



universität
wien

Exposé

For the doctoral thesis

Rethinking copyright in the context of Text and Data Mining and AI Training

Submitted by

Mag. iur. Christoph Korab

Aspired academic degree

Doktor der Rechtswissenschaften (Dr. iur.)

Vienna, June 2023

| | |
|------------------------|-------------------------------------|
| Field of Study | Rechtswissenschaften |
| Degree programme code: | A 783 101 |
| Field of Dissertation: | Immaterialgüterrecht (Urheberrecht) |
| Matriculation number | 11703257 |

1. An introduction to the problem

In the last years data has become an increasingly valuable resource. Every day, 2.5 quintillion bytes of data are created, and this number is expected to rise.¹ The data economy (meaning the overall impact of the data market on the economy) is expected to reach a 1.1. trillion € value in the UK by 2030 and around €1 trillion in the EU.² In 2020 already the EU therefore set forth a European strategy to harvest the value of the data created within the Union and assure access to data as well as a high safety standard.³ The legal landscape has, as a consequence of this plan of action, seen a number of legislation put forward, dealing with questions surrounding data, notably the access to data. The proposal for a Data Act⁴ aims to establish rules for promoting fairness in the allocation of value from data among actors in the data economy and to foster access to and use of data.⁵ The Data Governance Act⁶ aims to foster the availability of data from data intermediaries and to strengthen data-sharing mechanisms across the EU. The Digital Markets Act⁷ finally targeted large platforms, that are key players in the new data economy, and aimed to reduce their market power in an antitrust-related matter. All of these acts have in common that they deal with questions surrounding data (access) and explicitly state that they are without prejudice to the area of intellectual property law, namely copyright and neighboring rights.⁸ This is consequent in some perspectives, less so in others. It is consequent insofar as the regulation of copyright is a sector specific regulation with its own aims, rationales and theories that existed long before data economy. Additionally, copyright does not monopolize factual information, but only protects the

¹ Eleonora Rosati, 'The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market- Technical Aspects. In-Depth-Analysis' 2.

² Mike Glennon and others, 'European Data Market Study 2021-2023 : Second Report on Facts and Figures' 21.

³ European Commission, 'A European Strategy for Data' <<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066>> accessed 14 May 2023.

⁴ European Commission, '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)'.

⁵ *ibid* explanatory memorandum 2.

⁶ 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) 2022 (OJ L 152/1).

⁷ 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) 2022 (OJ L 265).

⁸ 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) recital 17; 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) recital 12; European Commission, 'COM/2022/68 Final' (n 4) explanatory memorandum 4.

expression of creativity.⁹ Therefore it has been argued by some authors in literature that copyright regimes should not make the legality of extracting information by automated means dependent on the authorization of rightsholders because such activities lie outside the scope of copyright.¹⁰ The silence of the European legislator on the issue of data access in the field of copyright is however inconsequential insofar, as the only rules that it put forward on the matter – Art. 3 and 4 of the DSM-Directive¹¹ - do not deliver a stringent answer to the pressing question, whether the extraction of information by automated processes should constitute an activity infringing on the exclusive rights of authors.

The proposed thesis wants to firstly examine the non-consequential status quo of the European Copyright regime when it comes to the automated extraction of information. It then proposes to explore the rationale and different theories behind copyright in order to give an answer to the core question, whether extraction of information from protected subject matter should constitute an infringing activity requiring (any) consent from the author. Finally, it will suggest how to alter the current legal regime in order to render it more stringent in answering the questions identified above.

2. The current legal situation

Copyright – and by its extension neighboring rights – has evolved into a regime of very broad, technical exclusive rights and very narrow exceptions.¹² As a reaction to quick and possibly disruptive technological advancements the CJEU and the EU have consequently broadened the scope of protection of copyright, in order to give authors the possibility to economically benefit from new forms of uses.¹³ This was accompanied

⁹ Thomas Margoni and Martin Kretschmer, 'A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology' (2022) 71 GRUR International 685, 689.

¹⁰ Margoni and Kretschmer (n 9); Matthew Sag, 'The New Legal Landscape for Text Mining and Machine Learning' (2019) 66 Journal of the Copyright Society of the USA 291, 301–309; Benjamin Raue, 'Free Flow of Data? The Friction Between the Commission's European Data Economy Initiative and the Proposed Directive on Copyright in the Digital Single Market' (2018) 49 International Review of Intellectual Property and Competition Law 379; Eleonora Rosati, 'Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity' (2019) 27 Asia Pacific Law Review 198.

¹¹ 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 2019 (OJ L 130).

¹² Matthias Leistner, 'Europe's Copyright Law Decade: Recent Case Law of the European Court of Justice and Policy Perspectives' (2014) 51 Common Market Law Review 559, 569.

¹³ Simone Schroff, 'The Purpose of Copyright-Moving beyond the Theory' (2021) 16 Journal of Intellectual Property Law & Practice 1262, 1266–1267.

by a constant lowering of the threshold requirements for originality and creativity¹⁴ and consistently strict interpretations of a closed list of exceptions, leading in combination to a situation where a vast number of actions in the digital domain are considered infringing.¹⁵ Especially the reproduction right has evolved into an all-engulfing instrument of copyright exclusivity, because it has been interpreted very technical without consideration to the intended purpose of the reproduction act at hand.¹⁶ The large scope of protection might in parts also be in violation of Article 11 of the charter of fundamental rights of the European Union, which sets forth that the free flow of information is the default setting and any limitation has to satisfy the conditions set out in Article 11 and 52 of the charter.¹⁷ Currently, the copyright and neighboring rights regime however works reversed to the situation envisaged by Art. 11 of the charter: it introduces a default setting of protection and then goes on to stipulate a catalogue of exceptions that is to be narrowly interpreted.¹⁸ Under this legal regime of broad and sometimes seemingly random protection, complex data mining activities necessary for extracting information and relevant for techniques such as machine learning encounter legal challenges, because they almost always include infringing activities of some form in the course of the process, possibly even multiple. Especially acts of reproduction are of central importance in the context of data extrapolation from protected subject matter via automated means but constitute an infringing activity regardless of the intended purposes. The DSM directive therefore recognized the necessity to create more legal certainty for economic operators undertaking automated information extraction, because the steps carried out in the process have multiple points of contact with the exclusive rights regime.¹⁹ However its approach remains somewhat inconsequent, because it seems to imply, that the process of extracting information from protected subject matter to gain new information is not a relevant use under the

¹⁴ see for example *Painer* [2011] CJEU C-145/10 ECR 798 on the low originality threshold for photographs.

¹⁵ Margoni and Kretschmer (n 9) 697–698.

¹⁶ Martin Senftleben, ‘Flexibility Grave - Partial Reproduction Focus and Closed System Fetish in CJEU, Pelham’ (2020) 51 IIC-International Review of Intellectual Property and Competition Law 751.

¹⁷ P Bernt Hugenholtz, ‘Data Property in the System of Intellectual Property Law’ in Sebastian Lohsse, Schulze Reiner and Dirk Staudenmayer (eds), *Trading data in the digital economy: Legal concepts and tools* (Nomos 2017) 94–97.

¹⁸ Robert P Merges, ‘The Economic Impact of Intellectual Property Rights: An Overview and Guide’ (1995) 19 Journal of Cultural Economics 103, 110; P Bernt Hugenholtz, ‘Neighbouring Rights Are Obsolete’ [2019] IIC-International Review of Intellectual Property and Competition Law 1006.

¹⁹ 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 recital 8; European Commission, ‘COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT on the Modernisation of EU Copyright Rules’ 104–106.

copyright regime but at the same time creates a limited exception for reproductions made in the course of data mining. This leads to a factual restriction of data mining outside the scope of the exception, because such processes almost always include making reproductions of some form²⁰ and therefore, following *e contrario* from the existence of the exception, require a license outside its scope. Since in alignment with the existing jurisprudence of the CJEU existing exceptions have to be interpreted narrowly, the European legislator has defined the scope of the exception – as the thesis intends to show – too narrowly if it is taken into consideration that it is exhaustive. As *Margoni* and *Kretschmer* have pointed out, despite the legislative indications, that information extraction is not a relevant use,²¹ the DSM directive through introducing an exception subjects the majority of data mining activities to the regime of author’s exclusive rights and thereby creates an opposite situation. This is due to the occurring of reproduction (or other, under copyright relevant) acts within the process of data mining. The thesis wants to show that this approach is not consequent and alien to the nature of copyright and that instead of introducing fragmented exceptions that threaten legal certainty, rather a holistic rethinking of the copyright regime in the context of text and data mining and AI training would be beneficial.

Another legal basis for justifying the reproductions made in the course of automated information extraction is Art. 5 Par. 1 of the InfoSoc Directive pertaining to transient copies without economic significance. However, its scope remains too narrow to effectively cover the needs for carrying out automated information extraction, especially because training data sets usually are intended to be reused.

A factor often overlooked within the relation of copyright and automated information extraction are the so-called moral rights. These rights aim to protect the authors “personal sphere” and award them with tools to protect their moral interests in the integrity of the work or their relation to the work.²² What rights exactly are accorded to the authors differs depending on the jurisdiction.²³ Major international copyright treatise have only set out minimum harmonization guidelines on moral rights.²⁴ Also on a

²⁰ Margoni and Kretschmer (n 9) 687–688.

²¹ see 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 recital 9.

²² Mira T Sundara Rajan, *Moral Rights: Principles, Practice and New Technology* (1st edn, Oxford University Press 2011) 7–14; Toms, ‘§19 Schutz Der Urheberschaft’ in Guido Kucsko and Christian Handig (eds), *urheber.recht* (2nd edn, Manz 2017) marginal number 5-10.

²³ Rajan (n 22) 11.

²⁴ *ibid* 12.

European level, the discretion has been left to the member states, with the CJEU for example clarifying that the right of adaptation has not been harmonized.²⁵ However, moral rights can be of relevance when using protected subject matter for the purpose of automated information extraction, especially where the information is used for generative AI.²⁶ A current class action against GitHub and Microsoft based on the defendants altering of copyright meta information identifying the author of training data serves as an example of their relevance.²⁷ The thesis proposes to examine how moral rights – especially the right of adaptation and the right of integrity – relate to automated information extraction processes and whether they could be an obstacle not sufficiently considered in the legal debate. Because of the non-harmonized nature of these rights, it proposes to focus on the legal situation in Germany and France as being central jurisdiction for the European market and subsequently pursue a comparative approach of the continental moral rights to moral rights of authors in a common law jurisdiction.

3. Should Copyright monopolize automated information extraction?

The thesis proposes to look at this core question firstly from a historical point of view and secondly to analyze it within existing theories pertaining to the rationale and fundamental principles of copyright law.

Historically two essential tendencies can be identified, when relating to the proposed question of whether copyright should monopolize automated information extraction. The first is, that awarded copyright monopolies have always been regarded from a utilitarian perspective as the most effective way to stimulate education and the progress of intellectual activity. Already the Statue of Anne – by scholars regarded as the first copyright text in history – clearly enunciates this objective.²⁸ Even in jurisdiction, where a natural right approach was more dominant, the aim of promoting education and intellectual activity was among the main factors to shape the scope of protection for authors, at least when determining the exceptions from the protection.²⁹ The thesis proposes to demonstrate how from this historical tendency an operating rationale can be deduced that is relevant for a European copyright system, that

²⁵ Margoni and Kretschmer (n 9) 701.

²⁶ Rita Matulionyte, 'Can AI Infringe Moral Rights of Authors and Should We Do Anything about It? An Australian Perspective' [2023] Law, Innovation and Technology.

²⁷ Matthew Butterick, 'GitHub Copilot Litigation' <<https://githubcopilotlitigation.com/>>.

²⁸ Gillian Davies, *Copyright and the Public Interest* (VCH 1994) 2.

²⁹ *ibid* 82–83.

understands itself as being in the tradition of its member states and international agreements on copyright. The second tendency that can be identified is, that throughout history the extension of copyright privileges was in most cases brought on by a need to adapt to technological development, that threatened the authors' control over works and their participation in its economic exploitation.³⁰ The thesis wants to shape out the key factors, that historically led to the introduction of new rights as a response to technological development and transfer the found insights to the case at hand.

Legal scholars have put forward different theories justifying copyright from a theoretical perspective and analyzing its objectives. A substantial number among those use an economic approach, which considers copyrights objective to lie in assuring the production of creative content in a society and the spreading of knowledge by providing economic incentive to those who create the content. It achieves its purpose through a monopoly, that constantly needs to be reevaluated for its function, because it seeks to consolidate the interest of the public in intellectual creation and authors' interest of being rewarded. As Macauley put it, the monopoly of author is an evil, that for the sake of good society must submit to, but that should not last a day longer than necessary.³¹ Automated information extracting shifts weights in an equilibrium established over hundreds of years and only recently subjected to fast, disruptive developments with the resurfacing of digital technologies. Besides economic theories that have been developed to explore and explain the function of copyright as an incentive for creatives, copyright has also been examined through the lens of its function of enabling cultural expression in democratic societies.³² As a middle ground between these approaches, consequent sensitive theories aim to take into account both rationales by taking other values than wealth maximization into consideration but at the same time attending to the multiple impacts that changing a rule of law has.³³ The use of copyright protected subject matter for automated information extraction or the training of AI touches upon many key principles of the legal regime. The present thesis therefore proposes to

³⁰ João Pedro Quintais and Joost Poort, 'A Brief History of Value Gaps: Pre-Internet Copyright Protection and Exploitation Models' in P Bernt Hugenholtz (ed), *Copyright Reconstructed: Rethinking Copyright's Economic Rights in a Time of Highly Dynamic Technological and Economic Change* (Kluwer Law International 2018).

³¹ Davies (n 28) 24–25.

³² see most famously Neil Weinstock Netanel, 'Copyright and a Democratic Civil Society' (1996) 106 *The Yale Law Journal* 283.

³³ Oren Bracha and Talha Syed, 'Beyond Efficiency: Consequence-Sensitive Theories of Copyright' (2014) 29 *Berkely Technology Law Journal* 229, 231–235.

assume a holistic, theoretical approach to rethink copyrights position towards and objectives for this new technology, seeking in that to find especially an answer as to how the European copyright regime can be aligned in this new environment with the traditional objectives and functions of copyright. To answer this question, it will not be sufficient to analyze the economic aspects but also be necessary to consider the impact on other values as well such as the freedom of expression, the public interest or the advancement of technology.

4. What is the problem with the current legal situation?

The current situation – as described above – is inconsequential and undesirable for a plethora of reasons. Firstly, the general restriction on automated information extraction paired with limited, and restrictive exceptions is inconsistent with the Europeans ambition to liberalize the access to data in the digital single market, especially if copyright should not protect the data contained within the works but is only relevant in these processes because of the wide, technical interpretation of the reproduction right and other exclusive rights.³⁴ By choosing the current path, the copyright regime could also give rise to an imbalanced, unfair market situation for small and medium sized enterprises, since “gatekeepers” within the meaning of the Digital Markets Act or other large companies in relevant sectors will often have the means to lever a license from rightsholders, while at the same time having no incentive to share their access with smaller operators.³⁵ The restricted access is also likely to aggravate a bias problem in AI, since it excludes a considerable amount of high-quality data from AI system developers, that will very likely resort to possibly incomplete datasets to avoid the risk of legal prosecution, if rightsholders seek to enforce their rights vigorously.³⁶ From a dogmatic point of view, the current legal regime is not consequent, because it does not provide a clear policy perspective for the issues relating to automated information extraction from copyrighted subject matter.

³⁴ Margoni and Kretschmer (n 9); Rosati, ‘Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity’ (n 10).

³⁵ Jenny Quang, ‘Does Training AI Violate Copyright Law?’ 36 Berkely Technology Law Journal 1407, 1425–1427.

³⁶ Amanda Levendowski, ‘How Copyright Law Can Fix Artificial Intelligence’s Implicit Bias Problem’ (2018) 93 Washington Law Review 579.

Lastly the thesis wants to submit a proposition of what are more consequent approach to the legal issues relating to automated information extraction could look like and show how a reshaping of copyright law, that goes beyond the introduction of an exception, could help in establishing a more stringent legal regime. It should reflect a suggestion for both possible answers to the core question, whether automated information extraction should be as a default exclusively reserved to the rightsholder, because in the end this decision is a policy decision and the ambition of the thesis is to show how the current, ambiguous position to it does not fit into the existing legal and theoretical framework, and not to suggest a policy decision.

5. Preliminary table of contents

1. Identification of the problem
 - a. Increasing value and importance of data
 - b. Automated information extraction and its use for AI
 - c. Copyright, related rights and the training of AI – technical aspects
 - d. A new form of use or no use at all?
2. Current legal situation
 - a. Copyright
 - i. Originality requirements under European copyright law
 - ii. Exclusive rights and their technical, broad interpretation
 - iii. Moral rights and automated information extraction
 1. Germany
 2. France
 3. Comparative Approach
 - b. Neighboring rights
 - i. Neighboring rights harmonized in European Law
 1. Performers

2. Phonogram producers
 3. Producers of first fixation of films
 4. Broadcasting organizations
 5. Press publishers
 - ii. Database sui-generis right
 - iii. Non-harmonized neighboring rights
 - c. Exceptions provided for in European Law relating to automated information extraction
 - i. Art 5 (1) Information Society Directive
 - ii. Art 3 and Art 4 Copyright in the Digital Single Market Directive
 - d. Preliminary conclusion concerning the current legal situation
3. Should Copyright monopolize the automated extraction of information?
 - a. Outline of the historical development of copyright and neighboring rights
 - b. Theories on the nature and reasons of copyright
 - i. Copyright as economic incentive
 - ii. Copyright as enabler of cultural expression in democratic societies
 - iii. Holistic approach
4. Where are the problems with the current regime?
 - a. Monopolization of data
 - b. Aggravation of bias problem by limiting access to copyrighted subject matter
 - c. No consequent answer to essential questions
 - d. Alignment with other IP-rights
5. What could a more consequent approach look like?

Literature

Bracha O and Syed T, 'Beyond Efficiency: Consequence-Sensitive Theories of Copyright' (2014) 29 Berkely Technology Law Journal 229

Cleve J and Lämmel U, *Data Mining* (3rd edn, De Gruyter 2020)

Davies G, *Copyright and the Public Interest* (VCH 1994)

Dreier T and Hugenholtz PB (eds), *Concise European Copyright Law* (2nd edn, Kluwer Law International 2016)

Drexl J and Hilty RM, 'Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective' (2019)
<<https://ssrn.com/abstract=3465577>>

European Commission, 'COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT on the Modernisation of EU Copyright Rules'

—, 'A European Strategy for Data' <<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066>> accessed 14 May 2023

—, 'Study on Copyright and New Technologies. Copyright Data Management and Artificial Intelligence.' (Publications Office of the European Union 2022)

—, 'Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on Harmonised Rules on Fair Access to and Use of Data (Data Act)'

Glennon M and others, 'European Data Market Study 2021-2023 : Second Report on Facts and Figures'

Grätz A, *Künstliche Intelligenz Im Urheberrecht* (1st edn, Springer 2021)

Han J, Pei J and Kamber M, *Data Mining: Concepts and Techniques* (Elsevier Science 2011)

Heath C, Sanders AK and Moerland A (eds), *Intellectual Property Law and the Fourth Industrial Revolution* (Kluwer Law International 2020)

—, 'Neighbouring Rights Are Obsolete' [2019] IIC-International Review of Intellectual Property and Competition Law 1006

Ian Goodfellow, Yoshua Bengio and Courville A, *Deep Learning. Das Umfassende Handbuch* (1st edn, mitp 2018)

Kucsko G and Handig C (eds), *urheber.recht* (2nd edn, Manz 2017)

Kuschel L and Dolling J, 'Access to Research Data and EU Copyright Law' (2022) 13 JIPITEC 247

Lee J-A, Hilty RM and Liu K-C (eds), *Artificial Intelligence and Intellectual Property* (1st edn, Oxford University Press 2021)

Leistner M, 'Europe's Copyright Law Decade: Recent Case Law of the European Court of Justice and Policy Perspectives' (2014) 51 Common Market Law Review 559

Levendowski A, 'How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem' (2018) 93 Washington Law Review 579

Linke D, *Künstliche Intelligenz Und Urheberrecht- Quo Vadis?* (1st edn, Nomos 2021)

Maamar N, *Computer Als Schöpfer. Der Schutz von Werken Und Erfindungen Künstlicher Intelligenz* (1st edn, Mohr Siebeck 2021)

Margoni T and Kretschmer M, 'A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology' (2022) 71 GRUR International 685

María Jesús G-E and Pavón J, *An Introductory Guide to Artificial Intelligence for Legal Professionals* (Kluwer Law International 2020)

Matulionyte R, 'Can AI Infringe Moral Rights of Authors and Should We Do Anything about It? An Australian Perspective' [2023] Law, Innovation and Technology

Merges RP, 'The Economic Impact of Intellectual Property Rights: An Overview and Guide' (1995) 19 Journal of Cultural Economics 103

Meys R, 'Data Mining under the Directive on Copyright and Related Rights in the Digital Single Market: Are European Database Protection Rules Still Threatening the Development of Artificial Intelligence?' (2020) 69 GRUR International 457

Netanel NW, 'Copyright and a Democratic Civil Society' (1996) 106 The Yale Law Journal 283

Quang J, 'Does Training AI Violate Copyright Law?' 36 Berkeley Technology Law Journal 1407

P Bernt Hugenholtz (ed), *Copyright Reconstructed: Rethinking Copyright's Economic Rights in a Time of Highly Dynamic Technological and Economic Change* (Kluwer Law International 2018)

Rajan MTS, *Moral Rights: Principles, Practice and New Technology* (1st edn, Oxford University Press 2011)

Raue B, 'Free Flow of Data? The Friction Between the Commission's European Data Economy Initiative and the Proposed Directive on Copyright in the Digital Single Market' (2018) 49 International Review of Intellectual Property and Competition Law 379

Rosati E, 'The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market- Technical Aspects. In-Depth-Analysis'

——, *Copyright and the Court of Justice of the European Union* (Oxford University Press 2019) <<https://doi.org/10.1093/oso/9780198837176.001.0001>> accessed 4 January 2023

——, 'Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity' (2019) 27 *Asia Pacific Law Review* 198

——, *Copyright in the Digital Single Market: Article-by-Article Commentary to the Provisions of Directive 2019/790* (Oxford University Press 2021)

Sag M, 'The New Legal Landscape for Text Mining and Machine Learning' (2019) 66 *Journal of the Copyright Society of the USA* 291

Schroff S, 'The Purpose of Copyright-Moving beyond the Theory' (2021) 16 *Journal of Intellectual Property Law & Practice* 1262

Lohsse S, Schulze R and Staudenmayer D (eds), *Trading data in the digital economy: Legal concepts and tools* (Nomos 2017)

Senftleben M, 'Flexibility Grave - Partial Reproduction Focus and Closed System Fetish in CJEU, Pelham' (2020) 51 *IIC-International Review of Intellectual Property and Competition Law* 751

Von Lewinski S and Walter M (eds), *European Copyright Law* (Oxford University Press 2010)

Wandtke A-A and Bullinger W (eds), *UrhR: Praxiskommentar Zum Urheberrecht* (CH Beck Verlag 2016)

Painer [2011] CJEU C-145/10 ECR 798

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 2019 (OJ L 130)

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) 2022 (OJ L 152/1)

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) 2022 (OJ L 265)