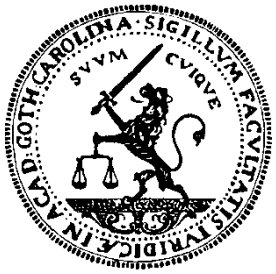

FACULTY OF LAW

LUND UNIVERSITY

Amanda Rydberg

Byte the Beat: Copying or Creating?

JURM02 Graduate thesis

Graduate thesis, Master of Laws program

30 higher education credits

Supervisor: Ana Nordberg

Semester: Spring 2024

Table of contents

Summary

Artificial intelligence (AI) and text and content-generating tools have been around for a while, bringing significant changes, challenges, and opportunities, especially regarding copyrighted materials such as music. As AI systems have become more sophisticated in generating new creative works, questions about these systems' potential copyright infringement have gained prominence.

This thesis delves into the legal landscape where creativity and technology intersect and explores the complexities surrounding copyrighted music's use in training AI models and the subsequent creation of new musical outputs. The central focus is whether it constitutes a copyright infringement under EU law when generative AI mimics an artist's sound and style.

To address this, the thesis will examine how EU law handles the use of copyrighted musical works in training AI models and how EU law addresses the copyright status of AI-generated musical outputs that imitate existing artists.

These questions are explored through a legal dogmatic method. A comparative perspective is also incorporated to a lesser extent to examine how different legal frameworks address these issues.

Musical works and their related right, such as phonograms, are generally protected by law, granting rightsholders certain exclusive rights, such as the right to reproduce the work. However, when AI generates music, the musical pieces are often reproduced multiple times in the training phase of the AI system. The thesis finds that these reproductions are not covered by the exemption for temporary acts of reproduction in Article 5(1) InfoSoc Directive. The exemption for text and data mining in Article 3 and Article 4 of the Directive (2019/790) on copyright and related rights in the Digital Single Market (DSM Directive) can potentially cover some reproductions of copyrighted musical works. Still, these exemptions are narrow in scope. Article 3 applies research organisations and cultural heritage institutions that conduct text and data mining activities on a non-commercial basis. Moreover, the general exemption in Article 4 faces practical hindrances, as rightsholders can reserve their right to prevent their works from being mined. Therefore, the mentioned provisions are not sufficient for AI music generators to train their generative models on copyright-protected music in most cases.

AI-generated output may constitute a derivative work. However, each case must be assessed individually, as some AI-generated musical outputs might include samples of existing works, whilst others may only be similar in style, which generally is not protected by copyright. The AI-generated voice is also typically not protected by copyright, as the voice is not considered fixed in a tangible medium or reflective of the author's own intellectual creation in most

circumstances. Thus, while the output may stay within the legal framework, the input often does not comply with EU copyright laws, rendering the overall process of generative AI mimicking an artist's sound and style unlawful.

Sammanfattning

Artificiell intelligens (AI) och dess genererande verktyg har funnits ett tag och medför med detta betydande förändringar, utmaningar och möjligheter, särskilt för upphovsrättsskyddat material såsom musik. I takt med att AI-system har blivit mer sofistikerat i att generera nya kreativa verk, har frågor om dessa system och deras potentiella upphovsrättsintrång fått ökad uppmärksamhet.

Den här uppsatsen undersöker därav det juridiska landskapet där kreativitet och teknik möts för att utforska hur användningen av upphovsrättsskyddad musik vid träning av AI-modeller samt det efterföljande alstret av nya musikaliska verk förhåller sig till upphovsrättslagstiftning. Fokus ligger på huruvida det utgör ett upphovsrättsintrång enligt EU-lagstiftningen när generativ AI imiterar en artists stil?

För att besvara denna fråga kommer uppsatsen att undersöka hur EU-lagstiftningen hanterar användningen av upphovsrättsskyddade musikaliska verk vid träning av AI-modeller. Samt hur EU-lagstiftningen hanterar AI-genererade musikaliska alster som imiterar befintliga artister.

Dessa frågor utforskas genom en rättsdogmatisk metod samt ett komparativt perspektiv i mindre utsträckning för att undersöka hur olika rättssystem hanterar dessa frågor.

Musikverk och deras närstående rättigheter, såsom fonogram, är i allmänhet skyddade av lag som ger rättighetsinnehavare vissa exklusiva rättigheter såsom rätten att mångfaldiga verket. När AI genererar musik, reproduceras musikstyckena oftast i flera steg under träningsfasen av AI-systemet. Uppsatsen konstaterar att dessa reproduktioner inte täcks av undantaget för tillfälliga exemplarframställningar i Artikel 5(1) InfoSoc-direktivet. Undantaget för text- och datautvinning i Artikel 3 och Artikel 4 i direktivet (2019/790) om upphovsrätt och närstående rättigheter på den digitala inre marknaden (DSM-direktivet) kan potentiellt omfatta vissa reproduktioner av upphovsrättsskyddade musikverk, men dessa undantag är snäva i sin omfattning. Artikel 3 omfattar forskningsorganisationer och kulturarvsinstitutioner som bedriver text- och datamining på icke-kommersiell basis. Det allmänna undantagen i artikel 4 möter förövrigt praktiska hinder då rättighetsinnehavare kan förbehålla sig rätten att deras verk utvinns. Dessa bestämmelser är därav inte tillräckliga i de flesta fall för att AI-musikgeneratorer ska kunna träna sina modeller på upphovsrättsskyddad musik.

AI-genererad alster kan även utgöra en bearbetning. Varje fall måste dock bedömas individuellt, eftersom vissa AI-genererade musikaliska alster kan innehålla sampling av befintliga verk, medan andra endast liknande tidigare verk i dess stil, vilket generellt sett inte är skyddat av upphovsrätt. Den AI-

genererade rösten är vanligtvis inte heller skyddad av upphovsrätt, eftersom rösten inte anses vara fixerad eller reflekterande av författarnas egna intellektuella skapelser. Således, även om det genererade alstret kan hålla sig inom det rättsliga ramverket så överensstämmer träningsprocessen oftast inte med EU:s upphovsrättslagar, vilket gör den övergripande processen där generativ AI imiterar en artists ljud och stil olaglig.

Preface

Writing this thesis has been an incredible journey with many ups and downs, marked by countless hours of research, reflection and writing. This would not have been possible without the support and encouragement of my friends and family. I would also like to give many thanks to my supervisor, Ana, for your valuable insight and feedback when writing this thesis.

With this, my studies and many years in Lund have finally come to an end. A special thanks goes out to all the amazing individuals I have met during my studies, who have enriched my experience and made my tough times more bearable and good times more enjoyable.

Amanda Rydberg

June, 2024

Abbreviations

AI	Artificial intelligence
AI Act	Artificial Intelligence Act Com 2021/206
AI-HLEG	High-Level Expert Group on Artificial Intelligence
Berne Convention	erne Convention for the Protection of Literary and Artistic Works (adopted 14 July 1967, entered into force 29 January 1970) 282 UNTS 221
CDPA	Copyright, Designs and Patents Act
CJEU	Court of Justice of the European Union
Database directive	Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases
DSM Directive	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
EU	European Union
EWC	European Writers Council
GAN	Generative Adversarial Network
GPAI	General Purpose Artificial Intelligence
GPT	Generative Pre-trained Transformer
InfoSoc Directive	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
IP	Intellectual property
LLM	Large Language Model
Rome Convention	International Convention for the Protection of Performers, Producers of Phonograms and Broadcasting Organisations (adopted 26 October 1961, entered into force 18 May 1961) 496 UNTS 43
TDM	Text and Data Mining
TRIPS Agreement	Agreement on Trade-Related Aspects of Intellectual Property Rights
U.S.C.	United States Code
VAE	Variational Autoencoder
WIPO	World Intellectual Property Organization

1 Introduction

1.1 Background

Artificial Intelligence (AI) plays a bigger part in today's society than ever before and is central to technological development, where processes such as generative AI and machine learning have made it possible to enjoy media in new ways.[1] These rapid advancements have ushered in a new era of digital creativity that affects the music industry, as AI has learnt to mirror the sound and style of human artists. One example of this is the AI-generated song "Heart on My Sleeve", which mimicked a collaboration between Drake and the Weeknd.[2] This phenomenon has led commentators to question the potential copyright infringement that AI may cause by either making copies in the AI system's initial training phase or by generating output that resembles existing works.[3]

The intersection between technology and creativity is becoming increasingly tense and the potential copyright infringement caused by generative AI systems has recently been brought to courts around the world, where rightsholders have challenged AI developers.[4] However, none of these cases are yet resolved, and none of them are within the EU or concern generative music.

The technological leap has furthermore sparked societal fears that AI could harm the artists and authors who make a living out of their craft, as well as damage the creative industry.[5] The creative sector generally requires innovation and human imagination to generate unique ideas drawn from a lifetime of experience.[6] The EU aims to foster innovation by using copyright as a tool

[1] Andrew R. Chow `AI's Influence on Music Is Raising Some Difficult Questions´ *TIME* (New York, 4 December 2023). < https://time.com/6340294/ai-transform-music-2023/> accessed 16 May 2024 and Written evidence from UK Music (CRF0037).

[2] Rachel Reed 'AI Created a Song Mimicking the Work of DRAKE and The Weeknd. What does that Mean for Copyright Law?' (Harvard Law School, 02 May 2023) <https://hls.harvard.edu/today/ai-created-a-song-mimicking-the-work-of-drake-and-the-weeknd-what-does-that-mean-for-copyright-law/> accessed 3 May 2024.

[3] Christopher T. Zirpoli, *Generative Artificial Intelligence and Copyright Law* (Congressional Research Service Legal sidebar LSB10922, 2023).

[4] See for instance The New York Times Company v Microsoft Corporation and OpenAI, Inc., Case No 1:23-cv-11195 (United States District Court Southern District of New York, filed 27 December 2023); Getty Images (US) Inc and others v Stability AI Ltd (2023) IL-2023-000007 (preliminary ruling, Mrs Justice Joanna Smith, 1 December 2023); Andersen v. Stability AI Ltd. (3:23-cv-00201).

[5] Cameron Shackell, `Will AI kill our creativity? It could – if we don't start to value and protect the traits that make us human´ (The Conversation, 27 September 2023) <https://theconversation.com/will-ai-kill-our-creativity-it-could-if-we-dont-start-to-value-and-protect-the-traits-that-make-us-human-214149#:~:text=We%20can%20expect%20an%20explosion,an%20increase%20in%20cultural%20tightness> accessed 4 May 2024.

[6] Nantheera Anantrasirichai & David Bull, `Artificial intelligence in the creative industries: a review´ (2022) 55 Artif Intell Rev 589.

to stimulate and reward creativity and provide returns on investments,[7] as creativity is considered to benefit society at large.[8] However, nowadays, AI has the capability to create multiple similar creations from one successful musical piece.[9] Therefore, AI can contribute to both innovative opportunities and challenges for artists, authors, and other creators in the way their artistic work is used, created, distributed, and consumed.[10]

This is similar to the fear sparked by the streaming revolution, which brought about significant changes and affected the music industry, and the way music and media are enjoyed. This later led to disputes such as the Hollywood writers' strike, where the shifting nature and economic uncertainty forced studios into mass layoffs that affected the writers. This caused the writer to go on strike, demanding, among others, that AI only be used as a research tool and not to replace creatives.[11] This fear is once again on the agenda, as AI is anticipated to get even more intelligent and take on roles of human creators, thereby making professionals redundant.[12]

These historic events and reactions to technology underline the ongoing tension between technological development and the rights of the creators. In response to this evolving landscape, the EU has been proactive in creating a common legislation to ensure safe AI systems that are transparent and respect fundamental rights and values. The implementation of the AI Act aims, among other things, to address the challenges of AI, including copyright.[13]

Despite the new act, many things are left in the open, and AI is anticipated to change the music industry further, potentially leading to a decline in music spaces, as well as other small and medium enterprises in the music industry.[14] Between EU copyright law and the new difficulties posed by generative AI, which leads to potential ethical issues of authenticity, creative implications, and economic implications, this thesis aims to explore what actually applies when creative generative AI mimics an artist's sound and style and to answer

[7] Tobias Kempas, *Artificiell Intelligens och immaterialrätt i Sverige och EU* (Norstedts juridik 2023) 18, 111; and Recital 10 InfoSoc Directive.

[8] Jim Jesse, `*The Music Copyright Manual: The Definitive Guide to Music Copyright Law in the Digital Age´* (Rock N Roll Law 2016) 8.

[9] Andersen (n 4).

[10] Recital 105 AI Act; discussion on TDM activities in chapter 4.1.2.

[11] Noma Bar, `Striking at the heart of Hollywood´ (*Los Angeles Times*, 10 April 2023) <https://www.latimes.com/entertainment-arts/business/story/2023-04-10/wga-writers-strike-story-collection> accessed 4 May 2024; and Shackell (n 5).

[12] Bob L. T. Sturm and others, `Artificial Intelligence and Music: Open Questions of Copyright Law and Engineering Praxis´ (2019) 8(3) Arts 1, 2.

[13] European Parliament, `EU AI Act: first regulation on artificial intelligence´ (European Parliament, 8 June 2023) < https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence> accessed 8 May 2024.

[14] Written evidence from UK Music (CRF0037).

the question of whether generative AI that mimics an artist's sound and style constitute copyright infringement.

1.2 Purpose and Research Question

The general purpose of this thesis is to explore what applies when creative generative AI mimics an artist's sound and style and to answer the question of whether generative AI that mimics an artist's sound and style constitute copyright infringement. This will be explored by looking at the copyright-protected music used in the initial training of an AI system as well as exploring if the AI-generated outputs infringe on the copyright-protected music that is used in the training and prompting or if these new AI-generated musical outputs instead are considered independent creations.

This purpose will be achieved by examining and analysing whether current copyright exceptions can be applied to the copyright-protected materials used in the training of the AI system, as well as to the AI-generated musical output.

To achieve the purpose of the thesis, the following main questions will be answered:

1. According to EU law, is it a copyright infringement when generative AI mimics an artist's sound and style?

    a. How does EU copyright law address the potential issue of copyright infringement when generative AI systems use copyright-protected music in the training of the AI model?

    b. How does EU copyright law address the potential issue of AI-generated musical outputs that mimic the sound and style of existing artists?

Thus, the thesis will focus on analysing both the input and the output. The input refers to the content used in training the AI system, and the output refers to the content generated by the AI system. The thesis will furthermore explore if there are any alternative solutions to address the potential issues that can occur in the intersection of creative generative AI and copyright law in the music sector.

1.3 Delimitations

A significant part of the current debate regarding copyright and generative AI has been focused on whether AI-generated output can be granted copyright protection. This is not something this thesis intends to explore any further, as it has been discussed extensively elsewhere. Instead, the focus of the thesis will be on analysing whether the exclusive rights of rightsholders are violated when AI mimics an artist's sound and style.

As the thesis explores AI-generated music from a copyright perspective, other legal fields, such as data protection laws and other IP regulations, will be excluded. The thesis will, however, draw analogies from previous trademark discussions surrounding counterfeit products to explore potential solutions to AI-generated outputs and copyright. Nonetheless, trademark law itself is not the object of analysis.

Despite the interesting nature of liability questions regarding AI, the narrow scope of the thesis unfortunately forces this element to be excluded.

In alignment with the thesis's purpose and research questions, the primary focus will be on EU law. However, due to music's international character and cross-border usage, copyright laws from other jurisdictions will be discussed briefly to present alternative solutions and interpretations to unregulated issues.

This thesis solely focuses on copyright issues related to AI-generated music. Other AI-related issues, such as explainability and ethical issues, will be outside the scope of the thesis and hence not discussed.

1.4 Method

This thesis will examine the research questions using a legal dogmatic method. The legal dogmatic method can be categorised as a study that attempts to provide a methodical exposition of the principles, concepts and rules that are guiding a particular area of law. By examining how these principles relate to one another, it is possible to solve ambiguities and gaps in the existing law.[15] The scientific method, thereby, has a subject matter that consists of legal phenomena governed by cultural values.[16] In our case, it is possible to methodologically look at the rules and principles that govern generative AI systems and their encounter with copyright-protected musical works.

The method is most often based on a specific legal issue; thereafter, the applicable law is interpreted, in which the legal dogmatic can be used to critique the current state of the law, suggest modifications, and propose future changes.[17] This thesis has therefore focused on the specific issue of AI-generated music that mimics an artist's sound and style and whether this constitutes copyright infringement. Thereafter, as the method suggests, the current

[15] Jan Smits, ´What is Legal Doctrine? On the Aims and Methods of Legal-Dogmatic Research´ In R. van Gestel, H. Micklitz, & E. L. Rubin (Eds.), Rethinking Legal Scholarship: A Transatlantic Dialogue (Cambridge University Press, 2017).

[16] Teruo Minemura, `Dogmatic Legal Science and Sociology of Law´ (1970) 56(3) ARSP: 351.

[17] Jan Kleineman `Rättsgodmatisk metod´ in Maria Nääv and Mauro Zamboni (eds) *Juridisk Metodlära* (MTM 2018) 23-25.

state of law within the EU, such as directives, treaties, case law, and reports, has been analysed.

The EU consists of multiple individual countries with their own legal systems, so when a legal field such as copyright and AI is studied, some differences in detail may exist between different national copyright legislations, depending on the country's historical development and legal systems. It is important to mention that EU copyright law is, in fact, a bundle of different national copyright laws operating separately from one another. Generally, there are difficulties in adapting a common directive, as the members first need to be convinced to agree to a directive by the commission. Thereafter, the national rules must be amended correspondingly, which may lead to different national rules due to the different national implementation techniques. Ensuring the same scope is, therefore, not an easy task as language and legal traditions vary, and the interpretation of differences can only be clarified by the CJEU.[18] Despite this, the national copyright laws within the EU are substantially similar in structure.[19]

Some parts of the legislation discussed throughout the thesis, such as Article 5(1) in the InfoSoc Directive, are also mandatory, making the interpretation of these regulations even more comparable within the jurisdiction and, hence, suitable for legal doctrine regarding these particular issues within the EU.

Compared to copyright law, AI is currently unregulated, which has contributed to a current proposal for an AI Act within the EU. The regulations regarding AI and creative generative AI systems and their relationship to copyright law have, therefore, been examined as one sole jurisdiction due to the lack of materials both within the EU as a union and for the individual nations.

As the thesis aims to discuss generative AI and copyright for music, and music has a very international character, some aspects of copyright law and music-related laws, cases, and doctrine from other foreign jurisdictions have been explored synoptically. This has mainly been done with respect to the USA, as this is where the main rightsholders of the world reside.[20] Other jurisdictions, such as the UK and Japan, have also been discussed to help draw conclusions and reasonings of how the current unregulated issues can be interpreted and resolved. As the few lawsuits related to AI can be found in the USA and UK, the thesis will, in some part, have a comparative perspective, as these lawsuits will be discussed to potentially see if they can be resolved similarly within the EU legal framework.

[18] Annette Kur, Thomas Dreier & Stefan Luginbuehl, *European intellectual property law: text, cases and materials* (2nd edn, MTM 2021).

[19] Kur (n 18).

[20] James Boyle, `(When) is Copyright Reform Possible?´ in Ruth Okediji (edt), *Copyright Law in an Age of Limitation and Exceptions.* (Cambridge University Press 2017).

Despite this, the thesis does not have a true comparative focus as these foreign policies are only used to contextualise and delimited to certain aspects. A comparative method normally requires more than the study of foreign law, and instead involves a comparison to better understand differences and similarities between legal systems.[21] A comparative method, though the functional model, can however be understood as every legal system facing similar issues; however, these are resolved through different means yet often end up with similar results.[22] Still, it is important to keep in mind that when we look at case law from other jurisdictions such as the USA, there are many differences, as the US, for instance, instead has federal and state rules that cover certain parts of their country. So, when we look at copyright law, there can be some differences in state laws; for instance, the similarity test used to establish a copyright infringement can vary across US courts.[23]

1.5 Material and Existing Research

Directives, doctrine, and EU legal precedents are the primary sources of information discussed in addressing the thesis question. The InfoSoc Directive, the DSM Directive, and the AI Act will be reviewed throughout the thesis. Some international treaties and reports from various EU bodies will also be examined more trivially. The selection of research material is based on authoritative sources of law, including laws, directives, treaties, and jurisprudence. The legal doctrine has been selected from acknowledged journals and legal professionals.

When addressing copyright and AI in jurisdictions outside the EU, legal cases, doctrine, and precedents from US and UK courts will be considered primarily.

Scientific information, articles, and legal doctrine have been used to provide a better understanding of AI technology and music copyright. Popular science and newspaper articles have also been used to merge the fields and gain a better insight into the general debate.

The current research regarding the training of AI systems on copyright-protected materials is quite extensive, considering the relatively new and recent issues that have arisen with AI. Much of this material is written from an American, British and EU copyright perspective. However, it is important to acknowledge that despite the many articles analysing the field of AI and copyright in the input stage, there is current worldwide uncertainty about how

[21] Kleineman (n 17) 143.

[22] K. Zweigert and H. Kötz, An Introduction to Comparative Law (2nd edn, 1998), at 34. Cited in Jaye Ellis, General Principles and Comparative Law, European Journal of International Law, Volume 22, Issue 4, November 2011, Pages 949–971.

[23] Zirpoli (n 3).

this will play out in the future, as there are no finalised lawsuits or clear laws at the moment.

Research on the interplay between protected works and AI-generated outputs, on the other hand, is very sparse. Therefore, related fields in the music sector, such as sampling and derivative works, have been analysed, mainly using legal doctrine and cases. Cases from other jurisdictions outside of the EU and other IP fields have also been examined to provide alternative insight into whether the output can infringe on existing rights.

Given the rapid progression of creative AI and the new concerns that have recently emerged regarding creative AI and music in both the input and output stage, it is a topical and highly debated subject at the moment.

1.6 Structure

The second chapter of the thesis aims to give a theoretical introduction to copyright legislation in the EU and its connection to the music industry and provide a brief introduction to copyright in other jurisdictions. This is intended to give a better understanding of the different elements that affect creative generative AI.

The third chapter introduces AI, different generative AI systems, and practical elements related to AI systems. Thereby, this chapter aims to provide an understanding of how copyright can be related to and affect generative AI and AI-generated musical outputs.

In the fourth chapter of the thesis, the copyright-protected materials used in the training of the generative AI model will be discussed and analysed from current exceptions and limitations. The chapter delves deeper into and critically analyses some existing exemptions in the EU, as well as other jurisdictions, that could potentially justify the digital reproductions and adaptations of copyright-protected musical works that are created during the training phase of AI systems.

The AI-generated musical outputs that are created from the training of the AI system will be analysed in the fifth chapter of the thesis. The different aspects of a musical piece will be examined through a copyright perspective, exploring how each of these aspects can potentially fall under the exclusive rights of the rightsholder. Previous similar cases in the EU, UK and USA will be referenced to provide the reader with a hypothetical answer to the unclear legal position for creative generative AI.

The sixth and seventh chapters of the thesis will provide a final discussion and analysis of whether generative AI that mimics an artist's sound and style infringe on existing copyright or neighbouring rights. These chapters will also

provide potential solutions to ensure a sustainable copyright system and evolving AI technology in the future.

2 Copyright and Music

2.1 EU Copyright

To begin with, there are two major copyright law traditions: the Anglo-American common law system and the continental European author's rights civil law system. While many aspects of copyright law have been internationally harmonized, most countries have some unique features.[24] Copyright within the EU is based on the principle of territoriality, meaning that the EU does not refer to union-wide copyright but instead, legislation that attempts to harmonize national laws, along with CJEU´s interpretation of primary and secondary legislation in the union.[25] Being a member country of the EU requires national legislation to be interpreted in accordance with EU law.[26]

Copyright is a limited term protection for authors that safeguards their literary and artistic works.[27] In the EU, the term of protection lasts for 70 years after the death of the authors.[28] This protection is often shared among multiple actors who hold a set of rights. These rights protect against both unlawful economic use, such as distribution, reproduction, and control over derivative works, as well as safeguard the moral rights of the author, such as the right for the authors to be linked to their creations.[29] Copyright, therefore, protects against different forms of copying and derivation or other unlawful use and imitation.[30] It should be noted that the economic rights can be transferred, while moral rights remain for the author. Copyright protection covers both the author's original and derivative works.[31]

To be granted copyright protection, a work must be an original creation by the author and reflect the author's own intellectual creation.[32] A work is

[24] Francisco Javier Cabrera Blázquez, Maja Cappello, Gilles Fontaine & Sophie Valais, *Exceptions and limitations to copyright*, IRIS Plus, European Audiovisual Observatory, Strasbourg, 2017.

[25] Kur (n 18).

[26] Kempas (n 7) 41.

[27] 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 (InfoSoc Directive) Article 2 (a).

[28] Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (Copyright Term Directive) Article 1.1- 1.2.

[29] Cabrera Blázquez (n 24); Berne Convention for the Protection of Literary and Artistic Works (adopted 14 July 1967, entered into force 29 January 1970) 282 UNTS 221 (Berne Convention) Article 6bis (1).

[30] Kempas (n 7) 38-40.

[31] WIPO, `Understanding Copyright and Related Rights´ (2nd edt, WIPO 2016) 20; European Commission, `Copyright´ (European IP Helpdesk) <https://intellectual-property-helpdesk.ec.europa.eu/ip-management-and-resources/copyright_en> accessed 27 February 2024.

[32] Case C- 5/08 Infopaq International A/S v Danske Dagblades Forening (2009) ECR I-6569, para 37.

considered the author's own if it reflects the author's personality. To be regarded as an intellectual creation, the author must have expressed their personality by making free and creative choices.[33] A work is not considered original if the choices have been dictated by technical deliberations, rules, and requirements that don't leave space for creative choices.[34] Furthermore, in Levola Hengelo, the court indicated that for there to be copyright protection, a work has to be fixed in a tangible medium. The CJEU stated in this case that a particular flavour of cheese could not be granted copyright protection as it, unlike literary or musical works, cannot be in a "precise and objective form of expression".[35] The CJEU furthermore expressed in Brompton Bicycle that a work must, in an original manner, encompass the author's own intellectual creation and that this is expressed in the creation.[36] Thus, to receive copyright protection, a work must be original and fixed in a tangible medium. Furthermore, copyright does not protect the idea itself, only the expression of those ideas.[37]

So, to summarize, a four-step test can be identified for assessing a copyrighted work: 1. It has to be in the literary, scientific, and artistic domain; 2. The work is the product of human intellectual efforts; 3. The work has been guided by creative choices; and 4. The creative choices shall be reflected in the output.[38]

Copyright legislation does not require any formalities but instead emerges when a work is created. Therefore, copyright legislation normally does not require registration; however, under some national laws, registration can be essential regarding the licensing of the work.[39] Legislation that is of considerable importance in the field of copyright in the EU includes the InfoSoc Directive and the DSM Directive[40]. Internationally, the Berne convention is also of great significance. The EU is not a contracting party of the Berne Convention, but many EU directives and obligations under other treaties, such as the TRIPS Agreement, reflect member states obligations to the Berne and Rome Convention.[41] All EU member states are contracting parties to the

[33] Case C-145/10 Eva-Maria Painer v Standard VerlagsGmbH and Others (2011) ECR I-12533, para 87-89.

[34] Case C-403/08 and C-429/08 Football Association Premier League Ltd and Others v QC Leisure and Others (C-403/08) and Karen Murphy v Media Protection Services Ltd (C-429/08) (2011) ECR I-9083, para 98.

[35] Case C-310/17, Levola Hengelo BV v. Smilde Foods BV. (2018). ECLI:EU:C:2018:899, para 39-42.

[36] Case C-833/18, Brompton Bicycle v Chedech/Get2Get. (2020) ECLI:EU:C:2020:461, para 22.

[37] WIPO (n 31) 6.

[38] Cabrera Blázquez (n 24).

[39] Kur (n 18).

[40] 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 (DSM Directive).

[41] European Commission `The EU copyright legislation´ (European Commission, 11 March 2024) <https://digital-strategy.ec.europa.eu/en/policies/copyright-legislation> accessed 17 May 2024, see Article 9.1 TRIPs that say member states shall comply with articles 1-21 of the Berne Convention.

Berne Convention,[42] and most member states are contracting parties to the Rome Convention.[43]

Works can furthermore be protected as a database, which is a collection of independent work, data, or other content that is arranged methodically and accessible through electronic or other means.[44] The term database can include artistic or musical collections of works or other materials.[45] These collections can, under certain circumstances, be granted copyright protection if the collection of data is original and expresses the author's own intellectual creation.[46] However, in general, the database is most often covered by the sui generis right for databases, which does not have the same requirements for originality.[47] This right can prevent the extraction and re-utilization of the whole content or parts of the content from the database.[48]

2.2 Neighbouring Rights

Neighbouring rights are a copyright-like protection that is related to the works of authorship.[49] In addition to authors' rights, performing artists, broadcasting organizations, and phonogram and film producers are generally protected by neighbouring rights.[50] These rights protect the interests of organizations or people who communicate the work to the public, as these are necessary to give life to some works.[51] The term performers include singers, actors, musicians, and others who act, sing, or otherwise perform literary and artistic works. A phonogram producer is simply someone who produces a fixation of a sound or performance.[52]

Unlike the Berne Convention, which protects authors, the Rome Convention accords protection to those who facilitate a work by communicating it to the public. The Convention states that performers are protected against unauthorized fixation of their performance, and phonogram producers are protected against reproduction of their phonograms.[53] The protection lasts for 50 years

[42] WIPO, `WIPO-Administered Treaties´ (WIPO) <https://www.wipo.int/wipolex/en/treaties/ShowResults?search_what=C&treaty_id=15> accessed 17 May 2024.

[43] All EU member states are contracting parties excluding Malta see WIPO, `WIPO-Administered Treaties´ (WIPO) <https://www.wipo.int/wipolex/en/treaties/ShowResults?search_what=C&treaty_id=17> accessed 17 May 2024.

[44] Article 1 (2) Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases (Database Directive).

[45] Recital 17 Database Directive.

[46] Article 3(1) Database Directive.

[47] Kempas (n 7) 205.

[48] Article 7(1) Database Directive.

[49] WIPO (n 31) 27.

[50] Kur (n 18).

[51] Recital 17 Copyright Term Directive; WIPO (n 31) 27.

[52] International Convention for the Protection of Performers, Producers of Phonograms and Broadcasting Organisations (adopted 26 October 1961, entered into force 18 May 1961) 496 UNTS 43 (Rome Convention) Article 3 (a)-(c).

[53] Kur (n 18).

after the publication of the fixation or performance.[54] The neighbouring rights are only partly harmonized within the EU, and member states have some possibilities to enforce more far-reaching protection.[55]

2.3 Exceptions and Limitations

To determine when an unauthorised use of a copyrighted work is lawful, the Berne Convention has instituted a "three-step test". The test is not just applicable to the Berne Convention but has been extended to cover exclusive rights under other international treaties as well.[56] According to Article 9(2) of the convention, it is a matter for the nations of the union to legislate for the reproduction of protected work in 1. special cases, 2. as long as it does not interfere with the normal use of the work, and 3. does not prejudice the lawful interest of the author.

Many of these exceptions and limitations in the EU can be found in the InfoSoc Directive Article 5, the Software Directive[57] Articles 5-6, the Database Directive Articles 6 and 9, and the DSM Directive, to name a few. The main objectives of the limitations and exceptions in the EU are linked to attaining public policy goals, and it is possible to divide exemptions and limitations under EU law into three categories, which are exceptions and limitations in support of fundamental freedoms, such as freedom of the press, and freedom of expression; exceptions and limitations in favour of public interest, for example, to benefit access to information and knowledge; and lastly, exceptions and limitations connected to the benefit of private use.[58] In the InfoSoc Directive, we can find the possibility for nations to create exceptions and limitations for private copying,[59] the parody exemption,[60] and the exemption for quotation for criticism or review.[61] The only mandatory exception in the InfoSoc Directive is the exception of temporary acts of reproduction, which grants permission for temporary acts of reproduction in certain circumstances.[62] However, the CJEU has repeatedly stated that exceptions and limitations must be strictly interpreted, as the restrictions described in the directives are a deviation from the general rule of exclusive rights.[63] The DSM

[54] Article 3 Copyright Term Directive.

[55] Case C-279/13 C More Entertainment AB v Linus Sandberg (2015) ECLI:EU:C:2015:199, para 37.

[56] Cabrera Blázquez (n 24).

[57] Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Software Directive).

[58] Cabrera Blázquez (n 24).

[59] Article 5.2 (b) InfoSoc Directive, which is limited to the reproduction of the content and for use in the private sphere; see Cabrera Blázquez (n 24).

[60] Article 5 (3) (k).

[61] Article 5.3 (d) InfoSoc Directive.

[62] Article 5.1 InfoSoc Directive.

[63] FAPL (n 34) para 162.

Directive contains two exemptions that regard text and data mining for scientific research found in Articles 3 and 4.

2.3.1 The Fair Use Doctrine, Fair Dealing, and Japan's TDM Exception

The US copyright law states that protection shall be granted to "original works of authorship fixed in any tangible medium of expression".[64] This gives the author certain exclusive rights, such as the right to reproduce and distribute their work.[65] However, there is an exception to this in form of the fair use doctrine, which can be described as a general clause that outlines exceptions to copyright.[66] It's found in 17 U.S.C. § 107. The Doctrine balances the copyright holders' rights against the purpose and character of the use, the nature of the copyrighted work, the amount of copyrighted work used, and the impact this has on the original work.[67]

In the UK, we can find exemptions to copyright law in the form of fair dealing in CDPA[68] Chapter III 29-30. Similar to copyright law in the EU, the UK system consists of exemptions such as 29 (1) research for non-commercial purposes, 29 (1C) for private study, 30 (1) for criticism or review, and 30 (2) reporting of current events.[69] The UK has a legal term, fair dealing, that is used to determine if the use of copyright-protected works is lawful. There is no statutory definition of fair dealing, but it can be described as a standard in that is based on how an honest and fair-minded person would have handled the work.[70] This forms part of the assessment of the potential infringement. For there to be fair dealing, the defendant must prove that (1) the dealing fits into one of the mentioned exemptions, (2) The dealing is fair, and (3) there is adequate recognition.[71]

Fair dealing in the UK is more limited than the US fair use doctrine. Generally, if the use of a copyright-protected material does not affect the sale of the work, and if the amount of copied work is considered appropriate and reasonable, then this use can be deemed fair. In these circumstances, the copyright-

[64] 17 U.S.C.§ 102(a).
[65] 17 U.S.C.§ 106.
[66] Cabrera Blázquez (n 24).
[67] Zirpoli (n 3).
[68] Copyright, Designs and Patents Act 1988.
[69] Gluseppina D'Agostino, 'Healing Fair Dealing - A Comparative Copyright Analysis of Canada's Fair Dealing to U.K. Fair Dealing and U.S. Fair Use' (2008) 53 McGill L J 309.
[70] UK Intellectual Property Office, *Exceptions to Copyright: Guidance for Consumers* (2018) <https://assets.publishing.service.gov.uk/media/5a80f292ed915d74e6231597/Exceptions_to_copyright_-_Guidance_for_consumers.pdf> accessed 13 May 2024.
[71] D'Agostino (n 69).

protected material can be used without the necessary permission from the rightsholder.[72]

Similar to the EU, Japan has a set of copyright exemptions that allow for the use and reproduction of copyrighted works.[73] These exemptions include, among others, private use, quoting, and library archiving.[74] Japan has, however, introduced a more flexible provision to allow for digital innovation in the form of Article 30.4 of the Japanese Copyright Act, which is especially interesting for TDM purposes. It is considered the broadest TDM exception that exists at the moment.[75]

2.4 Music Copyright

A musical work can be defined as musical compositions with or without words and can include elements such as the songs, the scores, the melody, operas and musicals.[76] Both author's rights and neighbouring rights may exist within a musical work. Copyright protects the original works, such as the music, the sheet music, and lyrics, while the neighbouring rights can protect the performance, interpretation, or the fixation of the sound recording. Therefore, involved parties like composers, songwriters, musicians, performers, and phonogram producers can each benefit from certain protections.[77] In the US, on the other hand, both the phonogram as well as the musical composition can get copyright protection.[78]

As the music industry often takes place in a transnational arena, some legislations have been established, such as Directive 2014/26[79], to create a pan-European contractual system to grant community-wide licenses to online music platforms.[80] Furthermore, the EU copyright framework provides a collective licensing mechanism.[81] In some instances, the Berne Convention has

[72] University of Edinburgh, 'Copyright Exceptions and Fair Dealing' (University of Edinburgh) <https://www.ed.ac.uk/information-services/library-museum-gallery/library-help/copyright/copyright-exceptions-and-fair-dealing> accessed 20 April 2024.

[73] Japan Patent Office, 'Copyright Law of Japan' (2008) 14 <https://www.jpo.go.jp/e/news/kokusai/developing/training/textbook/document/index/Copyright_Law.pdf> accessed 21 May 2024.

[74] Chie Kasahara, 'Getting the Deal Through: Copyright 2012 - Japan' (2012) 80 <https://www.aplawjapan.com/archives/pdf/file/GettingtheDealThroughCopyright2012Japan.pdf> accessed 21 May 2024.

[75] Artha Dermawan, 'Text and Data Mining Exceptions in the Development of Generative AI Models: What the EU Member States Could Learn from the Japanese "No Enjoyment" Purposes?' (2024) 27 The Journal of World Intellectual Property 44–68.

[76] Copyright.eu, 'Protection of Music and Songs' (Copyright.eu, 22 August 2022) <https://www.copyright.eu/docs/protection-of-music-and-songs/> accessed 8 May 2024

[77] Sturm (n 12) 4.

[78] Jesse (8) 17.

[79] Directive 2014/26/EU of the European Parliament and of the Council of 26 February 2014 on collective management of copyright and related rights and multi-territorial licensing of rights in musical works for online use in the internal market.

[80] Kur (n 18).

[81] Cabrera Blázquez (n 24).

allowed for two non-voluntary licenses that make mechanical reproduction of musical works and broadcasting possible.[82] Orphan works can also create issues. Orphan works are identified as works in which the rightsholder cannot be identified or located. Therefore, in certain circumstances, protected works can be transformed into orphan works if the data on the author or any other rightsholder goes missing or becomes outdated.[83] Article 4 of Directive 2012/28/EU[84] states that a work classed as an orphan work in one member state shall be considered an orphan work in all other member states as well. These works can, therefore, be used without further search or permission.[85]

[82] WIPO (n 31) 17.

[83] Kur (n 18).

[84] Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works (Orphan Works Directive).

[85] Kur (n 18).

3 Artificial Intelligence

Artificial Intelligence has no straightforward definition, but can be understood as machines approximating some human behaviours.[86] According to the AI Act, AI systems are defined as software that can interfere by producing outputs, such as recommendations, decisions or predictions, that affect the environment they interact with and are created with one or more techniques such as machine learning, statistical approaches, and logic- and knowledge-based approaches.[87] WIPO defines AI as a computer science discipline aimed at developing systems to perform tasks that normally require human intelligence, with little to no human intervention.[88] Just like the human brain, AI systems achieve rationality by perceiving their environment, thereby collecting and interpreting the data, processing this perceived information, deciding the best action and then acting accordingly.[89]

AI systems can be either software-based, including voice assistance or image analysis software or they can be embedded in hardware devices, under which advanced robots or the internet can be found.[90] AI systems can furthermore be categorised based on their capabilities, including reasoning, learning, and robotics.[91] Going forth, the thesis will focus on the learning category, as it is the most relevant for generative AI systems.

Within the learning category, we can find machine learning. In essence, machine learning is a self-learning process that allows machines to learn from a large number of examples or phenomena and then build mental models to create output based on the new input. These algorithms then enable the system to learn from future data intake and experience.[92] One key benefit of AI systems is their adaptivity, understanding of correlations, and capacity to improve their performance by learning from experience.[93]

[86] Ryan Calo, 'Artificial Intelligence Policy: A Primer and Roadmap' (2017) 51 UC Davis Law Review 399, 404.

[87] Laying down harmonised rules on artificial intelligence Proposal 2021/206 COM (AI Act) Recital 12.

[88] WIPO, 'Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence' (WIPO/IP/AI/2/GE/20/1 Rev, 21 May 2020) prepared by the WIPO Secretariat 3.

[89] AI-HLEG, 'A Definition of AI: Main Capabilities and Disciplines' (European Commission, 8 April 2019) 1, 4 < https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines> accessed 4 March 2024.

[90] AI-HLEG (n 89) 1.

[91] AI-HLEG (n 89) 3.

[92] Theodoros Chiou, 'Copyright Lessons on Machine Learning: What Impact on Algorithmic Art?' (2019) 10(3) JIPITEC 398, 399.

[93] Kempas (n 7) 28.

3.1 Generative Artificial Intelligence

Generative modelling is a subgroup of machine learning aimed at generating new data indistinguishable from the original input.[94] WIPO defines the term "AI-generated" as the "generation of an output by AI without human intervention".[95] By examining training examples and learning their patterns and distribution, generative AI can create artificial outputs such as videos, texts, audio, and images.[96] A high-level generative AI system encodes a streamlined representation of its training data to generate new work that is comparable but not identical to the original work.[97] This creative capability is not dependent on programming with a specific result in mind but rather on algorithms teaching the AI to master a specific art style. The system then generates new works following the learnt style.[98] The portrait "The Next Rembrandt" is a famous example of an AI-generated artistic work.[99]

Common network techniques for generative learning are Generative Adversarial Network (GAN), Generative Pre-trained Transformer (GPT), diffusion models, and Variational autoencoders (VAEs). In a nutshell, GAN uses two different networks. Firstly, the generator, which generates new data instances, and secondly, the discriminator, which tries to classify this data and determine if the data belongs to the training set or not. The generator tries to "lie" by producing data indistinguishable from the original content, and the discriminator tries to identify fake data.[100] This process repeats and improves over time.[101] The VAE similarly use two networks: the encoder that converts input into a denser representation of the data and the decoder that tries to reconstruct this to the original input data.[102]

3.2 Data Mining and Data Scraping

The generative process requires a large amount of training data, and Text and Data Mining (TDM) may be used to train AI for the purpose of AI-driven creativity.[103] Data mining involves the extraction of knowledge from large amounts of data and text to identify patterns from these pieces of

---

[94] Prafulla Dhariwal and others, 'Jukebox: A Generative Model for Music' (2020) arXiv 1 <https://arxiv.org/pdf/2005.00341.pdf> accessed 2 February 2024.

[95] WIPO 2020 (n 88) 4.

[96] Axham (n 119).

[97] Kim Martineau, `What is generative AI´ (IMB Blog 20 April 2023) < https://research.ibm.com/blog/what-is-generative-AI> accessed 13 March 2024.

[98] Kempas (n 7) 80-81.

[99] ibid 81.

[100] Chris V. Nicholson, ´A Beginner´s Guide to Generative AI´ (Pathmind wiki). <https://wiki.pathmind.com/generative-adversarial-network-gan> accessed 13 March 2024

[101] NVIDIA, ` What is Generative AI´ (NVIDIA) <https://www.nvidia.com/en-us/glossary/generative-ai/> accessed 13 March 2024.

[102] NVIDIA (n 102).

[103] 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(2) Asia Pac. Law Rev. 198, 198.

information.[104] The DSM Directive defines TDM as automated analytical methods used to analyse digital data and text in order to produce information such as correlations, patterns, and trends.[105] For example millions of The New York Times newspaper articles were used to train ChatGPT.[106] In the UK lawsuit between Stability AI and Getty, more than 12 million photographs were used in the training of the AI.[107] Data scraping, on the other hand, is the technique in which programs extract data from human-readable outputs.[108]

3.3 Legal Framework

In 2021, the European Commission proposed an act to set a common legal framework for AI within the EU. The aim is to support the development of trustworthy AI that ensures the fundamental rights of people and businesses while supporting innovation in conformity with union values.[109] Although the act has not yet been published, a final draft has been laid down.[110]

Copyright is mentioned in some parts of the Act. The recital states that providers of GPAI models shall "put in place a policy to comply with Union copyright law" and make a "sufficiently detailed summary" of the content used to train the model available while still safeguarding trade secrets and business information.[111]

The only article in the act that discusses copyright is Article 53(1) (c), which states that providers of general-purpose AI models shall

> put in place a policy to comply with Union copyright law, and in particular to identify and comply with, including through state of the art technologies, a reservation of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790.

Furthermore, the commission has set up AI-HLEG, which has developed, among other things, ethics and policy recommendations for the use of AI.[112]

[104] Rosati 2019 (n 103) 199.
[105] Article 2 (2) DSM Directive.
[106] The New York Times v. OpenAI (n 4).
[107] Getty Images v Stability AI (n 4) para 48.
[108] Cambridge University Press, 'Data Scraping' (Cambridge Dictionary) <https://dictionary.cambridge.org/dictionary/english/data-scraping> accessed 13 March 2024.
[109] Recital 1 AI Act.
[110] This thesis will be based on the final draft version of the 16th of April 2024.
[111] Article 53.1 (a) and Recital 105-107 AI Act.
[112] AI-HLEG (n 89).

3.4 Artificial Intelligence and Music

AI technologies in the music sector have been available for more than two decades but have increased rapidly over the last few years.[113] Recent developments in AI-generated music suggest that neural networks may significantly impact music creation and consumption, as AI nowadays can learn patterns from data to synthesise similar data and create output in novel media, text, and audio.[114] AI generators can, for instance, produce musical outputs that capture timbres, melody, rhythm, and long-range composition for various instruments, as well as vocal styles and tones.[115] AI systems can learn to recognise physical and digital works of the system and generate an output that is not identical to the input data.[116] Furthermore, AI can also modify already existing songs and instruct the performer on what to sing.[117] Despite AI's ease in creating new pieces, music remains one of the most challenging types of audio to generate as it contains varied patterns with various instruments.[118]

Most music generator systems use deep-learning neural networks that rely on large volumes of existing musical works. To create new music, the algorithms, among other things, search for patterns in chords and relations between notes. However, different AI systems vary in their output methods, as some rely heavily on the input data, while others use principles derived from music theory.[119] No system can compose music fully autonomously yet. It is, therefore, common for the end user to rework the generated output.[120]

The concept of machines writing music is not new, as this has existed since the 1950s when Lejaren Hille used a computer to produce parts of "String Quartet No.4".[121] AI is now used by artists for music production, such as Sony´s Flow Machine and Google Magenta, which helps artists in their

[113] Nantheera Anantrasirichai and David Bull, 'Artificial Intelligence in the Creative Industries: A Review' (2021) 55 Artif. Intell. Rev. 589.

[114] Eric Sunray, 'Sounds of Science: Copyright Infringement in AI Music Generator Outputs' (2021) 29 Cath U J L & Tech 185; Patent and Trademark Office, 'Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation' (Docket No. PTO–C–2019–0038).

[115] Dhariwal (n 262) 2.

[116] Kempas (n 7) 80-81.

[117] Dhariwal (n 262) 2.

[118] Meta,` Introducing AudioCraft: A Generative AI Tool For Audio and Music´(Meta, 2 August 2023) <https://about.fb.com/news/2023/08/audiocraft-generative-ai-for-music-and-audio/> accessed 9 April 2024.

[119] Johan Axhamn, 'Artificial Intelligence and the Need for Legislation' in Morten Rosenmeier, Thomas Riis, Jens Schovsbo, & Henrik Udsen (eds.), *Festskrift till Jørgen Blomqvist* (Lund University 2021) 44.

[120] Axhamn (n 119) 64.

[121] Alex Marshall, `From Jingles to Pop Hits, A.I. Is Music to Some Ears´ (The New York Times, 23 January 2017) <https://www.nytimes.com/2017/01/22/arts/music/jukedeck-artificial-intelligence-songwriting.html> accessed 17 May 2024.

creative processes, including editing.[122] AI can contribute with competitive advantages for companies and socially beneficial outcomes in media and culture by enabling semi-automated music creations, as well as providing recommendations from previously archived recordings.[123] Consequently, several companies have emerged that seek to commercialise music AI, some of them targeted to consumers, and others catering to cultural producers by producing algorithmic made without royalties.[124]

An example of AI creating music is the system AIVA, which uses a network trained on roughly 30,000 classical pieces in different genres, allowing it to compose music in the form of notes for performing musicians. People can also provide AVIA with inspirational materials to generate similar output.[125] The AI music generator Jukebox uses conditioning information such as artist, genre, lyrics, and timing to start the process. These codes are then sampled, decoded and converted into audio.[126]

[122] Emmanuel Deruty and others, 'On the Development and Practice of AI Technology for Contemporary Popular Music Production' (2022) 5(1) Transactions of the International Society for Music Information Retrieval 35, 35-36.

[123] Recital 4 AI Act; Sturm (n 12) 1-2.

[124] Eric Drott, ´Copyright, compensation, and commons in the music AI industry´ (2020) 14(2) Creat. Ind. J. 190, 190.

[125] Kempas (n 7) 80 see also AVIA, `AI Generated songs in Seconds´ (AVIA) <https://www.aiva.ai/> accessed 8 March 2024.

[126] Dhariwal (n 94) 4.

4 Input and Paving the Path for Processing

In 2016 Sony CSL research lab released the AI-generated song Daddy´s Car, a song in the style of The Beatles. Composer Benoit Carre arranged and wrote the lyrics for the song, while the music was created through a system called Flow Machines, trained on music by The Beatles.[127] To raise a potential copyright issue; "How could it be possible for AI to create a song in the style of The Beatles if it did not also have access to The Beatles repertoire?".[128]

Music generator requires large amounts of pre-existing data, such as music, lyrics, audio recordings, and scores in their initial training.[129] For example, the AI music generator Jukebox was trained on a dataset consisting of 1,2 million songs, paired with matching lyrics from LyricWiki and metadata including artists, genre, year, and keywords.[130] As noted in Chapter 3.2, copyright-protected content from both The New York Times and Getty Images has been used to train various AI systems.[131] OpenAI has earlier acknowledged that public datasets are used in the training of its programs, some of which included copyright-protected works, and that the process of training the AI system involves making copies of the analysed data.[132] This data is often reproduced in several steps.[133]

Such reproduction can potentially be considered a copyright infringement, as a direct or indirect, temporary or permanent reproduction is an exclusive under Article 2 of the InfoSoc Directive. The use can also be protected by the sui generis right for databases. Training the generative model can also include making an adaptation of the work if the data is processed and altered.[134] As seen with the VAEs system, where the encoder network converts the input to a denser version before the decoder generates a new work.[135]

However, for a copyright infringement to occur, the work used in the model's training must be copyright-protected. A song is likely copyright-protected as it is usually original and the author's own intellectual creation fixed in a tangible medium. The master recording also receives protection when the sound

[127] Glen Tickle, `Daddy's Car: A Song Composed by Artificial Intelligence Created to Sound Like The Beatles' (Laughing Squid, 22 September 2016) < https://www.flow-machines.com/history/press/daddys-car-song-composed-artificial-intelligence-created-sound-like-beatles/> accessed 9 April 2024.

[128] Rosati 2019 (n 103) 200.

[129] Sturm (n 12) 5.

[130] Dhariwal (n 94) 5.

[131] The New York Times v. OpenAI (n 4) and Getty Images v Stability AI (n 4).

[132] Zirpoli (n 3).

[133] Kempas (n 7) 171-172.

[134] Philipp Hacker, 'A Legal Framework for AI Training Data—from First Principles to the Artificial Intelligence Act' (2021) 13(2) Law, Innovation and Technology 257, 277.

[135] NVIDIA (n 101).

recording is fixed.[136] Therefore, the music, the lyrics, and the phonogram are normally protected.[137] This risk of infringement is thereby especially apparent when AI systems are used to generate new music.[138]

It is important to recognise that making protected content available internally without permission constitutes copyright infringement, regardless of whether this work is accessed.[139] The content used in the training could also be covered by a Creative Commons licence[140], be an orphan work, or be in the public domain.

When a generative AI mimics an artist's sound and musical style, potential copyright infringements may occur. The key question is whether the input used to train the AI model constitutes copyright infringement or is covered by a copyright exemption.

4.1 Exceptions and Limitations

There is a possibility that copyright-protected work and databases may be utilised and reproduced under certain circumstances without constituting copyright infringement. The following section will detail if protected work or databases can be used as input data during the training phase of an AI system and present the most applicable exemptions relevant to the subject of the thesis.

4.1.1 Temporary Acts of Reproduction and Generative AI Systems

Digital use of works typically results in many instant copies, including intermediate storage on various servers and routers during data transfer over the internet. These instant reproductions are considered infringements under copyright law. But they are necessary for technology to function.[141] To address this issue, a mandatory exception was introduced in Article 5 (1) InfoSoc Directive. It explains that temporary acts of reproduction, mentioned in Article 2 InfoSoc Directive, that are “transient or incidental” and an essential component of a technological process are exempt if they enable (a) “a

[136] Jesse (n 8) 23, 194.

[137] Case C-476/17, Pelham and Others v Hütter and Others (2019) ECLI:EU:C:2019:624, para 29. The CJEU has clarified that the reproduction of a sound sample, regardless of how short, must be interpreted as a reproduction "in part" of the phonogram, and as a result, the reproduction falls within the exclusive rights granted to the phonogram's producer under that provision.

[138] Kempas (n 7) 172.

[139] Kempas (n 7) 158.

[140] Creative Commons licenses offer the possibility to use certain creative works without further authorisation. There are different creative common licenses, see Creative Commons, `About CC Licenses´ (CreativeCommons, 2019) < https://creativecommons.org/share-your-work/cclicenses/> accessed 18 May 2024.

[141] Kempas (n 7) 184.

transmission in a network between third parties by an intermediary", or (b) "a lawful use" of a work that lacks an independent economic value.[142]

The InfoSoc Directive further explained that if these conditions are met, the provision covers actions that permit browsing or caching, including those that facilitate the efficient function of transmission systems, as long as the intermediary doesn't alter the information or interfere with the legal use of the technology.[143]

This leads us to the question of whether the temporary acts of reproduction exemption in Article 5 (1) InfoSoc Directive can justify reproductions of copyright-protected materials in the training of generative AI models? The CJEU has ruled that specific temporary copies that have emerged during data-gathering processes might be supported with this exception if they have a secondary significance, are transient, and do not serve any independent purpose to the technical process.[144] Temporary acts of reproduction must be part of the technological process, and the reproduction shall be necessary for the process to function correctly, even though human intervention is required to start and stop the process.[145]

That the reproduction shall be a necessary part of a technological process suggests that certain reproduction made during the training of AI systems could be permitted.[146] This is because reproductions during the training are essential to produce the later musical creations. Therefore, training can be seen as a means to an end, and the more data the model is trained on, the better it will perform.[147]

The requirement for copies to be transient and aimed at enabling a technical process means a reproduction may only be deemed ephemeral if it doesn't exceed the time required for the procedure to function properly. Additionally, the procedure needs to be automated in the sense that reproductions are removed automatically and without human interaction when no longer needed.[148]

One relevant case is Infopaq, in which the company Infopaq compiled, extracted, indexed, and printed articles and keywords from different Danish publishers. In Infopaqs data capture process, they started by manually

[142] Article 5 (1) InfoSoc Directive.

[143] Recital 33 InfoSoc Directive.

[144] Case C-360/13, Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd [2014] ECLI:EU:C:2014:1195, para 43.

[145] Case C-302/10, Infopaq International A/S v Danske Dagblades Forening (Infopaq II) (2012) ECLI:EU:C:2012:16, para 29-39.

[146] Kempas (n 7) 185.

[147] See Christophe Geiger, "Elaborating a Human Rights friendly Copyright Framework for Generative AI" (2024). Joint PIJIP/TLS Research Paper Series. 123, 8, 23.

[148] Public Relations Consultants Association (n 144) para 34-40; Infopaq I (n 32) para 62-63.

registering newspaper publications in a database. Then, they scanned a selection of publications, creating TIFF files. These TIFF files were digitally processed and saved as text files before being deleted. The next step involved processing the text files to find search words and capture snippets. Lastly, a cover sheet was printed out, containing all the matching pages and the 11-word snippets.[149]

These data capture processes are similar to steps carried out in machine learning processes.[150] The initial registering of the publications is akin to "the corpus compilation phase where corpora are selected on the basis of their relevance and obtained from a variety of sources".[151] The conversion of data into other formats mirrors the data processing step in machine learning. Transferring files to text files is equivalent to converting data into a format readable by machine learning algorithms. Identifying search words is similar to training machine learning algorithms. Finally, printing the records corresponds to creating a model that contains the extracted features.[152]

In the case, the court found that the printout of an extract of 11 words during a data capture process was not considered transient, thus constituting an "article 2 reproduction in part or perhaps even an adaptation".[153] This suggests that many reproductions made during machine learning may not be classified as transient under Article 5(1) InfoSoc Directive.[154] This can be especially apparent when AI-generated outputs, such as music, are based on the input materials.[155] Moreover, the lack of transparency from AI companies regarding their use of copyrighted materials complicates the issue further.[156]

Even if the machine learning process could be considered transient, another hurdle is the criterion of lawfulness. The term lawful can include both the author's consent and situations not restricted by law.[157] Meaning that both authorisation and an exception or limitation fall within the scope of lawfulness.

[149] Infopaq I (n 32) 13-2.

[150] Thomas Margoni, 'Artificial Intelligence, Machine Learning and EU Copyright Law: Who Owns AI?' (2018) CREATe Working Paper 16 <https://zenodo.org/records/2001763#.YakLB1VBzIU> accessed 9 April 2024.

[151] Margoni (n 150) 17-18.

[152] ibid 17-18.

[153] Infopaq I (n 32) para 70; Margoni (n 150) 18.

[154] See European Commission, Directorate-General for the Internal Market and Services, Jean-Paul Triaille, Jerome de Meeûs d'Argenteuil & Amelie de Francquen, *Study on the Legal Framework of Text and Data Mining (TDM)* (Publications Office 2014) 46.

[155] See Public Relations Consultants Association (n 144) para 45-52.

[156] Rishi Bommasani, Kevin Klyman & Daniel Zhang and Percy Liang, 'Understanding the EU AI Act: A Comprehensive Overview' (Center for Research on Foundation Models, 15 June 2023) <https://crfm.stanford.edu/2023/06/15/eu-ai-act.html> accessed 11 April 2024.

[157] Recital 33 InfoSoc Directive.

The CJEU determined that temporary acts of reproduction allowing a satellite decoder and television screen to operate properly were legitimate.[158] However, temporary acts of reproduction made on a multimedia player obtained from unauthorised streaming websites were not.[159]

It is, therefore, questionable whether reproductions carried out during an AI systems training phase qualify as lawful, particularly concerning music AI generators, as these systems aim to generate music by technical means and are modelled after the copied material.[160]

The criterion that temporary reproduction must not have an independent economic value is another obstacle. If the reproduction generates profit beyond that from the lawful use of the work, the requirement is not fulfilled.[161] Even if a small profit might be acceptable, this profit has to be proportionally compared to the efficiency gains that the technological processes allow.[162] This would require an economic evaluation of the collected data, but it can be argued that all actions within the data collection process could have financial value, as the information collected would likely have monetary worth if it were transferred to someone else.[163] This is particularly true for musical compositions, whose value is distinct and significant. For instance, a study found that music in the generative AI market had an estimated value of 300-million-dollar in 2023.[164]

Lastly, the provision does not apply to situations where the reproduction leads to a modified version of the work.[165] For instance, the VAE network or GAN network modifies input data to generate new outputs, which is not permitted under the exemption.

In conclusion, temporary copies made during the machine learning process could be incidental or transient if they are deleted automatically after the process.[166] However, the exemption found in Article 5 (1) InfoSoc faces many obstacles in allowing reproductions during the training phase of generative models and the CJEU has made clear that the conditions set out in the directive should be interpreted strictly.[167] This makes it unlikely that temporary

[158] FAPL (n 34) para 170–172.
[159] Case C-527/15, Stichting Brein v Jack Frederik Wullems (Filmspeler) (2017) ECLI:EU:C:2017:300, para 69-72.
[160] Axhamn (n 119) 55.
[161] Infopaq II (n 145) para 50-54.
[162] Infopaq II (n 145) para 51.
[163] Triaille (n 154) 47; Margoni (n 150) 17- 19.
[164] Goldmedia, *AI and Music: Impacts of Artificial Intelligence on the Music Industry* (GEMA, SACEM 2023) 7.
[165] Infopaq II (n 145) para 48-53.
[166] Margoni (n 150) 17.
[167] FAPL (n 34) para 162.

acts of reproduction during machine learning are permissible under Article 5(1) InfoSoc.

4.1.2 Text and Data Mining

Although Article 5 (1) of the InfoSoc Directive likely does not justify the use of copyrighted works or related rights during the generative AI model's training phase, the exemptions regarding TDM activities found in the DSM directive may offer a solution.

The DSM Directive was adopted to ensure a functioning internal market for copyright law in the digital environment, where creativity and innovation are stimulated while maintaining high levels of protection for rightsholders.[168] The Directive includes two relatively new and mandatory exemptions in Article 3 and Article 4, that permits certain TDM activities.

TDM activities involve securing and acquiring data for further processing and analysis to generate knowledge and insight.[169] TDM processes generally start by accessing content, then extracting or copying it, and finally mining the data to enable the discovery of patterns and relationships, leading to knowledge discovery.[170] The DSM Directive describes that TDM should be understood as technologies that "enable the automated computational analysis of information in digital form", such as images or sound.[171]

To determine if the training of a generative AI model falls under TDM, we must explore its compatibility with TDM activities. Some argue that there is legal uncertainty surrounding the regulation of TDM for AI within both EU and national copyright laws.[172] While TDM and machine learning use the same type of key algorithms to discover patterns, their utilities differ.[173] However, the recital to the AI Act explained that training of large generative AI models requires access to vast amounts of data, and TDM techniques may, therefore, be used for the retrieval and analysis. This data, often protected by copyright and related rights, normally requires authorisation from the rightsholders unless copyright exemptions and limitations apply. The DSM Directive introduces exemptions that allow reproduction and extraction under certain conditions.[174] Thereby, by directly referencing the TDM exception in the DSM Directive regarding AI training, the AI Act clarifies the discussion, linking TDM activities to AI training and recognising that TDM exemptions

[168] Recital 2 DSM Directive.
[169] Kempas (n 7) 176.
[170] Rosati 2019 (n 103) 205.
[171] Recital 8 DSM Directive.
[172] Rosati 2019 (n 103) 201.
[173] ibid 199.
[174] Recital 105 AI Act.

in the DSM Directive apply to these uses.[175] Therefore, TDM activities can be used in AI training, for example, for AI-driven creativity.[176]

4.1.2.1 Text and Data Mining Activities for Research and Cultural Heritage Purposes

The second step is to see if the certain conditions set out by the TDM exemptions are met for the extraction of works or other subject matters. The first TDM exemption we will explore is found in Article 3 of the DSM Directive.

Article 3 DSM states that member states shall provide an exemption allowing research organisations and cultural heritage institutions to perform text and data mining for scientific research purposes, reproducing and extracting works or other subject matter to which they have lawful access. However, these must be stored with an "appropriate level of security", and rightsholders may implement measures to "ensure the security and integrity of the networks and databases where the works or other subject matter are hosted". Additionally, "Member States shall encourage rightholders, research organisations and cultural heritage institutions to define commonly agreed best practices concerning the application of the obligation".[177]

This section will examine how the exemption found in Article 3 of the DSM Directive applies specifically to AI-generated music. As understood from the article above, this provision normally applies to libraries, museums, and research organisations.[178] The term research organisation is an intricate one to determine, just like the term 'scientific research'. Therefore, we first need to determine if the training and development of an AI system can be classed as a research organisation or scientific research.

Article 2(1) of the DSM Directive defines a research organisation as a research institute or an entity whose primary goal is scientific research or educational activities for scientific research on a not-for-profit basis or for a public interest recognised by the Member state in a way that the results must not be used preferentially.[179]

Scientific research under the directive covers both the natural and human sciences as long as it is an activity aimed at uncovering new knowledge or

[175] Paul Keller, `A first Look at the copyright relevant parts in the final AI Act compromise´ (*Kluwer Copyright Blog*) <https://copyrightblog.kluweriplaw.com/2023/12/11/a-first-look-at-the-copyright-relevant-parts-in-the-final-ai-act-compromise/> accessed 27 March 2024.

[176] Rosati 2019 (n 103) 198.

[177] Article 3 DSM Directive.

[178] See Recital 13 DSM Directive.

[179] Article 2.1 DSM Directive.

insights.[180] Thus, the exemption benefits the research community by allowing the mining of copyright-protected works, leading to new knowledge.[181] However, the emphasis in the requirement is that TDM must be carried out on a non-profit basis, meaning that public-private partnerships are allowed, but the commercial element cannot be too large.[182] Often, in these public-private partnerships, the result will accrue to the private actor, while the public actor receives limited publication rights. If this is the case, then the TDM activities within the partnership won't be covered by Article 3 DSM Directive.[183] This can make the exemption narrow in scope, as many research organisations rely on these partnerships to fund research.[184]

The idea that scientific research should uncover new knowledge or insights seems applicable to AI systems, especially generative AI, which can develop new content. However, the non-commercial element complicates matters for music AI generators, as many operate on a commercial basis.[185] Even if there can be a commercial element, this cannot be too large. Since the focus of the exemption is on scientific research, the potential harm affecting the rightsholders should be minimal.[186] However, as discussed in Chapter 4.1.1, the significant value of music and the music industry does not align with this. Especially when music is mined to generate an output, as it can have a substantial value that competes with the original creation.

We can likely rule out the possibility of it being considered in the public interest, as this requires recognition by the Member State. According to the recital, a public interest mission could be reflected through funding, contracts, or provisions in national laws.[187] Whether this interest needs to be recognised beforehand is another question open to interpretation.

In most cases, AI music generators won't be considered a research organisation unless used solely for scientific, non-commercial purposes. If an AI music generator meets the requirements, the work or subject matter must be acquired legally for reproduction and pre-processing.

Legal access includes situations where the rights holder has consented to the access, contractual licences, or when the access is supported by law, for

[180] Recital 12 DSM Directive; Eleonora Rosati, *Copyright in the digital single market: article-by-Article commentary to the provisions of directive 2019/790 (*Oxford University Press 2021) 43.

[181] Recital 8 DSM Directive.

[182] Recital 11 DSM Directive; Recital 12 DSM Directive.

[183] Kempas (n 7) 200.

[184] Christopher Geiger, Giancarlo Frosio, Oleksandr Bulayenko, "Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations, IIC (2018) 49: 814, 826.

[185] See for instance, Soundraw, `Pricing´ (Soundraw) < https://soundraw.io/#pricing-section> accessed 11 May 2024.

[186] See Recital 17 DSM Directive.

[187] Recital 12 DSM Directive.

example, due to copyright restrictions.[188] Open access policies or content that is freely available online are also covered.[189] However, rightsholders can, according to the article, implement measures like user authentication to secure the networks where the works are hosted.[190] This can potentially hinder access to the desired content to some extent. Some have even argued that rightsholders might increase prices in their different licences by incorporating TDM costs.[191] This could further prevent start-ups and research organisations from entering the market or affect the availability of data, thereby negatively impacting AI development.[192]

The provision furthermore requires that the reproductions must be stored securely.[193] What this entails, besides deletion when the storage of them no longer can be justified for scientific research purposes,[194] is so far unclear.

In music, this exemption could apply to data collection for experiments, clinical trials for music therapy, or archiving for legitimate purposes. However, overall, the exception has a narrow scope and is unlikely to apply to most current music AI developers, especially when the mined data is processed and adapted, which doesn't comply with the requirements of this provision.

4.1.2.2 Text and Data Mining Activities for Other Purposes

The other TDM exemption is found in Article 4 DSM Directive. Unlike Article 3 DSM Directive, which only applies to research organisations and cultural heritage institutions, Article 4 DSM Directive is broader as it does not have any restriction in terms of beneficiaries and can be used by both commercial and non-commercial actors.[195] This exemption is, therefore, perhaps the most suitable exception and particularly interesting for AI-driven creativity.[196]

Article 4 DSM Directive states that member states shall make exceptions to rights for the "reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining." Reproduction and extraction may go on for as long as it is necessary. However, this only applies to the use of work that has "not been expressly reserved by their

[188] Kempas (n 7) 197-198.
[189] Recital 14 DSM Directive.
[190] Recital 16 DSM Directive.
[191] Geiger (n 184) 836.
[192] Charlotte Gerrish and Anders Mmolander Skavlan, 'European Copyright Law and Text and Data Mining Exceptions and Limitations: In Light of the Recent DSM Directive, is the EU Approach a Hindrance or Facilitator to Innovation in the Region?' (2019) 2(2) SIPLR 58, 62.
[193] Kempas (n 7) 198.
[194] Rosati 2021 (n 180) 57.
[195] ibid 74; Kempas (n 7) 200.
[196] Rosati 2019 (n 103) 200.

rightsholders in an appropriate manner".[197] This includes works, related rights, databases, and even computer programmes.[198]

However, similar to Article 3 DSM Directive, certain conditions must be met. One key requirement is that the work must be legally accessed. The criterion for lawfully accessed in Article 4 DSM Directive differs from the criterion in Article 3.[199] Lawful access in Article 4 of the DSM Directive covers situations when the work is accessed lawfully, or contain non-protected elements, or when other applicable exemptions and limitations are at hand.[200]

Recital 18 of the DSM Directive explains that rightsholders should be "able to licence the use of their works or other subject matter" that does not fall under the mandatory exemptions for TDM activities for scientific research purposes, as well as existing limitations and exceptions found in the InfoSoc Directive, lawful access can also include situations where the content have been made publicly available online.[201] This aligns with previous case law, which explains that use is lawful when authorised by the rightsholders or not restricted by legislation.[202] This indicates that lawfulness applies to situations where the content has been licensed, hence authorised, or is available online.

Recital 14 of the DSM Directive further explains that lawful access includes open access policies.[203] However, the recital in question only applies to research organisations and cultural heritage institutions. It remains unclear if open access policies also apply to the TDM exemption in Article 4 DSM for broader use.

The other situation of lawful access covers instances that fall under exemptions or limitations. This can include situations such as temporary acts of reproduction in Article 5(1) InfoSoc Directive, the exception for private use found in Article 5 (2) (b) InfoSoc, or for teaching or scientific research purposes found in Article 5(3)(a) InfoSoc, as well as Article 6(1) of the database Directive.[204]

Rightsholder consent through licences can create practical issues for AI developers, as there might be difficulties in correctly identifying the subjects from whom permission should be granted.[205] In reality, this gives an

[197] Article 4 DSM Directive.
[198] Rosati 2021 (n 180) 75.
[199] ibid 76.
[200] ibid 76-77.
[201] Recital 18 DSM Directive.
[202] See FAPL (n 34) para 168.
[203] Recital 14 DSM Directive.
[204] Rosati 2019 (n 103) 208.
[205] ibid 205.

advantage to larger AI companies, which generally have better resources for negotiation compared to start-ups.

Additionally, rightsholders can reserve their rights regarding TDM activities if done in an appropriate manner. What is considered "appropriate" is not strictly defined, but it is assumed that such reservations must be made by technical means. Although the provision's wording is less strict, the recital clarifies that other methods may be permitted for content made available by means other than online.[206] The assumption that reservations should be made online stems from CJEU case law that highlights the difficult of otherwise verifying whether the rightsholder intended to oppose the use.[207]

In practice, this means that the rightsholder's systems or database must technically integrate with the tools a third party uses for TDM. However, rightsholders are not required to take extensive technical measures to accommodate this.[208] Simply methods, such as using metadata, or terms and conditions on websites, can be sufficient.[209] Normally, robot.text files are placed on databases and systems to instruct about the opt-out. Therefore, it can be assumed that TDM activities that contravened the instructions in these files do not apply to Article 4 DSM.[210]

For example, Amazon's user agreement for its music service platform states, "You may not use the Services to store, transfer, or distribute content of or on behalf of third parties." It also specifies that users cannot redistribute, transmit, repurpose, adapt, transfer, or use purchased music or music service content in ways such as sale, reproduction, or distribution, as the streaming site does not grant these rights for any music content.[211]

The rightsholders' ability to oppose the use of their works severely limits the potential of Article 4 DSM Directive. This is especially true for music, as music streaming platforms often restrict TDM activities in their terms and conditions, and record labels such as Universal Music Group have asked streaming platforms like Spotify and Apple to block AI from scraping music from their platforms.[212]

---

[206] Rosati 2021 (n 180) 90.

[207] See Case C-392/19 VG Bild-Kunst v Stiftung Preußischer Kulturbesitz (2021) ECLI:EU:C:2021:181, para 46.

[208] Kempas (n 7) 202-203.

[209] Recital 18 DSM.

[210] Kempas (n 7) 202-203.

[211] See Amazon, `Amazone Music Terms of Use´ (Amazon, 1 May 2023, <https://www.amazon.com/gp/help/customer/display.html?nodeId=201380010> accessed 12 April 2024.

[212] Ephrat Livni, Lauren Hirsch and Sarah Kessler, ´Who Owns a Song Created by A.I.?´ (The New York Times, 17 April 2023) <https://www.nytimes.com/2023/04/15/business/dealbook/artificial-intelligence-copyright.html> accessed 18 January 2024.

The exceptions and limitations in the DSM directive aim to achieve a balance between the rightsholders and the users.[213] With this in mind, it seems fair that rightsholders can opt-out, and permission does not have to be sought as it does with other copyrighted works.

If a rightsholder does not appropriately reserve their rights, it will likely be harder to prevent unauthorised TDM activities.[214] This can be especially difficult for musicians who publish their own works independently on websites or blogs. However, as mentioned above, it should be possible to prevent TDM activities through the terms and conditions on the website.

However, how this works in reality is hard to predict, as AI systems generally cannot determine whether a website has been subject to contractual restrictions. Furthermore, it can be nearly impossible to monitor all the reservations due to the large volume of data being mined. [215]

The criteria of lawful access and the rightsholder's reservation rights are closely connected. However, the reservation right goes a step further by preventing TDM on works that have been lawfully accessed by the user.

Lastly, the mined data cannot be stored longer than necessary. Storage by electronic means includes, for instance, cloud storage.[216] Article 3 of the DSM Directive requires an appropriate level of security for storage, while Article 4 specifies that the data can only be stored as long as necessary. This means that the copies must be removed once the purpose of the TDM activity is complete.[217]

If all requirements are met, the DSM Directive could permit specific reproductions for automated computational analysis of music stored in digital repositories.[218] However, the DSM Directive is still quite new, and copyrighted works obtained before the TDM exception entered into force may be considered stolen intellectual property.[219] Additionally, only acts of reproduction are permissible, which further restricts the use of this exception in the context of AI-generated music, as generative AI systems likely modify mined content during the training.

[213] See Recital 6 DSM Directive.

[214] Rosati 2021 (n 180) 90.

[215] Gerrish (n 192) 61.

[216] Case C-433/20, AustroMechana Gesellschaft zur Wahrnehmung mechanisch-musikalischer Urheberrechte GmbH v Strato AG (2022) ECLI:EU:C:2022:217, para 16-18.

[217] Rosati 2021 (n 180) 89.

[218] Sturm (n 12) 5.

[219] EWC, `European Writers' Council (EWC) second Statement on the AI Act Proposal´ (EWC, 26 July 2023) <https://europeanwriterscouncil.eu/23ewc_on_aiact/> accessed 28 March 2024.

4.2 Input Interpretation from the Artificial Intelligence Act

In this section, only relevant provisions and recitals that can affect copyright for generative models and the interpretation of the lawfulness in the initial training phase will be discussed.

As mentioned above, AI model training generally requires large amounts of data, and TDM techniques may be used to retrieve and analyse this content. If the rightsholders have reserved their rights appropriately, providers of general-purpose AI models must first obtain authorisation from them.[220]

Furthermore, the Act has laid out a requirement of transparency that requires the providers to draw up a detailed public summary of the content used in the training of the AI model to facilitate the parties with legitimate interests, such as copyright holders, to be able to "exercise and enforce their rights",[221] limits the possibility for providers to train generative models on copyright protected content. This, in combination with the fact that the act makes a reference to the DSM Directive and the opt-out option, which severely limits the possibility to use TDM techniques in Article 4 DSM directive, gives the impression that the EU don't intend to allow the unauthorised use of copyright-protected materials in the training phase of AI systems.

The transparency requirement can seem aimed at larger commercial AI-model providers, as they have so far been reluctant to disclose descriptions of their training data. A study found that many foundation model providers do not comply with the act's requirements regarding the description of the use of copyrighted training data.[222]

The act further emphasises that high-risk AI systems must respect and observe intellectual property rights in accordance with national and Union law.[223] To clarify, the position on territoriality makes it possible to pre-empt situations like in the ongoing British case Getty v. Stability, where Stability AI unsuccessfully tried to question the court's jurisdiction.[224] It is, however, unclear whether the territoriality requirement can be implemented legally or effectively, as the realities of AI model training is that "copies take place in a context that can be far removed (both geographically and temporally) from the actual use of the resulting models, this seems like a necessary intervention".[225] This indicates that EUs intend not to let providers circumvent the

[220] Recital 105 AI Act.
[221] Article 53.1 (a) and (d) and Recital 107 AI Act.
[222] Bommasani (n 156).
[223] Article 25 (5) AI Act.
[224] See Getty Images v Stability AI (n 4).
[225] Keller (n 175).

provisions and allow for any possibility of training AI models on content covered by copyright or related rights.

The European Writers Council (EWC) has commented about the transparency requirement found in article 28b(4) and 52.3a of the draft from the14th of June 2023,[226] that these provisions were not sufficient enough to provide a sustainable future for authors and preforms who are the primary suppliers of all cultural work, nor are they sufficient enough make up for the past harm done by the illegal scraping and mining of authors works.[227] The EWC further explains that tracking the utilised work would be the only way to establish a required licensing regime and sources for generative AI machine learning. Which must be properly accounted for. Any prior programs that relied on theft must be discontinued immediately.[228]

Even though the proposal is a step in the right direction to regulate and clarify the use of AI, it comes across as a bit murky. Yet, the overall take is that the EU does not intend to accept generative AI models being trained on content covered by copyright and related rights without previous authorisation from the rightsholders.

4.3 Different Solutions to the Same Problem

This section presents and discusses how 3 other jurisdictions could potentially handle the question of copyright-protected works or content covered by related rights in the training of a generative AI system to see if this can give the EU copyright regime any future guidance.

4.3.1 USA and the Fair Use Doctrine

The question of whether the content covered by copyright or related rights used in the input of a generative AI model is considered a copyright infringement gets quite complicated if we look at the Fair Use Doctrine, as the training of an AI system on copyrighted materials may constitute fair use depending on "the purpose and character of the use", "the nature of the copyrighted work", the amount of copyrighted work used, and the effect the use has on

[226] European Parliament, `Artificial Intelligence Act Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))´ P9_TA(2023)0236; Article 28.b (4) explains that generative AI systems shall comply with transparency, ensure adequate safeguards against breaches, and make available a detailed summary of the used training data. Article 52.3a explains that deep fakes that should be disclosed in an appropriate manner do not apply if it is necessary to express freedom of expression.

[227] EWC (n 219) and See European Parliament (n 41) for a comparison with Article 53 (1) (d) AI Act. The input is referred to as content instead of data in the final draft. The transparency requirement has not changed much in substance from the draft in June 2023. Therefore, the EWC's opinion should stand.

[228] EWC (n 219).

the copyright-protected material.[229] The application of these different factors will depend on the specific facts of the case.[230] The AI company OpenAI have, for instance, argued that its purposes are transformative and that the reproductions that are made under the training of the model are not made available to the public, that the training itself should not harm the market or the value of the copyrighted work, and that this, therefore, should be considered fair use.[231]

There are so far no final cases regarding AI in the USA. However, a few are currently ongoing, such as The New York Times v. OpenAI and Microsoft, Andersen v. Stability AI Ltd, and Thomson Reuters v Ross Intelligence (Westlaw).[232] The closest case so far that can be of guidance in the field is the Authors Guild v Google, in which Google scanned and stored copies of a total of 15 million books, half of which were copyright protected; these were then offered via the internet in the form of snippets or full texts by google books.[233] In the case the Second Circuit found that the defendant's activities were fair use, as the search engine made new forms of research (text mining and data mining) possible by using the corpus to gain new statistical information. Furthermore, Google never intended to sell these books; instead, it only provided information about the books in the form of snippets.[234]

> When the case was filed, not a lot of people would have thought that putting millions of books in the database of a for-profit company would be fair use. The law evolved and by the time the case was decided, it was fair use.[235]

This is particularly interesting for training AI models, as the case shows that TDM activities for commercial purposes can fall under fair use.

Another case that can provide guidance is Andy Warhol v. Goldsmith, in which the artist Warhol was not protected by fair use when he made a portrait based on a photograph of Prince taken by Lynn Goldsmith. It was not

[229] Zirpoli (n 3).

[230] Evana Wright, 'Professionals Beware: The Opportunities and Risks of Generative AI in Legal Practice' (University of Technology Sydney) <https://opus.lib.uts.edu.au/bitstream/10453/176325/3/Professionals%20Beware%20The%20Opportunities%20and%20Risks%20of%20Generative%20AI%20in%20Legal%20Practice.pdf> accessed 14 April 2024.

[231] Patent and Trademark Office (n 114).

[232] The New York Times v. OpenAI (n 4) and Andersen v. Stability AI (n 4) and Thomson Reuters Enterprise Centre GmbH and West Publishing Corporation v Ross Intelligence Inc (WestLaw) Case No 1:20-cv-00613 (D Del, filed 6 May 2020).

[233] Authors Guild, Inc. v Google Inc., 804 F.3d 202 (2d Cir. 2015).

[234] ibid 1-6.

[235] Irina Ivanova, `Artists sue AI company for billions, alleging "parasite" app used their work for free´ (CBS news, 20 Janauary 2023) <https://www.cbsnews.com/news/ai-stable-diffusion-stability-ai-lawsuit-artists-sue-image-generators/> accessed 18 January 2024´ (Quoting Ginsburg from Columbia University).

considered fair use as the photographs served the same purpose and would, therefore, risk substituting the licensed version.[236]

In the ongoing Westlaw lawsuit, the court has so far found that while OpenAI's use was commercial, a jury needs to resolve whether the use is transformative to the extent the plaintiff's work favoured fair use or if more than necessary was copied and whether this constitutes a market substitute.[237] This can be particularly interesting if we discuss AI systems that generate musical outputs in the style of an artist.

Generative AI can force these issues to be legally addressed, as outputs "are meaningfully based upon input work".[238] Overall, many believe that the existing law doesn't have to be changed, as the fair use doctrine is flexible and hence capable of adapting to the new issue of copyrighted works in an AI context,[239] However, it is still unsure if this is permissible under fair use laws, particularly when it comes to data that has particular licences, to reproducing this data.[240]

The main lesson to be learned from the Google Books and Warhol case from an EU perspective is that the interference of the author's normal use is of high importance in the infringement assessment.

4.3.2 The UK and Fair Dealing

Just like in the US, there is also an ongoing Lawsuit in the UK against an AI developer for using copyright-protected materials to train an AI system.[241]

How the country will treat the training of AI systems on copyright-protected content is so far unclear. The UK have an exception in Sec 28A of the CDPA, similar to Article 5 (1) in the InfoSoc Directive, that states that it is not copyright infringed if the reproduction is temporary or incidental and if this copy is a necessary component of a technological process, whit the only purpose being to facilitate the transmission of the work or use if for a legitimate purpose, and the reproduction cannot have an independent economic value. It is

[236] Andy Warhol Foundation for the Visual Arts, Inc. v Goldsmith et al., No 21–869 (US Supreme Court, 18 May 2023.

[237] Thomson Reuters v ROSS Intelligence (n 232) , in which legal action has been brought by media giant Thomson Reuters against artificial intelligence (AI) company Ross Intelligence for its legal research. The owner of the legal research website WestLaw, Thomson Reuters, claims that Ross violated its copyright by unauthorised reproduction of WestLaw's concise summaries of legal issues found in court rulings.

[238] Ian Krietzberg, `Copyright expert predicts result of NY Times lawsuit against Microsoft, OpenAI´ (*TheStreet*, 7 January 2024) <https://www.thestreet.com/technology/copyright-expert-predicts-result-of-ny-times-lawsuit> accessed 14 April 2024 (quoting Grimmelmann from Cornell University).

[239] USPTO, `Public Views on Artificial Intelligence and Intellectual Property Policy´ (USPTO, 2020).

[240] Bommasani (n 156).

[241] Getty Images v Stability AI (n 4).

likely that the generative AI model does not fulfil this requirement, just like Article 5 (1), as it might not be considered lawful use or incidental or without economic significance. However, this issue will likely be argued in future copyright litigations.[242]

The UK also has an exception for TDM activities for non-commercial purposes.[243] TDM activities can benefit scientific research in general. Just like the TDM exceptions found in the DSM Directive, this provision comes with questions about the definition of research and "whether a dataset collected for research can be reused by a commercial entity".[244] Even if TDM activities occur outside these jurisdictions, publishers' policies still apply, and just like the provision in the DSM Directive, the performer of the TDM activity must have lawful access to the artistic work via subscriptions, for instance.[245]

So far, this is quite similar to the EU copyright legislation. However, the country did have a proposal to change the TDM exception to include a broader spectrum of beneficiaries. This suggestion has since been withdrawn due to criticism from the creative sector and the House of Lords, which states that this change would be unnecessary and harmful and would affect the music heritage.[246] They affirmed that developing AI is important. However, this should not transpire at unnecessary cost, and if the legislator were to find any negative effects in relation to the creative industries, other alternatives should be pursued instead.[247] As the proposal did not go any further, it can be assumed that the effects on the creative industries, such as the music industry, were considered too large.

4.3.3 Japan

Japan has announced that copyright on data used in AI training cannot be enforced.[248] Japan's policy for generative models permits processing data

[242] See Andres Guadamuz, `A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs´ (2024) 73(2) GRUR International 111, 117.

[243] 29A UK CDPA.

[244] Guadamuz (n 242) 119.

[245] Rosati 2019 (n 103) 204; University of Edinburgh, 'Copyright Exceptions and Fair Dealing' (University of Edinburgh, 15 June 2021) https://www.ed.ac.uk/information-services/library-museum-gallery/library-help/copyright/copyright-exceptions-and-fair-dealing accessed 18 May 2024.

[246] Rachel Montagnon and Sungmin Cho, ` UK withdraws plans for broader Text and Data Mining (TDM) copyright and database right exception´ (Herbert Smith Freehills, 1 March 2023) <https://hsfnotes.com/ip/2023/03/01/uk-withdraws-plans-for-broader-text-and-data-mining-tdm-copyright-and-database-right-exception/> accessed 13 May 2023; Communications and Digital Committee, At Risk: Our Creative Future (2nd Report of Session 2022-23, HL Paper 125, 17 January 2023) ch 2, para 31; UK Music (n 14).

[247] Communications and Digital Committee (n 246) para 34-35.

[248] James G. Gatto, ´Getty Image's AI Model Training Lawsuit in UK Against Stability to Proceed´ (Art Law Gallery Blog, 7 December, 2024) <https://www.artlawgallery.com/2023/12/articles/artificial-intelligence/getty-images-ai-model-training-lawsuit-in-uk-against-stability-to-proceed/> accessed 18 May 2024.

regardless of purposes and content, if the content is obtained unlawfully, or if the sourcing extends to exclusive rights other than reproductions.[249]

However, the country differentiates between model usage and model training. The model training involved the training of the LLM, while the model usage involves usage prompts[250] and prompt-based outputs. If the training is used from information analysis, the usage most likely won't be subject to Japanese copyright law, but this will not apply to the prompt input stage. Therefore, users who use copyright-protected works in the prompt input stage could be held liable for violating copyright law. It is furthermore unclear how this applies to the generated outputs, but there is a possibility that it can be an infringement if the output resembles and is reliant on an existing work.[251]

Japan's legislation makes it lawful to train generative models, but it can be hard in practice to make music generators lawful as this extends beyond information analysis. There is also a fear that this could make AI companies move to Japan to avoid the potentially harsher EU legislation. This leads to questions on whether the EU can fall behind in the development of AI.

[249] John Donegan, `The US should look at Japan's unique approach to generative AI copyright law´ (ManageEngine, 24 February 2024) <https://insights.manageengine.com/artificial-intelligence/the-us-should-look-at-japans-unique-approach-to-generative-ai-copyright-law/#:~:text=Unlike%20the%20U.S.%20and%20E.U.,AI%20(gen%20AI)%20models> accessed 18 May 2024.

[250] Instructions to the AI model that can include voice, text, or image inputs.

[251] See Donegan (n 249).

5 Output

Another potential copyright issue can arise with the AI-generated output, as these outputs can resemble existing works, which can constitute an infringement on distribution and reproduction rights.[252] It can also potentially violate the author's moral rights. There are, however, uncertainties regarding the unauthorised reproduction of copyright-protected works and how these cohere with moral rights, especially the right of attribution of authorship as well as the right of integrity of authorship.[253]

In the ongoing Getty lawsuit, the plaintiff claimed that the AI generator produced nearly identical images and derivative works of the Getty images.[254] A study found that less than 2 % of the images produced by Stable Diffusion contained significant copying. However, the study's authors later revealed that their methodology probably underestimated the true rate of copying.[255]

The New York Times lawsuit evidence highlights examples of ChatGPT reproducing output that was almost identical to the copyright-protected articles from The New York Times.[256] In a motion to dismiss, OpenAI claimed that the chatbot was hacked by exploiting a bug and violating the service's terms of use to generate "highly anomalous results."[257]

Bug or not, the cases mentioned above speak for the possibility of infringement on the author's exclusive rights, such as the distribution right. This can include communication to the public via retransmission works included in terrestrial television broadcasting by organisations other than the original broadcaster via an internet stream.[258]

On the contrary, one AI company has expressed that overfitting, which could make outputs similar, is a very unlikely accidental outcome of a well-designed generative AI system.[259]

---

[252] Sunray (n 114) 175.

[253] Wright (n 230) 7.

[254] Getty Images v Stability AI (n 4).

[255] Zirpoli (n 3).

[256] The New York Times v. OpenAI (n 4). The plaintiff claimed that the LLM can generate output that recites the New York Times articles both literally, as well as summarises them and mimics the style which can be seen in Exhibit J.

[257] Cade Metz and Robert Wade, `OpenAI Seeks to Dismiss Parts of The New York Times Lawsuit´ (*The New York Times*, 28 February 2024) <https://www.nytimes.com/2024/02/27/technology/openai-new-york-times-lawsuit.html> accessed 15 April 2024.

[258] Case C-607/11, ITV Broadcasting Ltd and Others v TVCatchup Ltd (2013) ECLI:EU:C:2013:147. Para 48; Kur (n 18).

[259] Patent and Trademark Office (n 114).

5.1 The Music and the Phonogram

Firstly, when it comes to human-created works like text, sound, and images, it should be presumed that these, or at least parts of these, are covered by copyright or related rights. That also applies to material in big data, such as sound recordings and performances.[260]

Therefore, when AI constructs a song based on the input data, both the musical composition and the sound recordings that musical works are tied to can be infringed, thereby, an infringement analysis must be conducted for these elements separately.[261]

Rather than using the general composition structure in musical works, an AI generator is more likely to mimic the sound recording for its performance details when creating a song, as it cannot replicate the full range of human musical expression. However, the generator will still analyse the underlying musical work during the training phase to understand how the music styles are composed.[262]

Regarding AI-generated songs and music in general, it is less evident what constitutes a reproduction. It has been described that music is more similar than different nowadays, as people are using the same plug-ins or sample packs. There is also an issue of a limited number of chords and melodies available.[263] This will, however, vary between musical categories as genres such as Jazz are less restrictive compared to pop music. Pop music is therefore often created in teams, which complicates the identification of the creators.[264]

There is also a vastly different approach to how music infringement cases are assessed in the UK, US, and EU. In the UK, a musicologist often plays an important part in these cases, whereas in the USA, they are decided by a jury.[265]

The Universal Music Group, representing both artists that the song "Hart on My Sleeve" mimicked, claimed that AI companies infringe on their copyright by utilising their music.[266] Therefore, the most suitable way to address the

---

[260] Kempas (n 7) 171.

[261] Newton v. Diamond, 388 F.3d 1189, 1193 (9th Cir. 2004) in Sunray (n 114) 206-207.

[262] Sunray (n 114) 208; Prafulla Dhariwal, Heewoo Jun, Christine McLeavey Payne, ´Jukebox´ (OpenAI, 30 April 2020) < https://openai.com/index/jukebox/> accessed 17 May 2024.

[263] Amy X. Wang, ` How Music Copyright Lawsuits Are Scaring Away New Hits´ (*Rolling Stones*, 9 January 2020) <https://www.rollingstone.com/pro/features/music-copyright-lawsuits-chilling-effect-935310/> accessed 23 January 2024.

[264] Rick Fulker, `Kraftwerk case: EU top court defines sampling´ (DW, 30 July 2019) <https://www.dw.com/en/kraftwerk-case-eu-top-court-defines-rules-of-sampling/a-46916334> Accessed 23 January 2024.

[265] Wang (n 263).

[266] Zirpoli (n 3).

issue regarding AI-generated musical outputs could be to look at legislation and case law for derivative work and sampling.

5.1.1 Sampling

Sampling can be described as the act of taking a part of a sound recording to use in a new song.[267] The CJEU has defined sampling as the "user taking a sample from a phonogram, most often by means of electronic equipment, and using the sample for the purposes of creating a new work".[268]

With the development of digital technology and home recordings, the act of sampling from other musicians has become as easy as copying and pasting documents.[269]

Today's remixes and samples require both the master's use and the sync licenses.[270] The sound recordings are often owned by the record label if an artist is signed to a bigger record label, whereas if the artist is independent, they often own the sound recording themselves. The copyright owners in the music and lyrics are often the artist or the music publishing company.[271] However, the ownership can be transferred through music copyrights.[272] If this is not done correctly, a copyright infringement may occur. This was the case when The Verve accidentally applied for the wrong license and, therefore, were forced to give up their rights, which meant that the band could not collect any royalties for their hit song "Bittersweet Symphony" for 22 years, after which the song was returned to them.[273] Similarly, the file-sharing company Napster also used unauthorised music and got sued by the band Metallica and later on by A&M Records, which led to the shutdown of the site.[274] Issues also arose

[267] Jesse (n 8) 95.

[268] Pelham (n 137) para 35.

[269] Jesse (n 8) 95.

[270] Soundcharts Team, `6 Basics of Music Copyright Law: What It Protects and How to Copyright a Song´ (Soundchars Blog, 1 January 2024) <https://soundcharts.com/blog/music-copyrights> accessed 17 April 2024.

[271] Jesse (n 8) 17.

[272] Jesse (n 8) 17.

[273] Jesse (n 8) 3-4; See Sky News, `Bitter Sweet Symphony: Rolling Stones return royalties and rights to Richard Ashcroft after 22-year row´ <https://news.sky.com/story/bitter-sweet-symphony-rolling-stones-return-royalties-and-rights-to-richard-ashcroft-after-22-year-row-11727342#:~:text=The%20song%20itself%20was%20a,nominated%20for%20a%20Grammy%20award.&text=But%20after%20almost%2022%20years,following%20a%20%22magnanimous%20gesture%22> accessed 14 May 2024.

[274] See A&M Records, Inc. v Napster, Inc., 239 F.3d 1004 (9th Cir. 2001) and Metallica v Napster, Inc., No C 00-04083 (ND Cal, 2000) and Stephen Robertson, 'How Napster Transformed a Copyright Shutdown to a $70m Price Tag' (Metis Partners) <https://metispartners.com/thought-leadership/napster/> accessed 18 May 2024.

with the streaming platform Spotify, which instead had to get a licence from all record labels presented on their platform.[275]

The use of samples is, therefore, generally not permitted without authorisation, and in the US case Bridgeport Music, the Court of Appeals for the Sixth Circuit ruled that sampling, regardless of length, can violate copyright law.[276] This case later influenced the CJEU in Pelham, in which the court ruled on the copyright implication of sampling and established the boundaries between sampling and artistic freedom. In this case, a two-second sample from Kraftwerk's song "Metall auf Metall" was used in the hip-hop track "Nur mir" that the producer Pelham created for Sabrina Setlur.[277]

In the case, the CJEU stated that article 2 (c) in the InfoSoc Directive should be interpreted in light of the charter of fundamental human rights in the European Union to mean that the phonogram producer's exclusive right to reproduce and distribute their phonogram also includes the right to prevent others from "taking a sound sample, even if very short of the phonogram to "including that sample in another phonogram unless that sample is included in the phonogram in a modified form unrecognisable to the ear."[278]

Thus, if a creator modifies a song by the method of 'sampling' to a degree that it is no longer recognisable to the ear, then this can be considered a form of artistic expression. This leads us to the question of when a song is unrecognisable.

The court leaves this question open, together with the unclarities regarding whom the sample shall be unrecognisable.[279]

If we circle back to the US and the fair use doctrine, the court decided that the 0.23-second horn blast from the song Ciccone was deemed unrecognisable when it was used in the song Vogue.[280]

This can perhaps guide us regarding the musical output created by AI. Logically, this can be interpreted as: If the AI-generated output uses a short part

---

[275] Jeff Price, `The Definitive Guide to Spotify Royalties´ (Medium, 17 February 2020) <https://medium.com/@JPriceOfMusic/the-definitive-guide-to-spotify-royalties-dc5960862c00#:~:text=If%20the%20record%20label%2Fdistributor,differ%20from%20agreement%20to%20agreement> accessed 14 May 2024.

[276] Bridgeport Music, Inc. v Dimension Films, 410 F.3d 792 (6th Cir. 2005), abstract.

[277] Fulker (n 264).

[278] Pelham (n 137) para 39.

[279] Benjamin Beck, `Court of Justice of the EU Rules on Copyright Implications of Sampling in Music´ (MayerBrown, 5 September 2019) <https://www.mayerbrown.com/en/perspectives-events/blogs/2019/09/court-of-justice-of-the-eu-rules-on-copyright-implications-of-sampling-in-music> accessed 23 January 2024.

[280] VMG Salsoul, LLC v Ciccone, 824 F.3d 871 (9th Cir. 2016). In the case, it was found that a "reasonable jury could not conclude that an average audience would recognize an appropriation of the Love Break composition".

of the music from the artist or phonogram, this can be a copyright infringement if it is not unrecognisable. If the sample is very short, it is less likely to be recognisable, hence lawful.

Therefore, there could be a risk that AI-generated music contains substantial samples, especially if the model replicates a specific dataset, for instance. The discussion on sampling is still highly debated, and some argue that good composers of serious music are highly creative in the way they steal music from others.[281]

Even if the sample used can be recognised to the ear, there is a possibility that this part might be covered by the quotation exemption if the sample used intends to converse with the work that served as the source of the sample.[282] However, for the quotation exemption to be applicable, the quotation cannot be extended to situations in which the quoted work is unidentifiable.[283] The EWC has also expressed that allowing machines to have freedom of expression would cause the hiding of AI information and dehumanise and remove the rights reserved for human beings and their intellectual creations.[284]

Overall, it can be concluded that an assessment must be conducted on a case-by-case basis, as this is the only way to ascertain the degree to which the original work has potentially been misappropriated and sampled in the output.[285]

5.1.1.1 Derivative Works

Even if sampling is a type of derivative work,[286] this section will explore situations in which the AI-generated output is more transformative.

Copyright protection covers both the author's original work and derivative works. Unfortunately, the copyright regulations on adaptations are not fully harmonised within the EU, and there are no standards on how much alterations are required for a work to be considered novel.[287]

The question is whether an AI-generated musical output is considered a derivative work or an adaptation when only the style of artists is mimicked and when AI-generated outputs show enough independence from the input to prevent an infringement?

[281] Fulker (n 264).
[282] Beck (n 279).
[283] Pelham (n 137) para 74.
[284] EWC (n 219).
[285] Sunray (n 114) 208.
[286] Jesse (n 8) 96.
[287] Kur (n 18) 316; European Commission (n 41); Sturm (n 12) 9.

One guideline is that the work should be created independently and not excessively influenced by other works. However, this should not be interpreted too strictly, as it is acceptable to be inspired by others.[288] To complicate things further, authors are not obligated to disclose their creative process.[289] Human creations are furthermore often based upon other works that the author, for instance, has seen, read about, or been influenced by in some way.[290] So, there can be a fine line between inspiration and plagiarism. Isaac Newton once stated, "If I have seen further than others, it is by standing upon the shoulders of giants".[291]

As mentioned above, copyright generally only exists in the tangible medium of expression, leaving the techniques, styles, ideas, and purpose behind the work unprotected.[292]

This assessment can, therefore, be especially complicated when AI is involved as it is not entirely clear if the creation of an AI output involves copying a work in the legal sense. It is more likely that the generator uses the data from the work to identify patterns and correlations, which in turn are utilised to generate a new output.[293]

The assessment can also depend on how many sources the AI output has been influenced by. If there, for instance, is a smaller number of identified input works that the output has been based upon, and the creative choices are limited to the same author, then it could be argued that there is a larger likelihood of infringing on an existing work and vice versa.[294] If a smaller pool of input data is used to generate the output, then it can be argued that the "creative choices were transferred" from the author to the machine-produced output. This could potentially overlap with the author's rights.[295] A detailed assessment will thereby be necessary to compare the final AI-generated output with the works used in the input data.[296] The findings that establish the nature and origin of the generated decision could then be a part of the legal process.[297]

Questions have also arisen about whether AI should be interpreted the same as humans regarding infringement assessments. AI are by some comparable to human artists, and as it is allowed for a humans to be inspired by others, the same should apply to AI, as AIs are learning in similar ways as humans.[298]

[288] Axhamn (n 119) 68.
[289] Cabrera Blázquez (n 24).
[290] Daniel J. Gervais, `AI Derivatives: The Application to the Derivative Work Right to Literary and Artistic Productions of AI Machines´ (2022) 53 Seton Hall Law Review 8.
[291] Fulker (n 264).
[292] Axhamn (n 119) 70.
[293] Cabrera Blázquez (n 24).
[294] Axhamn (n 119) 69.
[295] Gervais (n 290).
[296] Axhamn (n 119) 69.
[297] Cabrera Blázquez (n 24).
[298] Ivanova (n 235).

Others claim that the situation of AI inspiration is unlike that of humans that instead store "works as electromagnetic traces in the brain".[299] The AI-generated painting "The Next Rembrandt" is, for instance, solely based on Rembrandt's works (346 paintings),[300] this is most likely "remembered" in a different manner for AI than the human brain.

Furthermore, copyrighted music gets even more complicated as there is a naturally restricted music palette, so songwriters frequently have to incorporate similar elements in their works, such as melodic intervals, rhythms and chords.[301] Others, on the other hand, claim that there is "an infinite number of creative choices in creating a song".[302] In the US, lawyers have long argued that judges are too reluctant to dismiss cases because of the complexity of music and that juries find it difficult to distinguish between parts of a song that are or are not covered by copyright.[303]

In 2015, Williams and Thicke's song "Blurred Lines" was considered to be too similar to Gaye's song "Got to Give It Up". The verdict has been highly controversial as elements such as the groove and feeling were discussed in the infringement analysis.[304] In the case, the jury had to look at the signature phrases, the hooks line, the bass line, the harmonic structure, the keyboard chords, and the vocal melodies and decide whether these layers of music were considered to go beyond normal copyright-ineligible building blocks.[305] Judge Nguyen wrote about the case that

> "Blurred Lines" and "Got To Give It Up" were not objectively similar as a matter of law under the extrinsic test because they differed in melody, harmony, and rhythm, and the majority's refusal to compare the two works improperly allowed the defendants to copyright a musical style.[306]

Many fail to recognise that works within the same genre share some protectable elements, and there is plenty of evidence that Williams and Thicke attempted to evoke Marvin Gaye's style. The question is whether they took too

---

[299] Axhamn (n 119) 52.

[300] See Axhamn (n 119) 43.

[301] Sunray (n 114) 196.

[302] Ben Sisario, `The `Blurred Lines´ Case Scared Songwriters. But Its Time May Be Up´ (*The New York Times,* 24 March 2020) <https://www.nytimes.com/2020/03/24/arts/music/blurred-lines-led-zeppelin-copyright.html> Accessed 20 April 2024.

[303] Sisario (n 302) (Quoting Christopher J. Buccafusco, a professor at Cardozo Law School).

[304] Rick Fulker, `Kraftwerk case: EU top court defines sampling´ July 30 2019, <https://www.dw.com/en/kraftwerk-case-eu-top-court-defines-rules-of-sampling/a-46916334> Accessed 23 January 2024.

[305] Juhasz Law, 'Blurred Lines: Copyright Infringement, Marvin Gaye, Patent Preemption, and a New Frontier in Copyright Litigation' (Patent Horizon, 28 June 2015) <https://www.patenthorizon.com/blurred-lines-copyright-infringement-marvin-gaye-patent-preemption-and-a-new-frontier-in-copyright-litigation/> accessed 20 April 2024.

[306] Williams v Gaye, 895 F.3d 1106 (9th Cir. 2018).

much of it.[307] It might be easier to consider the different components in a musical piece as layers of sound, and in this case, Gaye´s song included layers like the melody line, chords, bass introduction and lyrics. The other layers of the sound that were not covered by copyright were the backing vocals, percussion parts, background noises and the singer's voice. For there to be an infringement, the sound layers of Williams and Thicke's song must be substantially similar to the sound lawyers in Gaye´s work.[308]

This lawsuit and argumentation could perhaps provide guidance when interpreting whether imitation by generative AI outputs infringes on existing copyright. This could be the case if an AI-generated song is too heavily inspired by another smaller sample of the song, and the underlying layers, such as the melody line, lyrics and chords, follow a structure that is too similar to the copyright-protected work.

However, the Gayes lawsuit is from the US, where the fair use doctrine is dominating, and the EU has generally been opposed to the idea of exception structures similar to Fair Use. Yet, many companies can operate successfully in different jurisdictions despite differences in copyright law; this suggests that the divisive copyright laws might be less concerning than people think.[309] This would not be the first time the EU and CJEU got influenced by US copyright lawsuits.[310]

However, many might argue that copyright only prohibits the use of a specific work and not an artist's overall style, meaning that when an AI system utilises the style of an existing work from the input data, and the output is just an imitation, there should not be an infringement.[311] Furthermore, it is not uncommon for the output to require human reworking when it is finished. "The technology may thus be regarded, wholly or partly, as a tool or extension of human creativity".[312] However, similar to other copyright infringement cases, the AI-generated output must be assessed in each case to determine if it's a free use, not subject to copyright or if it is an adaptation.[313]

Another thing that needs to be considered regarding AI-generated output is when a work has been copied for parodic purposes, as this can fall under an exemption.[314] The parody exemption can, however, only be considered if the

[307] Williams v Gaye (n 306) 61-64.

[308] Juhasz Law (n 305).

[309] See Boyle (n 20) the article discusses differences between the Doctrine and EU copyright exceptions and explains that the doctrine covers the same areas such as legalising TDM for research purposes, the parody example, private copying and format shifting, educational exceptions, library archiving and even support for an "EU-wide change to introduce more general technology sensitive flexibilities into a copyright?

[310] See chapter 5.1.

[311] Axhamn (n 119) 70.

[312] Axhamn (n 119) 48.

[313] Axhamn (n 119).

[314] Article 17 (7) (b) DSM Directive.

audience can recognise the comedic version from its original source. This means that a significant enough amount must be appropriated to be recognised. In one case, a DJ was refused permission to use a song, but this did not stop the DJ from using parts of the song anyway and transforming it into a parody. The artist of the original work claimed that this was a copyright infringement. The court, on the other hand, deemed this as Fair Use.[315] Similarly, the CJEU has held that the parody must be distinctly different from the original works and contain expressions of humour or mockery, and it doesn't need to be attributed to someone other than the original works.[316] Unless there is a case in which an AI-generated output mimics an artist's sound and style, the parody exception might be less relevant, but it can be important to keep in mind for certain generated musical pieces that might display elements of humour or mockery.

Furthermore, Article 17(7) (b) DSM directive provides an exemption for pastiches, which could be especially interesting for AI-generated musical outputs as these are creations that imitate existing artists.[317] Unfortunately, there is no common understanding of the term so far, but it can be understood as a communicative artefact that borrows and adopts elements of other works.[318] This could potentially benefit AI-generated outputs if only the style is mimicked and not the underlying music work, the sound samples or perhaps even the musical layers.

5.2 The Voice

With the development of AI technologies, speech synthesis has become possible. This technology simulates the likeness of a person’s attributes, such as their sound and voice, which creates numerous opportunities to use synthetic voices in the music industry.[319]

The technology works by first being trained on audio recordings of a human voice, after which that voice can be replicated. Once the speech is developed, it can later be used to recite any text. This makes it possible for an artist to record a song without using their own voice or to perfect their own pitches.[320]

---

[315] Fisher v. Dees, 794 F.2d 432 (9th Cir. 1986) para 2.

[316] Johan Deckmyn and Vrijheidsfonds VZW v Helena Vandersteen and Others (C-201/13) [2014] ECLI:EU:C:2014:2132 (Deckmyn) para 33.

[317] See for instance Aisis a pastiche to Oasis <https://www.youtube.com/watch?v=TW9y570uuXw> accessed 18 May 2024.

[318] Till Kreutzer and Felix Reda, `The Pastiche in Copyright Law – Towards a European Right to Remix´ (*Kluwer Copyright Blog*, 13 March 2023) <https://copyrightblog.kluweriplaw.com/2023/03/13/the-pastiche-in-copyright-law-towards-a-european-right-to-remix/> accessed 14 May 2024.

[319] WIPO 2020 (n 88) 9; Cabrera Blázquez (n 24).

[320] Bobby Owsinski, ´Songwriters are using AI Generated voices for song pitches´, (Hypebot) <https://www.hypebot.com/hypebot/2023/08/songwriters-are-using-ai-generated-voices-for-song-pitches.html> accessed 23 April 2024; Cabrera Blázquez (n 24).

There is also a possibility for others to mimic an artist's voice, known as deepfakes.[321]

An example of this is the YouTube channel "Vocal Synthesis", which allows for the artificial creation of voice synthesis of an artist singing a song.[322] As a result of this, some record labels chose to issue copyright strikes to remove some of these audio deepfakes, but after initially removing these videos, YouTube later chose to reinstate them as the takedown requests were incomplete.[323] Similarly, in 2016, Mark Zuckerberg created a voice assistant using the voice of Morgan Freeman.[324] Despite the lack of consent from the actor, he cannot have a valid copyright claim for the misappropriation of his voice.[325] The logic behind this is that a voice is not protected by related rights, as this right is reserved for specific performance.[326]

In US copyright law, voice synthesis is considered transformative as no original work is sampled in any way because the vocal samples are taken out of their instrumental context and often used in a parodic manner.[327] Furthermore, a voice cannot be copyrightable as it is not fixed in a tangible medium.[328] This was later reaffirmed in a case where the singer Waits' raspy singing voice was imitated in a commercial; the court stated in this case that the use of a voice is a different kind than copyright cases challenging unauthorised use, and therefore, the misappropriation claim for his voice where not covered by federal copyright law.[329]

As mentioned, this applies in the US, and so far, there have been no lawsuits in the EU where the copyright of a voice has been challenged. However, it can be assumed that similar reasoning would be used in the EU, especially considering the Levola Hengelo lawsuit.

---

[321] Cabrera Blázquez (n 24) deepfakes can be used with both malicious intent and as a creative artistic expression. In this thesis I will solely focus on the copyright aspect of the voice synthesis.

[322] Cabrera Blázquez (n 24).

[323] Nick Statt, 'Jay-Z Gets YouTube to Remove Deepfake Videos Impersonating Him' (The Verge, 28 April 2020) <https://www.theverge.com/2020/4/28/21240488/jay-z-deepfakes-roc-nation-youtube-removed-ai-copyright-impersonation> Accessed 23 April 2024.

[324] Darrell Etherington, `Watch Mark Zuckerberg's Morgan Freeman-voiced Jarvis AI in action´ (Techcrunch, 20 December 2016) <https://techcrunch.com/2016/12/20/watch-mark-zuckerbergs-morgan-freeman-voiced-jarvis-ai-in-action/ > accessed 23 April 2024.

[325] Eastwood v. Superior Court, 149 Cal. App. 3d 409, 417, 198 Cal. Rptr. 342, 347 (1983) cited in Zack Naqvi, 'Artificial Intelligence, Copyright, and Copyright Infringement' (2020) 24 Marq Intell Prop L Rev 15**,** 43.

[326] Cabrera Blázquez (n 24).

[327] Andy Baio, 'With Questionable Copyright Claim, Jay-Z Orders Deepfake Audio Parodies off YouTube' (Waxy, 28 April 2020) <https://waxy.org/2020/04/jay-z-orders-deepfake-audio-parodies-off-youtube/> accessed 23 April 2024.

[328] Midler v. Ford Motor Co., 849 F.2d 460 (9th Cir. 1988).

[329] Waits v Frito-Lay, Inc., 978 F.2d 1093 (9th Cir. 1992).

Therefore, if AI can demonstrate that it is adopting an "accent" of an artist's voice rather than repeating a recorded sample, then the AI won't infringe on any copyright.[330] Despite this being a way to circumvent the performer's right.

Yet, the copyrightability of the voice can depend on the circumstances. If someone, for instance, uses their voice in a unique way that shows originality and expresses the author's own creative choices and variations in a recognisable way, this could potentially, in very special cases, be subject to copyright protection.

Regarding the criteria that copyright must be fixed in a tangible medium, it can be argued that previous recordings on phonograms can make the sound of the voice fixed in a tangible medium. Therefore, the possibility of copyrighting a voice should not be completely ruled out.

5.3 Consumer and Artist Protection

Since there is no clear case or legal framework regarding whether AI-generated output infringes on a copyright-protected work or a related right, some analogies could perhaps be drawn from trademark discussions.

An argument can be made that auditorial attribution should be protected as it could protect the customers,[331] who, in the case of AI-generated musical outputs, are the listeners of the musical output. The discussion on auditorial attributions can be compared to a trademark as the work and the author or artist can signal a certain quality and contribute to the author´s identity as an author.[332] The authorship indicators such as sound, style, and voice must provide the listener with some useful information; if not, it would be questionable why a claim of authorship even exists.[333] Some might even say that the AI generator has failed with its primary function if the generated output cannot be recognised.[334] Furthermore, AI-generated musical outputs often show little originality and musical value as the AI songs are built from short sound recordings and are, therefore, more qualitatively dense than the corresponding section in the original underlying musical work.[335]

In the proposal to the AI Act, it is described that when people interact with AI-generated audio systems, this information shall be disclosed so people can

[330] Naqvi (n 325) 44.

[331] Greg Lastowka, 'The Trademark Function of Authorship' (2005) 85 BU L Rev 1171, 1179-1180.

[332] Mark Rose, *Authors and Owners (*Harvard University Press, 1993) 1-2.

[333] Lastowka (n 345)1179-1180 . For other intellectual property rights, such as some patents, it is not necessary to know who the creators are. The author discusses the unimportance of knowing who the creators of the Post-it note are.

[334] Sunray (n 114) 211.

[335] ibid 208.

make informed choices as it for some AI-generated outputs can be difficult to identify whether the output was created by AI or a human.[336]

There can also be instances in which an AI mimics an artist and conveys a message that the artist or author doesn't support. In the New York Times v. OpenAI lawsuit, it has been emphasised that the GPT model has fabricated articles in the past, which can affect the newspaper´s goodwill as a referral site.[337] This should, therefore, affect the author or artist's moral rights.

From the aftermath of the song "Heart on My Sleeve," some artists have expressed concern that AI programs can mass-produce works in their style, thereby potentially undercutting the value of their work. The EWC has expressed that the past and present use of human-created works and labour is being misused to produce competitive outputs that will replace the works the generator was trained on.[338]

[336] Article 50.4 AI Act; Cabrera Blázquez (n 24).
[337] The New York Times v OpenAI (n 4).
[338] EWC (n 219).

6 Discussion

The rapid development of AI and AI-generated music has created legal uncertainty. The AI-generated song “Heart on My Sleeve,” which has recently sparked much debate, is an example of this.

AI can contribute with competitive advantages for companies and create socially beneficial outcomes in media and culture, this suggests that generative networks might have a big effect on future music consumption. The tension has begun to raise many lawsuits in recent times of the potential copyright infringement of the AI-generated input and output, making private actors such as record take precautionary steps to prevent this further.

However, it can be tricky to balance protecting and rewarding creativity while still ensuring that AI development respects existing copyright laws, which can restrict AI development and digital creativity.

Therefore, the thesis has explored the legal framework for copyright and neighbouring rights within the EU and whether or not these AI creations that imitate artists might infringe on an existing copyright, as well as compared potential solutions found in other jurisdictions.

6.1 Input

AI-generative technologies learn to produce new musical outputs by training on extensive datasets that often include content covered by copyright, neighbouring rights, or the sui generis database right. For the purpose of this thesis, we will focus solely on the protected works and exclude consideration of orphan works, Creative Commons licensed materials, and work in the public domain, although AI systems might be trained on these.

One key challenge when creative, generative systems create music is how existing protected works can be introduced into the training phase of the AI system to allow for the creation of the final output. When the system is trained, it looks for patterns in the music, such as chords, tempo, and related notes. The work is, however, also reproduced in several steps during the training phase. This reproduction is an exclusive right to the rightsholder, and the CJEU has concluded that even part of a reproduction shall be considered a reproduction. The exemption for transient copies in article 5 (1) InfoSoc directive has been explored to see if this can justify the reproductions created in the training phase. However, this exemption can be disregarded as the transient requirement, and the requirement of limited profit beyond the lawful use cannot be met.

Based on the analysis above and the AI Acts' direct reference to the DSM directive and TDM activities, it can be concluded that these will be the most

beneficial for developers of generative AI systems. The exemption in Article 3 DSM directive can be used for reproductions depending on the beneficiary performing the TDM and what is done with the collected material. For mining activities to be allowed under Article 3, a combination of when, whom, and why the data was extracted will affect whether the mining activity can be lawful or not. However, the lack of transparency from AI companies regarding their use of copyright-protected materials raises concerns about the requirement of lawfulness being fulfilled. Furthermore, the article leaves room for interpretation, particularly when distinguishing between non-profit and commercial actors. It should be noted that the exceptions should be interpreted strictly.

Since many AI systems that create generative music are commercial, they are excluded from Article 3 and must instead rely on the general provision in Article 4. This provision is not limited to certain beneficiaries but doesn't have the same mandatory element as Article 3, as the rightsholders can opt-out. This is particularly problematic for the music industry, and it can, therefore, be argued that this provision is somewhat back-handed. The sorting and separation of mined content can further create problems for AI developers due to the vast volume of data and practical issues for machines to read these reservations. Moreover, these two provisions contain unclarities that will need to be addressed by the CJEU.

Furthermore, the provisions don't allow for adaptations of the content that has been mined, making it unlawful, especially for generative models. The AI Act's silence on this issue leaves AI developers unable to find any new leverage to train their systems.

The better option for the AI industry would perhaps be to acquire licences or establish a revenue-sharing system, as the economic rights can be transferred. However, this, too, poses challenges since it can often be difficult to acquire permission from all involved rightsholders, especially for musical works that often have multiple authors for the overall composition as well as phonogram producers. Identifying the rightsholders becomes even more challenging if the works are extracted from databases.

While this approach leads to increased cost, it would result in a long-lasting system that benefits and incentivises creativity as well as benefits technical development. Previous revolutionary technology in the music industry, such as Spotify´s online streaming platform, acquired licenses through record labels. This could be a solution for building generative models in a sustainable manner without forcing AI companies to enter into an agreement with each individual party.

The fact that the AI Act has addressed increased transparency from AI developers might indicate that the reproductions in the training phase as it is

conducted so far, are not permitted. Considering that Stability AI wanted to move their trial to the US further indicated that the US legal exemption is more favourable to AI input than the UK and EU.

TDM is important for the development of technology, and a lot of research is financed by private actors. Therefore, An option could be to follow Japan's lead, which could benefit AI development but prevent creative AI as the model usage and prompt will most likely fall outside the exemption. This can then benefit both actors and prevent narrow and biased training sets. This can be in line with Recital 6 DSM directive as well as the three-step test. Extending Article 3 to involve commercial actors faced lots of backlash in the UK due to the harm creative industries and their investments would face. An open-ended provision like the fair use doctrine sounds good with its 4 pillars that asses the purpose, the nature, the amount used, and the effect the use has on the copyright-protected material, hence resembling the three-step test in the Berne convention to some extent. However, this will create uncertainties and could instead make the somewhat harmonised legislation in the EU drift apart.

However, if the EU is too strict against developers, developers may leave the EU. The AI Act has highlighted that anyone operating on an EU market shall comply with EU law, which means respecting copyright in the training phase; this can force AI developers to adjust or simply remove their tech from the EU market. On the other side of the coin, protecting and preserving artists and authors' rights within the Union can instead foster creativity and preserve cultural heritage unless it is too strict, which in turn can prevent musicians from using creative tools to develop their own music and prevent innovation that might lead to new musical creations and styles via AI. It all comes down to what approach the EU want to take when balancing these interests.

6.2 Output

The AI-generated output and whether this could be a copyright infringement pose different challenges than the input analysis, as the line that separates original independent work from derivative works can be hazy. This will make it harder for the authors and neighbouring rights holders to prove a reproduction as copyright intends to protect original expressions and not the idea itself. It is more likely in most instances that the AI-generated output instead mimics the sound recording for details to get insight into the style and how the song is composed.

Therefore, each case must be assessed independently to examine how the content used in the input is reflected in the final output. This is not cost-efficient, but a general rule is that less input material collected from a narrower source is more likely to infringe on someone's copyright or related rights.

It all comes down to when a work is considered transformative enough. Issues regarding inspiration can transpire. This can lead to further questions regarding whether AI should be assessed in a different way in the infringement assessment compared to human creators, as AI can easily recreate works and remember larger quantities of content.

Overall, copyright in the music industry is a complicated field as it, in theory, contains unlimited options. However, in reality, some musical genres are more restrictive. As the law is not yet in place, an option for EU courts could be to follow the sound-layer argument found in the Bridgeport lawsuit from the USA. This would not be the first time American courts have inspired the European legal system in difficult copyright cases.

In some cases, it could be possible to use existing precedents on sound sampling for samples that are recognisable to the ear. However, these sampling lawsuits can also create difficult questions, such as to whom the sample should be unrecognisable and whether the generated part can be covered by the quotation exemption, parody exemption, or pastiches.

There are currently no cases in the EU regarding pastiches that are especially interesting for musical outputs; however, it can be assumed that they, just like parodies, should be recognised and communicated to the original work or rightsholders. Furthermore, the AI Act should need to clarify exemptions such as parody, caricature, or pastiche to prevent future uncertainties.

The voice itself can, in most cases, not be covered by copyright despite this being a way to circumvent perform rights and can, from a consumer protection point of view, mislead the listeners as well as harm the personal rights of the author and destroy the rightsholders moral reputation as the artist style, sound and name can be associated with a certain artistic style and quality, as well as financially harm the rightsholders by having output directly competing with them.

It is, therefore, important for the EU to clarify inconsistencies and its position regarding the output, as AI can harm consumers and current and future rightsholders financially, and by generating content with poorer quality. Art should be something inspired by life and part of a creative process. However, AI also has the possibility to create something new and unique in terms of music and style, as well as increase access to music.

7 Conclusion

The thesis has been looking at AI-generated music that mimics the sound and style of an artist and whether or not this is a copyright infringement in either the input or the output stage of the process.

The thesis concludes that if an AI generator were to adapt work in its initial training phase, this would be unlawful; reproductions of work that take place in the training phase will also, in many circumstances, be considered to infringe on existing copyright despite the current TDM exceptions, which aren't broad enough to cover many instances of AI training in reality, especially regarding AI-generated music. This means that they often infringe on the rightsholder's exclusive rights.

AI-generated output, on the other hand, is less likely to infringe as sound and style are generally not protected subject matter, and the same goes for the synthesised voice. However, even if the output is lawful, the process can be considered unlawful because the input, in many circumstances, will not meet the criteria to fall under an exemption.

It can furthermore be concluded that there currently exist some grey areas in the legal framework, where the combination of the rapid evolution of AI and the slow-moving legal system can create questions on how this shall be regulated and the future needs of all parties involved.

Therefore, the overall recommendation is that further research is needed to explore the existing copyright exceptions and limitations and determine whether a new legal framework is needed to address these challenges. These issues will be increasingly relevant and require careful consideration by lawmakers, artists, and developers.

Bibliography

Books and Articles

Anantrasirichai N and Bull D, 'Artificial Intelligence in the Creative Industries: A Review' (2022) 55 Artif Intell Rev 589.

Axhamn J, 'Artificial Intelligence and the Need for Legislation' in Morten Rosenmeier, Thomas Riis, Jens Schovsbo, and Henrik Udsen (eds), *Festskrift till Jørgen Blomqvist* (Lund University 2021).

Boyle J, '(When) is Copyright Reform Possible?' in Ruth Okediji (ed), *Copyright Law in an Age of Limitation and Exceptions* (Cambridge University Press 2017).

Cabrera Blázquez FJ, Cappello M, Fontaine G, and Valais S, *Exceptions and Limitations to Copyright*, IRIS Plus, European Audiovisual Observatory, Strasbourg, 2017.

Calo R, 'Artificial Intelligence Policy: A Primer and Roadmap' (2017) 51 UC Davis Law Review 399, 404

Chiou T, 'Copyright Lessons on Machine Learning: What Impact on Algorithmic Art?' (2019) 10(3) JIPITEC 398.

Dhariwal P and others, 'Jukebox: A Generative Model for Music' (2020) arXiv 1 <https://arxiv.org/pdf/2005.00341.pdf> accessed 2 February 2024.

Drott E, ´Copyright, Compensation, and Commons in the Music AI Industry´ (2020) 14(2) Creat Ind J 190.

Dermawan A, 'Text and Data Mining Exceptions in the Development of Generative AI Models: What the EU Member States Could Learn from the Japanese "Nonenjoyment" Purposes?' (2024) 27 The Journal of World Intellectual Property 44–68.

Deruty E, 'On the Development and Practice of AI Technology for Contemporary Popular Music Production' (2022) 5(1) Transactions of the International Society for Music Information Retrieval 35,

D'Agostino G, 'Healing Fair Dealing - A Comparative Copyright Analysis of Canada's Fair Dealing to U.K. Fair Dealing and U.S. Fair Use' (2008) 53 McGill L J 309.

Ellis J, General Principles and Comparative Law, European Journal of International Law, Volume 22, Issue 4, November 2011, Pages 949–971,

Gervais DJ, 'AI Derivatives: The Application to the Derivative Work Right to Literary and Artistic Productions of AI Machines' (2022) 53 Seton Hall Law Review 8.

Goldmedia, 'AI and Music: Impacts of Artificial Intelligence on the Music Industry' (GEMA, SACEM 2023).

Guadamuz A, 'A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs' (2024) 73(2) GRUR International 111.

Hacker P, 'A Legal Framework for AI Training Data—from First Principles to the Artificial Intelligence Act' (2021) 13(2) Law, Innovation and Technology 257.

Jesse J, *The Music Copyright Manual: The Definitive Guide to Music Copyright Law in the Digital Age* (Rock N Roll Law 2016).

Kasahara C, 'Getting the Deal Through: Copyright 2012 - Japan' (2012) <https://www.aplawjapan.com/archives/pdf/file/GettingtheDealThrough-Copyright2012Japan.pdf> accessed 21 May 2024.

Kempas T, *Artificiell Intelligens och immaterialrätt i Sverige och EU* (Norstedts juridik 2023).

Kleineman J, 'Rättsgodmatisk metod' in Maria Nääv and Mauro Zamboni (eds), *Juridisk Metodlära* (MTM 2018)

Kur A, Dreier T, and Luginbuehl S, *European Intellectual Property Law: Text, Cases and Materials* (2nd edn, MTM 2021).

Lastowka G, 'The Trademark Function of Authorship' (2005) 85 BU L Rev 1171.

Margoni T, 'Artificial Intelligence, Machine Learning and EU Copyright Law: Who Owns AI?' (2018) CREATe Working Paper <https://zenodo.org/records/2001763#.YakLB1VBzIU> accessed 9 April 2024.

Naqvi Z, 'Artificial Intelligence, Copyright, and Copyright Infringement' (2020) 24 Marq Intell Prop L Rev 15, 43.

Minemura T, 'Dogmatic Legal Science and Sociology of Law' (1970) 56(3) ARSP 351.

Rosati E, '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(2) Asia Pac Law Rev 198.

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

Rose M, *Authors and Owners* (Harvard University Press 1993)

Smits J, 'What is Legal Doctrine? On the Aims and Methods of Legal-Dogmatic Research' in R van Gestel, H Micklitz, and EL Rubin (eds), *Rethinking Legal Scholarship: A Transatlantic Dialogue* (Cambridge University Press, 2017).

Sturm BL and others, 'Artificial Intelligence and Music: Open Questions of Copyright Law and Engineering Praxis' (2019) 8(3) Arts 1.

Sunray E, 'Sounds of Science: Copyright Infringement in AI Music Generator Outputs' (2021) 29 Cath U J L & Tech 185.

Wright E, 'Professionals Beware: The Opportunities and Risks of Generative AI in Legal Practice' (University of Technology Sydney) <https://opus.lib.uts.edu.au/bitstream/10453/176325/3/Professionals%20Beware%20The%20Opportunities%20and%20Risks%20of%20Generative%20AI%20in%20Legal%20Practice.pdf> accessed 14 April 2024.

Zirpoli CT, *Generative Artificial Intelligence and Copyright Law* (Congressional Research Service Legal Sidebar LSB10922, 2023).

Websites and Blogs

Amazon, 'Amazon Music Terms of Use' (Amazon, 1 May 2023) <https://www.amazon.com/gp/help/customer/display.html?nodeId=201380010> accessed 12 April 2024.

AVIA, 'AI Generated Songs in Seconds' (AVIA) <https://www.aiva.ai/> accessed 8 March 2024.

Baio A, 'With Questionable Copyright Claim, Jay-Z Orders Deepfake Audio Parodies off YouTube' (*Waxy*, 28 April 2020) <https://waxy.org/2020/04/jay-z-orders-deepfake-audio-parodies-off-youtube/> accessed 23 April 2024.

Beck B, 'Court of Justice of the EU Rules on Copyright Implications of Sampling in Music' (MayerBrown, 5 September 2019)

<https://www.mayerbrown.com/en/perspectives-events/blogs/2019/09/court-of-justice-of-the-eu-rules-on-copyright-implications-of-sampling-in-music> accessed 23 January 2024.

Bommasani R, Klyman K, Zhang D, and Liang P, 'Understanding the EU AI Act: A Comprehensive Overview' (Center for Research on Foundation Models, 15 June 2023) <https://crfm.stanford.edu/2023/06/15/eu-ai-act.html> accessed 11 April 2024.

Cambridge University Press, 'Data Scraping' (Cambridge Dictionary) <https://dictionary.cambridge.org/dictionary/english/data-scraping> accessed 13 March 2024.

Copyright.eu, 'Protection of Music and Songs' (Copyright.eu, 22 August 2022) <https://www.copyright.eu/docs/protection-of-music-and-songs/> accessed 8 May 2024.

Creative Commons, 'About CC Licenses' (Creative Commons, 2019) <https://creativecommons.org/share-your-work/cclicenses/> accessed 18 May 2024.

Donegan J, 'The US should look at Japan's unique approach to generative AI copyright law' (ManageEngine, 24 February 2024) <https://insights.manageengine.com/artificial-intelligence/the-us-should-look-at-japans-unique-approach-to-generative-ai-copyright-law/> accessed 18 May 2024.

Etherington D, 'Watch Mark Zuckerberg's Morgan Freeman-voiced Jarvis AI in Action' (*TechCrunch*, 20 December 2016) <https://techcrunch.com/2016/12/20/watch-mark-zuckerbergs-morgan-freeman-voiced-jarvis-ai-in-action/> accessed 23 April 2024.

European Commission, 'Copyright' (European IP Helpdesk) <https://intellectual-property-helpdesk.ec.europa.eu/ip-management-and-resources/copyright_en> accessed 27 February 2024.

Gatto JG, 'Getty Image's AI Model Training Lawsuit in UK Against Stability to Proceed' (*Art Law Gallery Blog*, 7 December 2024) <https://www.artlaw-gallery.com/2023/12/articles/artificial-intelligence/getty-images-ai-model-training-lawsuit-in-uk-against-stability-to-proceed/> accessed 18 May 2024.

Geiger C, "Elaborating a Human Rights friendly Copyright Framework for Generative AI" (2024). Joint PIJIP/TLS Research Paper Series. 123.

Geiger C, Frosio G and Bulayenko O, "Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? Legal Analysis and Policy Recommendations, IIC (2018) 49: 814.

Gerrish c, Skavlan AM, ‘European Copyright Law and Text and Data Mining Exceptions and Limitations: In Light of the Recent DSM Directive, is the EU Approach a Hindrance or Facilitator to Innovation in the Region?’ (2019) 2(2) SIPLR 58.

Glen T, 'Daddy's Car: A Song Composed by Artificial Intelligence Created to Sound Like The Beatles' (*Laughing Squid*, 22 September 2016) <https://www.flow-machines.com/history/press/daddys-car-song-composed-artificial-intelligence-created-sound-like-beatles/> accessed 9 April 2024.

Juhasz Law, 'Blurred Lines: Copyright Infringement, Marvin Gaye, Patent Preemption, and a New Frontier in Copyright Litigation' (Patent Horizon, 28 June 2015) <https://www.patenthorizon.com/blurred-lines-copyright-infringement-marvin-gaye-patent-preemption-and-a-new-frontier-in-copyright-litigation/> accessed 20 April 2024.

Keller P, 'A First Look at the Copyright Relevant Parts in the Final AI Act Compromise' (*Kluwer Copyright Blog*) <https://copyrightblog.kluweriplaw.com/2023/12/11/a-first-look-at-the-copyright-relevant-parts-in-the-final-ai-act-compromise/> accessed 27 March 2024.

Kreutzer T and Reda F, 'The Pastiche in Copyright Law – Towards a European Right to Remix' (*Kluwer Copyright Blog*, 13 March 2023) <https://copyrightblog.kluweriplaw.com/2023/03/13/the-pastiche-in-copyright-law-towards-a-european-right-to-remix/> accessed 14 May 2024.

Martineau K, 'What is Generative AI' (*IBM Blog*, 20 April 2023) <https://research.ibm.com/blog/what-is-generative-AI> accessed 13 March 2024.

Maybe Definitely, `Oasis vs. Aisi: Can Ai Recreate the Oasis Sound?´ (youtube, 8 May 2023) <https://www.youtube.com/watch?v=TW9y570uuXw > accessed 18 May 2023

Montagnon R and Cho S, 'UK Withdraws Plans for Broader Text and Data Mining (TDM) Copyright and Database Right Exception' (Herbert Smith Freehills, 1 March 2023) <https://hsfnotes.com/ip/2023/03/01/uk-withdraws-plans-for-broader-text-and-data-mining-tdm-copyright-and-database-right-exception/> accessed 13 May 2023.

Meta, 'Introducing AudioCraft: A Generative AI Tool For Audio and Music' (Meta, 2 August 2023) <https://about.fb.com/news/2023/08/audiocraft-generative-ai-for-music-and-audio/> accessed 9 April 2024.

NVIDIA, 'What is Generative AI' (NVIDIA) <https://www.nvidia.com/en-us/glossary/generative-ai/> accessed 13 March 2024.

Nicholson CV, 'A Beginner´s Guid to Generative AI' (Pathmind Wiki) <https://wiki.pathmind.com/generative-adversarial-network-gan> accessed 13 March 2024.

Owsinski B, 'Songwriters Are Using AI Generated Voices for Song Pitches' (*Hypebot*) <https://www.hypebot.com/hypebot/2023/08/songwriters-are-using-ai-generated-voices-for-song-pitches.html> accessed 23 April 2024.

Reed R, 'AI Created a Song Mimicking the Work of DRAKE and The Weeknd. What does that Mean for Copyright Law?' (Harvard Law School, 2 May 2023) https://hls.harvard.edu/today/ai-created-a-song-mimicking-the-work-of-drake-and-the-weeknd-what-does-that-mean-for-copyright-law/ accessed 3 May 2024.

Robertson S, 'How Napster Transformed a Copyright Shutdown to a $70m Price Tag' (Metis Partners) <https://metispartners.com/thought-leadership/napster/> accessed 18 May 2024.

Soundcharts Team, '6 Basics of Music Copyright Law: What It Protects and How to Copyright a Song' (*Soundcharts Blog*, 1 January 2024) <https://soundcharts.com/blog/music-copyrights> accessed 17 April 2024.

Soundraw, 'Pricing' (Soundraw) <https://soundraw.io/#pricing-section> accessed 11 May 2024.

University of Edinburgh, 'Copyright Exceptions and Fair Dealing' (University of Edinburgh, 15 June 2021) <https://www.ed.ac.uk/information-services/library-museum-gallery/library-help/copyright/copyright-exceptions-and-fair-dealing> accessed 18 May 2024.

WIPO, 'WIPO-Administered Treaties' (WIPO) <https://www.wipo.int/wipolex/en/treaties/ShowResults?search_what=C&treaty_id=15> accessed 17 May 2024.

'WIPO-Administered Treaties' (WIPO) <https://www.wipo.int/wipolex/en/treaties/ShowResults?search_what=C&treaty_id=17> accessed 17 May 2024.

Cases

EU Cases

Infopaq International A/S v Danske Dagblades Forening (C-5/08) [2009] ECR I-6569.

Eva-Maria Painer v Standard VerlagsGmbH and Others (C-145/10) [2011] ECR I-12533.

Football Association Premier League Ltd and Others v QC Leisure and Others (C-403/08) and Karen Murphy v Media Protection Services Ltd (C-429/08) [2011] ECR I-9083.

Levola Hengelo BV v Smilde Foods BV (C-310/17) [2018] ECLI:EU:C:2018:899.

Brompton Bicycle v Chedech/Get2Get (C-833/18) [2020] ECLI:EU:C:2020:461.

C More Entertainment AB v Linus Sandberg (C-279/13) [2015] ECLI:EU:C:2015:199.

AustroMechana Gesellschaft zur Wahrnehmung mechanisch-musikalischer Urheberrechte GmbH v Strato AG (C-433/20) [2022] ECDR 10.

Stichting Brein v Jack Frederik Wullems (Filmspeler) (C-527/15) [2017] ECLI:EU:C:2017:300.

Infopaq International A/S v Danske Dagblades Forening (Infopaq II) (C-302/10) [2012] ECLI:EU:C:2012:16.

ITV Broadcasting Ltd and Others v TVCatchup Ltd (C-607/11) [2013] ECLI:EU:C:2013:147.

Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd (C-360/13) [2014] ECLI:EU:C:2014:1195, para 43.

Johan Deckmyn and Vrijheidsfonds VZW v Helena Vandersteen and Others (C-201/13) [2014] ECLI:EU:C:2014:2132.

Pelham and Others v Hütter and Others (C-476/17) [2019] ECLI:EU:C:2019:624.

VG Bild-Kunst v Stiftung Preußischer Kulturbesitz (C-392/19) [2021] ECLI:EU:C:2021:181.

US Cases

A&M Records, Inc. v Napster, Inc., 239 F.3d 1004 (9th Cir. 2001).

Metallica v Napster, Inc., No 2:00-cv-03914 (CD Cal, filed 2000).

Midler v Ford Motor Co, 849 F.2d 460 (9th Cir. 1988).

Authors Guild, Inc. v Google Inc., 804 F.3d 202 (2d Cir. 2015).

Andy Warhol Foundation for the Visual Arts, Inc. v Goldsmith et al., No 21–869 (US Supreme Court, 18 May 2023).

Williams v Gaye, 895 F.3d 1106 (9th Cir. 2018).

Waits v Frito-Lay, Inc., 978 F.2d 1093 (9th Cir. 1992).

Bridgeport Music, Inc. v Dimension Films, 410 F.3d 792 (6th Cir. 2005).

VMG Salsoul, LLC v Ciccone, 824 F.3d 871 (9th Cir. 2016).

Fisher v Dees, 794 F.2d 432 (9th Cir. 1986).

Thomson Reuters Enterprise Centre GmbH and West Publishing Corporation v Ross Intelligence Inc, Case No 1:20-cv-00613 (D Del, filed May 2020).

The New York Times Company v Microsoft Corporation and OpenAI, Inc., Case No 1:23-cv-11195 (United States District Court Southern District of New York, filed 27 December 2023).

Sarah Andersen, Kelly McKernan, and Karla Ortiz v Stability AI Ltd, Stability AI, Inc., DeviantArt, Inc., and Midjourney, Inc., Case No 3:23-cv-00201 (ND Cal, filed 13 January 2023).

UK Cases

Getty Images (US) Inc and others v Stability AI Ltd [2023] IL-2023-000007 (preliminary ruling, Mrs Justice Joanna Smith, 1 December 2023).

Newspaper Articles

Bar N, 'Striking at the heart of Hollywood' (*Los Angeles Times*, 10 April 2023) <https://www.latimes.com/entertainment-arts/business/story/2023-04-10/wga-writers-strike-story-collection> accessed 4 May 2024.

Chow AR, 'AI's Influence on Music Is Raising Some Difficult Questions' (*TIME*, 4 December 2023) <https://time.com/6340294/ai-transform-music-2023/> accessed 16 May 2024.

Fulker R, 'Kraftwerk case: EU top court defines sampling' (*DW*, 30 July 2019) <https://www.dw.com/en/kraftwerk-case-eu-top-court-defines-rules-of-sampling/a-46916334> accessed 23 January 2024.

Ivanova I, 'Artists Sue AI Company for Billions, Alleging "Parasite" App Used Their Work for Free' (*CBS News*, 20 January 2023) <https://www.cbsnews.com/news/ai-stable-diffusion-stability-ai-lawsuit-artists-sue-image-generators/> accessed 18 January 2024.

Krietzberg I, 'Copyright expert predicts result of NY Times lawsuit against Microsoft, OpenAI' (*TheStreet*, 7 January 2024) <https://www.thestreet.com/technology/copyright-expert-predicts-result-of-ny-times-lawsuit> accessed 14 April 2024.

Livni E, Hirsch L, and Kessler S, 'Who Owns a Song Created by A.I.?' (*The New York Times*, 17 April 2023) <https://www.nytimes.com/2023/04/15/business/dealbook/artificial-intelligence-copyright.html> accessed 18 January 2024.

Marshall A, 'From Jingles to Pop Hits, A.I. Is Music to Some Ears' (*The New York Times*, 23 January 2017) <https://www.nytimes.com/2017/01/22/arts/music/jukedeck-artificial-intelligence-songwriting.html> accessed 17 May 2024.

Metz C and Wade R, 'OpenAI Seeks to Dismiss Parts of The New York Times Lawsuit' (*The New York Times*, 28 February 2024) <https://www.nytimes.com/2024/02/27/technology/openai-new-york-times-lawsuit.html> accessed 15 April 2024.

Price J, 'The Definitive Guide to Spotify Royalties' (*Medium*, 17 February 2020) <https://medium.com/@JPriceOfMusic/the-definitive-guide-to-spotify-royalties-dc5960862c00> accessed 14 May 2024.

Shackell C, 'Will AI kill our creativity? It could – if we don't start to value and protect the traits that make us human' (*The Conversation*, 27 September 2023) https://theconversation.com/will-ai-kill-our-creativity-it-could-if-we-dont-start-to-value-and-protect-the-traits-that-make-us-human-214149 accessed 4 May 2024.

Sky News, 'Bitter Sweet Symphony: Rolling Stones return royalties and rights to Richard Ashcroft after 22-year row' (*Sky News*, 24 May 2019) <https://news.sky.com/story/bitter-sweet-symphony-rolling-stones-return-

royalties-and-rights-to-richard-ashcroft-after-22-year-row-11727342#:~:text=The%20song%20itself%20was%20a,nominated%20for%20a%20Grammy%20award.&text=But%20after%20almost%2022%20years,following%20a%20%22magnanimous%20gesture%22> accessed 14 May 2024.

Sisario B, 'The `Blurred Lines´ Case Scared Songwriters. But Its Time May Be Up' (*The New York Times*, 24 March 2020) <https://www.nytimes.com/2020/03/24/arts/music/blurred-lines-led-zeppelin-copyright.html> accessed 20 April 2024.

Statt N, 'Jay-Z Gets YouTube to Remove Deepfake Videos Impersonating Him' (*The Verge*, 28 April 2020) <https://www.theverge.com/2020/4/28/21240488/jay-z-deepfakes-roc-nation-youtube-removed-ai-copyright-impersonation> accessed 23 April 2024.

Wang AX, 'How Music Copyright Lawsuits Are Scaring Away New Hits' (*Rolling Stone*, 9 January 2020) <https://www.rollingstone.com/pro/features/music-copyright-lawsuits-chilling-effect-935310/> accessed 23 January 2024.

Legal Sources

AI-HLEG, 'A Definition of AI: Main Capabilities and Disciplines' (European Commission, 8 April 2019) <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines> accessed 4 March 2024.

Berne Convention for the Protection of Literary and Artistic Works (adopted 14 July 1967, entered into force 29 January 1970) 282 UNTS 221 (Berne Convention).

Communications and Digital Committee, *At Risk: Our Creative Future* (2nd Report of Session 2022-23, HL Paper 125, 17 January 2023) ch 2.

Copyright, Designs and Patents Act 1988.

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 (DSM Directive).

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 (InfoSoc Directive).

Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (Copyright Term Directive).

Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Software Directive).

Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works (Orphan Works Directive).

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

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases (Database Directive).

European Commission, Directorate-General for the Internal Market and Services, Triaille J-P, d'Argenteuil JdM & de Francquen A, *Study on the Legal Framework of Text and Data Mining (TDM)* (Publications Office 2014).

European Parliament, 'Artificial Intelligence Act Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))' P9_TA(2023)0236.

European Commission, 'The EU Copyright Legislation' (European Commission, 11 March 2024) <https://digital-strategy.ec.europa.eu/en/policies/copyright-legislation> accessed 17 May 2024.

European Parliament, 'Artificial Intelligence Act Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))' P9_TA(2023)0236.

International Convention for the Protection of Performers, Producers of Phonograms and Broadcasting Organisations (adopted 26 October 1961, entered into force 18 May 1961) 496 UNTS 43 (Rome Convention).

Japan Patent Office, 'Copyright Law of Japan' (2008) <https://www.jpo.go.jp/e/news/kokusai/developing/training/textbook/document/index/Copyright_Law.pdf> accessed 21 May 2024

Patent and Trademark Office, '*Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation'* (Docket No. PTO–C–2019–0038).

UK Intellectual Property Office, Exceptions to Copyright: Guidance for Consumers (2018) https://assets.publishing.service.gov.uk/media/5a80f292ed915d74e6231597/Exceptions_to_copyright_-_Guidance_for_consumers.pdf accessed 13 May 2024.

USPTO, '*Public Views on Artificial Intelligence and Intellectual Property Policy'* (USPTO, 2020).

WIPO, *'Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence'* (WIPO/IP/AI/2/GE/20/1 Rev, 21 May 2020) prepared by the WIPO Secretariat 3.

WIPO, *'Understanding Copyright and Related Rights'* (2nd edn, WIPO 2016).

World Intellectual Property Organization (WIPO), TRT/WPPT/002, *Agreed Statements Concerning WIPO Performances and Phonograms Treaty*, adopted by the Diplomatic Conference on December 20, 1996 <https://www.wipo.int/wipolex/en/text/295

Other sources

European Writers' Council (EWC), 'EWC Second Statement on the AI Act Proposal' (EWC, 26 July 2023) <https://europeanwriterscouncil.eu/23ewc_on_aiact/> accessed 28 March 2024.

Written evidence from UK Music (CRF0037).