reproduction right extends to transient fragments of the works within the memory of a satellite decoder and on a television screen [...]”.¹⁷

Copyright Law appears then an essential component of this emerging “Data Law” governing technologies at the micro-level. It possibly applies everywhere. This is actually the mere result of the interaction between digital and copyright architectures, since the transfer of data pre-supposes their copying, which may trigger the application Copyright Law, including database protection.¹⁸ This incidental conflation of architectures might have expanded the reach of copyright and would explain how, at the turn of the 21st century, calls to “rebalance” Copyright Law became increasingly frequent. In the EU, that rebalancing was largely the endeavour of the CJEU, interpreting the copyright directives in the light of the Charter of Fundamental Rights, a movement that is known as the “constitutionalisation” of copyright.¹⁹

It comes with little surprise that most of the copyright cases dealt by the CJEU through the lens of Fundamental Rights were concerned with digital issues, especially the Internet. The “fair balance” principle that directs modern interpretation of Copyright Law (that enjoys the protection of

15. Art. 5(2) Berne Convention.

16. ECJ, 16 July 2009, *Infopaq International*, Case C-5/08, EU:C:2009:465, at [51].

17. *Football Association Premier League*, joined Cases C-403/08 and C-429/08, EU:C:2011:631, at [159].

18. This idea has been put forward elsewhere in another context, see L. LESSIG, *Remix – Making Art and Commerce Thrive in the Hybrid Economy*, New York, The Penguin Press, 2008, p. 103.

19. See e.g. C. GEIGER and E. IZUMENKO, “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 - International Review of Intellectual Property and Competition Law*, 2020, Vol. 51, No. 3, pp. 282-306.

Art. 17(2) of the Charter) in case it conflicts with other Fundamental Rights came out in *Promusicae*,²⁰ in which the CJEU addressed for the first-time enforcement of copyright on the internet. In *GS Media*, where the CJEU explicitly stated that the Directive 2001/29 [*InfoSoc*] aims at maintaining, in particular in the electronic environment, a “fair balance” between interests protected by different Fundamental Rights, the facts were concerned with the qualification of hyperlinking as a communication to the public.²¹ The Grand Chamber of the CJEU addressed the general relation between Copyright and Fundamental Rights in the triptych *Funke Medien*²²/*Spiegel Online*²³/*Pelham*²⁴, where the first two cases were concerned with online communications to the public (the latter concerning the technique of “sampling”). And the last decision issued in the *Poland*²⁵ case was entirely dedicated to the analysis of the compatibility of Article 17 of DSM Directive, namely the (algorithmic) regulation by online content-sharing service providers, with the right to freedom of expression and information, bringing into discussion previous related cases on filtering measures²⁶ and blocking injunctions.²⁷

In so doing, step by step the CJEU is drawing at the macro-level the general framework into which copyright related technologies can develop. As the President Lenaerts **once** put:

“[...] the margin of discretion available to Member States when transposing copyright directives is ‘framed’ by the requirement that a fair balance be struck between the fundamental rights applicable. In our opinion, this requirement suggests that, in copyright matters, Union Law imposes ‘a certain uniformity in the abstract,’ while allowing ‘constitutional pluralism *in concreto*’.”²⁸

20. ECJ, 29 January 2008, *Promusicae*, Case C-275/06, EU:C:2008:54, at [68].

21. CJEU, 8 September 2016, *GS Media*, Case C-160/15, EU:C:2016:644, at [31].

22. CJEU, 29 July 2019, *Funke Medien*, Case C-469/17, EU:C:2019:623.

23. CJEU, 29 July 2019, *Spiegel Online*, Case C-516/17, EU:C:2019:625.

24. CJEU, 29 July 2019, *Pelham e.a.*, Case C-476/17, EU:C:2019:624.

25. CJEU, *Poland*, Case C-401/19, EU:C:2022:297.

26. CJEU, 16 February 2012, *Sabam*, Case C-360/10, EU:C:2012:85. See also CJEU, 24 November 2011, *Scarlet Extended*, Case C-70/10, EU:C:2011:771.

27. CJEU, 27 March 2014, *UPC Telekabel*, Case C-314/12, EU:C:2014:192.

28. K. LENAERTS, “Le droit d’auteur dans l’ordre juridique de l’Union européenne: une perspective constitutionnelle”, in J. CABAY, V. DELFORGE, V. FOSSOUL AND M. LAMBRECHT (eds), *20 ans de nouveau droit d’auteur – 20 jaar nieuw auteursrecht*, Limal, Anthemis, 2015, p. 245: “la marge d’appréciation dont les États membres disposent lors de la transposition des directives en matière de droit d’auteur est ‘encadrée’ par l’exigence que soit assuré un juste équilibre entre les droits fondamentaux applicables. À notre avis, cette exigence suggère qu’en matière de droit d’auteur, le droit de l’Union impose ‘une certaine uniformité dans l’abstrait,’ tout en permettant ‘un pluralisme constitutionnel *in concreto*’” (our translation).

It follows from all this that Copyright Law appears essential to both the development of Data Law and Fundamental Rights in the EU. Hence, we suggest that regulating Copyright Law in the digital world is actually meaningful beyond the sole issue of copyright. More specifically, as the Digital Single Market²⁹ can exist only through technologies, in particular the Internet as an infrastructure for sharing data amongst individuals and undertakings, we wonder whether copyright regulation could not be the way forward for building such market.

In order to substantiate our claim and inform our hypothesis, in this contribution we will explore two recent developments brought by the Directive (EU) 2019/790 on copyright and related rights in the Digital Single Market. Addressing copyright at the micro-level, we will analyse the Text and Data Mining exceptions (Art. 3 and 4) and their role in ~~F~~ostering Artificial Intelligence (§ 2). Considering then copyright at the macro-level, we will evaluate the idea of resorting to technology ~~for fighting against~~ copyright infringement (Art. 17) and its impact on the development of Algorithmic Regulation (§ 3). We will then conclude in the sense that copyright does not only bridge the gap between two ~~L~~AWS from a normative standpoint: it also proves likely to push EU integration in practice on both levels of the operators and the infrastructures.

II. FOSTERING AI DEVELOPMENT THROUGH TDM EXCEPTIONS

As discussed above at § 1, ~~c~~opyright and ~~r~~elated ~~r~~ights play a substantial role in our data-driven economy and can shape current and future trends in technological advancement. If we take machine learning and AI-based processes as a substantial example, the relevance of ~~c~~opyright and ~~r~~elated ~~r~~ights is self-evident, including at the very outset, that is when it comes to the selection and use of materials that are required for machine learning and AI-training purposes. It is also worth noting that, as it will be discussed below at § 3, technologies based on machine learning and AI have been also having an increasingly central role in the very protection of ~~i~~ntellectual ~~p~~roperty ~~r~~ights, including ~~c~~opyright and ~~r~~elated ~~r~~ights, on the ~~i~~nternet. As both the Advocate General and the Grand Chamber reflected in *Poland*, C-401/19, for the time being the use of automatic recognition and filtering

29. EU Commission, Communication, “A Digital Single Market Strategy for Europe”, Brussels, 6 May 2015, COM(2015) 192 final, at [3]: “A Digital Single Market is one in which the free movement of goods, persons, services and capital is ensured and where individuals and businesses can seamlessly access and exercise online activities.”

tools appears not to have valid alternatives when user-uploaded content needs to be moderated at a scale.³⁰ In this sense, it is more than evident that a close link does exist between the obligations imposed upon online content-sharing service providers (OCSSPs) under Article 17(4) of the DSM Directive³¹ and the exceptions and limitations (E&L) for text and data mining (TDM) as also introduced by that directive under Articles 3 and 4.

Besides issues of appropriateness and accuracy of tools developed through machine learning and AI-based processes, which the Grand Chamber tackled to an extent in *Poland*, C-401/19, a fundamental question is thus that concerning the very lawfulness of undertaking machine learning and AI “training” processes without a licence and what role copyright and related rights play in all this. Following a brief overview of the growing relevance of machine learning and AI to today’s society and a mapping of some of the main approaches to unlicensed TDM in different regions of the world (§ 2.1), the sections below discuss the history, rationale, and content of the provisions under Articles 3 and 4 of the DSM Directive (§ 2.2) and note that, despite the adoption of such provisions, legal restrictions may still be imposed on the undertaking of TDM processes without a licence (§ 2.3).

A. *The growing relevance of TDM processes, including from a copyright regulatory standpoint*

While TDM may be performed in different ways, the key value of predictive TDM processes lies in facilitating the treatment, recombination, and extraction of further knowledge from large amounts of data and text, thus allowing the identification of patterns and associations between seemingly unrelated pieces of information.³² From a technical and commercial standpoint, despite that classical TDM and machine learning have different

30. CJEU, *Poland*, Case C-401/19, EU:C:2022:297, at [54], referring to Opinion of Advocate General SAUGMANDEGAARD *ØE in Poland*, Case C-401/19, EU:C:2021:613, at [57-69].

31. Directive (EU) 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 L 130 of 17 May 2019, pp. 92-125.

32. See Article 2, No. 2, of the DSM Directive (n 2). See also A. BENSAMOUN and Y. BOUQUEREL, “Transposition des Exceptions de Fouille de Textes et de Données: Enjeux et Propositions”, 18 December 2020, pp. 12-14, available at <https://www.culture.gouv.fr/Sites-thematiques/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux/Missions/Mission-du-CSPLA-sur-les-exceptions-de-fouille-de-textes-et-de-donnees-text-and-data-mining>.

utility, both use the same key algorithms to discover patterns in data. TDM also plays a significant role in the advancement of AI applications.³³

As it is discussed at greater length elsewhere,³⁴ TDM is an example of an area in which legislative intervention has been broadly justified by reference to the need of freeing up certain copyright-covered spaces to facilitate research and increase innovation and competitiveness. It should be noted at the outset that, on the one hand, some commentators hold the view that TDM would not even be covered by Copyright Law.³⁵ On the other hand, the debate around TDM has not developed in a context devoid of licensing practices, at least in Europe. Especially in the aftermath of a 2013 stakeholder-led dialogue, Licences for Europe,³⁶ scientific, technical, and medical publishers included TDM for non-commercial purposes in their subscription licences for academic institutions and developed common infrastructures to facilitate access to the content to be mined. This said, different contractual conditions and policies were found leading to uncertainty and, as a result, giving rise to transaction costs.³⁷

In some countries, existing systems of E&L, including fair use under § 107 of the US Copyright Act,³⁸ have been deemed likely to accommodate certain unlicensed TDM activities, although recent and – at the time of writing – pending litigation will require a more substantive assessment as to whether that is in fact the case.³⁹ In other legal systems, specific E&L

33. See the discussion in E. ROSATI, *Copyright in the Digital Single Market. Article-by-Article Commentary to the Provisions of Directive 2019/790*, Oxford, OUP, 2021, pp. 68-71.

34. E. ROSATI, “Copyright reformed: The narrative of flexibility and its pitfalls in policy and legislative Initiatives (2011 – 2021)”, *APLR*, 2023 (available on advance access), pp. 9-11.

35. Recently, see the discussion in M. SENFTLEBEN, “Compliance of national TDM rules with international copyright law – An overrated nonissue?”, 2022, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4134651, considering that in any event the international three-step test would not constrain E&L for TDM.

36. See <https://digital-strategy.ec.europa.eu/en/library/licences-europe-stakeholder-dialogue>.

37. European Commission, “Impact Assessment on the Modernisation of EU Copyright Rules – Part 1”, Commission Staff Working Document, SWD(2016) 301, § 4.3.1. But cf. European Commission, Directorate-General for the Internal Market and Services, “Assessing the economic impacts of adapting certain limitations and exceptions to copyright and related rights in the EU”, 2013, available at <https://data.europa.eu/doi/10.2780/90295>, in particular p. 81, finding that standardized and ‘one-stop-shop’ solutions would be increasingly available and that, as a result, an E&L for TDM for scientific purposes would be only appropriate insofar as no licensing was available.

38. See, e.g., *Authors Guild v. Google, Inc.*, No. 13-4829 (2d Cir. 2015), finding that the possibility to mine the Google Books Library corpus would *inter alia* support a finding of fair use. In scholarly literature, see in particular M. SAG, “The new legal landscape for text mining and machine learning”, *J Copyr Soc USA*, 2019, Vol. 66, 291.

39. A class action has been recently filed before the US District Court for the Northern District of California, alleging infringement of copyright in the development and functioning of AI image generator Stable Diffusions: *Andersen and Others v Stability AI Ltd and Others*, Case 3:23-cv-00201, filed 13 January 2023.

relating to content to which lawful access has been secured have been adopted instead. This has been for example the case of Japan,⁴⁰ some EU Member States individually at first and then through action at the EU level and, more recently, Singapore. The introduction of a specific E&L for TDM has also featured in Hong Kong copyright reform discourse. However, the most recent governmental position is that, given the diversity of views expressed by concerned stakeholders, “rushing into incorporating these issues in the amendment bill” is not recommended.⁴¹

During its tenure as an EU Member State, the UK was the first to rely on the EU copyright *acquis* as it existed in 2014 – specifically: the research E&L in Article 5(3)(a) of the InfoSoc Directive⁴² – to legislate and adopt an express defence, which cannot be overridden by contract, allowing text and data analysis for non-commercial research.⁴³ The Hargreaves Review, from which that reform stemmed, expressly noted how “the law can block valuable new technologies like text and data mining, simply because those technologies were not imagined when the law was formed”⁴⁴ and all this whilst the resulting activities would “not prejudice the central objective of copyright, namely the provision of incentives to creators.”⁴⁵ Recently, a debate emerged as to whether the scope of the 2014 E&L should be broadened.⁴⁶ Such plans have been however axed in early 2023.⁴⁷

40. Since 2011 Japan has had in force an E&L (originally introduced in 2009) specifically allowing TDM: see Art. 47(7) of the Copyright Law of Japan.

41. Legislative Council Panel on Commerce and Industry, “Updating Hong Kong’s Copyright Regime – Outcomes of Public Consultation and Proposed Way Forward”, 19 April 2022, p. 19, available at <https://www.cedb.gov.hk/citb/en/consultations-and-publications/consultation-papers.html> (Legislative Council Panel). Indeed, such a reform does not feature in the Copyright (Amendment) Bill 2022 (n 29).

42. 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 L* 167 of 22 June 2001, pp. 10-19 (InfoSoc Directive).

43. Section 29A CDPA. In mid-2022, UK Government announced that it would broaden the scope for unlicensed TDM activities and introduce a new E&L that would allow TDM for any purpose, subject to a lawful access requirement: see UK Intellectual Property Office, “Artificial Intelligence and Intellectual Property: Copyright and Patents: Government Response to Consultation”, 28 June 2022, available at <https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation>.

44. I. HARGREAVES, “Digital Opportunity. A Review of Intellectual Property and Growth”, 2011, § 5.3, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/32563/ipreview-finalreport.pdf (Hargreaves Review).

45. *Ibid.*, § 5.40.

46. Communications and Digital Committee, “Corrected oral evidence: A creative future”, 22 November 2022, available at <https://committees.parliament.uk/oralevidence/11918/pdf/>.

47. See E. ROSATI, “UK Government axes plans to broaden existing text and data mining exception”, *The IPKat*, 3 February 2023, available at <https://ipkitten.blogspot.com/2023/02/uk-government-axes-plans-to-broaden.html>.

Further to the UK initiative, other EU Member States (France, Estonia, Germany, and Ireland) also considered legislating or legislated in the field of TDM. In 2019, two new mandatory EU-wide E&L for TDM were however adopted as part of the DSM Directive. It was within the Council – that is where EU Member States find their representation in the EU Law-making process – that the introduction of a further E&L (besides the one now found under Article 3) without restrictions in terms of beneficiaries and purposes of the TDM (now Article 4) initially emerged.⁴⁸

B. *The TDM E&L in the DSM Directive*

Further to a complex legislative process⁴⁹ and persisting uncertainties regarding the scope thereof⁵⁰, the DSM Directive includes both an exception on TDM for research purposes (Art. 3)⁵¹ and an E&L not circumscribed to such purposes and available without particular restrictions in terms of beneficiaries (Art. 4). The rationale of EU intervention in relation to unlicensed TDM is found in the preamble to the DSM Directive. Recital 8 acknowledges, on the one hand, the value and potential of TDM but, on the other hand, notes the restrictions that copyright and related rights pose to the doing of TDM activities without a licence. Further to the latter, recital 10 highlights the insufficiency of the existing framework, due to both the optional nature of exceptions and limitations to copyright and related rights for scientific research purposes and the limitations of licensing agreements. As such, the intervention of the EU legislature would serve to remedy the legal uncertainties surrounding TDM activities (recital 11) through the introduction of a mandatory, non-compensated exception for

48. See further E. ROSATI, *Copyright in the Digital Single Market. Article-by-Article Commentary to the Provisions of Directive 2019/790*, Oxford, OUP, 2021, pp. 67-68. In any case, Article 25 and recital 5 of the DSM Directive expressly allow EU Member States to adopt or maintain broader provisions, compatible with the E&L provided for in the Database Directive (Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, *OJ L 77* of 27 March 1996, pp. 20-28) and the InfoSoc Directive, including exceptions and limitations allowing TDM pursuant to Article 6(2)(b) of the former and Article 5(3)(a) of the latter.

49. See further E. ROSATI, *Copyright in the Digital Single Market. Article-by-Article Commentary to the Provisions of Directive 2019/790*, Oxford, OUP, 2021, pp. 63-68.

50. Insofar as TDM is concerned a question relates to the scope of the E&L, including whether they only concern acts of extraction and reproduction, or also subsequent communication of the results. See, having specific regard to the Italian transposition of Article 3, B. CALABRESE, “Scientific TDM exception and communication to the public: did Italians do it better... or at least not worse?”, *JIPLP*, 2022, Vol. 17, No. 5, 399.

51. See further E. ROSATI, *Copyright in the Digital Single Market. Article-by-Article Commentary to the Provisions of Directive 2019/790*, Oxford, OUP, 2021, pp. 40-41, discussing the implications of characterizing Article 3 as an exception rather than an E&L.

the benefit of research organizations and cultural heritage institutions and a mandatory exception or limitation, the one in Article 4, without any particular restrictions in terms of beneficiaries.

The specific rationale of Article 4, which is of particular relevance in the context of the present discussion given the absence of restrictions in terms of beneficiaries and purposes of the TDM, is explained in recital 18 and is linked to a threefold consideration. First, awareness of the relevance of TDM practices outside the context and purpose of scientific research, including, *e.g.*, for government services, complex business decisions and the development of new applications or technologies. Secondly, awareness that licensing opportunities should not be unduly affected by the introduction of an additional TDM E&L besides the one in Article 3: rightsholders should remain able to license the uses of their works or other subject matter falling outside the scope of that provision and of the existing exceptions and limitations provided for in the EU *acquis*. Thirdly, the need to remove legal uncertainty in connection with reproductions and extractions made for such purposes and whether they can be carried out on lawfully accessed works or other subject matter, in particular when the reproductions or extractions made for the purposes of the technical process do not fulfil all the conditions of the existing exception for temporary acts of reproduction provided for in Article 5(1) of the InfoSoc Directive. It follows, that Article 4 seeks to both provide greater legal certainty for TDM practices that do not fall within the scope of application of Article 3 and encourage innovation, by means of engagement with TDM practices, also in the private sector.

C. Restrictions to the undertaking of TDM processes in a post-DSM Directive world

Besides the (unexplainable) broader scope of application of Article 4 compared to Article 3⁵², the different beneficiaries and purpose of the authorized TDM activities, a key feature is that – unlike Article 3 – that provision allows rightsholders to exclude the application of the E&L through an appropriate reservation. The specific rationale of this is explained in the already mentioned recital 18 and is *inter alia* linked to the consideration

52. Unlike Article 3, the E&L in Article 4 also encompasses the rights of reproduction, translation, adaptation, arrangement and any other alteration of a computer program under Article 4(1)(a)-(b) of the Software Directive (Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs [Codified version], *OJ L* 111 of 5 May 2009, pp. 16-22).

that licensing opportunities should not be unduly affected by the availability of an E&L. This is in line with the three-step test, which also bounds E&L under the DSM Directive, as it is evident from Article 7(2) therein.

While the discourse around contractual override of E&L is significantly broader and more complex than what it is possible to consider here,⁵³ in the context of the present contribution it is evident that whether and to what extent rightholders will exercise their reservation right under Article 4(3), which in any case needs to be done in an appropriate manner, is likely to have an impact, not only on unlicensed machine learning and AI training processes, but also the development of content recognition and filtering tools that might serve to assist those very rightholders that exercise such a right.

Article 4(3) allows rightholders to reserve the right to perform TDM activities, insofar as such reservation is done, as indicated in recital 18 and as stated, in an appropriate manner. To this end, that recital distinguishes between two different scenarios. In the case of content that has been made publicly available online, it should only be considered appropriate to reserve the rights in Article 4(1) by the use of machine-readable means, including metadata and terms and conditions of a website or a service. In any event, other uses shall not be affected by the reservation of rights for the purposes of TDM. In other cases, it appears that it *might* be appropriate to reserve the rights by other means, such as contractual agreements or a unilateral declaration. ~~However~~, in light of the CJEU judgment in *VG Bild-Kunst*, C-392/19, it appears preferable to adopt a corrected reading of the provision, in the sense that reservation by rightholders shall be only possible if done by adopting effective technological measures within the meaning of Article 6(1) and (3) of the InfoSoc Directive. This modality, according to the Court, is the one that ensures legal certainty and the smooth functioning of the internet. This so also considering that, in the absence of such effective technological measures, it might prove difficult for individual users to ascertain whether the concerned rightholders intended to reserve the doing of TDM activities in relation to their copyright works and other protected subject matter, including where these are subject to sub-licences.⁵⁴

53. See further E. ROSATI, "Copyright Reformed: The Narrative of Flexibility and Its Pitfalls in Policy and Legislative Initiatives (2011 – 2021)", *APLR*, 2023 (available on advance access), pp. 20-21.

54. CJEU, 9 March 2021, *VG Bild-Kunst*, Case C-392/19, EU:C:2021:181, at [46]. See further E. ROSATI, "Linking and copyright in the shade of *VG Bild-Kunst*", *CML Rev.*, 2021, Vol. 58, No. 6, p. 1890.

The DSM Directive does not specify what the consequences of making the reservation in other ways, *e.g.*, through means other than machine-readable ones for content made publicly available online, would be. In such instances the reservation made would fall short of the requirement of appropriateness and the result would be that it would be unenforceable against those who have performed the acts that Article 4(1) restricts without the prior authorization of the relevant rightholder and in compliance with all the other requirements for the application of the exception or limitation therein. In any case, the activities described in Article 4(1) cannot be restricted by relevant rightholders when: (i) they are performed in relation to non-protected elements of a work or other subject matter (in this sense, recital 7 states that, when TDM activities are carried out in relation to mere facts or data that are not protected by copyright or related rights, no prior authorization is required under Copyright Law or related rights); (ii) they do not involve the doing of acts restricted by copyright or related rights; (iii) the reproductions made fall under other applicable exceptions and limitations, in respect of which contractual override is prohibited.

On a final note, it is worth noting that – despite the availability of two new mandatory EU TDM E&L and leaving the issue of reservation under Article 4(3) of the DSM Directive aside – other legal regimes, as well as contractual restrictions, can also restrict the undertaking of TDM without a licence. Among other things, the application of data protection and privacy laws to the realm of text and data extraction might come into consideration, including relevant provisions on the processing of personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in the GDPR.⁵⁵ Article 28 of the DSM Directive indeed specifically states that the processing of personal data carried out within the framework therein shall be carried out in compliance with both the GDPR and the ePrivacy Directive.⁵⁶ Questions concerning data ownership, this

55. Following the adoption of the Directive, in an op-ed published by *Politico*, the former Rapporteur on the Proposal on behalf of the Committee on Legal Affairs (JURI) of the Parliament (see Report on the Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market (COM(2016)0593 – C8-0383/2016 – 2016/0280(COD))) (Rapporteur: MEP Axel Voss) lamented that provisions in the General Data Protection Regulation 2016/679 “make it impossible to properly use or even develop” several of the “important technologies of the future – such as artificial intelligence” and noted that “[i]n order to train AI systems, you need to be able to process large amounts of data, particularly if you want to avoid discrimination in their algorithms”: A. Voss, “How to bring the GDPR into the digital age”, *Politico*, 25 March 2021, available at <https://www.politico.eu/article/gdpr-reform-digital-innovation/>.

56. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation), *OJ L 119 of 4 May 2016*, pp. 1-88.

being an area surrounded by a degree of legal uncertainty and subject to diverging national approaches, may also arise.⁵⁷ Furthermore, restrictions may subsist in relation to databases that are protected by neither copyright nor the *sui generis* (database) right. As the CJEU has clarified, the holder of a database of this kind is not subject to the limitations to copyright and the *sui generis* right set in the Database Directive. Hence, they are free to determine by contract and in compliance with the applicable national law the conditions of use of their database.⁵⁸

III. THE DEVELOPMENT OF ALGORITHMIC REGULATION THROUGH THE FIGHT AGAINST COPYRIGHT INFRINGEMENT

The fight against copyright infringement online has loaded the CJEU jurisprudence with cases that are important, not only in relation to the specific issue of copyright protection, but for the development as well of a set of principles, framing the use of technologies for the purposes of monitoring, moderating and controlling contents online. Indeed, some of the decisions issued in that particular contexts have broader implication from a normative standpoint and are therefore cornerstones of “Algorithmic Regulation” in the EU (§ 3.1). Those developments are mirrored in the Article 17 of the DSM Directive, which was adopted with the aim of fixing the “Value Gap,” namely tackling the somewhat undue profits made by legitimate operators of the Information Society, an issue that is actually slightly different from traditional counterfeiting (§ 3.2). The decision rendered by the CJEU in the *Poland* case confirming the validity of that provision, though not surprising, has definitely made the EU endorsing the idea of “Algorithmic Regulation” and will certainly serve as a guide for the general provisions to be found in the recent Digital Services Act (DSA) (§ 3.3). A cautious approach needs however to be adopted, as expanding the reasoning can have unforeseen consequences with regards to harmonization (§ 3.4).

57. See M. KOP, “The right to process data for machine learning purposes in the EU”, *Harv JL & Tech*, 2021, Vol. 34, No. 1, pp. 6-7, available at <https://jolt.law.harvard.edu/digest/the-right-to-process-data-for-machine-learning-purposes-in-the-eu>.

58. CJEU, 15 January 2015, *Ryanair*, Case C-30/14, EU:C:2015:10.

A. *The premises of “Algorithmic Regulation”
in the CJUE Copyright Case-Law*

The *Promusicae*⁵⁹ case of the CJEU can be seen as the first and foremost contribution of copyright to the development of a normative framework for “Algorithmic Regulation.” In that case, the rightsholders were seeking before the national judge disclosure of information (that qualifies for personal data) on certain users of KaZaA – a peer-to-peer program that was widely used for illegally sharing copyrighted works –, in order to be able to sue them for copyright infringement. According to the Grand Chamber of the CJEU, that situation raised “[...] the question of the need to reconcile the requirements of the protection of different fundamental rights, namely the right to respect for private life on the one hand and the rights to protection of property and to an effective remedy on the other.”⁶⁰ To answer this question, the CJEU set what can be seen as the most important safeguard, namely the “fair balance” principle. That is to say that when transposing EU Law and implementing such transposition measures, the Member States shall make sure that a “fair balance” is struck between the fundamental rights in conflict.⁶¹

Whereas in this case the “fair balance” principle was set to address the limits of the obligation to communicate personal data of users of an infringing technology, in *Scarlet*⁶² and *Sabam*⁶³ the same principle was then relied on by the CJEU for addressing the limits of resorting to technology for fighting copyright infringement. In those twin cases, Sabam was seeking from the Belgian judge an order requiring an Internet service provider and a hosting service provider (respectively) to install a filtering system for bringing to an end copyright infringement by their customers. Under the specific facts of this case, especially the extremely invasive nature of filtering system envisaged by Sabam, the CJEU considered that an injunction addressed to such service providers to put in place the contested filtering system “[...] would not be respecting the requirement that a fair balance be struck between the right to intellectual property, on the one hand, and the freedom to conduct business, the right to protection of personal data and the freedom to receive or impart information, on the other.”⁶⁴

59. ECJ, 29 January 2008, *Promusicae*, Case C-275/06, EU:C:2008:54.

60. At [65].

61. At [68].

62. CJEU, 24 November 2011, *Scarlet Extended*, Case C-70/10, EU:C:2011:771.

63. CJEU, 16 February 2012, *Sabam*, Case C-360/10, EU:C:2012:85.

64. CJEU, 24 November 2011, *Scarlet Extended*, Case C-70/10, at [53]; CJEU, 16 February 2012, *Sabam*, Case C-360/10, at [51].

On the one hand, it seems appropriate to allow rightholders and intermediaries – including platforms – to rely on technology to protect their rights and interests against infringers, that commonly make use of technologies with this aim. On the other hand, it was made clear in those two judgments that rightholders should not expect the legal system to abide an unconditional use of technological measures for protecting their legitimate interests, although protected by Article 17(2) of the Charter. As the CJEU emphasized, “[...] nothing whatsoever in the wording of that provision or in the Court’s case-law [...] suggest that that right is inviolable and must for that reason be absolutely protected.”⁶⁵ Later on, *UPC Telekabel* would add that such measures must only “be sufficiently effective to ensure genuine protection of the fundamental right at issue, that is to say that they must have the effect of preventing unauthorised access to the protected subject-matter or, at least, of making it difficult to achieve and of seriously discouraging internet users [from such access] [...]”⁶⁶

The *Scarlet* and *Sabam* cases clearly stated what was too far of a use of technology for protecting copyright. They however did not draw a clear line as to what uses of technology, with the aim of protecting legitimate interests, are acceptable under the “fair balance” principle. In that regard, the *UPC Telekabel* case brought an essential development. Here, the injunction issued obliged the Internet service provider to block the access of its customers to a website making available infringing contents, leaving this provider with the freedom to decide how to in practice. In the national proceedings, the appellate court indeed considered that the first court erred in specifying the means that UPC Telekabel had to introduce in order to block the website at issue and thus execute the injunction. Instead, UPC Telekabel could only be required to achieve a particular result (forbidding its customers to access the website at issue), remaining free to decide the means to be used.

Following the reasoning of the CJEU in that case, the injunction was likely to impose a limitation on several fundamental rights (freedom to conduct business, freedom of information, intellectual property). However, those limitations could be accepted provided some procedural guarantees were in place, engaging the CJEU in some sort of “proceduralisation” of the “fair balance” principle.⁶⁷

65. CJEU, 24 November 2011, *Scarlet Extended*, Case C-70/10, at [43]; CJEU, 16 February 2012, *Sabam*, Case C-360/10, at [41].

66. CJEU, 27 March 2014, *UPC Telekabel*, Case C-314/12, EU:C:2014:192, at [62].

67. [at 54; 57].

This procedural way of striking a fair balance can be seen as a counterpart of another milestone set by the CJEU in this case – and reflecting the prohibition of a general monitoring obligation under Article 15 of the e-Commerce Directive –, namely that the measures must be “strictly targeted.” Though the CJEU means by that that “[...] they must serve to bring an end to a third party’s infringement of copyright or of a related right but without thereby affecting internet users who are using the provider’s services in order to lawfully access information”⁶⁸, it also assumes that it will not be always possible for the technology to distinguish adequately between unlawful and lawful content.⁶⁹ From there derives the need for an effective procedural safeguard upstream, since a simple “restoration” downstream of the necessary balance would not be complying with the case-LAW on fundamental rights.⁷⁰

From those four cases emerges the sketch of a normative framework for “Algorithmic Regulation” in the field of copyright, but that could arguably apply to other fields as well. Embedded in the Charter of fundamental rights and the overarching principle of “fair balance,” it assumes that content regulation online could be operated through technologies, provided that they are strictly targeted towards unlawful contents and offer procedural safeguards in case they would reach lawful contents. Some technologies, such as broad filtering measures, would not pass the bar, whereas balanced blocking measures might. But in any case, they shall never be seen as offering absolute protection of the rightsholders.

B. Article 17 DSM Directive: from fighting Copyright Infringement to fixing the Value Gap

It is on the basis of those premises that the Article 17 of the DSM Directive was adopted, a provision that has been largely commented elsewhere.⁷¹ The objective of this provision is fixing the “Value Gap,” namely “the alleged

68. [at 56].

69. Comp. CJEU, 24 November 2011, *Scarlet Extended*, Case C-70/10, at [52]; CJEU, 16 February 2012, *Sabam*, Case C-360/10, at [50].

70. See the Opinion of Advocate General CRUZ VILLALON in CJEU, 27 March 2014, *UPC Telekabel*, Case C-314/12, EU:C:2013:781, at [88].

71. See J. CABAY, “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique: Vers une obligation de filtrage limitée par la CJUE, garante du ‘juste équilibre’”, in J. DE WERRA (ed.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain – Intellectual Property in the era of Big Data and Blockchain*, Geneva/Zürich, Schulthess Editions Romandes, 2020, pp. 169-273, and cited references.

mismatch between the value that online sharing platforms extract from creative content and the revenue returned to the copyright holders.”⁷²

Indeed, under the law as it stood at the time of its adoption, those “online content-sharing service providers” (OCSSP)⁷³ would have not themselves performed an act of communication to the public under the Article 3 InfoSoc Directive unless certain factors – which the CJEU listed in *YouTube*, C-682- and C-683/18 – could be established,⁷⁴ and could in principle benefit from the safe harbour provisions of the Article 14 e-Commerce Directive for the acts accomplished by their users.⁷⁵ Though some of those providers entered into agreements with rightholders for monetizing their content,⁷⁶ this state of the law created a clear asymmetry of information and power, with the rightholders having very small bargaining power in the negotiation.⁷⁷ This situation is referred to in recital 61 of the DSM Directive,⁷⁸ where it also made clear that the EU legislator wished to “foster the development of the licensing market between rightholders and online content-sharing service providers” in order for them to receive an “appropriate remuneration for the use of their works or other subject matter.”

72. C. ANGELOPOULOUS and J.P. QUINTAIS, “Fixing Copyright Reform – A Better Solution to Online Infringement”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 2019, p. 148.

73. See definition in Art. 2(6).

74. After the adoption of the DSM Directive, the CJEU clarified that under certain circumstances, an OCSSP makes an act of communication to the public in the meaning Art. 3 InfoSoc Directive, see CJEU, 22 June 2021, *YouTube v. Cyando*, Cases C-682/18 and C-683/18, EU:C:2021:503 (discussed hereafter). But at the time of the adoption, there was much more uncertainty, despite the evolution of the case law, in particular in CJEU, 14 June 2017, *Ziggo*, Case C-610/15, EU:C:2017:456. The CJEU held that, in circumstances such as those at issue in the main proceedings, comes within the concept of ‘communication to the public’ the making available and management, on the internet, of a sharing platform which, by means of indexation of metadata relating to protected works and the provision of a search engine, allows users of that platform to locate those works and to share them in the context of a peer-to-peer network.

75. See however the evolution of the case law, in particular in CJEU, 12 July 2011, *L’Oréal*, Case C-324/09, EU:C:2011:474. The CJEU held that when the operator of an online marketplace has provided assistance which entails, in particular, optimising the presentation of the offers for sale in question or promoting those offers, it must be considered not to have taken a neutral position between the customer-seller concerned and potential buyers but to have played an active role of such a kind as to give it knowledge of, or control over, the data relating to those offers for sale. It cannot then rely, in the case of those data, on the exemption from liability referred to in Article 14(1) e-Commerce Directive.

76. See Y. LEV-ARETZ, “Second Level Agreements”, *Akron Law Review*, 2012, Vol. 45, No. 1, p. 139.

77. J. CABAY, “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique: Vers une obligation de filtrage limitée par la CJUE, garante du ‘juste équilibre’”, in J. DE WERRA (ed.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain: Intellectual Property in the era of Big Data and Blockchain*, Geneva/Zürich, Schulthess Éditions Romandes, 2020, pp. 185-191.

78. “That uncertainty affects the ability of rightholders to determine whether, and under which conditions, their works and other subject matter are used, as well as their ability to obtain appropriate remuneration for such use.”

In order to reach this, Article 17 brings two major changes. Firstly, paragraph (1) clarifies that OCSSPs perform an act of communication to the public and are therefore required to obtain an authorisation. Secondly, paragraph (3) excludes the benefit of the safe harbour provision from the e-Commerce Directive for those OCSSPs, and substitutes a new regime in case no authorisation was granted. Pursuant to paragraph (4), in order to escape from liability OCSSPs will have to demonstrate that they have:

- “(a) made best efforts to obtain an authorisation, and
- (b) made, in accordance with high industry standards of professional diligence, best efforts to ensure the unavailability of specific works and other subject matter for which the rightholders have provided the service providers with the relevant and necessary information; and in any event
- (c) acted expeditiously, upon receiving a sufficiently substantiated notice from the rightholders, to disable access to, or to remove from their websites, the notified works or other subject matter, and made best efforts to prevent their future uploads in accordance with point (b).”

In its original version (Art. 13 in the proposal), the provision explicitly put forward the use of “effective content recognition technologies” by those service providers as an avenue for ensuring the “functioning of agreements concluded with rightholders,” or to “prevent the availability on their services of works or other subject-matter identified by rightholders through the cooperation with the service providers.”⁷⁹ Though it has been waived in the wording of Article 17(4)(b) and (c), it is admitted that ~~it~~ the use of such technologies is actually the only workable possibility to comply with the obligations ~~they contain~~.⁸⁰ The main objective of this provision is to “foster the development of the licensing market between rightholders and online content-sharing service provider,”⁸¹ and the obligations under Article ~~7~~(4)(b) and (c) only arise when – despite best efforts being made by the concerned OCSSP – an authorization has not been obtained.

This is a subtle change in the role given to technology in Copyright Law. Here, it shall ensure the negotiation of a licence and the payment of an “appropriate remuneration” to rightholders in the sense that if the parties did not succeed in reaching an agreement in the first place, then the OCSSP shall put in place such technological measures, it being understood that the

79. Art. 13 Proposal for a Directive of the European Parliament and the Council on copyright in the Digital Single Market, Brussels, 14 September 2016, COM(2016) 593 final.

80. Such a reading, dominant in the literature, and has been confirmed by the CJEU, *Poland*, Case C-401/19, EU:C:2022:297, at [54].

81. Recital 61.

scope of such obligation is partly defined by the “relevant and necessary information” (b) or notification (c) they received from the rightholders. Now, as it was stressed in the Impact Assessment, some platforms had already “[...] voluntary taken measures to help rightholders in identifying and monetising the use of content on their services, in particular through content identification technologies. [...] They can be applied at the time of upload of the content or later on to verify through an automated procedure whether the content uploaded by users is authorized or not, based on data provided by rightholders.”⁸² In other words, Article 17 is pushing towards making this voluntary measure an actual rule. Technology in that sense is seen as supporting the “preventive”⁸³ nature of the exclusive right, and therefore encouraged to be put in place upfront. As such, it is a clear example of what “Algorithmic Regulation” would look like.

C. Paving the way of “Algorithmic Regulation” with the blessing of the CJEU: from DSM to DSA

That change was seen as a necessity, given the transformation brought by the rapid technological developments, including new ways of creating, producing, distributing, as well as the emergence of new business models and new actors (recital 3 DSM Directive). Such a general view, combined with the idea (expressed in the same recital) that the “[r]elevant legislation needs to be future-proof so as not to restrict technological development,” makes this provision having potential broader implication for the whole society.

Strikingly enough for a piece a legislation in an “innocent” field, its adoption was all but smooth. In particular, until the very end the issue of “upload filters” raised many concerns amongst the public and some governments, as to the balance between fundamental rights,⁸⁴ eventually making the Republic of Poland disputing before the CJEU the compatibility of Article 17(4)(b) and (c) of the DSM Directive with the freedom of expression and information, guaranteed in Article 11 of the Charter.

82. Commission Staff Working Document, “Impact Assessment on the modernization of EU copyright rules”, 14 September 2016, SWD(2016) 301 final, Part 1/3, p. 140.

83. See *e.g.* CJEU, 16 November 2016, *Soulier*, Case C-301/15, EU:C:2016:878, at [33].

84. See in particular the Statement by Germany on the Draft Directive of the European Parliament and the Council on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (first reading), Brussels, 15 April 2019, 2016/0280(COD), 7986/19 ADD 1 REV 2.

Trying to figure out what would be the likely interpretation by the CJEU of this Article 17, we expressed the view that this was a “textbook case” for applying the “fair balance” principle.⁸⁵ Indeed, every paragraph of [Article 17 DSM Directive](#) embodies one or more provisions of the Charter, [which](#) can summarised as follows:

Charter	Article 17 DSM Directive
Protection of personal data (Art. 8)	<ul style="list-style-type: none"> – No general monitoring obligation (8) – No identification of individual users / processing of personal data, except in accordance with GDPR <i>e.a.</i> (9)
Freedom of expression and information (Art. 11) Freedom of the arts and science (Art. 13)	<ul style="list-style-type: none"> – Authorization OCSSP shall also cover acts carried out by users (2) – No prevention of legitimate uses, including copyright exceptions (7)
Freedom to conduct business (Art. 16)	<ul style="list-style-type: none"> – Taking into account features of service, availability/costs means (5) – Taking into account market position (6)
Intellectual property (Art. 17(2))	<ul style="list-style-type: none"> – Extension communication to the public right (1)
Right to an effective remedy (Art. 47)	<ul style="list-style-type: none"> – Complaint and redress mechanism (9)

In addition, close scrutiny evidences that the design of some of those paragraphs mirrors the case-law of the CJEU discussed above,⁸⁶ which in turn reinforces the status of these cases as premises of the “Algorithmic Regulation.”

In its decision in this *Poland* case, the Grand Chamber of the CJEU dismissed the action for annulment, reaching the conclusion that Article 17(4) DSM Directive “[...] has been accompanied by appropriate safeguards by the EU legislature in order to ensure, in accordance with Article 52(1) of the Charter, respect for the right to freedom of expression and information of the users of those services, guaranteed by Article 11 of

85. J. CABAY, “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique: Vers une obligation de filtrage limitée par la CJUE, garante du ‘juste équilibre’”, in J. DE WERRA (ed.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain – Intellectual Property in the era of Big Data and Blockchain*, Geneva/Zürich, Schulthess Éditions Romandes, 2020, p. 196.

86. Comp. for example Art. 17(4) to (6) with CJEU, 27 March 2014, *UPC Telekabel*, Case C-314/12, EU:C:2014:192, at [52; 53]. See J. CABAY, “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique: Vers une obligation de filtrage limitée par la CJUE, garante du ‘juste équilibre’”, in J. DE WERRA (ed.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain – Intellectual Property in the era of Big Data and Blockchain*, Geneva/Zürich, Schulthess Éditions Romandes, 2020, pp. 232-233.

the Charter, and a fair balance between that right, on the one hand, and the right to intellectual property, protected by Article 17(2) of the Charter, on the other.”⁸⁷ That was in line with the conclusion we reached.⁸⁸

In its reasoning, the CJEU upheld the premises of “Algorithmic Regulation” found in its previous case-law. In light of the “fair balance” principle⁸⁹ and referring to the triptych *Funke Medien*⁹⁰/*Spiegel Online*⁹¹/*Pelham*⁹², it emphasized once again that the exceptions and limitations to copyright confer rights on the users of work⁹³. Consequently, safeguarding them is therefore a “specific result to be achieved” by virtue of this law,⁹⁴ which entails that the measures adopted by service providers must be “strictly targeted,”⁹⁵ that a “filtering system which might not distinguish adequately between unlawful content and lawful content” could not be accepted,⁹⁶ and that “procedural safeguards” must be in place.⁹⁷ Importantly, the CJEU borrowed from *Glawischnig-Piesczek*⁹⁸ one additional safeguard, namely that the service providers “cannot be required to prevent the uploading and making available to the public of content which, in order to be found unlawful, would require an independent assessment of the content by them in the light of the information provided by the rightholders and of any exceptions and limitations to copyright.”⁹⁹

But even more important, through this decision the CJEU endorsed the very idea of “Algorithmic Regulation,” beyond the particular context of copyright. Indeed, the whole decision is based on the twin assumption that, on the one hand, “online content-sharing service providers are [...] required to use automatic recognition and filtering tools,” since no one (defendant

87. CJEU, *Poland*, C-401/19, EU:C:2022:297, at [98].

88. J. CABAY, “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique: Vers une obligation de filtrage limitée par la CJUE, garante du ‘juste équilibre’”, in J.-B. WERRA (ed.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain – Intellectual Property in the era of Big Data and Blockchain*, Geneva/Zürich, Schulthess Éditions Romandes, 2020, pp. 225-237.

89. At [66].

90. CJEU, 29 July 2019, *Funke Medien*, Case C-469/17, EU:C:2019:623.

91. CJEU, 29 July 2019, *Spiegel Online*, Case C-516/17, EU:C:2019:625.

92. CJEU, 29 July 2019, *Pelham e.a.*, Case C-476/17, EU:C:2019:624.

93. At [87].

94. At [76-80; 87].

95. At [81]. Compare CJEU, 27 March 2014, *UPC Telekabel*, Case C-314/12, EU:C:2014:192, at [56].

96. At [86]. Compare CJEU, 24 November 2011, *Scarlet Extended*, Case C-70/10, EU:C:2011:771, at [52]; CJEU, 16 February 2012, *Sabam*, Case C-360/10, EU:C:2012:85, at [50].

97. At [93-95]. Compare CJEU, 27 March 2014, *UPC Telekabel*, Case C-314/12, EU:C:2014:192, at [54; 57].

98. CJEU, 3 October 2019, *Glawischnig-Piesczek*, Case C-18/18, EU:C:2019:821, at [45].

99. At [90].

institutions, interveners) has been “[...] able, at the hearing before the Court, to designate possible alternatives to such tools,”¹⁰⁰ and, on the other hand, that “[s]uch a prior review and prior filtering [...] constitute a limitation on the right guaranteed by Article 11 of the Charter.”¹⁰¹ That means that as a matter of EU Law, since the respect of the fundamental rights recognized by the Charter is a “condition of the lawfulness of EU acts,”¹⁰² a limitation on one of those rights through the use of technology is acceptable in principle. This is a blessing.

It is true that already in *Glawischnig-Piesczek* the CJEU admitted that Facebook could recourse to “automated search tools and technologies” for the monitoring of and the searching for defamatory statements made against a specific person, and that had been previously declared illegal.¹⁰³ But in that case, rather than reaching that conclusion on the formal basis of the Charter, the third chamber of the CJEU did so in light of Article 15 e-Commerce Directive prohibiting the imposition of general monitoring obligation on service providers.¹⁰⁴ Besides, the targeted information that could be subject to search monitoring and blocking had already been declared illegal by a national judge. From a normative standpoint, the decision in *Poland* therefore distinguishes from this previous case-law and is arguably more authoritative since it was issued by the Grand Chamber, interpreting the Charter of fundamental rights, in relation to technological measures that would mainly target content that were not previously declared infringing by a judge.

That being said, one may wonder the extent of this normative approach, as such a blessing could rapidly turn into a curse. In that regard, the decision indeed contains three elements that should not be undermined.

First, the CJEU repeated that “the legislation which entails an interference with fundamental rights must lay down clear and precise rules governing the scope and application of the measure in question and imposing minimum safeguards,” and that “the need for such safeguards is all the greater where the interference stems from an automated process.”¹⁰⁵ Here, such safeguards have been found appropriate in the EU Secondary Law, but it was stressed out elsewhere by President Lenaerts that “[...] the

100. CJEU, *Poland*, C-401/19, EU:C:2022:297, at [54].

101. At [55].

102. Opinion 2/13, EU:C:2014:2454, at [169].

103. CJEU, 3 October 2019, *Glawischnig-Piesczek*, Case C-18/18, EU:C:2019:821.

104. At [31-47]. Compare with the Opinion of Advocate General SZPUNAR who explicitly referred to the ‘fair balance’ principle, EU:C:2019:458, at [62; 74].

105. At [67].

way in which a conflict between fundamental rights is resolved *in concreto* is not a matter for Union Law, but for national law. Union Law provides only an analytical framework which Member States must respect.”¹⁰⁶

Second, as an opener for the discussion as to the justification for the limitation on the exercise of the freedom of expression, the CJEU emphasized that “[...] it must be borne in mind that, in accordance with a general principle of interpretation, an EU measure must be interpreted, as far as possible, in such a way as not to affect its validity and in conformity with primary law as a whole and, in particular, with the provisions of the Charter. Thus, if the wording of secondary EU Law is open to more than one interpretation, preference should be given to the interpretation which renders the provision consistent with primary law rather than to the interpretation which leads to its being incompatible with that law.”¹⁰⁷ Given the traditional accusation of “judicial activism” addressed to the CJEU, and the complex debates that eventually led to the adoption of this Article 17, we can’t help but see in that general disclaimer the hint of a decision meant – at least in part – to protect the EU as a political institution, rather than a blanket endorsement of “Algorithmic Regulation.”

Third, also in line with what we expected,¹⁰⁸ the CJEU made clear that its examination is limited to Article 17(4) DSM Directive and “[...] does not prejudice any examination which may subsequently be carried out in relation to the provisions adopted by the Member States for the purposes of transposing that directive or of the measures determined by those providers in order to comply with that regime.”¹⁰⁹ In other words, this is not the final word of the CJEU. This last point is crucial, since it means the apparent embracing of “Algorithmic Regulation” by the CJEU is entirely dependent of the possibility to strike a “fair balance” in practice, which remains to be addressed (what we do in part at § 3.4).

106. K. LENAERTS, “Le droit d’auteur dans l’ordre juridique de l’Union européenne: une perspective constitutionnelle”, in J. CABAY, V. DELFORGE, V. POSSOUL and M. LAMBRECHT (eds), *20 ans de nouveau droit d’auteur – 20 jaar nieuw auteursrecht*, Limal, Anthemis, 2015, p. 245: «la façon dont un conflit entre droits fondamentaux est résolu *in concreto* n’incombe pas au droit de l’Union, mais au droit national” (our translation).

107. At [70].

108. J. CABAY, “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique: Vers une obligation de filtrage limitée par la CJUE, garante du ‘juste équilibre’”, in J. DE WERRA (ed.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain – Intellectual Property in the era of Big Data and Blockchain*, Geneva/Zürich, Schulthess Éditions Romandes, 2020, p. 237.

109. At [71].

Despite those possible limitations that ~~can~~ derive from the wording, with this decision CJEU has given the EU legislator free rein for adopting the Digital Services Act (DSA),¹¹⁰ that updates the e-Commerce Directive and explicitly allows providers of intermediary services in general to “carry out voluntary own-initiative investigations into, or take other measures aimed at detecting, identifying and removing, or disabling access to, illegal content,” without risking losing the benefit of the safe harbour provisions it contains (Art. 7). The DSA also defines obligations that rest upon providers of very large online platforms and of very large online search engines to manage systemic risks (Art. 33 and following), which notably impose those providers to carry out a risk assessment including “any actual or foreseeable negative effects for the exercise of fundamental rights” (Art. 34(1)(b)) and to “put in place reasonable, proportionate and effective mitigation measures, tailored to the specific systemic risks identified pursuant to Article 34, with particular consideration to the impacts of such measures on fundamental rights” (Art. 35(1)). Those last provisions entrench deeply the idea of “Algorithmic Regulation,” and delegates algorithms and companies running them the task of ensuring EU citizens’ fundamental rights. This, admittedly and despite all safeguards, might be a curse.

It goes without saying that if the CJEU had annulled Article 17 DSM Directive, then the adoption of those rules would have been highly hazardous. Besides, though the DSA is “without prejudice” to EU Copyright Law (Art. 2(4)(b)) and should leave unaffected the specific rules and procedures established of the DSM Directive,¹¹¹ the *Poland* case (and possible further preliminary rulings on other aspects of Article 17 DSM Directive and its implementation) will also prove useful for interpreting the provisions of the DSA. In that sense, copyright case-law has been and will contribute to shaping “Algorithmic Regulation” in the EU.

D. *Regulating technology or harmonizing through technology?*

In fancy wording, we are tempted to see the contribution of copyright case-law as identifying, through the lens of the Charter, “Algorithmic

110. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), *OJ L 277* of 27 October 2022, pp. 1-102.

111. See further on the relation between the rules in the DSA and the DSM Directive, J.P. QUINTAIS and S.F. SCHWEMMER, “The Interplay between the Digital Services Act and Sector Regulation: How Special Is Copyright?”, *European Journal of Risk Regulation*, 2022, Vol. 13, pp. 191-217.

Regulation” principles that have to be implemented in the technology, “by design.”¹¹² Already present in particular in Art. 25 GDPR in relation to protection of personal data (“privacy by design”), such an idea to build-in fundamental rights safeguards is clearly to be found in the Article 35(1) of the DSA, since the mitigating measures that providers may adopt include “(a) adapting the design, features or functioning of their services, including their online interfaces; [...] (c) adapting content moderation processes [...] as well as adapting any relevant decision-making processes and dedicated resources for content moderation; (d) testing and adapting their algorithmic systems, including their recommender systems [...]” The Artificial Intelligence (AI) Act Proposal,¹¹³ through prohibiting some AI practices for reason of their contradiction with fundamental rights¹¹⁴ and regulating AI systems that qualify as “high-risk,”¹¹⁵ is in the same vein, in particular when one looks at provisions such as Article 9(4) (risk management and mitigation measures),¹¹⁶ Article 15(3) (biases in feedback loops),¹¹⁷ Articles 53(3) and 54(1)(c) (regulatory sandboxes).¹¹⁸ Those provisions from the DSA and AI Act embody a sort of idea that, “by design,” mitigating measures of adverse effects on fundamental rights should be implemented in the technology.

112. See in particular in the field of Copyright Law M. LAMBRECHT, “Free Speech by Design – Algorithmic protection of exceptions and limitations in the Copyright DSM Directive”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 2020, Vol. 11, No. 1, pp. 68-94; N. ELKIN-KOREN, “Fair Use by Design”, *UCLA Law Review*, 2017, Vol. 64, No. 5, pp. 1082-1101.

113. 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), Brussels, 23 February 2022, COM(2022) 68 final.

114. “[...] Such practices are particularly harmful and should be prohibited because the contradict [...] Union fundamental rights [...]” (recital 15; see practices listed in Art. 5).

115. “[...] common normative standards for all high-risk AI systems should be established. Those standards should be consistent with the Charter of fundamental rights [...]” (recital 13; see AI systems in Art. 6(1) and Annex III referred to in Art. 6(2)).

116. “[...] In identifying the most appropriate risk management measures, the following shall be ensured: (a) elimination or reduction of risks as far as possible through adequate design and development; (b) where appropriate, implementation of adequate mitigation and control measures in relation to risks that cannot be eliminated; [...]”

117. “[...] High-risk AI systems that continue to learn after being placed on the market or put into service shall be developed in such a way to ensure that possibly biased outputs due to outputs used as an input for future operations (‘feedback loops’) are duly addressed with appropriate mitigation measures.”

118. “[...] Any significant risks to health and safety and fundamental rights identified during the development and testing of such systems shall result in immediate mitigation and, failing that, in the suspension of the development and testing process until such mitigation takes place” (Art. 53(3)); “In the AI regulatory sandbox personal data lawfully collected for other purposes shall be processed for the purposes of developing and testing certain innovative AI systems in the sandbox under the following conditions: [...] there are effective monitoring mechanisms to identify if any high risks to the fundamental rights of the data subjects may arise during the sandbox experimentation as well as response mechanism to promptly mitigate those risks and, where necessary, stop the processing” (Art. 54(1)(c)).

Close scrutiny of technologies that are employed in the field of Copyright Law actually reveals that safeguarding fundamental rights “by design” might actually be difficult (even impossible), with the result that a cautious approach seems warranted. Moreover, it appears that regulating technology through Copyright Law *in abstracto* might actually have unintended consequences *in concreto*, as it would lead to *de facto* harmonization through technology. In other words, “Algorithmic Regulation” ~~through~~ a “by design” approach shall be carried out carefully. The following reflections, enshrined in the copyright case-law of the CJEU, are intended to enucleate all this more clearly.

First and foremost, technology is not neutral. Interpreting the private copying exception as applying to copies made through a cloud computing service, the CJEU in *Strato* referred to “the principle of technological neutrality, according to which the law must specify the rights and obligations of persons in a generic manner, so as not to favour the use of one technology to the detriment of another.”¹¹⁹ The CJEU in the case of *Poland* reached the conclusion that the obligation made upon OCSSPs to use technologies has been accompanied by appropriate safeguards by the legislator. In its Opinion, the Advocate General Saugmandsgaard Øe, in a closer reading of the underlying technology, suggested however to differentiate between technologies following a distinction between “false positives” (consisting of blocking legal content) and “false negatives” (consisting of not blocking illegal content).¹²⁰ According to him, “[...] the error rate should be as low as possible. It follows that, in situations in which it is not possible, in the current state of technology, for example as regards certain types of works and protected subject matter, to use an automatic filtering tool without resulting in a ‘false positive’ rate that is significant, the use of such a tool should, in my view, be precluded under paragraph 7.”¹²¹ This part of the Opinion has not been endorsed by the CJEU, but it warrants a deeper reflection on the appropriateness to discriminate between different technologies and their results.

Second, in relation thereto, one should not think of “Algorithmic Regulation” without duly taking into account the state of the art. Whilst Article 25 GDPR (“privacy by design”) explicitly states that it should be taken into consideration,¹²² the DSM Directive only puts this forward in

119. CJEU, 24 March 2022, *Austro-Mechana*, Case C-433/20, EU:C:2022:217, at [27].

120. Opinion of Advocate General SAUGMANDSGAARD ØE in *CJEU, Poland*, C-401/19, EU:C:2021:613, at [207].

121. At [214].

122. See also Art. 9(3) AI Act Proposal.

recital 66 in ambiguous terms, with “the evolving state of the art as regards existing means” being one of the elements to consider in assessing whether an OCSSP has made its “best efforts” pursuant to Article 17(4). Recalling the position of the Advocate General Saugmandsgaard Øe mentioned above, we consider that the state of the art is determining whether an obligation to recourse to technology can be imposed in the first place. Contrasting with the case of *Poland*, in *YouTube/Cyando* the CJEU generally stated that the operator of a video-sharing platform or a file-hosting and -sharing platform does not make a communication to the public within the meaning of Article 3 InfoSoc Directive, “[...] unless it contributes, beyond merely making that platform available, to giving access to such content to the public in breach of copyright [...]”, which is the case where, “[...] despite the fact that it knows or ought to know, in a general sense, that users of its platform are making protected content available to the public illegally via its platform, refrains from putting in place the appropriate technological measures that can be expected from a reasonably diligent operator in its situation in order to counter credibly and effectively copyright infringements on that platform [...].”¹²³ In so doing, the CJEU has been making the scope of a right, and more specifically the possibility to avoid liability for its infringement in one particular case, entirely dependent of the supposed existence of a technology. From the documentation reviewed, it does not seem, however, any analysis of the state of the art was conducted at the time.

Third, in the line of this criticism, it is not enough to assume the existence of such a technology or, though it exists, to assume that it can “by design” integrates appropriate safeguards, in particular as to fundamental rights. In *YouTube v. Cyando*, the existence of such technologies was “apparent” to the CJEU,¹²⁴ which apparently means that Google was convincing as to the efficiency of its system Content ID, even for complex tasks.¹²⁵ But we should be sceptical on the reliability of the quality assessment of one’s own technology, in particular where the investment in developing such technology

123. CJEU, 22 June 2021, *YouTube v. Cyando*, Cases C-682/18 and C-683/18, EU:C:2021:503, at [102].

124. “[...] YouTube has put in place various technological measures in order to prevent and put an end to copyright infringements on its platform, such as [...] content recognition software for facilitating the identification and designation of such content. Thus, it is apparent that that operator has adopted technological measures to counter credibly and effectively copyright infringements on its platform” (at [94]).

125. See the Opinion of Advocate General SAUGMANDSGAARD ØE, EU:C:2020:586, at [222]. See in particular footnote 209: “According to the explanations given by Google, once a reference file for a work has been entered into the Content ID database, that software would automatically recognise (nearly) all files containing that work when they are uploaded.”

were important,¹²⁶ even the more if we consider there is a market for such type of technologies.¹²⁷ Neither the technology or its providers are neutral. In that context, it must be stressed that, as the Commission itself noticed, there are governance issues related to the “[...] training and testing data sets, and the impossibility to compare accuracy and performance across different tools due to inaccessibility of the systems.”¹²⁸ If we are to leave under “Algorithmic Regulation,” a caution approach would be not to take this at face value, despite the incredible progresses of such technologies.

Fourth, in order to substantiate this criticism in the field of intellectual property, one of these authors is currently running a research project (IPSAM) focusing in particular on the Trademark Image Search Engines operated through AI.¹²⁹ Through a comparison of the ability of the tools provided by the Benelux Office for Intellectual Property (BOIP) and the European Union Intellectual Property Office (EUIPO), based on a large-scale statistical analysis, he could recently evidence the very different outcomes of those two systems and seriously question their reliability.¹³⁰ For the purpose of our discussion here, findings of this research have two implications. On the one hand, they show that “efficiency rate,” “error rate,” or “false positives rate” means very little, as it is entirely dependent on the assessment method. As it follows, further refining the approach by the Advocate General Saugmandsgaard Øe, a technology that would evidence a “low error rate” should not *per se* be considered building-in the necessary safeguards as to fundamental rights. On the other hand, the poor (or at least

126. *Ibid.*, at footnote 210: “According to the information provided by Google, it has spent more than 100 million United States dollars (around EUR 88 million) developing Content ID.”

127. See Annex 11 (referring *inter alia* to Audible Magic, Amazon Rekognition, Clarifai, NanoNets) in Commission Staff Working Document, “Impact Assessment Report – Annexes”, Brussels, 15 December 2020, SWD(2020) 348 final, Part 2/2, p. 195.

128. *Ibid.*

129. <https://droit-prive.ulb.be/ipsam-adressing-intellectual-property-relevant-similarities-in-images-through-algorithmic-decision-systems/>.

130. T. VANDAMME, J. CABAY and O. DEBEIR, “A Quantitative Evaluation of Trademark Search Engines’ Performances through Large-Scale Statistical Analysis”, *ICAIL ’23*, 19-23 June 2023, Braga, Portugal (acceptance pending). In this research, thanks to automated procedures we put in place, we have uploaded in the two systems Trademarks (TM) which registration has been opposed (the Applicant’s TM) for reason they were triggering a likelihood of confusion with a prior TM (the Opponent’s TM), eventually leading to a decision by the EUIPO Opposition Division, and verified if and where in the displayed results would appear the TM that had served as a basis for the opposition. Out of 5.757 queries that actually led to a result in both systems, we could show that the Match Ratio (M) for the EUIPO system was comprised between 62.01% and 99.16%, whereas for the BOIP system it was comprised between 7.36% and 11.72%. Though this could have led to the conclusion that the EUIPO was extremely efficient and the BOIP one poorly efficient, a more thorough analysis suggest that the good performances of the EUIPO system are only a result of our method for testing it (the main hypothesis being that the dataset we used for evaluating its performance is actually the very same set, or at least a major subset thereof, that was used for its training).

questionable) efficiency of the technologies that have been subject to the large-scale statistical analysis referred above, in relation to their capacity in identifying not identical signs but relevant similarities between signs under the “Likelihood of confusion” test of Trademark Law, suggest that we should question uses of the technology with the aim of identifying, not “identical” contents but, “similar” contents. In *Glawischmig-Piesczek* the CJEU admitted that the technology used could target not only the “identical” contents but “equivalent” contents as well, provided it would not suppose for the operator concerned to carry out an “independent assessment” of that content.¹³¹ This last element is found in the *Poland* decision, not the distinction between “identical” and “equivalent” contents (though it was seemingly adopted by the Advocate General).¹³² In any case, as we suggested elsewhere, the latter category should be refined distinguishing between “equivalent” and “similar” contents¹³³, which in the case of copyright is in our view more subtle than the general opinion of limiting to what is “manifestly infringing.”¹³⁴

Which leads us to our fifth and last point. Though fingerprinting seems in the particular area covered by Article 17 DSM Directive to be the most widely used technology,¹³⁵ it is one amongst others which, most likely, are going to rely increasingly on AI.¹³⁶ The recent introduction of TDM exceptions in the DSM Directive (discussed above at § 2) will probably contribute to foster its development. But TDM will only permit AI to develop in contexts where relevant data exist. If we do consider the example of Copyright Law, in order to build a content recognition system based on AI able to identify

131. CJEU, 3 October 2019, *Glawischmig-Piesczek*, Case C-18/18, EU:C:2019:821, at [38-46].

132. Opinion of Advocate General SAUGMANDSGAARD ØE in *CJEU, Poland*, Case C-401/19, EU:C:2021:613, at [200]; in CJEU, 22 June 2021, *YouTube v. Cyando*, Cases C-682/18 and C-683/18, EU:C:2020:586, at [220-221].

133. J. CABAY, “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique: Vers une obligation de filtrage limitée par la CJUE, garante du ‘juste équilibre’”, in J. DE WIJCKA (ed.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain – Intellectual Property in the era of Big Data and Blockchain*, Geneva/Zürich, Schulthess Éditions Romandes, 2020, pp. 238-240.

134. See Opinion of Advocate General SAUGMANDSGAARD ØE in *Poland*, Case C-401/19, EU:C:2021:613, at [201]; Communication of the EU Commission, *Guidance on Article 17 of Directive 2019/790 on Copyright in the Digital Single Market*, Brussels, 4 June 2021, COM(2021) 288 final, p. 20; J.P. QUINTAIS, G. FROSIO, S. VAN GOMPEL, P.B. HUGENHOLTZ, M. HUSOVEC, B.J. JÜTTE and M. SENFLEBEN, “Safeguarding User Freedoms in Implementing Article 17 of the Copyright in the Digital Single Market: Recommendations from European Academics”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 2019, Vol. 3, p. 280.

135. Communication of EU Commission, “Guidance on Article 17 of Directive 2019/790 on Copyright in the Digital Single Market”, Brussels, 4 June 2021, COM(2021) 288 final, p. 12.

136. See for a mapping of relevant technologies (including AI) in the field of intellectual property EUIPO Study, *Automated Content Recognition: Discussion Paper – Phase 1 ‘Existing technologies and their impact on IP’*, November 2020.

infringing contents which are not limited to “equivalent” ones but extend to “similar” ones, we should at least have a common ground for assessing similarities. Yet, there is so far no (or very limited) harmonization of the substantive requirements for copyright infringement, in particular, the relevant criterion for assessing similarities when the alleged infringing work is transformative.¹³⁷ Which means that if, as we believe, not enough legal data is available as to what are the relevant similarities for copyright infringement, or if it is not possible identify regular patterns as to this similarity assessment,¹³⁸ or if the data are affected with biases,¹³⁹ it is just not possible to build such a technology. Which conversely means that building a technology pretending to be able to do so would lead to *de facto* harmonization, based on the standard it implements.¹⁴⁰ If such a standard

137. This was one of the question referred to the CJEU in *Levola* (Case C-310/17, EU:C:2018:899), that has not been answered since it found that a taste could not enjoy copyright protection in the EU (see question 2(d): “How should the court in infringement proceedings determine whether the taste of the defendant’s food product corresponds to such an extent with the taste of the applicant’s food product that it constitutes an infringement of copyright? Is a determining factor here that the overall impressions of the two tastes are the same?”). Currently, there is absolutely no harmonization in relation to the relevant ‘similarity’ criteria. As an example, whereas the Dutch Hoge Raad has considered that copyright infringement should be addressed in the light of the ‘overall impression’ between the works, the French Court of cassation has ruled to the contrary that the judge cannot subject a finding of infringement to such an ‘overall impression’, see Hoge Raad, 12 April 2013, ECLI:NL:HR:2013:BY1532, at [5.3]; Cass. (com.) (France), 8 April 2014, *P.I.*, 2014, p. 271, obs. A. LUCAS. See further on that issue J. CABAY, *L’objet de la protection du droit d’auteur – Contribution à l’étude de la liberté de création*, Ph.D. Thesis, Université libre de Bruxelles, 2016, pp. 309-313.

138. See for example, identifying ~~the many different criteria in Belgian case law~~, J. CABAY, “Qui veut gagner son procès?” – Lavis du public dans l’appréciation de la contrefaçon en droit d’auteur”, *A&M*, 2012, p. 13-29. See also in the USA, where the ‘substantial similarity’ is often criticized as lacking uniformity, a recent study substantiating these claims through reviewing a random sample of 1.005 substantial similarity opinions issued between 1978 and 2020, C.D. ASAY, “An Empirical Study of Copyright’s Substantial Similarity Test”, *UC Irvine Law Review*, 2022, Vol. 13, No. 1, pp. 35-102.

139. See in particular in the USA, where the studies have shown cognitive biases in the assessment of the substantial similarities, C. KLONICK, “Comparing Apples to Applejacks: Cognitive Science Concepts of Similarity Judgment and Derivative Works”, *Journal of the Copyright Society of the U.S.A.*, 2013, Vol. 60, No. 3, pp. 365-386; I.D. MANTA, “Reasonable Copyright”, *Boston College Law Review*, 2012, Vol. 53, No. 4, pp. 1303-1356; L. MCKENZIE, “Drawing Lines: Addressing Cognitive Bias in Art Appropriation Cases”, *UCLA Entertainment Law Review*, 2013, Vol. 20, No. 1, pp. 83-106. Such biases have been confirmed through empirical studies, see S. BALGANESH, I.D. MANTA and T. WILKINSON-RYAN, “Judging Similarity”, *Iowa Law Review*, 2014, Vol. 100, No. 1, pp. 267-290 (influence of various elements brought to the knowledge of the relevant public on its appraisal); J. LUND, “An Empirical Examination of the Lay Listener Test in Music Composition Copyright Infringement”, *Virginia Sports & Entertainment Law Journal*, 2011, Vol. 11, No. 1, pp. 137-177 (influence of musical performance on musical infringement assessment).

140. Compare with the idea that a ‘fair use’ standard translated into algorithms would change its nature, see D.L. BURK, “Algorithmic Fair Use”, *University of Chicago Law Review*, 2019, Vol. 86, No. 2, pp. 283-307 (concluding that “attempting to incorporate fair use into enforcement algorithms threatens to degrade the exception into an unrecognizable form. Worse yet, social internalization of a bowdlerized version of fair use deployed in algorithmic format is likely to become the new legal and social norm”, at [306]).

does not reflect EU Law, then harmonization could not be seen as the deliberated product of EU legislature, but rather the result of the interaction between the architectures of Copyright Law and AI technology. Findings of our IPSAM research project evidence that in the field of Trademark, despite the relatively set case-law on the “Likelihood of confusion,” the number and stable formatting of decisions of the EUIPO, ~~the task seems elusive~~. In comparison, we cannot imagine the task being easier in the field of copyright, to the contrary.¹⁴¹ The Commission admitted that “[i]n the present state of the art, no technology can assess to the standard required in law whether content, which a user wishes to upload, is infringing or a legitimate use.”¹⁴² We posit that even in the future it will never be possible, for the simple reason that the data needed for building such technology does not and will not exist. Failing to admit this would mean accepting technologies which are not compliant with fundamental rights.¹⁴³

All that said leads to one conclusion. If we were to leave “Algorithmic Regulation” without discriminating amongst technologies, we would open the door to technologies that might not, “by design,” include the necessary fundamental rights safeguards, for the simple reason that such a designing is sometimes impossible. In the field of copyright, the Advocate General Saugmandsgaard Øe saw here a danger, that is if content recognition tools were to reach uses that are not included in the monopoly of rightholders, it would be to “to the detriment of other forms of creativity which are also positive for society.”¹⁴⁴ But if one think of outside of copyright, no doubt the danger might be much greater.

Hence, we cannot think of regulating technologies without considering the risk that, in return, it will harmonize our laws and society to an extent

141. The IPSAM research project aims amongst others at substantiating this concern, through empirically verifying some of the concerns expressed in the literature as to the possibility to implement the ‘Likelihood of confusion’ standard in a AI technology, see in particular D. S. GANGJEE, “Eye, Robot: Artificial Intelligence and Trade Mark Registers”, in N. BRUUN, G.B. DINWOODIE, M. LEVIN and A. OHLY (eds), *Transition and Coherence Intellectual Property Law*, Cambridge, Cambridge University Press, 2021, pp. 185-190.

142. Communication of EU Commission, *Guidance on Article 17 of Directive 2019/790 on Copyright in the Digital Single Market*, Brussels, 4 June 2021, COM(2021) 288 final, p. 20.

143. J. CABAY, “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique: Vers une obligation de filtrage limitée par la CJUE, garante du ‘juste équilibre’”, in J. DE WERRA (ed.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain – Intellectual Property in the era of Big Data and Blockchain*, Geneva/Zürich, Schulthess Éditions Romandes, 2020, pp. 244-349.

144. Opinion of Advocate General SAUGMANDSGAARD ØE in *Poland*, Case C-401/19, EU:C:2021:613, at [190]; in *CJEU*, 22 June 2021, *YouTube v. Cyando*, Cases C-682/18 and C-683/18, EU:C:2020:586, at [243].

we do not foresee. Having in mind the “Collingridge Dilemma,”¹⁴⁵ we hope those few reflections enshrined in Copyright Law will contribute to keeping “Algorithmic Regulation” under control.

IV. CONCLUSION: THE TWO TALES OF COPYRIGHT LAW IN THE DIGITAL SINGLE MARKET

The new regimes brought by the provisions discussed offer two readings as to the role of Copyright Law in bringing the Digital Single Market to reality. Both TDM exceptions and Content recognition systems are far from limited to the copyright realm and, to the contrary, have a much broader reach from a normative standpoint. One would remember that the InfoSoc Directive led to debates as to whether its focus was on “Copyright” or on the “Information Society.”¹⁴⁶ Arguably, the richness of the CJEU case-LAW interpreting this Directive and exhuming fundamental concepts (starting with the originality criterion) stressed out the copyright aspect of the law. Conversely, as we evidenced in this contribution, the DSM Directive is of so much importance for the development of technologies, AI in particular, that one can argue its main focus is not copyright anymore, but the Digital Single Market. These have always been the two tales of Copyright Law in the EU.

Through regulation of the technologies envisaged and the regime put in place, Copyright Law bridges the gap between two micro and macro levels and appears a way forward for building a Single Digital Market.

Obviously, operators on the Single Market ~~will~~ can make their own arrangements for operating on a cross border market through contracts. Both Article 4 (licence or opt-out) and Article 17 (licence or filter) DSM Directive actually promote the conclusion of licenses as a first option. Hence, if Copyright Law is an essential component of the emerging “Data Law” governing technologies at the micro-level, on that level copyright Contracts are one of the building blocks of a Digital Single Market.

But such a Single Market cannot grow without solid digital infrastructures. Once again, both Article 4 (supporting AI) and Article 17

145. D. COLLINGRIDGE, *The Social Control of Technology*, London, Frances Pinter, 1980, p. 19: “[...] attempting to control a technology is difficult, and not rarely impossible, because during its early stages, when it can be controlled, not enough can be known about its harmful social consequences to warrant controlling its development; but by the time these consequences are apparent, control has become costly and slow.”

146. See e.g. L. BENTLY, “The return of industrial copyright?”, *European Intellectual Property Review*, 2012, Vol. 34, No. 10, p. 661.

(possibly supported by AI) DSM Directive actually regulate important part of these infrastructures. In particular, since the use thereof in the realm of copyright is intrinsically linked with important values in a democratic society, such as freedom of expression and information, it makes Copyright Law a vehicle for developing the general framework of Fundamental Rights which, at the macro-level, governs the very same technologies. Fundamental Rights having on Copyright Law a certain harmonizing effect,¹⁴⁷ we see here again how essential it is to the development of a Digital Single Market.¹⁴⁸

Reconciling the “Laws of everything” at both micro- and macro- level in the Digital world, what Physics cannot in the Real world, that is one of a pleasant copyright story. And even if that wasn’t true, that would still be a great copyright story which, after all, is more about Art and Literature, rather than Science and Technology. Or is it really so? Well, that’s another tale of two Copyrights¹⁴⁹...

147. J. GRIFFITHS, “Constitutionalising or harmonizing? The Court of Justice, the right to property and European copyright law”, *European Law Review*, 2013, Vol. 38, No. 1, pp. 65-78.

148. See further on this J. CABAY, “La Cour de justice, le droit d’auteur et le marché unique numérique: voyage intertextuel au pays des hyperliens”, in N. BERTHOLD (coord.), *Droit de la propriété intellectuelle – Actualité législative et jurisprudence récente de la Cour de justice de l’Union européenne*, Bruxelles, Larcier, 2019, pp. 51-82.

149. See J.C. GINSBURG, “A Tale of Two Copyrights: Literary Property in Revolutionary France and America”, *RIDA*, 1991, p. 124.