AI and the audiovisual sector: navigating the current legal landscape

IRIS

A publication
of the European Audiovisual Observatory

Observatoire européen de l'audiovisuel
European Audiovisual Observatory
Europäische Audiovisuelle Informationsstelle

COUNCIL OF EUROPE
75
1949 - 2024
CONSEIL DE L'EUROPE

IRIS
AI and the audiovisual sector: navigating the current legal landscape
European Audiovisual Observatory, Strasbourg, 2024
ISSN 2079-1062

Director of publication – Susanne Nikoltchev, Executive Director
Editorial supervision – Maja Cappello, Head of Department for Legal Information
European Audiovisual Observatory
Editorial team – Justine Radel-Cormann, Sophie Valais

Authors (in alphabetical order)
Malte Baumann, Judit Bayer, Mira Burri, Gianluca Campus, Mark Cole, Kelsey Farish, Philipp Hacker, Elodie Migliore, Jan Bernd Nordemann, Justine Radel-Corman, Sandra Schmitz-Berndt, Bart van der Sloot

Proofreading
Anthony Mills, Aurélie Courtinat, Udo Lücke

Translation
Julie Mamou, Erwin Rohwer, Nathalie Sturlèse

Editorial assistant – Sabine Bouajaja
Press and Public Relations – Alison Hindhaugh, alison.hindhaugh@coe.int
European Audiovisual Observatory

Publisher
European Audiovisual Observatory
76, allée de la Robertsau, 67000 Strasbourg, France
Tel.: +33 (0)3 90 21 60 00
Fax: +33 (0)3 90 21 60 19
iris.obs@coe.int
www.obs.coe.int

Cover layout – ALTRAN, France

Please quote this publication as
Cappello M. (ed.), *AI and the audiovisual sector: navigating the current legal landscape,* IRIS, European Audiovisual Observatory, Strasbourg, October 2024
© European Audiovisual Observatory (Council of Europe), Strasbourg, 2024

Opinions expressed in this publication are personal and do not necessarily represent the views of the Observatory, its members or the Council of Europe.

AI and the audiovisual sector: navigating the current legal landscape

Malte Baumann, Judit Bayer, Mira Burri, Gianluca Campus, Mark Cole, Kelsey Farish, Philipp Hacker, Elodie Migliore, Jan Bernd Nordemann, Justine Radel-Corman, Sandra Schmitz-Berndt, Bart van der Sloot

Observatoire européen de l'audiovisuel
European Audiovisual Observatory
Europäische Audiovisuelle Informationsstelle

COUNCIL OF EUROPE
75
1949 - 2024
CONSEIL DE L'EUROPE

Foreword

When asked whether AI-assisted creation is becoming a new genre of film-making, the project creators Pierre Zandrowicz and Matt Tierney answered, "It just gives us more paint brushes in our bucket and I [Matt Tiernay] mean essentially what we get to do is take every frame of the film and paint into it through text, through prompting."[1]

This statement illustrates the role of AI, particularly generative AI, as a tool that assists various professionals in enhancing their work. Generative AI can enable workers to become more multidisciplinary; for example, authors might use a generative AI tool to create visuals for promoting their work. However, the rise of generative AI in the audiovisual sector brings also new issues, such as job disruptions and copyright concerns, which decision-makers must address.

In response, the European Audiovisual Observatory (EAO) has reopened its 2020 AI file to explore the intersection of technological innovation and legislative frameworks. This report confronts some of AI's difficulties in the AV sector with existing regulations, asking whether they are future-proof and adaptable to evolving technological landscapes. Conceived, shaped and coordinated by the EAO's legal department, the report is divided into four parts.

Part 1 introduces readers to AI in the audiovisual sector, highlighting both its benefits and complications that the (quite) fragmented existing regulatory framework will have to tackle. **Chapter 1**, authored by Justine Radel-Cormann (EAO), sets the stage for this discussion.

The second part delves into legal questions surrounding AI and data feeding the machine. **Chapter 2**, by Philipp Hacker (Yale University), explores data protection and privacy implications, and the impact of regulations like the GDPR and AI Act. It also examines international data transfers and comparisons with US law. **Chapter 3**, by Gianluca Campus (PwC Digital Innovation), analyses the use of copyrighted works for AI training, the creation of derivative works, and the legal framework for using copyrighted data.

The third part addresses five key issues AI poses to the audiovisual sector. **Chapter 4**, by Malte Baumann and Jan Nordemann (law firm NORDEMANN, Berlin), discusses authorship, liability, and transparency in the generative AI era. **Chapter 5**, by Kelsey Farish (Reviewed & Cleared, London), considers the protection of actors' images, voices and personality rights against AI replication. **Chapter 6**, by Elodie Migliore (University of Strasbourg), examines AI's impact on labour law, referencing recent US strikes and legislative initiatives. **Chapter 7**, by Judit Bayer (University of Münster), investigates AI's role in disinformation and regulatory measures to combat it. **Chapter 8**, by Mira Burri (University of Lucerne), explores AI's impact on media pluralism and cultural diversity (e.g. content personalisation and bias) and possible regulatory measures to mitigate these effects and promote diverse content consumption.

---

[1] Helisek with Breezeway Productions interviewing the creators Pierre Zandrowicz and Matt Tierney at the 2023 Tribeca film festival. Their anime "In Search of Time" was presented at the Tribeca Festival 2023.

Part 4 looks to the future, evaluating whether recent AI regulations are ready for AI challenges brought to the AV sector. **Chapter 9**, by Mark Cole and Sandra Schmitz-Berndt (EMR), offers a forward-looking perspective on how future regulations can better address the evolving difficulties and opportunities brought about by AI, ensuring a balanced approach that fosters innovation while protecting the rights and interests of all stakeholders in the audiovisual industry. **Chapter 10**, by Bart van der Sloot (University of Tilburg), rounds off the publication discussing ethical dilemmas such as authenticity, the potential for AI to distort reality, and broader societal impacts of AI-generated content.

The introductory texts and concluding remarks, authored by Justine Radel-Cormann (EAO), aim to contextualise these diverse legal and policy issues.

I extend my warmest thanks to the brilliant authors who contributed to this rich report. To our readers, I will just say: enjoy the read!

Strasbourg, October 2024

Maja Cappello
IRIS Coordinator
Head of the Department for Legal Information
European Audiovisual Observatory

Table of contents

1. Artificial intelligence in the audiovisual sector 2

1.1. Defining “AI” and “audiovisual” 3

1.2. The transformative advantages of AI in the audiovisual industry 6

1.3. Examples of AI uses in the audiovisual industry 7

1.3.1. Case study 1: Claude, a conversational assistant to help with project development 8

1.3.2. Case study 2: DiversityCatch, measuring diversity in content 9

1.3.3. Case study 3: Midjourney and DALL.E, AI tools for creating images and videos 9

1.4. Challenges posed by AI in the audiovisual industry 10

1.4.1. What challenges lie ahead for the audiovisual sector? 10

1.4.2. The legislative framework surrounding AI: a complex puzzle 11

2. AI and Data Protection in Audiovisual Media 16

2.1. Introduction 16

2.2. Audiovisual material as personal data 17

2.3. Selected data protection and privacy concerns 18

2.3.1. Legal basis for training 19

2.3.2. Hallucinations 20

2.3.3. LLMs as personal data 21

2.3.4. Sensitive data 22

2.3.5. Information and user control 23

2.3.6. Automated decision making 23

2.4. The AI Act 24

2.5. International data transfers 25

2.6. Comparison with US and international law 26

3. AI & Copyright Protection when Feeding the Machine 28

3.1. Introduction 28

3.1.1. Overview of AI systems and their processing of copyrighted data 28

3.1.2. Considerations on derivative works 29

3.2. Text and data mining exception for training data 31

3.2.1. Examination of the applicability of TDM exemption to AI training data 31

3.2.2. TDM and the impact on reproduction and extraction rights 32

3.3. AI relevant legislations 34

3.3.1. EU AI Act and copyright: transparency rules and measures for TDM 34

3.3.2. AI and TDM exception: some national law proposals in Italy and Poland 35
3.4. Impact of case law 36
3.4.1. Overview of relevant cases on training data (USA and Europe) 36
3.5. Some (preliminary) conclusions on the case law 37

4. Authorship, Liability and Transparency in relation to AI-generated content .. 40
4.1. Authorship 40
4.1.1. The human creator as author 40
4.1.2. AI-assisted creation of works 42
4.1.3. Who is the author? 44
4.1.4. Protection through neighbouring rights 44
4.2. Liability for AI output 45
4.2.1. When is an infringement deemed to have occurred? 45
4.2.2. Exceptions and limitations to copyright applicable to AI output 47
4.2.3. Responsibility of the user 47
4.2.4. providers' terms of use 50
4.2.5. Reducing potential liability 50
4.2.6. Transparency 51

5. Personality Rights & Transparency 53
5.1. Setting the Scene 53
5.2. Commercial Drivers 55
5.2.1. The Evolution of Digital Doubles 55
5.2.2. Performers' Perspectives: Empowerment or Exploitation? 56
5.2.3. Regulatory Gap 57
5.3. Transparency in European Instruments 58
5.3.1. European AI Act 58
5.3.2. Transparency in the Framework Convention on AI 60
5.3.3. Different angles: The United States and the United Kingdom 61
5.4. Transparency as a keystone to uphold personality rights 63

6. Impact of AI on the audiovisual labour market in Europe 65
6.1. Introduction 65
6.2. Impacts of AI on labour law in the audiovisual sector in the US 66
6.2.1. The WGA and SAG- AFTRA strikes 66
6.2.2. The WGA agreement after the strike 67
6.2.3. The SAG-AFTRA agreement 67

6.3. Impacts of AI on labour law in the audiovisual sector in the EU 70
6.3.1. European Union policy 70
6.4. Analysis of the different initiatives of selected stakeholders 71
6.4.1. Collective management organisations (CMOs) 71
6.4.2. Associations and federations 73
6.4.3. Trade unions 74
6.5. Concluding remarks: remaining gaps and path forward 75

7. Disinformation and AI in the AV Sector 77
7.1. Defining Disinformation 77
7.2. . AI applications in the disinformation industry 78
7.2.1. Generative AI 78
7.2.2. Text and images 78
7.2.3. Deepfakes 79
7.2.4. Audio 80
7.2.5. Bots 81
7.2.6. AI and misinformation 82
7.3. The fight against disinformation 82
7.3.1. Regulation 82
7.3.2. Factchecking with the help of AI 85
7.4. Conclusion 88

8. Diversity and Pluralism 89
8.1. Setting the scene: AI as a disruptive technology 89
8.2. AI's impact on freedom of expression, media pluralism and cultural diversity 91
8.2.1. Introductory remarks: what's different? 91
8.2.2. Implications for content distribution and consumption 92
8.2.3. Implications for content creation 94
8.2.4. Additional aspects to consider 96
8.3. Addressing the real and potential effects of AI systems on pluralism and diversity 97
8.4. Concluding remarks 98

9. The world of tomorrow: are the texts AI-proof and ready for the AV challenges? 100
9.1. Recap: Existing and forthcoming regulatory approaches 100
9.1.1. Starting with Recommendations: Early approaches by the OECD and UNESCO 100
9.1.2. Moving towards binding law: Developments in the Council of Europe and the EU 101

9.2. Reality bites!? Applicability and limitations of regulatory approaches to the specifics of the AV sector 106
9.2.1. Data protection aspects 106
9.2.2. Intellectual property rights aspects 107
9.2.3. Personality rights aspects 109
9.2.4. Disinformation as an important challenge 111
9.2.5. The reach of the AI Act: geographical scope 112
9.3. Looking ahead: On "future-proofness" and global standards 113
9.3.1. Towards global and flexible risk-based standards in specific legislation 113
9.3.2. Considering sector-specific aspects for (AV) media 115

10. Ethical Dilemmas and Societal Challenges Raised by Generative AI 117

10.1. Introduction 117
10.2. Ethical foundations 118
10.3. How AI-generated content affects these concepts 119
10.4. Ethical dilemmas within the audiovisual sector 120
10.5. Societal challenges raised by AI 122
10.6. Conclusion 125

11. Concluding remarks 128

Figures

Figure 1. From regulatory semantic to key concepts and applications 3
Figure 2. Examples of AI applications across the audiovisual value chain 8
Figure 3. AI: example of a variety of legislations 13

Tables

Table 1. Definitions 4

PART I - Generative artificial intelligence and its potential to transform the audiovisual sector

Generative artificial intelligence (genAI) is the core of this new wave of frenetic regulatory activity. While discussions in the EU on adoption of the AI Act began in April 2021, they gained momentum following the release of open genAI software to the general public at the end of 2022. GenAI can generate new content, such as text, images, audio, videos, etc. based on sentences (prompts) provided by users in the genAI tool. The quality of the prompt influences the quality of the output.

The possibilities introduced by genAI are infinite, offering not only creative opportunities but also efficiency gains. In the audiovisual sector, AI could prove useful at various stages of the value chain. With genAI, there is a world of possibilities where roles may overlap, allowing individual creators to take on tasks beyond their traditional scope, fostering a more multidisciplinary approach. For instance, could an author create a music sketch for their script? Might the tasks of a scriptwriter intersect with those of an editor? Could these roles eventually merge?

Or, on the contrary, could this multi-disciplinarity be merely a myth, ultimately unhelpful to creators?

1. Artificial intelligence in the audiovisual sector

Justine Radel-Cormann, Legal Analyst, European Audiovisual Observatory

The audiovisual sector has long been at the front line of technological and digital advancements, continuously evolving to meet the changing needs and preferences of audiences. From the earliest cameras capturing silent black-and-white films to the modern era of ultra-high-definition streaming on portable devices, the industry has embraced innovation to enhance both content creation and distribution.

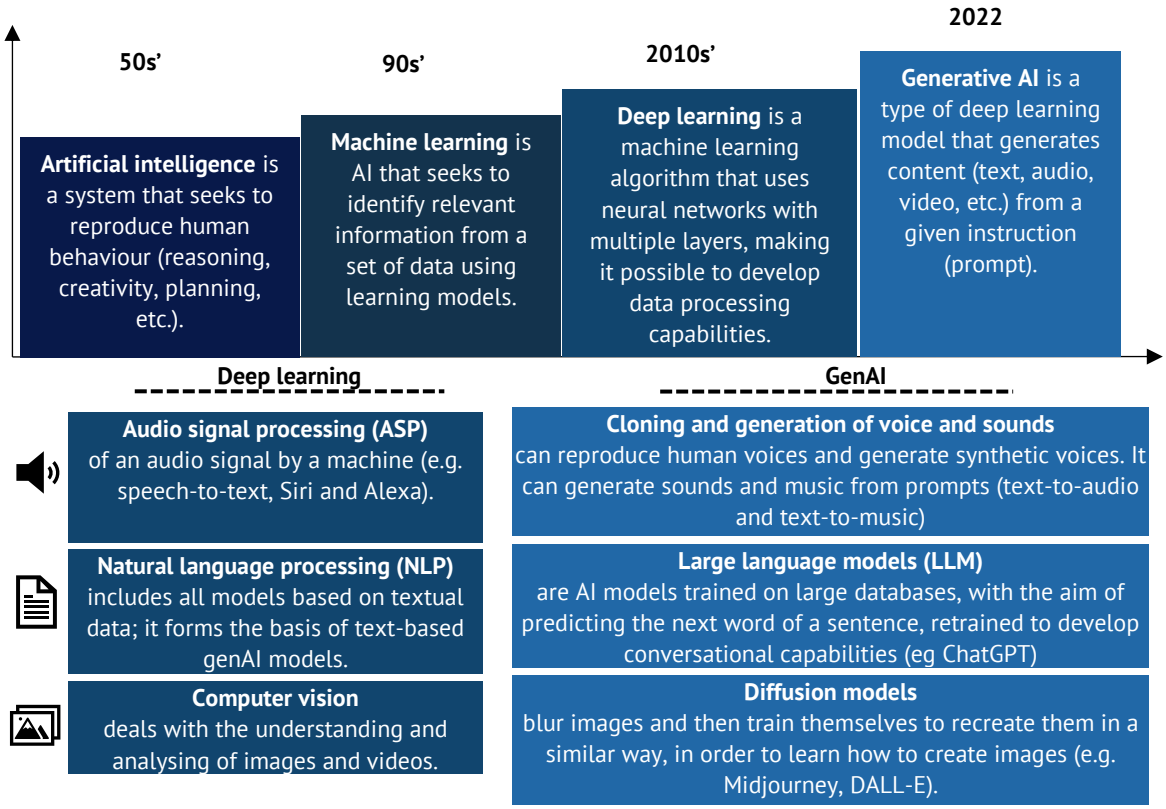
The most recent developments in the sector are the new functions that artificial intelligence (AI) is bringing: deep learning has been progressing for a decade, culminating with the recent rise of generative artificial intelligence (genAI). The potential applications it offers to the audiovisual sector have sparked both excitement and concerns.

Figure 1 below shows the evolution of AI, its different technologies, and applications. [2]

[2] This visual was created based on the IBM blog "AI, machine learning and deep learning: what's the difference?" and the CNC report "Quel impact de l'IA sur les filières du cinéma, de l'audiovisuel et du jeu vidéo," 8 April 2024

© European Audiovisual Observatory (Council of Europe) 2024

Figure 1. From regulatory semantic to key concepts and applications



Source: European Audiovisual Observatory

GenAI is reshaping the audiovisual industry, rapidly impacting everything from content creation to distribution, while the current regulatory landscape has to adapt to its fast-evolving AI nature. Section 1 will lay out the definitions of "audiovisual" and "AI" as understood throughout the report. Section 2 will explore what advantages AI could bring to the industry, with specific examples provided in Section 3. Finally, Section 4 will address the various challenges ahead and examine the existing legislative framework and its implications.

1.1. Defining "AI" and "audiovisual"

The term "audiovisual" essentially refers to all media except the printed press: Cinema, television, radio, video and the various on-demand services (such as video on demand or catch-up TV) are all sectors of the audiovisual industry. Additionally, when looking at the

© European Audiovisual Observatory (Council of Europe) 2024

value chain, we mean the various branches of the audiovisual industry such as film production, distributors, exhibitors, and public and private broadcasters.[3]

The notion of "AI" is more complex; there is no widespread consensus on a definition.[4] It is a broad phenomenon that different parties are trying to understand, and thus there are various definitions at international (OECD, Council of Europe), European Union, national (USA, China, and UK) and industry (OpenAI, MetaAI, Gemini) levels.

Table 1. Definitions

Text	Article	Quote
International texts		
OECD Council Recommendation[5]	Point 1	An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.
Council of Europe Framework Convention[6]	Article 2	An artificial intelligence system is a machine-based system that for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment.
European Union texts		
AI Act[7]	Article 3(1)	AI system means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.
Proposal for an AI liability Directive[8]	Article 2(1)	"AI system" means an AI system as defined in the AI Act.

[3] See Recital 23 of the Directive 2010/13/EU of 10 March 2010 on the coordination of certain provisions laid down by law (Audiovisual Media Services Directive): "For the purposes of this Directive, the term 'audiovisual' should refer to moving images with or without sound, thus including silent films but not covering audio transmission or radio services."

[4] "One of the biggest problems in regulating AI is agreeing on a definition," Carnegie Endowment for International Peace, 2022

[5] OECD Council recommendation on Artificial Intelligence, adopted on 22 May 2019, and amended on 3 May 2024

[6] Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, adopted on 17 May 2024 by the Committee of Ministers of the Council of Europe

[7] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act))

[8] Proposal for an AI liability Directive, proposed by the European Commission on 28 September 2022

© European Audiovisual Observatory (Council of Europe) 2024

National texts		
USA **The Department of State on AI**[9]	/	The term artificial intelligence means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments.
UK **AI Regulation White Paper**[10]	Point 3.2.1	The paper refrains from offering a fixed definition of AI due to its rapid evolution. Instead, it focuses on two defining characteristics necessitating regulatory attention: • Adaptivity: AI's ability to train on data and make inferences which may result in outcomes that are challenging to explain or predict. • Autonomy: Some AI systems can make decisions without the express intent or ongoing control of a human.
China **Proposal for the AI Law of the People's Republic of China**[11]	Article 94 (i)	AI means technology that utilises computers to simulate human intelligent behaviour for use in prediction, recommendation, decision-making, or content generation, etc. for specialised or general purposes.
Industry		
OpenAI[12]	Charter	OpenAI's Charter "OpenAI's mission is to ensure that artificial general intelligence (AGI)—by which we mean highly autonomous systems that outperform humans at most economically valuable work—benefits all of humanity."
Meta AI[13]	Meta AI page	Meta AI is an intelligent assistant that is capable of complex reasoning, following instructions, visualizing ideas, and solving nuanced problems.
Gemini Google[14]	Introducing Gemini	Gemini was built from the ground up to be multimodal, which means it can generalize and seamlessly understand, operate across and combine different types of information including text, code, audio, image and video.

Source: European Audiovisual Observatory

Based on these nine examples, it is clear that there are common criteria in the definitions set by the different entities (terminology, scope), but their focus areas may vary depending on the context and objectives of the entity proposing a definition:

[9] Quote on the website of the US Department of State on what AI is in 2020
[10] A pro-innovation approach to AI Regulation, presented to Parliament by the Secretary of State for Science, Innovation and Technology by Command of His Majesty on 29 March 2023
[11] Proposal for the AI Law of the People's Republic of China, preliminary document that has circulated among scholars, hosted on the website of the Centre for Security and Emerging Technology' (Georgetown University's Walsh School of Foreign Service) and translated into English
[12] OpenAI charter
[13] Meta AI service description
[14] Introducing Gemini, by Demis Hassabis, CEO and co-founder of Google DeepMind, December 2023

© European Audiovisual Observatory (Council of Europe) 2024

Common criteria:

- Objective-driven: AI systems are designed to achieve explicit or implicit objectives (e.g. making predictions, generating content, making decisions)
- Training the machine: AI systems receive input and generate output that can influence the physical or virtual environment
- Autonomy and adaptiveness: AI systems vary in their levels of autonomy and adaptiveness after their development; such variation implies the AI systems may evolve or learn from their interactions with data and the environment

Divergent elements:

- Terms: most refer to "AI systems", or "artificial intelligence" but a few refer to "machine-based system" or "technology"
- Uses and influences: the impact on decision-making is mentioned once, with some sources referring only to predictions, recommendations, and content generation
- Reference to humans: OpenAI refers to "general AI" capable of outperforming humans; the other sources give examples of capabilities (predictions, recommendations, and content generation)
- AI capabilities: definitions range from systems simulating human intelligence to those solving nuanced problems

While the definition and technical aspects of AI can be complex for non-scientists, AI's applications may be more intuitive to understand for non-experts: AI advantages become more apparent when contextualised within the audiovisual sector.

1.2. The transformative advantages of AI in the audiovisual industry

AI has the potential to positively impact the audiovisual industry along its entire value chain: (from the initial content concept to production, distribution, and protection) by assisting in the creative process, automating tasks, promoting linguistic diversity, enhancing content distribution, combating piracy, and reinforcing democratic values.[15]

Creativity and idea generation: GenAI systems can boost creativity by assisting in content creation and production. Writers can use AI to generate alternative ideas, overcoming writer's block. AI can also suggest design concepts and visuals for shooting sets and film posters. Although these AI-generated suggestions may not be

[15] For further reading on the various advantages see: i) the result of a survey conducted with the EAO's advisory committee members in March 2024; ii). "BBC's plans for GenAI and how we plan to use AI tools responsibly"; BBC, 28 February 2024; iii) CNC report "Quel impact de l'IA sur les filières du cinéma, de l'audiovisuel et du jeu vidéo", 8 April 2024; iv) DACS survey of artists on AI, "AI and artists' work", DACS, 18 January 2024; v) "AI is transforming the entertainment business", The Economist, 4 January 2024.

© European Audiovisual Observatory (Council of Europe) 2024

perfect, they can help advance the creative process. Additionally, many AI tools are currently free or low-cost, providing broader access to a large number of users with Internet and computer access. Such accessibility allows creators with limited budgets to pitch ideas to producers, helping them kickstart potential development processes.

Automating administrative tasks: AI can also automate time-consuming tasks with little creative added value, such as analysing audience data to understand content preferences. Additionally, AI can save time on administrative tasks, like creating and managing shooting schedules and coordinating crew logistics.

Content curation and personalisation: AI-powered tools can curate content by automatically filtering, categorising, and ranking it to match audience interests. This improves content targeting and can also increase discoverability by suggesting new content to different audiences.

Translation and linguistic diversity: AI-powered translation tools can increase linguistic diversity by making audiovisual content available in more languages, and promote accessibility. The use of avatars for sign language translation can improve accessibility for viewers with hearing impairments These AI tools can also speed up content dissemination by translating it more quickly, allowing content to reach a broader audience.

Anti-piracy and content protection: AI tools can track the use of copyrighted works, ensuring proper remuneration for authors, and detect unauthorised use, allowing to fight infringement. AI-based anti-piracy tools can help locate and address sources of piracy.

Promoting media pluralism: AI has the potential to promote democratic values by connecting newsrooms with audiences who might otherwise not engage with traditional media. AI tools can provide access to reliable, diverse information and foster media pluralism by offering content that resonates with a broader audience.

Enhancing audience experience and preserving heritage content: AI tools can facilitate the restoration of old movies and improve their image quality by adding more pixels or colors to an image. Sound restoration is also possible. These restorations can even upgrade the image quality of content for higher quality broadcasts on TV (such as 4K).

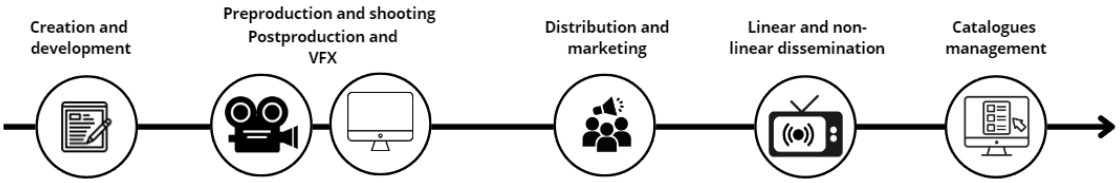
1.3. Examples of AI uses in the audiovisual industry

A variety of AI applications in the audiovisual industry are possible along the entire value chain of content: from creation and development to release on linear/non-linear

© European Audiovisual Observatory (Council of Europe) 2024

platforms. The visual below summarises the various (but non-exhaustive) uses of AI.[16] It is followed by three concrete examples of AI tools.

Figure 2. Examples of AI applications across the audiovisual value chain



Creation and development	Preproduction and shooting Postproduction and VFX	Distribution and marketing	Linear and non-linear dissemination	Catalogues management
Assistance with script writing and creating shooting schedules	Smart cameras, AI-assisted video editing, Automated sound effects	Creation of trailers and promotional content	Creation of metadata	Virtual product placement
Analysis of project potential	Foreign language dubbing automation, Subtitling automation	New marketing possibilities	Content recommendation	Restoration of works in the catalogue
Project illustration and storyboarding	Automated audio description	Assistance with office tasks	Programme schedule and video stream optimisation	Clipping and resale of automated extracts
	Face capture and simulation: face swapping, rejuvenation/ageing, digital clones	Production and operation reporting	Verification of compliance with broadcasting regulations	Fighting piracy
	Video generation from prompts, AI-assisted simulations			

Source: European Audiovisual Observatory

1.3.1. Case study 1: Claude, a conversational assistant to help with project development

Claude, an AI conversational assistant developed by Anthropic, is designed to assist with brainstorming and idea development.[17] The tool is free to use with certain limitations, and additional features are available with a Claude Pro subscription. By feeding Claude with data such as a script or story, users can receive insights on various aspects of content production, including the need for rewrites, shooting budgets (including detailed chart breakdowns), cost-saving suggestions, the number of extras required, and identification of scenes requiring special preparation or visual effects. It can also offer sales estimates by

[16] CNC report "Quel impact de l'IA sur les filières du cinéma, de l'audiovisuel et du jeu vidéo", 8 April 2024, and "How genAI tools like Lore machine revisualize storyboarding", Variety, 15 March 2024,
[17] https://www.anthropic.com/claude

© European Audiovisual Observatory (Council of Europe) 2024

territory and assist with distribution by providing lists of top foreign distributors and their contact details.

While Claude's recommendations may not be 100% accurate, they offer valuable alternative perspectives. One downside is the lack of data transparency, as the sources of Claude's information are not disclosed.

1.3.2. Case study 2: DiversityCatch, measuring diversity in content

Developed by MediaCatch in collaboration with a Danish university, DiversityCatch is an AI-driven software solution designed to measure diversity in various types of content, including broadcasts, social media, feature films, and radio.[18] It extracts and analyses data in real-time, providing insights into diversity metrics such as gender, ethnicity, and age.

DiversityCatch's advanced AI capabilities enable it to process and analyse large volumes of content quickly, outperforming traditional human data collection methods. This allows producers to develop strategies for more inclusive content creation. The software is currently employed by major industry players, including Netflix, Danish broadcasters, and the European Broadcasting Union.

Recognising the growing demand for diverse content and the existing data gaps, DiversityCatch offers a valuable solution to promote inclusivity in the media landscape.

1.3.3. Case study 3: Midjourney and DALL.E, AI tools for creating images and videos

While some AI tools can help create images for marketing purposes, some can even generate videos with a storyline. AI tools like Midjourney[19] and DALL.E[20] can assist in designing film posters or in transforming existing movie scenes into animations. Midjourney realised the first-short-generated-AI film “In search of time”.[21]

However, generating high-quality images requires mastering detailed prompt techniques.

Besides, there are concerns about the rights involved in exploiting AI-generated images, as the legal framework, at the time of writing, remains uncertain. Producers using such images may face risks of infringement procedures due to the legal ambiguity surrounding AI-generated content.

---

[18] https://mediacatch.io/solution/diversitycatch
[19] https://www.midjourney.com/showcase
[20] https://openai.com/index/dall-e-3/
[21] https://tribecafilm.com/films/in-search-of-time-2023

© European Audiovisual Observatory (Council of Europe) 2024

1.4. Challenges posed by AI in the audiovisual industry

With the rise of genAI, attention has been drawn to challenges for the audiovisual industry (subsection 1). Though a range of legislation is already shaping the use of AI, outside and within the audiovisual industry and may help overcome some of the challenges, this regulatory landscape appears fragmented (subsection 2).

1.4.1. What challenges lie ahead for the audiovisual sector?

The integration of AI into the audiovisual industry presents myriad challenges requiring careful consideration.[22] Some associations representing the audiovisual industry have voiced concerns over AI developments,[23] but what are the main challenges AI poses to the sector? They include for instance:

Jobs disruption: AI threatens to disrupt traditional job roles within the AV industry, potentially leading to job losses for professionals such as voice actors and production staff. This not only impacts livelihoods but also raises concerns about the loss of creative input and diversity in the workforce.

Preserving the human touch in creativity: While AI can enhance efficiency in production and editing processes, there is a need to preserve the human touch and creativity that are integral to the artistic processes. Questions arise about the balance between AI assistance and human creativity, particularly in the context of funding and support from public institutions.

Competition issues: Most AI tools on the market are developed and based in the USA. Their development is not within the EU scope, and the European audiovisual industry may lack the geographic scope of action to enforce its rights across the Atlantic.

Data input and copyright: The use of copyrighted data to train AI models without explicit consent from rightsholders poses legal and ethical challenges. Additionally, the scraping of data from the Internet for content creation raises concerns about data protection and privacy laws.

---

[22] For further reading on the various challenges, see: i) the result of a survey conducted with the EAO's advisory committee members in March 2024; ii). DACS survey of artists on AI, "AI and artists' work", DACS, 18 January 2024; iii)) "AI is transforming the entertainment business", The Economist, 4 January 2024; iv) Society of Audiovisual Authors' Policy Paper, "AI must serve society and enhance human creativity", 4 October 2023; v) "The impact of AI technologies on the writing profession", The Authors Guild; and vi) "The AI data scraping challenge: how can we proceed responsibly?", OECD.AI, Lee Tiedrich, 5 March 2024

[23] For further reading on the various challenges voiced by associations, see also i) SAA, "EU AI Act: joint statement from European creators and rightsholders", policy position published on 13 March 2024, ii) ACT, "ACT Response to the EC Call for contribution on competition in virtual worlds and generative AI", policy position published on 15 March 2024, iii) FERA, "Authors' performers' and other creative workers' organisations joint statement on generative AI and the EU AI Act", policy position published on 25 April 2024 and iv) EBU, "EBU welcomes the European Parliament's vote on the AI Act", policy position published on 13 March 2024

© European Audiovisual Observatory (Council of Europe) 2024

Personality rights: The scraping of data raises personality rights concerns, as photos, voices or videos could be used to create AI-generated content.

Impact on newsrooms: The use of generative AI tools in newsrooms raises questions about journalistic integrity and the role of newsrooms in collaborating with AI companies. Concerns about maintaining human-centred journalism and media pluralism underscore the need for careful consideration.

Disinformation: The proliferation of AI-generated content raises concerns about the spread of disinformation and misinformation, challenging the credibility of media sources and public trust.

Environmental cost: The increasing reliance on AI technologies has environmental implications, including energy consumption and electronic waste generation, which must be addressed for sustainable development.

Ethical dilemmas: All the above involve ethical challenges. One may explore the implications of AI-generated actors for the industry, including questions about their rights, audience perception, and the future of cinema. One may also question the cultural implications of AI-generated content versus human creativity, and how this relates to the concept of cultural diversity and whether it affects democracy in the audiovisual sector. Discussions around the role of AI in journalism and its potential impact on news media, without forgetting consideration of the balance between automation and the human touch in reporting is another angle one may explore.

To determine if the legislation presented in the next section 1.4.2. will address these issues, the following chapters (from 2 to 10) will delve into the challenges raised and question whether the regulations are AI-future-proof and capable of adapting to evolving technological landscapes within the audiovisual industry.

1.4.2. The legislative framework surrounding AI: a complex puzzle

European legislation related to AI forms a complex and interconnected framework, where each piece influences and complements the others. It reflects the multifaceted nature of AI's impacts and challenges.

The Directive on liability for defective products, originally enacted in 1985,[24] is being revised to address AI advancements. The European Commission's proposal, unveiled on 28 September 2022, highlights the need for these updates.[25] Alongside this revision, the AI Liability Directive was proposed to specifically address liability issues unique to AI

[24] Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products

[25] European Commission's proposal for a Directive on liability for defective products (28 September 2022)

© European Audiovisual Observatory (Council of Europe) 2024

systems.[26] Despite its critical role, the legislative process for the AI Liability Directive has been slow, with little progress since the European Parliament's JURI Committee was appointed in October 2022.

The AI Act, formally approved by the Council of the EU on 21 May 2024 imposes transparency obligations on GPAI providers.[27] Additionally, they must ensure compliance with Union copyright laws, as outlined in Article 53(1) of the AI Act. The Act references the Directive on Copyright in the Digital Single Market (CDSM), mandating that AI providers respect the rights of content creators, particularly in scenarios involving text and data mining (TDM).[28] [29]

Data mining is critical for AI development, but it must comply with several data protection regulations. The General Data Protection Regulation (GDPR), enacted in 2016, sets the baseline for data protection across the EU.[30] This was followed by the Data Governance Act in 2022, which underscores the pivotal role of data in the rapid development of AI technologies (Recital 2).[31] More recently, the Data Act of 2023, although not exclusively linked to AI, impacts the use of data in AI systems, (e.g. those involving AI-based IoT devices).[32] These regulations collectively ensure that the processing and use of data for AI applications respect privacy and data protection standards.[33]

When data processing becomes an essential infrastructure, competition law (e.g. Article 102 TFEU)[34] can prevent dominant undertakings from abusing their power by retaining control over this crucial infrastructure within the EU internal market. Competition law now includes the Digital Markets Act (DMA),[35] part of the Digital Services Package alongside the Digital Services Act (DSA).[36] The DMA specifically regulates how designated "gatekeepers" manage data, a vital resource for AI systems (Article 5). In contrast, the DSA calls for algorithmic transparency and accountability requirements from providers of very large online platforms (VLOPs) (see for instance Article 33).

---

[26] Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), 28 September 2022

[27] GPAI model means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market, Article 3(63) of the AI Act (ibid).

[28] Directive (EU) 2019/790 on copyrights and related rights in the Digital Single Market, 17 April 2019. According to Article 2(2), TDM means any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations.

[29] See Chapters 3 and 4 of this publication.

[30] Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, 27 April 2016

[31] Regulation (EU) 2022/868 on European data governance, 30 May 2022

[32] Regulation (EU) 2023/2854 on harmonised rules on fair access to and use of data, 13 December 2023

[33] See Chapters 2 and 5 of this publication.

[34] Treaty on the Functioning of the European Union (TFEU, Article 102)

[35] Regulation (EU) 2022/1925 on contestable and fair markets in the digital sector, 14 September 2022

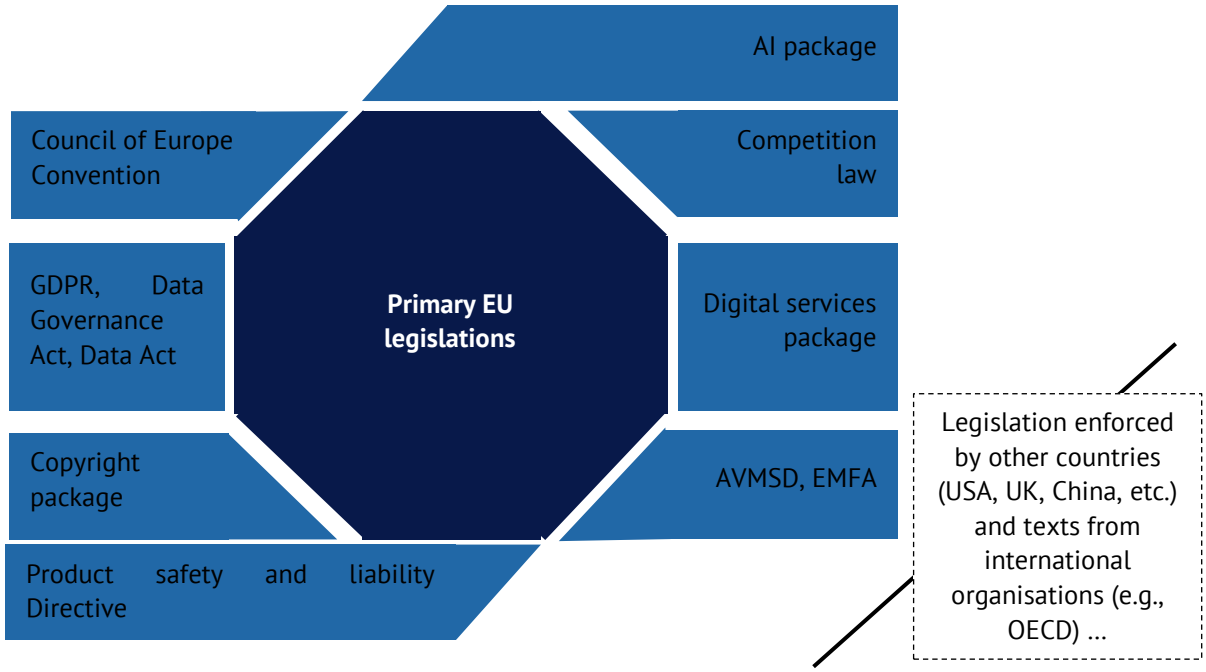
[36] Regulation (EU) 2022/2065 on a single market for digital services, 19 October 2022

© European Audiovisual Observatory (Council of Europe) 2024

Specific regulations are in place for the audiovisual sector per se. The Directive on Audiovisual Media Services provides a regulatory framework for audiovisual content, ensuring diversity and fairness.[37] The recently enacted European Media Freedom Act (EMFA) includes provisions for VLOPs, mandating functionalities for recipients to declare AI-generated content has been subject to human review or editorial control (Art. 18(1)(e)).[38] These measures aim to maintain the integrity and quality of audiovisual content in the age of AI.

Beyond EU regulations, international instruments play a crucial role. The Council of Europe's Convention on Artificial Intelligence and Human Rights, Democracy, and the Rule of Law ensures that AI development and deployment respect fundamental human rights and democratic values.[39] This Convention, set to open for signature on 5 September 2024, underscores the global dimension of AI governance and the need for international cooperation.[40]

Figure 3. AI: example of a variety of legislations



Source: European Audiovisual Observatory

[37] Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services, 10 March 2010, amended in 2018

[38] Regulation (EU) 2024/1083 establishing a common framework for media services in the internal market, 11 April 2024

[39] Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, adopted on 17 May 2024 by the Committee of Ministers

[40] See Chapter 9 of this publication.

© European Audiovisual Observatory (Council of Europe) 2024

The heterogeneity of legislative tools, encompassing both Directives and Regulations, implies that member states of the European Union may achieve the goals of Directives through varying methods. For example, in February 2024, Poland's proposal to transpose the latest Copyright Directive included an exclusion for the creation of generative AI from the scope of the TDM exception. [41]

Furthermore, the true test of legislation lies in its implementation and adaptability to ongoing developments, as shown by recent events which will be further discussed in the next chapters.

There is no doubt that the future will see more cases, both advantageous and challenging for the industry, necessitating clear legislative frameworks around the world.

[41] "TDM: Poland challenges the rule of EU copyright law", Kluwer Copyright Blog, 20 February 2024

© European Audiovisual Observatory (Council of Europe) 2024

Part II – GenAI and data transparency

"Prompt engineering is an amazingly high-leverage skill" stated OpenAI's CEO Sam Altman in 2023.[42]

The formulation of the prompt directly influences the quality of the resulting output. While prompt engineering is increasingly recognised as a top new job, some believe the contrary as AI becomes better at understanding natural language without meticulous engineered prompts.[43] However, one should not forget that prompts trigger a system trained on a vast amount of data.

One challenge for open genAI is the lack of data transparency. Users often remain unaware of the data sources used to train the machines.

There is limited information about data sourcing when using genAI, including whether this data is protected. For instance, the scraping of voice data could trigger data protection regulations like the GDPR.

Copyright concerns are critical when training genAI to assist creativity in the audiovisual industry, which might be rich in copyrighted works. Data serves as the new gold for training AI, yet it could also be a revenue source for rightsholders. Without sufficient transparency and disclosure of data sources, rightsholders may be unable to track the use of their works, give consent, or receive royalties.

[42] https://x.com/sama/status/1627796054040285184
[43] AI Prompt Engineering isn't the Future, Oguz A. Acar, Harvard Business Review, 6 June 2023

2. AI and Data Protection in Audiovisual Media

Prof. Dr. Philipp Hacker, LL.M., Yale University

2.1. Introduction

AI has a significant impact on the audiovisual sector, transforming content creation, distribution, and personalisation. GenAI, in particular, makes use of existing images, videos, and audio material – often scraped from the Internet – to create audiovisual content. However, this technological advancement introduces significant data protection challenges that must be addressed to comply with existing regulations and protect individual privacy.

It goes without saying that data serves as the cornerstone of AI development, particularly in the audiovisual sector. AI technologies rely heavily on large datasets to train models that power recommendation systems, automate content moderation, and analyse audience behaviors. Various types of data feed AI training within the audiovisual sector, extending beyond copyrighted content to include raw and processed data, metadata, user-generated content, and public domain materials. This data diversity allows AI systems to learn and adapt to different contexts. However, it also contributes to the proliferation of falsehoods, biases and information covered by data protection regimes.

As a response, a vast regulatory landscape has evolved in the audiovisual sector to tackle these data protection and related challenges. Key regulations include the EU General Data Protection Regulation (GDPR), the recently enacted AI Act, and other non-EU frameworks, for example in the US or UK and at the international level.

The machine learning pipeline in the audiovisual sector encompasses several stages, each with distinct data protection challenges:

- Datasets: Large datasets are essential for training AI models, but they raise significant privacy implications. The collection, storage, and use of extensive personal data must be diligently managed to avoid data protection violations – which may not in every case be feasible.
- Training: The legal basis for AI training must be clearly defined, and provisions specifically protecting sensitive data be respected.
- Model: Once trained, AI models must address issues such as model inversion and data leakage, which can expose personal data. The right to erasure under GDPR is

© European Audiovisual Observatory (Council of Europe) 2024

also crucial, allowing individuals to request the removal of their data from AI systems – or even the deletion of the entire model in extreme cases.

- Deployment: During deployment, AI systems must adhere to legal requirements for processing data, ensuring the accuracy of outputs, and preventing the dissemination of misinformation or "hallucinations."[44] Additionally, the use of AI for automated decision-making must consider transparency provisions and specific prohibitions. Furthermore, the protection of minors and other vulnerable groups remains a key concern.

These elements collectively underscore the intricate relationship between AI development and data protection in the audiovisual sector, in the inherent tensions between an accelerating technological environment, particularly since the advent of genAI, and the legal obligations centering on purpose limitation, data minimisation and storage limitation.

2.2. Audiovisual material as personal data

Audiovisual data, such as images, videos, and voice recordings, count as personal data under the GDPR if they relate to an identified or identifiable natural person (Article 4 GDPR). Under similar conditions, they qualify as personally identifiable information in other data protection frameworks, such as the US.[45]

Hence, photographs and video recordings fall under the category of personal data if they can identify an individual. The Italian Data Protection Authority ruled as much concerning photographs in its injunction against Clearview AI.[46] For example, if an image or video shows a person's face or other identifiable features, it is generally considered personal data, as the UK Information Commissioner's Office has mentioned.[47] As Recital 51 notes, when these images or videos undergo specific technical processing, such as for facial recognition, they may even fall into the category of biometric data, which is specifically protected under Art. 9 GDPR. As the Irish Data Protection Commission has pointed out, once pictures are shared online, the household exemption, which examines certain private processing activities from the scope of the GDPR (Article 2(2)(c)), does not apply anymore.[48]

---

[44] This refers to information that is nonsensical or unfaithful to the provided source content, see 2.3.

[45] See, e.g., Erika McCallister, Tim Grance and Karen Scarfone, "Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). Recommendations of the National Institute of Standards and Technology", National Institute of Standards and Technology Special Publication 800-122, 2010, sec. 2-1; Usercentrics, "Personally Identifiable Information (PII) vs. Personal Data – What's the difference?", Usercentrics CMP, Munich, 3 March 2021.

[46] Ordinanza ingiunzione nei confronti di Clearview AI, 10 February 2022, Case 9751362, Point 3.4.

[47] UK Information Commissioner's Office, "Taking photographs: data protection advice for schools", Cheshire.

[48] Irish Data Protection Commission, "What is the position regarding individuals taking photographs/videos in a public place?", Dublin.

© European Audiovisual Observatory (Council of Europe) 2024

Voice recordings generally qualify as personal data because an individual can be identified through their unique vocal characteristics,[49] and may even constitute sensitive data, as attributes such as age or gender can be inferred from it.[50]

Advanced methods for voice-based speaker anonymisation aim to suppress the speaker's identity.[51] The first strategy performs voice transformation techniques that alter the source or filter characteristics of the speech.[52] Recent research has proposed the use of x-vector speaker representations to suppress the timbre of a speaker, thereby preventing speaker identification.[53] However, it should be borne in mind that many re-identification techniques exist,[54] and may even arise over time, converting non-personal data into personal data (Recital 26 GDPR).[55] Overall, the vast majority of audiovisual material will therefore qualify as personal data/personally identifiable information and fall under the scope of the data protection laws of the respective countries.

2.3. Selected data protection and privacy concerns

Different data protection policy regimes will raise different challenges. However, several problems will likely be germane to many data protection laws existing in various Council of Europe countries, as recent publications by data protection authorities show.[56] These

[49] Cf. Nora Ni Loideain and Rachel Adams, "From Alexa to Siri and the GDPR: the gendering of virtual personal assistants and the role of data protection impact assessments", Computer Law & Security Review 105366, 2020, 10.

[50] Andreas Nautsch and others, "The GDPR and Speech Data: Reflections of the Legal and Technology Communities: First Steps towards a Common Understanding", Interspeech: Crossroads of Speech and Language, 2019, p. 3.

[51] Ingo Siegert and others, "Personal data protection and academia: GDPR issues and multi-modal data-collections "in the wild"", Online Journal of Applied Knowledge Management, 2020, p. 20.

[52] Miran Pobar and Ivo Ipšić, "Online speaker de-identification using voice transformation", 37th International convention on information and communication technology, electronics and microelectronics, 2014, p. 1264.

[53] Fuming Fang and others, "Speaker Anonymization Using X-vector and Neural Waveform Models", 10th ISCA Workshop on Speech Synthesis (SSW 10), 2019.

[54] Luc Rocher, Julien M Hendrickx and Yves-Alexandre De Montjoye, "Estimating the success of re-identifications in incomplete datasets using generative models", Nature Communications 10, 2019, pp. 1-9; see also Paul Ohm, "Broken promises of privacy: Responding to the surprising failure of anonymization", UCLA Law Review, 2009, pp. 1701-1777; Manon Oostveen, "Identifiability and the applicability of data protection to big data", International Data Privacy Law, 2016, pp. 299-309.

[55] Michèle Finck and Frank Pallas, "They who must not be identified—distinguishing personal from non-personal data under the GDPR", International Data Privacy Law, 2020, pp. 11-36; Philipp Hacker and Jürgen Neyer, "Substantively smart cities–Participation, fundamental rights and temporality", Internet Policy Review, 2023, pp. 1-30.

[56] See, e.g., guidelines by the European Data Protection Board, "Report of the work undertaken by the ChatGPT Taskforce", 23 May 2024; German data protection authorities, "Orientierungshilfe der Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder", Künstliche Intelligenz und Datenschutz, Version 1.0, 6 May 2024; Bavarian Data Protection Authority, "the data protection checklist for AI", 24 January 2024; French data protection authority, "Self-assessment guide for artificial intelligence (AI) systems"; UK Information Commissioner's Office, "Guidance on AI and Data Protection", 15 March 2023; Italian Data Protection Authority, "Instructions against web scraping", 20 May 2024.

© European Audiovisual Observatory (Council of Europe) 2024

include:[57] a legal basis for including AV material in a training data set, including scraping; the problem of hallucinations and factually incorrect personal data; LLMs as personal data; the treatment of sensitive data; information provision and user control; and automated decision-making.

2.3.1. Legal basis for training

To train an AI model, vast amounts of audiovisual material are typically processed. To the extent that those images, videos or sounds constitute personal data (see above), data protection law kicks in: Any action involving the processing of personal data, such as scraping, storage, transfer, or copying, necessitates a legal basis under Article 6 GDPR. This regulation extends to companies outside the EU that provide services within the EU, encompassing many major AI companies. Utilising personal data for AI training, including fine-tuning, is unlawful under the GDPR unless a specific legal basis is applicable.

Obtaining valid consent from the numerous individuals whose data is incorporated into large datasets is generally infeasible due to the high transaction costs involved.[58] Consequently, AI training often relies on the balancing test of Article 6(1)(f), which justifies data processing if the developer's legitimate interests outweigh the data subjects' rights and freedoms.[59] The outcome of the balancing test must be evaluated individually. However, some general indications can be given.

If an AI model has socially beneficial applications or if the data usage was reasonably anticipated by the data subjects (Recital 47), the balance might favor the developers. However, the latter criterion is seldom fulfilled. Moreover, privacy-enhancing measures like pseudonymisation, transparency, or encryption can also support the legality of AI training. On the other hand, the nature and scope of processing, the type of data (especially sensitive data), and the level of transparency and control offered to data subjects might tip the balance against legality.[60]

In the context of narrowly tailored AI models using supervised learning, it might be argued that AI training does not significantly harm data subjects, especially if the model is not widely disseminated and data breaches are unlikely due to robust IT security.[61] However, justifying this for genAI is more difficult. These models are often

[57] See also Claudio Novelli and others, "Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity", arXiv preprint arXiv:240107348, 2024, pp. 1-36.
[58] Miranda Mourby, Katharina Ó Cathaoir and Catherine Bjerre Collin, "Transparency of machine-learning in healthcare: The GDPR & European health law", Computer Law & Security Review, 2021, 105611.
[59] Frederik J Zuiderveen Borgesius and others, "Tracking walls, take-it-or-leave-it choices, the GDPR, and the ePrivacy regulation", European Data Protection Law Review, 2017, pp. 353-368.
[60] Philipp Hacker, Andreas Engel and Marco Mauer, "Regulating ChatGPT and other Large Generative AI Models", ACM Conference on Fairness, Accountability, and Transparency (FAccT '23) 1112, Technical Report, 2023, pp. 1-22.
[61] Tal Z Zarsky, "Incompatible: The GDPR in the age of big data", 47 Seton Hall L Rev 995, 2016, pp. 995-1018; Philipp Hacker, "A legal framework for AI training data—from first principles to the Artificial Intelligence Act", 13 Law, Innovation and Technology, 2021, pp. 257-301.

© European Audiovisual Observatory (Council of Europe) 2024

widely used. And recent studies show that they are prone to revealing personal data through data leakage and model inversion (see below, 3.).[62] This challenge is further exacerbated in fine-tuning scenarios.[63]

Reflecting these concerns, a recent restrictive guideline from the Dutch Data Protection Authority highlights that mass web scraping of personal data is almost always illegal unless specifically tailored to narrow purposes.[64] Additionally, the Italian Data Protection Authority has ruled that web scraping by Clearview AI for general face recognition purposes lacks a legal basis and cannot be justified by the balancing test.[65] Overall, the mass collection and processing of personal data for large language models, particularly from the Internet, is difficult and in some cases impossible to reconcile with data protection laws that demand specific legal bases for processing activities, such as the GDPR.

2.3.2. Hallucinations

Beyond requiring a legal basis, data protection laws generally enshrine a set of principles that the processing of personal data needs to adhere to. As has been noted repeatedly,[66] big data analytics and AI are not easily squared with principles such as purpose limitation, storage limitation, or data minimisation. One principle that has assumed particular urgency with the advent of genAI, however, is the principle of data accuracy; it is found, for example, in the GDPR, but also in the UK GDPR.[67] In the AV context, AI-generated movie summaries may provide inaccurate information about actors and directors; or deepfakes suggest certain actions or words by data subjects that they never made or spoke. Overall, due to its reliance on probabilistic methods, genAI is prone to hallucinations–content that is factually incorrect, nonsensical or unfaithful to the provided source content.[68] While new tools are being developed to detect hallucinations,[69]

---

[62] See, e.g., Stella Biderman and others, "Emergent and predictable memorization in large language models", 36 Advances in Neural Information Processing Systems, 2024, pp. 1-9; Nicholas Carlini and others, "Quantifying Memorization Across Neural Language Models", The Eleventh International Conference on Learning Representations, 2023, pp. 1-19; Nicholas Carlini and others, "Extracting training data from large language models", 30th USENIX Security Symposium (USENIX Security 21) 2633, 2021, pp. 1-13; Eric Lehman and others, "Does BERT pretrained on clinical notes reveal sensitive data?", arXiv preprint arXiv:210407762, 2021, pp. 1-10; Nicholas Carlini and others, "Extracting Training Data from Diffusion Models" (2023) arXiv preprint arXiv:230113188, 2023, pp. 1-16.

[63] Jaydeep Borkar, "What can we learn from data leakage and unlearning for law?", arXiv preprint arXiv:230710476, 2023, pp. 1-3

[64] Autoriteit Persoonsgegevens, "AP: scraping bijna altijd illegal", 1 May 2024, pp. 3-26

[65] Garante per la protezione dei dati personali, "Injunction against Clearview AI, Case 9751362", Point 3.6.2, 10 February 2022, pp. 1-30

[66] See, e.g., Zarsky, "Incompatible: The GDPR in the age of big data"; Novelli and others, "Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity", 14

[67] ICO, Guidance on AI and Data Protection, 2023, p. 38

[68] See only Ziwei Ji and others, "Survey of hallucination in natural language generation", ACM Computing Surveys, 2023, pp. 1-3

© European Audiovisual Observatory (Council of Europe) 2024

they operate probabilistically, too, and are unlikely to catch and remove all hallucinations in critical scenarios.[70]

Therefore, while the accuracy principle is crucial, it is subject to balancing against other rights. In practice, only significant false information is likely to mandate correction.[71] However, preventing even this more limited set of hallucinations will prove challenging for the LLM developers and deployers.[72]

2.3.3. LLMs as personal data

Modern data protection laws like the GDPR include the right to erasure of personal data, which becomes complex with AI due to issues like model inversion and data leaks. Model inversion can reconstruct training data, including censored audiovisual materials, and memorisation may cause AI to output personal data included in training data, even via simple prompts. This suggests that LLMs themselves might be considered personal data. If so, merely updating or downloading LLMs would require a legal basis, and individuals could potentially request model deletion under Article 17 GDPR. If LLMs are indeed classified as personal data, it could imply a deluge of data protection breaches by entities developing or using these models.

Recent guidance from the Hamburg Data Protection Authority on 15 July 2024 seeks to reassure users that LLMs are generally not considered personal data.[73] However, this decision does not end the debate.[74] Rather, LLMs can be likened to compressed and encrypted data; hence, they may still be personal data if certain conditions are met: this depends on the technical ability to link the model to specific individuals, the likelihood of the controller using this method, and ongoing legal debate about the impact of the method's legality on this classification.[75]

---

[69] Sebastian Farquhar and others, "Detecting hallucinations in large language models using semantic entropy", Nature, 2024, pp. 625-630
[70] Cf. ibid., 629
[71] Cf. again ICO, Guidance on AI and Data Protection, 2023, 39
[72] Cf. also EDPB Report, para. 29-31
[73] https://datenschutz-hamburg.de/news/hamburger-thesen-zum-personenbezug-in-large-language-models
[74] Conceiving LLMs as personal data, e.g., Michael Veale, Reuben Binns and Lilian Edwards, "Algorithms that remember: model inversion attacks and data protection law", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2018, 20180083; Paulina Jo Pesch and Rainer Böhme, "Verarbeitung personenbezogener Daten und Datenrichtigkeit bei großen Sprachmodellen" Multimedia und Recht, 2023, p. 920; negating, e.g., Flemming Moos, "Personenbezug von Large Language Models", Computer und Recht, 2024, para. 27 et seqq.; cf. also EDPB Report, para. 25
[75] See Patrick Breyer, Judgment of 19 October 2016, C-582/14

© European Audiovisual Observatory (Council of Europe) 2024

2.3.4. Sensitive data

Another pressing challenge under data protection law involves audiovisual material that can reveal sensitive information such as age, racial or ethnic origin, political opinions, religious or philosophical beliefs, and trade union membership; this may be the case, for example, with photographs (age, racial or ethnic origin, religious background) and even voice recordings (age).[76] A key case highlighting this issue is the *Meta v Bundeskartellamt* case, where the court ruled that data does not need to directly refer to sensitive attributes to be protected under Article 9 GDPR. It is enough "that data processing allows information falling within one of those categories to be revealed".[77] With advanced analytics, this will often be the case. For example, the AI-part content recommendation engine may, deliberately or inadvertently, process sensitive data within this framing, such as information about the age, ethnic origin, religion or political opinions of recommendees. Moreover, biometric data, such as images or videos used for identification purposes in facial recognition, also falls under Article 9 GDPR.[78]

Article 9(2) GDPR outlines exceptions for processing sensitive data, but these exceptions are limited. One such exception, under Article 9(2)(e), is when the data has been "manifestly made public by the data subject". However, voluntary publication by the data subject does not legitimise the use of the data for purposes beyond the original intent of the publication.[79] The Italian Data Protection Authority ruled that no exception applies to the indiscriminate scraping of images from the web for face recognition purposes, even if they were published voluntarily by the data subjects, in its ruling against Clearview AI.[80]

Consequently, except for explicit consent, which is challenging to obtain, no clear exception exists for using sensitive data in general generative models and audiovisual materials. Specific contexts, such as health-related scenarios, may have individual exceptions enshrined in national laws with significant safeguards. However, these exceptions are narrowly defined and do not broadly apply to generative AI models and the processing of audiovisual materials.

---

[76] See Ordinanza ingiunzione nei confronti di Clearview AI, Injunction of 10 February 2022, Case 9751362, Point 3.4

[77] Meta Platforms and Others, Judgment of 4 July 2023, C-252/21, para. 73

[78] Recital 51 GDPR and Ordinanza ingiunzione nei confronti di Clearview AI, Injunction of 10 February 2022, Case 9751362, Point 3.4

[79] Article 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 9 April 2014, 39; Ordinanza ingiunzione nei confronti di Clearview AI, Injunction of 10 February 2022, Case 9751362, Point 3.4, "Likewise, it is noted that the Internet publication of personal data by the person to whom they refer, for example in the context of a social media network, does not, in itself, entail a sufficient condition to legitimise its free reuse by third parties." [automated translation].

[80] Ordinanza ingiunzione nei confronti di Clearview AI, Injunction of 10 February 2022, Case 9751362, Point 3.6.3.

© European Audiovisual Observatory (Council of Europe) 2024

2.3.5. Information and user control

The next significant challenges for ensuring GDPR compliance in LLMs (or other genAI models) are primarily found in requirements to provide notice and information to data subjects, for example pursuant to Articles 12-15 of the GDPR. They pose unique difficulties due to the extensive and varied nature of the data processed by genAI.[81]

Article 14 of the GDPR is particularly pertinent when considering data harvested from the internet for training purposes. However, the requirement to inform each individual whose data is included in the training set can be impractical due to the significant effort involved. This is where Article 14(5)(b) GDPR comes into play, which provides for exemptions when the effort is disproportionate. Key factors in this assessment, as noted in Recital 62 of the GDPR, include the number of data subjects, the age of the data, and the safeguards implemented. The Article 29 Working Party has also highlighted the impracticality of informing individuals when data is aggregated from numerous sources and contact details are unavailable.[82]

In contrast, personal data submitted by users via chat interfaces (prompts) does not benefit from such exemptions. Article 13 of the GDPR explicitly requires informing data subjects about several key aspects, including the purposes of processing, the legal basis for processing, and any legitimate interests pursued by the data controller. This also holds for any audiovisual materials that data subjects may upload.

The balance between practical compliance challenges and the rights of data subjects is delicate. Although Article 14(5) GDPR offers a potential exemption for cases of disproportionate effort, this remains contentious, especially when it comes to scraping and processing data for commercial purposes. The data controller, as defined in Article 4(7) of the GDPR, must meticulously document their considerations under this provision to ensure compliance with the accountability principle enshrined in Article 5(2) of the GDPR. Furthermore, making documents regarding the methods of collecting training data publicly accessible would reinforce a commitment to data protection principles and enhance transparency.

2.3.6. Automated decision making

Significantly, the use of AI models, such as LLMs, might also be classified under automated decision-making processes scrutinised by the GDPR. Article 22 generally prohibits decisions solely based on automated processing, including profiling, that have legal or similarly significant effects on individuals unless specific exceptions apply. This is particularly relevant in contexts like recruitment or credit scoring, where automated

[81] Hacker P., Engel A. and Mauer M., "Regulating ChatGPT and other Large Generative AI Models", ACM Conference on Fairness, Accountability, and Transparency (FAccT '23), 5 Feb 2023, 2-3

[82] Article 29 Data Protection Working Party, "Guidelines on Transparency under Regulation 2016/679", WP260 rev.01, Brussels, 2018, para. 63

© European Audiovisual Observatory (Council of Europe) 2024

evaluations can significantly influence outcomes; however, significant effects may also arise in the context of content recommendation engines, deepfakes, or automated movie summaries. The recent SCHUFA case by the CJEU lowered the bar for finding automated decision-making:[83] it is sufficient for a probability value generated by one party (e.g., AI provider) to significantly influence a third party's decision (e.g., an employer, bank, or store) to enter into, execute, or terminate a contractual relationship with the data subject.

Exemptions to this prohibition are limited and include scenarios where explicit consent is obtained, the processing is necessary for a contract, or specific legal provisions exist. However, obtaining valid consent can be challenging due to power imbalances, and arguments based solely on efficiency are unlikely to suffice (Recital 43 GDPR). Instead, companies must demonstrate tangible benefits to data subjects.

These cases and regulatory insights again showcase the growing need for transparency and legal compliance in the use of automated systems and AI to ensure that individuals' rights are protected in increasingly digital and automated environments.

2.4. The AI Act

The recently enacted EU AI Act[84] imposes several significant obligations on both AI providers and deployers when processing audiovisual material, whether for training or inference. It establishes a comprehensive framework for managing the risks associated with AI systems processing audiovisual material. Providers must implement robust risk management, data governance, and transparency measures, while deployers have monitoring, documentation, and impact assessment responsibilities. Transparency is further emphasised through clear disclosure and labeling requirements. This reinforces the transparency mandates under the GDPR.[85]

However, tensions exist between data protection law and the AI Act, too.[86] The AI Act introduces new roles and terminologies, such as “providers” (developers) and “deployers” (professional users) of AI systems, which do not perfectly align with the GDPR’s categories of “controllers” and “processors”. This divergence could lead to complexities in determining compliance responsibilities, especially in cases where the

[83] CJEU, SCHUFA Holding (Scoring), judgment of 7 December 2023, C-634/21, para. 73

[84] See, e.g., Michael Veale and Frederik Zuiderveen Borgesius, “Demystifying the Draft EU Artificial Intelligence Act—Analysing the good, the bad, and the unclear elements of the proposed approach”, Computer Law Review International, otto schmidt, Cologne, 2022, p. 97; Martin Ebers and others, “The European commission’s proposal for an artificial intelligence act—a critical assessment by members of the robotics and AI law society (RAILS)”, j, MDPI, Basel, 2021, p. 589

[85] See, e.g., Philipp Hacker and Jan-Hendrik Passoth, “Varieties of AI Explanations under the Law. From the GDPR to the AIA, and Beyond”, xxAI – Beyond Explainable AI, Springer, Cham, 2022, p. 343

[86] See, e.g., James Clark, Muhammed Demircan & Kalyna Kettas, “Europe: The EU AI Act’s relationship with data protection law: key takeaways”, Privacy Matters, DLA Piper, 25 April 2024; Sergio Barezzani, “Artificial Intelligence Act (AI Act) and the GDPR”, Encyclopedia of Cryptography, Security and Privacy, Springer, Cham, 2024, pp. 1-6; Christiane Lawson-Hetchely, “The Potential Impact of the Future AI Act on the GDPR”, University of Oslo, Oslo, 2022

© European Audiovisual Observatory (Council of Europe) 2024

same entity may be considered both a mere deployer under the AI Act but a controller under the GDPR – as will often be the case.[87] Additionally, practical challenges in enforcement and cooperation between different regulatory authorities remain. For instance, both acts have distinct supervisory frameworks, which might result in overlapping or conflicting regulatory actions.[88]

2.5. International data transfers

Yet other data protection obligations come into play when audiovisual, or other, data is sent outside of the EU, for example to enable cloud-based analytics of processing. For example, broadcasters using AI for multiple language versioning may transfer videos from the EU to a cloud system based in the US; and smart home devices may send voice recordings to non-EU servers for natural language processing. Articles 44 and following of the GDPR address the rules and safeguards required for international data transfers from the EU to third countries, such as the USA. These articles aim to ensure that personal data transferred outside the EU receives a level of protection essentially equivalent to that guaranteed within the EU. Most importantly, an adequacy decision by the European Commission allows for data transfers to countries deemed to provide adequate data protection levels, simplifying the compliance process for organisations operating internationally.

The EU-US Data Privacy Framework (DPF) is the basis of the latest adequacy decision designed to replace the Privacy Shield invalidated by the Schrems II decision.[89] That decision ruled that the bulk collection and processing of personal data by US authorities for national security reasons is incompatible with the proportionality principle and an adequate level of privacy; and that EU citizens lack an effective judicial remedy to challenge potential violations. Against this background, the DPF introduces enhanced safeguards, including stricter oversight and enforcement mechanisms, and new redress avenues for EU citizens.[90]

The DPF explicitly emphasises the necessity and proportionality principles, seeking to ensure that access to data by US authorities is strictly limited to what is necessary and proportionate for national security purposes.[91] Additionally, the framework establishes the Data Protection Review Court (DPRC), an independent and impartial body that provides EU individuals with a mechanism to seek redress regarding the collection

[87] Sebastião Barros Vale, "GDPR and the AI Act interplay: Lessons from FPF's ADM Case-Law Report", Future of Privacy Forum, 3 November 2022
[88] Paweł Hajduk, "AI Act and GDPR: On the Path Towards Overlap of the Enforcement Structures", RAILS Blog, RAILS, Berlin, 1 October 2023
[89] Schrems II, Judgment of 16 July 2020, CJEU Case C-311/18
[90] See, e.g., David Michael Watry, "The transatlantic data privacy framework: Schrems II, GDPR and American national security", University of Malta 2023; Linda Kidwell, "GDPR Compliance in EU-US Data Transfers", University of Lund 2023
[91] Alex Wodi, "The EU General Data Protection Regulation (GDPR):Five Years After and the Future of Data Privacy Protection in Review", Working Paper, 2023, 9.

© European Audiovisual Observatory (Council of Europe) 2024

and use of their data by US intelligence agencies.[92] This is important if, for example, investigative journalists from the EU use a US-based AI company to verify the authenticity of a video depicting a relevant event: they can now challenge access to the video by US intelligence services. The DPRC even has the authority to order the deletion of data if it determines that the data was collected in violation of the established safeguards.[93] For example, if a US-based post-production service provider (e.g., AI-based movie subtitling; voice translation) fails to comply with the DPF's principles, the affected EU company can seek enforcement through the DPRC.

Its impact on GDPR compliance is significant as it seeks to address the concerns raised by the CJEU in the Schrems II decision. However, the DPF might eventually be invalidated, too, as the mandate to engage in bulk data processing is broad: it may be authorised when "it is determined to be necessary to engage in bulk collection in order to advance a validated intelligence priority".[94] The US understanding of necessity, in this context, may be broader than the strict necessity and proportionality requirements in the CJEU doctrine.[95] This raises the specter of a potential Schrems III decision and further uncertainty concerning international data transfers between the EU and the US.

2.6. Comparison with US and international law

The GDPR, HIPAA, and various state laws in the US all aim to protect personal data but operate under different frameworks and scopes. The GDPR provides comprehensive data protection across the EU, ensuring robust safeguards for all personal data, including a particularly stringent regime for sensitive data, including medical information. For example, any use of audiovisual materials in medical AI training under the GDPR must adhere to strict transparency and, typically, consent requirements. This is similar to HIPAA in the US, which mandates protections for medical data. When using medical images for AI training, HIPAA requires de-identification of data or obtaining explicit patient consent to ensure privacy and security are maintained.[96]

US state legislations, such as the California Consumer Privacy Act (CCPA) and the Colorado Privacy Act (CPA) replicate in parallel form many GDPR principles by enforcing strict data protection measures, including rights to access, delete, and opt-out of data processing.[97] These laws provide additional layers of protection similar to the comprehensive GDPR approach. As a consequence, companies processing audiovisual data

[92] Ibid.
[93] European Commission, "Questions & Answers: EU-US Data Privacy Framework", 10 July 2023.
[94] 50 U.S.C. § 3001, Ex. Ord. No. 14086, Oct. 7, 2022, 87 F.R. 62283, Sec. 2(c)(ii)(A).
[95] Bjørn Aslak Juliussen and others, "The third country problem under the GDPR: enhancing protection of data transfers with technology", International Data Privacy Law 2023, pp. 225, 229.
[96] Steve Alder, "Editorial: HIPAA, Healthcare Data, and Artificial Intelligence", The HIPAA Journal, 16 December 2022; Becky Whittaker, "Healthcare AI and HIPAA privacy concerns: Everything you need to know", The Intake, 15 December 2022.
[97] Bloomberg Law, "Which States Have Consumer Data Privacy Laws?", 18 March 2024.

© European Audiovisual Observatory (Council of Europe) 2024

for AI training or other purposes must implement stringent privacy measures and, ideally, obtain explicit consent from individuals.

However, state-level initiatives, paired with sectoral approaches in the US (e.g., through the Biden Executive Order on AI), increasingly intricate data transfer rules, and comprehensive legislation in the EU and China threaten to create a patchwork of privacy, data protection and AI regulation applicable to AI training and deployment, particularly but not exclusively in the audiovisual sector.

Hence, international efforts are paramount to, potentially, mapping out a path through the growing maze. Initiatives like the UN Global Digital Compact and the G7 Hiroshima Process reflect a growing consensus on the need for responsible AI and data protection standards worldwide. These frameworks aim to harmonise AI regulations across borders, promoting core principles, such as transparency, accountability, and human rights protections, akin to those enshrined in the GDPR. Such global efforts are crucial for creating a cohesive approach to AI governance, ensuring that audiovisual data and other personal information are protected regardless of where they are processed – but also that effective compliance remains possible for companies using audiovisual and other data for societal benefit.

Ultimately, these international efforts will have to link up to the emerging international standards developed by standard-setting organisations such as ISO or CEN/CENELEC, in order to operationalise vague principles on the ground and in concrete machine learning systems. Simultaneously, this points to the pressing need to include a broad variety of stakeholders, beyond industry, in any standardisation efforts, and to create effective ways, through scholarships and other means, to enable civil society and academic participation in those endeavors.

© European Audiovisual Observatory (Council of Europe) 2024

3. AI & Copyright Protection when Feeding the Machine

Gianluca Campus[98]*, PwC Digital Innovation*

3.1. Introduction

3.1.1. Overview of AI systems and their processing of copyrighted data

Everybody has a clear perception of the relevance of AI as a disrupting technology, since it became capable of replicating (and even surpassing) human abilities, but with the introduction of the Generative AI new crucial legal challenges are posed from the IP perspective.

This section of the Report will focus on the potential risk of copyright infringement deriving from the use of works as training data for genAI systems and will analyse how legislator and courts are addressing such legal challenges.

First of all, it is useful to understand how the training data are treated within a genAI system. To understand more in detail how the AI-generated outputs are deriving from the works included in the training dataset, it was suggested to consider a sort of "generative-AI supply chain",[99] an interconnected set of stages that transform training

[98] Director of Legal Operations at PwC Digital Innovation Italy, PhD, Fellow of the University of Milan.

[99] See Katherine Lee, A. Feder Cooper and James Grimmelmann, Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain, 27 July, 2023, forthcoming, in Journal of the Copyright Society 2024. On the substantial difference in the creative process of the AI systems as compared to human creativity and on the impact that such differences have on the reconstruction of copyright aspects, see also Giancarlo Frosio, *Should we ban Generative AI, incentivise it or make it a medium for inclusive creativity?*, July 31, 2023, in Enrico Bonadio and Caterina Sganga (eds), A Research Agenda for EU Copyright Law (Edward Elgar, forthcoming), according to which "*One factor that calls for careful consideration when contemplating legal incentives for AI-generated creativity is the unique nature of machine-generated creativity, which differs significantly from human creative processes. In this context, it is crucial to reflect on the distinctive characteristics of creativity generated by machines, which excel in cumulative and combinatorial processes […] Unlike machines, humans do not recall the actual objects themselves but rather conceptual ideas of those objects.*"

© European Audiovisual Observatory (Council of Europe) 2024

data into generations (e.g. a new and hopefully never-before-seen picture of an item that may or may not ever have existed).

According to the authors' reconstruction, the supply chain starts with creative works: all of the books, artwork, software, and other products of human creativity that genAI seeks to learn from and emulate. Next, works and other information must be converted into data: digitally encoded files in standard, known formats. Individual items of data are useless for AI training by themselves. Instead, they must be compiled into training datasets: vast and carefully structured collections of related data. The process requires both extensive automation and thoughtful human decision-making.

To create a genAI model, its creator picks a technical architecture, assembles training datasets, and then runs a training algorithm to encode features of the training data in the model. Model training is both a science and an art, and it involves massive investments of time, money, and computing resources. The model that results from this initial training process is called a "base" or "pre-trained model", because it is often just a starting point. A model can also be fine-tuned to improve its performance or adapt it to a specific problem domain. This process, too, involves extensive choices — and it should not be carried out by the same entity that did the initial training.

A deployed system can be used to generate outputs: new creative works that are based on statistical patterns in the training dataset but combine them in new ways. An output — or "generation" — is based on a prompt supplied by the user: an input that describes the particular features they want the output to have. This is typically the only part of the supply chain that users see.

In such a reconstruction, the model is simply a different and complicated arrangement of training examples. But the model could be also seen as a derivative work of its training data, a work based upon one or more preexisting works that combines the authorship in an existing work with new authorship. Training datasets contain complete literal copies of millions of digitised copyrighted works. A model, as a collection of parameters, is different in kind from the copyrightable works it was trained on.

3.1.2. Considerations on derivative works

It is not simple and not obvious to understand whether there is a "derivative" relationship between the training dataset and the AI-generated output. It is crucial to understand whether the output generated via AI systems after data processing can be considered a derivative work and consequently whether the rightsholders of the training data must authorise the derivative work generated by AI. With regards to the US legal system, Professor Daniel Gervais[100] points out that the Copyright Act provides an exclusive right "to prepare derivative works based upon the copyrighted work" and defines "derivative

[100] See Daniel J. Gervais, AI derivatives: the application to the derivative work right to literary and artistic productions of AI machines, Seton Hall Law Review, Vol. 53, 2022 and Vanderbilt Law Research Paper No. 22-12.

© European Audiovisual Observatory (Council of Europe) 2024

work" in part as any work "based upon one or more preexisting works". Translated in the AI environment, it is necessary to take into account that AI systems can produce literary and artistic content (output) that is almost necessarily "based upon" a dataset consisting of preexisting works.

Moreover, derivative works must satisfy the original requirement to be eligible for copyright protection. "Originality" is not defined by the laws, but it was defined by the US Supreme Court as meaning that the derivative work must be independently created by its author and must embody expression that is at least minimally creative (i.e. the work is the result of creative choices made by the author).[101]

In addition, the notion of originality applied to the protection of derivative works requires that the person claiming to have authored a derivative work must have added or transformed one or more preexisting works in some way. The legal nature of the derivative work can stem from an authorisation from the copyright owner, from an exception such as fair use, or because the underlying work is no longer protected.

So, it has to be verified whether the creative choices made by the programme's author (or arguably by the user, if applicable) are present in the AI system's output. If not, protecting that output as the work of the programmer (or user) is incompatible with both fundamental doctrinal tenets of copyright and its policy purpose, and it would over-reward the programmer (or user).

With regards to the EU legal system, the principles of the Berne Convention are applicable, according to which "translations, adaptations, arrangements of music and other alterations of a literary or artistic work shall be protected as original works without prejudice to the copyright in the original work".[102] In addition, the CJEU, too, has indicated a requisite of "originality" for the derivative works and has clarified that the EU originality test requires more than skill, labor or effort and, more in detail, has dictated that technical considerations, rules and constraints do not confer originality.[103]

In the absence of clear indications from the legislation or from the case law, this would be most probably the subject of case-by-case analysis on the training of the AI system, so as to assess whether the outputs are elaborations close to forms of expression of the initial works used for training and/or whether the patterns used by the AI system for generating new works reproduce output hardly discernible from the original works of the author.

[101] See Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 346 (1991); Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53 (1884).
[102] See Article 2(3) Berne Convention.
[103] See E. Rosati, When is a derivative work original and thus protectable by copyright? Classicist's critical edition makes its way to Luxembourg in fresh Romanian CJEU referral. See also CJEU judgment (Fifth Chamber) 11 June 2020 in case C-833/18.

© European Audiovisual Observatory (Council of Europe) 2024

3.2. Text and data mining exception for training data

3.2.1. Examination of the applicability of TDM exemption to AI training data

AI in general and, above all, generative AI systems require large datasets for machine training and deep learning,[104] including copyrighted works such as music, images or text, depending on the planned output. Such requirements are usually satisfied via text and data mining (TDM), defined as the automated process of extracting information and insights from large amounts of text and data.[105] There are two types of data that can be handled via TDM: while data mining handles structured data coming from systems, such as databases, spreadsheets, etc., text mining deals with unstructured data found in documents, emails, social media, and the web, where the patterns are extracted from natural language text rather than from structured databases of facts.[106] Text mining benefits from the advances in natural language processing, particularly when transforming unstructured text into structured data suitable for analysis.

The TDM activities become critical when they imply the access and the extraction of data from copyrighted contents, whereby these activities may potentially infringe the exclusive rights recognised by national laws and international treaties of authors and related rights owners, essentially reproduction and adaptation rights. The relevance of the TDM activities is also related to the fact that they are at the core of the balance between the rights of rightsholders and the rights of innovators, who need large amount of data for developing technologies which can foster innovation.

The fundamental rule intended to pursue said balance according to the principles in the international legal framework is the so called three-step test,[107] highlighting the

---

[104] For a distinction between artificial intelligence, deep learning and machine learning see video "AI vs Machine learning vs. deep learning: know the differences", simplilearn, https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/ai-vs-machine-learning-vs-deep-learning

[105] A schematic overview of the processes involved in text mining of scholarly content can be found on https://libereurope.eu/topic/text-data-mining/ See S. Ercolani, *Text and data mining: the copyright connection*, in Campus G, Franzosi M. Pollicino O. "Digital Single Market and Artificial Intelligence", Aracne Ed., 2024, 799 ss.

[106] Hearst, M.A. *Text Data Mining*, Mitkov, R. (ed.), The Oxford Handbook of Computational Linguistics, Oxford University Press: Oxford, UK, 2005; pp. 616–662.

[107] The Three-Step Test is found specifically in Article 9(2) of the Berne Convention and Article 13 of the TRIPS Agreement. It states that any limitation or exception to copyright must satisfy three criteria:

a. Special Cases: The limitation or exception must apply to certain special cases that do not conflict with the normal exploitation of the work.

b. No Conflict: exceptions must not conflict with the normal exploitation of the work; and

c. No Unreasonable Prejudice: The limitation or exception must not unreasonably prejudice the legitimate interests of the rights holder.

The Italian Copyright Law 633 of April 22, 1941 incorporates literally the three criteria in article 69-bis, para. 5, art. 70-sexies, 71-bis para. 3-octies, 71-sexies, para.4 and 71-nonies.

© European Audiovisual Observatory (Council of Europe) 2024

criteria to be taken into account by states when introducing exceptions and limitations to the exclusive rights. The three-step test is not only in the Berne Convention (Article 9 (2)) but also in the Agreement on Trade Related Aspects of Intellectual Property Rights - TRIPs (Article 13),[108] the WIPO Copyright Treaty (WCT, Article 10)[109] and the WIPO Performances and Phonograms Treaty (WPPT, Article 16).[110] In the EU, the three-step test is enshrined in art. 5.5 of the Infosoc Directive,[111] as well as in other directives.

In other jurisdictions, for example in the United States, a different approach is adopted, with potentially broader exception – to be adopted in the light of the three-step test – according to the principle of fair use,[112] which allows assessment on a case-by-case basis of whether certain uses of copyright works are admissible for transformative and non-commercial purposes.

3.2.2. TDM and the impact on reproduction and extraction rights

With regards to Directive (UE) 2019/790 on Copyright in the Digital Single Market (CDSM Directive),[113] Articles 3 and 4 are dedicated to text and data mining (TDM), that is the use of automated analytical techniques to analyse large amounts of text and data for research, innovation, and other purposes, with the aim to generate new insights, knowledge, and potentially new outputs, possibly based on the analysis of copyrighted content. Given the rise of genAI starting from November 2022 (with the launch of ChatGPT), it is relevant to highlight that, when the EU legislator introduced the TDM exception, the technical landscape was not focused on the possibility to generate new content via AI starting from the training data potentially collected on the basis of the text-and-data-mining exception.

---

[108] The TRIPS Agreement is a Protocol to the GATT of the World Trade Organization. WTO Members must comply with the substantive law provisions of the Berne Convention, except the provisions on authors' moral rights. International agreements concluded by the Union are, as from their entry into force, an integral part of the legal order of the European Union (Judgments of 30 April 1974, *Haegeman* (181/73, EU:C:1974:41, paragraphs 2/6); of 30 September 1987, *Demirel* (12/86, EU:C:1987:400, paragraph 7); and of 8 March 2011, *Lesoochranárske zoskupenie* (C-240/09, EU:C:2011:125, paragraph 30). They are therefore binding upon the institutions of the Union and on its Member States pursuant to Article 216(2) TFEU.

[109] https://www.wipo.int/wipolex/en/text/295166#P83_10885

[110] https://www.wipo.int/wipolex/en/text/295578

[111] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32001L0029

[112] According to the US Copyright Office, "Fair use is a legal doctrine that promotes freedom of expression by permitting the unlicensed use of copyright-protected works in certain circumstances. Section 107 of the Copyright Act provides the statutory framework for determining whether something is a fair use." Section 107 calls for consideration of the following four factors in evaluating a question of fair use: 1. Purpose and character of the use, including whether the use is of a commercial nature or is for nonprofit educational purposes; 2. Nature of the copyrighted work; 3. Amount and substantiality of the portion used in relation to the copyrighted work as a whole. 4. Effect of the use upon the potential market for or value of the copyrighted work.

[113] Margoni T., Kretschmer M., 2018/04/25, *The Text and Data Mining exception in the Proposal for a Directive on Copyright in the Digital Single Market: Why it is not what EU copyright law needs*,

© European Audiovisual Observatory (Council of Europe) 2024

In Article 3 a broader exception for TDM is introduced for research and cultural institutions, while in Article 4 narrower conditions are established for the general TDM exception, also dedicated to potentially commercial purposes.[114] There are, however, some common aspects, such as the exempted exclusive rights covering reproduction and extraction. As to the reproduction right, copyright contents are possibly copied onto the miner's storage facilities and through the subsequent automatic selection, they are copied (and/or adapted) into a new dataset by means of the analysis software; such reproduction may be merely transient and only consist of fragments of works.[115] Also for fragments, absent a copyright exemption, TDM would require the rightsowners' authorisation.

The term "extraction" in the provisions on TDM seems a clear reference to the exemption of the TDM from the *sui generis* right that reserves for the maker the "extraction or re-utilization of a substantial part" of the database. No explicit reference is made to the applicability of the TDM to the rights on adaptations or alteration, which may be considered a restricted act in view of article 12 of the Berne Convention,[116] and would represent for sure the core aspect in considering the TDM exception as the rationale for justifying the training of AI systems with copyrighted contents.

In order to foster innovation via the TDM exception also for commercial purposes, Article 4 introduces a general exception for individuals or organisations engaging in TDM activities. Between copyright, on the one hand, and innovation and research on the other, achieving a fair balance is more complex than in the case of Article 3, which opens the possibility to license the use of copyright contents for TDM. Article 4 has identified such balance in the right of "opt out", the prerogative that rightsowners can exercise by means of a reservation expressed "in an appropriate manner". When the copyrighted contents are made available online, the reservation should be exercised by machine-readable means.

At present, a few licenses have been announced between rightsowners and platforms (between OpenAI and the Associated Press),[117] while The New York Times prohibits using its content to train AI models[118] and French media such as Radio France and France 24 are implementing anti-scraping tools.[119]

---

[114] Geiger C., Frosio G., Bulayenko O., *The exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market Legal Aspects*, in Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2018-02,

[115] According to the CJEU, 4 October 2011, Joined Cases C-403/08 and C-429/08 (Premier League), §159, partial reproductions are covered by the reproduction right of Article 2 of the Infosoc Directive, where the fragments "contain elements which are the expression of the authors' own intellectual creation, and the unit composed of the fragments reproduced simultaneously must be examined in order to determine whether it contains such elements".

[116] Article 12 - Right of Adaptation, Arrangement and Other Alteration - Authors of literary or artistic works shall enjoy the exclusive right of authorising adaptations, arrangements and other alterations of their works.

[117] https://apnews.com/article/openai-chatgpt-associated-press-ap-f86f84c5bcc2f3b98074b38521f5f75a

[118] https://www.theverge.com/2023/8/14/23831109/the-new-york-times-ai-web-scraping-rules-terms-of-service

[119] https://www.lesechos.fr/tech-medias/medias/ia-les-medias-francais-sorganisent-face-a-la-collecte-de-donnees-par-les-robots-1973079

© European Audiovisual Observatory (Council of Europe) 2024

3.3. AI relevant legislations

3.3.1. EU AI Act and copyright: transparency rules and measures for TDM

The Regulation laying down harmonised rules on Artificial Intelligence (the "AI Act") is part of a much broader and more ambitious project being carried out by the von der Leyen Commission since as early as 2019, which *inter alia* includes the White Paper on AI – A European approach to excellence and trust[120] as well as the Proposal for a Directive on adapting non-contractual civil liability rules to Artificial Intelligence.[121] At the same time, the European Parliament has also undertaken considerable endeavors in the area of AI, particularly with regard to issues such as ethics, responsibility and copyright,[122] confirming the EU's intention to take the lead in identifying and regulating the management issues and legal parameters of artificial intelligence for the future.

The choice of a regulation – and its consequent direct applicability in EU member states as set forth in Art. 288 TFEU – rather than a directive is a clear indication of the direction of travel of the EU. Through the AI Act, in fact, the EU will actually be able to deploy a uniform discipline directly injected into the respective legal frameworks of each member state, in theory without the need for local transposition or implementation.

On July 12, 2024, the Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence[123] (the AI Act), was published in the Official Journal of the European Union. The AI Act will enter into force 20 days after its publication.

Article 53 on "Obligations for providers of general-purpose AI models" was introduced, with two distinct requirements related to copyright: (i) Section 1(c) requires providers of GPAI models to:

> *put in place a policy to respect Union copyright law in particular to identify and respect, including through state of the art technologies where applicable, the reservations of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790" and (ii) Section 1(d) requires them to: "draw up and make publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model, according to a template provided by the AI Office.*[124]

---

[120] European Commission, *White paper on artificial intelligence - a European approach to excellence and trust*, COM(2020) 65 final, 2020.
[121] European Commission, *Proposal for a Directive on adapting non contractual civil liability rules to artificial intelligence* (AI Liability Directive), COM(2022) 496 final, 2022.
[122] European Parliament, *Resolution of 20 October 2020 on a framework of ethical aspects of artificial intelligence, robotics and related technologies*, 2020/2012(INL).
[123] Available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689
[124] Keller P., A first look at the copyright relevant parts in the final AI Act compromise.

© European Audiovisual Observatory (Council of Europe) 2024

What the "sufficiently detailed summary" will consist of will be determined by a template to be developed by the EU's AI Office. Recital 107 indicates that the summary should be generally comprehensive rather than technically detailed, e.g. by listing "the main data collections or sets that went into training the model". Before the template is available, operators will need to develop industry best practices.[125]

On the other hand, the reservation of rights under the TDM exception for being effective – especially in an online environment – could imply the development of adequate "state of the art technologies", which are likely part of the Standardization Request already submitted by the European Commission to the European Standards Organisations (ESOs).[126]

3.3.2. AI and TDM exception: some national law proposals in Italy and Poland

On 23 April 2024, the **Italian government** published the text of a draft law[127] introducing regulatory provisions, concerning the use of Artificial Intelligence systems, to the Italian legal system ("AI Law Proposal").[128] The text was approved by the Council of Ministers and then submitted to the Italian Parliament for discussion on 20 May 2024.[129] With regards to training data, Article 24 of the AI Law Proposal also introduces a new Article 70-*septies* in the Italian Copyright Law ("The reproduction and extraction of works or other materials through artificial intelligence models and systems, including generative ones, are permitted in accordance with articles 70-*ter* and 70-*quarter*."). This proposed Article appears intended to strengthen the principle according to which, save for the case of scientific research purposes, copyright holders can opt out from the use of their content for text-and-data mining for commercial purposes. This provision is consistent with the principle already expressed in the EU AI Act Article 53 co 1 lett. c.

Poland is still in the process of implementation of the provisions of the 2019 Copyright in the Digital Single Market Directive into national law. In this particular case, the Polish government claims that the delay allowed it to properly consider the impact of genAI on copyright and come to the conclusion that training generative AI systems on copyrighted works does not in fact fall within the scope of the text and data mining exceptions contained in the Directive, since this type of permitted use was not conceived for artificial intelligence.[130]

---

[125] See Frank C. and Schmid G., AI, the Artificial Intelligence Act & Copyright.
[126] See https://artificialintelligenceact.eu/standard-setting/ and https://www.etsi.org/newsroom/blogs/entry/standardization-request-in-support-of-safe-trustworthy-artificial-intelligence.
[127] See Campus G., Artificial Intelligence and copyright: the Italian AI Law Proposal.
[128] See https://www.governo.it/it/articolo/comunicato-stampa-del-consiglio-dei-ministri-n-78/25501
[129] See https://www.senato.it/service/PDF/PDFServer/DF/437373.pdf
[130] See Keller P., TDM: Poland challenges the rule of EU copyright law.

© European Audiovisual Observatory (Council of Europe) 2024

3.4. Impact of case law

3.4.1. Overview of relevant cases on training data (USA and Europe)

The US Copyright class action against OpenAI: this class action was filed on 28 June 2023,[131] in United States District Court - Northern District of California, San Francisco Division by two authors (Paul Tremblay and Mona Awad), on behalf of themselves and other parties in the class action complaint, against OpenAI Inc. and others. The plaintiffs demanded a jury trial to recover injunctive relief and damages as a result and consequence of defendants' alleged unlawful conduct. According to the claimant, "*a large language model's output is therefore entirely and uniquely reliant on the material in its training dataset*" (see § I.3). Much of the material in OpenAI's training datasets comes from copyrighted works – including books written by plaintiffs – that were copied by OpenAI without consent, without credit, and without compensation. OpenAI has never revealed what books are part of its Books1 and Books2 datasets, which are the "training dataset [that] came from two internet-based books corpora" (see § V.30). OpenAI has justified its lack of information on the provenance of the datasets due to both "the competitive landscape and safety implications of large-scale models" (see § V.35).

The US class action against Google Bard for web scraping: another class action[132] was filed against Google in the United States District Court - Northern District of California for alleged web scraping (covering both copyright and privacy aspects) in the training of its AI tools, Bard, Imagen, MusicLM, Duet AI, and Gemini.[133] For developing its products, Google's AI model was pre-trained on an estimated 1.56 trillion words of "public dialog data and web text," drawn from Infiniset, an amalgamation of internet content meticulously selected to improve the model's conversational abilities (§ I.76).[134] In addition, the origin of the data used to train LaMDA,[135] the language model behind Google Bard, includes the C4 dataset. The C4 dataset, created by Google in 2020, is taken from the Common Crawl dataset, which is an open-source dataset but it is intended to be used for research and education and, according to the plaintiffs, it was never intended to be turned into an AI product for commercial use.[136]

The US class action against Meta LLaMA: two class actions against Meta are promoted by some copyright holders (mainly book authors), with regard to an alleged

[131] See Tremblay P. and Awad M. v. OpenAI INC. et al, No. 3:23-cv-03223.
[132] See J.L. v. Alphabet Inc, U.S. District Court for the Northern District of California, No. 3:23-cv-03440.
[133] https://fingfx.thomsonreuters.com/gfx/legaldocs/myvmodloqvr/GOOGLE%20AI%20LAWSUIT%20complaint.pdf .
[134] See https://medium.com/@taureanjoe/what-sites-were-used-for-training-google-bard-ai-1216600f452d and https://www.searchenginejournal.com/google-bard-training-data/478941/#close .
[135] See https://arxiv.org/pdf/2201.08239.pdf .
[136] See https://commoncrawl.org/ and https://www.forbes.com/sites/kalevleetaru/2017/09/28/common-crawl-and-unlocking-web-archives-for-research/?sh=1e3d3c233b83.

© European Audiovisual Observatory (Council of Europe) 2024

infringement of IP in their books and written works as far as training materials for LLaMA (Large Language Model Meta AI) are concerned. Such case law is interesting with respect to the reconstruction of the technology deployed by Meta and of the training methodology (at least from the plaintiff's perspective) but also because the court has had the chance to preliminarily evaluate the robustness of the claims.[137] The first class action *Kadrey v Meta* was filed on 7 July 2023,[138] in U.S. District Court for the Northern District of California - San Francisco Division. The second class-action *Chabon v Meta* was filed on 12 September 2023 before the same court.[139] Both complaints are essentially based on the same arguments and factual allegations. Meta notes that 85 gigabytes of the training data come from a category called "Books". According to the plaintiffs, in such category is included Bibliotik, a "shadow library" that has long been of interest to the AI-training community because of the large quantity of copyrighted material it contains (including plaintiffs' written works).

The District Court of Hamburg on LAION case: With regards to the EU, there is a German court case currently pending before the Hamburg regional court. A stock photographer is suing the non-profit organization LAION, which offers the LAION-5B dataset used for the training of large image-text models. The lawsuit alleges unlawful copying and aims to have the images removed from the training set. LAION in contrast relies particularly on the general TDM exception under Art. 4 DSM Directive, but also on the TDM exception for purposes of scientific research under Art. 3 DSM Directive (due to its non-profit nature), which does not provide for an 'opt-out'.[140] Some preliminary findings from the hearing phase, as reported,[141] pointed out that the Court held the disputed images as "lawfully accessible" on the stock photo site and that under Section 44b German Copyright Law copies under TDM exception can only be made "for the purpose of gathering information, in particular regarding patterns, trends and correlations" (and the Court tended towards accepting a use for gathering correlations). Another relevant point debated relates to the proper way to opt out, since Section 44b German Copyright Law requires that this happen – when in the online environment – in a machine-readable format (this means "plain text" opt-out are not sufficient online; an opt-out expressed via robots.txt file is needed).

3.5. Some (preliminary) conclusions on the case law

The above-mentioned cases are mainly at an early stage. Nonetheless they appear relevant for a number of reasons. First of all, because in their factual reconstructions it

[137] Available at https://storage.courtlistener.com/recap/gov.uscourts.cand.415175/gov.uscourts.cand.415175.62.0.pdf
[138] Available at https://www.courtlistener.com/docket/67569326/kadrey-v-meta-platforms-inc/
[139] Available at https://www.courtlistener.com/docket/67785353/chabon-v-meta-platforms-inc/
[140] See https://cepic.org/news/an-up-date-on-the-robert-kneschke-v-laion-e-v and https://www.heise.de/hintergrund/Stock-photographer-sues-AI-association-LAION-The-crux-with-AI-training-data-8988690.html
[141] See Brüß M. here and Graef O.R. here.

© European Audiovisual Observatory (Council of Europe) 2024

appears evident what the crucial issue is with the training data for the most prominent LLMs. Therefore, some of the first rules in the AI Act dedicated to the training data are specifically transparency rules aimed at shedding some light on the training process.

The second point of relevance relates to the arguments used by the genAI providers to respond to the plaintiffs' allegations. They leverage the fact that the plaintiffs were not able to demonstrate how, based on the functioning of the LLMs, the training data are converted into outputs and whether they can be considered derivative works (mainly, the allegations note the similarity between works used for training and outputs).

In some cases, a *fair use* defense has also been introduced. *Fair use* is an exception to copyright law designed to allow limited use of copyrighted material without permission for purposes like commentary, criticism, news reporting, and scholarly reports.[142] But the counterargument is that the defendants' collection and use of copyrighted material, with no option for copyright owners to opt out, would exceed the legal interpretation of *fair use*, since copying an entire work militates against a finding of fair use.[143]

It will be interesting to note whether the US and EU case law will find coherent or divergent solutions on the issue of training data, taking into account that both US and EU approaches to copyright exceptions should be interpreted in line with the three-step test under the Berne Convention.

[142] See *McGucken vs Pub Ocean Limited*, 42 F.4th 1149 (9th Cir. 2022)
[143] See *VHT vs Zillow Group*, 918 F.3d 723, 743 (9th Cir. 2019); *Worldwide Church of God vs Phila. Church of God, Inc.*, 227 F.3d 110, 1118 (9th Cir. 2000).

© European Audiovisual Observatory (Council of Europe) 2024

PART III – Legal status of prompts in genAI

In discussions on AI-assisted or AI-generated content, the focus often centres on potential copyright infringement, especially within the audiovisual industry, where rightsholders are concerned about their works being used as training data.

But what about the prompt itself? Could a prompt be protected by law?[144] Might it qualify as a trade secret if its use proves to be significantly beneficial to a company? Since the prompt instructs the genAI and influences the resulting output, should it be given legal importance?

As to copyright infringement, prompts are used by individuals, making it difficult for rightsholders to detect whether prompts may lead to infringement. Should prompting activities be subjected to scrutiny? Such scrutiny could potentially conflict with users' freedom to express themselves through prompts. However, when balancing interests, is it justifiable to limit one's freedom to protect something greater? Could copyright protection have an effect on freedom of expression?

These questions remain, at the time of writing, still theoretical, and it is yet to be seen how human rights frameworks will address these challenges.

[144] Rethinking Copyright Law: The Case for Protecting AI-Generated Content and Rewarding Those Who Truly Know What They Want, Ziyong "Sean" Li, Benesch, 14 May 2024

4. Authorship, Liability and Transparency in relation to AI-generated content

Malte Baumann and Jan Bernd Nordemann[145]*, Attorneys at Law, Law firm NORDEMANN, Berlin;*

4.1. Authorship

4.1.1. The human creator as author

Under EU law, only a human creation can enjoy copyright protection. To satisfy the definition of a work, a given subject matter must reflect the personality of its author as an expression of his or her free and creative choices.[146] As such the focus is on the human creator and his or her actions in shaping the work. Authors are able to give their works a "personal touch" through their personal choices and their use of the freedom available to them.[147] A purely aesthetic effect that is not the result of a personal, creative choice is not sufficient to justify protection as a work.[148] Moreover, copyright protection cannot be afforded if the design of a product is dictated by technical considerations, rules or constraints.[149]

This anthropocentric approach of EU law can be seen not only in the criterion of originality but also in the term of protection, which is based on the life of the author.[150] In

[145] Prof. Dr. Jan Bernd Nordemann (Partner) and Dr. Malte Baumann (Associate), Attorneys at Law. Jan Bernd Nordemann is also honorary professor at the Humboldt University in Berlin
[146] *Cofemel*, Judgment of 12 September 2019, C-683/17; Eva-Marie Painer, Judgment of 1 December 2011, C-145/10.
[147] *Eva-Marie Painer*, Judgment of 1 December 2011, C-145/10.
[148] *Cofemel*, Judgment of 12 September 2019, C-683/17.
[149] *Football Dataco*, Judgment of 1 March 2012, C-604/10.
[150] Article 1 Directive 2006/116/EC.

© European Audiovisual Observatory (Council of Europe) 2024

the Berne Convention, the concept of moral rights underlines the human-based approach.[151]

Courts in EU member states (such as Czechia) have already applied this principle and found that only a human can be an author, AI cannot.[152] Some countries (like France) have begun producing legislative proposals for dealing with AI, which clarify that AI cannot itself be the author of a work.[153] In other countries (such as Spain), the copyright laws already leave no room for doubt by explicitly stipulating that only natural persons can be authors.[154]

Looking to the USA, the principle that only humans can be authors was already established in copyright law even before the new age of AI.[155] Accordingly, the District Court of Columbia ruled in 2023 that material the expression of which is solely attributable to an artificial system running on a machine does not enjoy copyright protection.[156] The US Copyright Office has maintained this principle and refuses to grant copyright protection to purely AI-generated material.[157] Only material that is the product of human creativity can be protected by copyright.[158] The Guild agreements of the Writers Guild of America (WGA) also follow this approach.[159]

China has also adopted the principle that AI models cannot themselves be authors.[160] A copyrightable work always requires an intellectual act on the part of a person.

Some jurisdictions (such as the United Kingdom and Ireland) are taking a different path by expressly recognising protection for computer-generated content.[161] However, even in those jurisdictions, authorship is attributed to the person who created the

---

[151] Article 6[bis] of the Berne Convention for the Protection of Literary and Artistic Works of 9 September 1886, WIPO Lex No. TRT/BERNE/009; also Hugenholtz, P.B. and Quintais J.P. "Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output?", *IIC - International Review of Intellectual Property and Competition Law 52*, 2021, pp. 1190-1216.
[152] Cerri A., "Czech court finds that AI tool DALL-E cannot be the author of a copyright work", *The IPKat*, 15 April 2024.
[153] Dreyfus, "Deciphering French Copyright Law in the Age of AI: A Critical Analysis of Recent Developments", *Dreyfus*, 19 January 2024
[154] Article 5, *Real Decreto Legislativo 1/1996*, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia (Spanish Copyright Act of 12 April 1996).
[155] United States Court of Appeals for the Ninth Circuit, judgment of 23 April 2018, No. 16-15469 [888 F.3d 418].
[156] United States District Court for the District of Columbia, judgment of 18 August 2023, Civil Action No. 22-1564 (BAH) [2023 WL 5333236 (D.D.C. Aug. 18, 2023)]
[157] US Copyright Office, "Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence", 16 March 2023.
[158] Ibid.
[159] Article 72 B, Memorandum of Agreement for the 2023 WGA Theatrical and Television Basic Agreement of 25 September 2023.
[160] Beijing Internet Court, Judgment of 27 November 2023, (2023) Jing 0491 Min Chu No. 11279.
[161] United Kingdom: Section 178, Copyright, Designs and Patents Act 1988, 15 November 1988, Ireland: Section 21 (f), Copyright and Related Rights Act, 2000, 10 July 2000.

© European Audiovisual Observatory (Council of Europe) 2024

conditions necessary for the material to be produced.[162] In Ukraine, a *sui generis* right in computer-generated content has recently been introduced, which is explicitly vested in the holder of the rights in the computer program.[163]

4.1.2. AI-assisted creation of works

Under EU law, output generated by a computer alone does not enjoy protection as a copyrighted work. In practice, however, there will often be some form of human influence.[164] This human contribution can theoretically suffice as a basis for copyright protection as the use of technical tools does not preclude copyright protection.[165]

What form must human influence take for the AI output to be attributed to the person as their creation of a work? There are three possible phases in which humans can exert a decisive influence on the generation of the AI output:

- In the creation and configuration of the AI model and AI system (selection of training data; programming of the AI system and setting its purpose; targeted training of the model);
- In the specifications made to AI through prompts;
- In the editing/reworking of the draft output produced by AI.

This corresponds to the creative phases developed by the CJEU, in the context of portrait photography, which offer areas for creative freedom: preparatory steps, the execution itself and the subsequent revision of the output.[166]

According to the CJEU, even a very limited degree of human creativity is sufficient to justify protection as a work. The crucial factor is that there is any freedom for individual choices at all.[167] This freedom does not have to be particularly great nor does it have to be used in a ground-breaking way: even just an extract of 11 words from a daily newspaper can constitute a protected work,[168] as can a quite simple portrait photograph.[169] In contrast, merely collating factual information does not suffice.[170] Most prompts will meet the requirements.

---

[162] United Kingdom: Section 9 (3), Copyright, Designs and Patents Act 1988; Ireland: Section 21 (f), Copyright and Related Rights Act, 2000.
[163] *Закон України № 2811-IX від 01.12.2022 Про авторське право і суміжні права* (Law No. 2811-IX of 1 December 2022 on Copyright and Related Rights), amended by No. 2974-IX of 20 March 2023.
[164] Militsyna K., "Human Creative Contribution to AI-Based Output - One Just Can('t) Get Enough", *GRUR Int.*, 2023 pp. 939-949.
[165] *Eva-Marie Painer*, Judgment of 1 December 2011, C-145/10.
[166] *Eva-Marie Painer*, Judgment of 1 December 2011, C-145/10; Hartmann C. et al., *Trends and developments in artificial intelligence*, Publications Office of the European Union, September 2020, p. 73.
[167] *Football Dataco*,, Judgment of 1 March 2012, C-604/10.
[168] *Infopaq*, Judgment of 16 July 2009, C-5/08
[169] *Eva-Marie Painer*, Judgment of 1 December 2011, C-145/10.
[170] *Funke Medien*, Judgment of 29 July 2019,C-469/17.

© European Audiovisual Observatory (Council of Europe) 2024

However, it is not only important that the scope for decision-making is used creatively but also that these personal decisions are reflected in the final AI output. The specific expression must reflect the free creative choices of the person.[171] The intervention of AI must therefore not completely overshadow the input by the human. This is in line with the general principle that a mere idea as such cannot enjoy copyright protection but only the concrete expression of it.[172]

The CJEU itself has not yet ruled on any AI-specific cases regarding the creation of works. However, there are some national judgments and decisions by public authorities. These show that the question as to what specific requirements should be placed on the human creative contribution can be answered with varying degrees of strictness. Ultimately, despite the efforts of legal experts to develop generalised assessment methods,[173] there will have to be a case-by-case assessment taking into account the standards that apply nationally. The AI tool used and the degree of automation will play just as important a role as the extent and quality of the specific human contribution.

French courts have so far applied a rather generous set of criteria. While lower courts established early on that computer-assisted creations can also be protected, the Cour d'appel de Bordeaux (Bordeaux Court of Appeal) only required a minimal degree of human originality.[174]

In China, a court decided a case in which a user made extensive and targeted specifications in over 100 prompts.[175] That was sufficient for the court to affirm copyright protection. According to the court, the clearer the specifications given in the prompts, the more likely the output will reflect the personal human expression.

The US Copyright Office, on the other hand, is stricter and sees the prompts purely as instructions to AI. It takes the view that AI is responsible for the specific expression and that this is generally not sufficient for copyright protection.[176] However, the US Copyright Office also points out that an artistic collage of AI-generated content or a human revision of AI content can justify protection.

---

[171] *Cofemel*, Judgment of 12 September 2019, C-683/17; Militsyna K., "Human Creative Contribution to AI-Based Output - One Just Can('t) Get Enough", *GRUR Int.*, 2023, pp. 939-949.
[172] Article 9(2), Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) of 15 April 1994,; Article 2, WIPO Copyright Treaty (WCT) of 20 December 1996; Article 1(2) of Directive 2009/24/EC.
[173] Militsyna K., "Human Creative Contribution to AI-Based Output - One Just Can('t) Get Enough", *GRUR Int.*, 2023, pp. 939-949; Hugenholtz P.B. and Quintais J.P., "Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output?", *The International Review of Intellectual Property and Competition Law 52*, 2021, pp. 1190-1216.
[174] Hartmann C. et al., *Trends and developments in artificial intelligence*, Publications Office of the European Union, September 2020, p. 82.
[175] Beijing Internet Court, Judgment of 27 November 2023, (2023) Jing 0491 Min Chu No. 11279.
[176] US Copyright Office, "*Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence*", 16 March 2023.

© European Audiovisual Observatory (Council of Europe) 2024

4.1.3. Who is the author?

In most cases, if there is deemed to be an author at all, it will be the user of the AI output. This applies, for example, when the users edit the AI output in such a way that it reflects their personality. As already mentioned, it is also conceivable that a prompter inputs instructions that are so specific that the concrete expression of the output sufficiently reflects their creative choices.

While creative choices made by developers during the development of an AI System can also lead to copyright protection, this protection will normally cover the software code. The choices made during the development of the AI system will usually not have a sufficient connection to the concrete expression of the output. This is because most generative AI systems are specifically intended to have a broad range of uses and are not created to produce one particular work.[177] The developers of the AI system create the tool but not the work.

4.1.4. Protection through neighbouring rights

Some neighbouring rights are not linked to a human creative achievement but protect investments or economic and organisational efforts. Particularly relevant in relation to AI output in the audiovisual sector is the neighbouring right of the film producer.

EU law has partially harmonised the neighbouring right of the film producer through directives.[178] According to Article 2(1)(c) of the Rental Directive, both cinematographic works and simple moving images that do not qualify as works fall under the definition of a 'film'. Beyond the EU, there are no international treaties that govern the neighbouring right of the film producer.[179]

Creating a video with the help of AI tools can require an economic and organisational effort that suffices to give rise to a neighbouring right.[180] The protection afforded may be justified on the basis of the effort required in the procurement of the software and hardware, the integration into the work processes and products, the conception of the content, the necessary rights clearance as well as the need for skilful prompting. As simple moving images are covered by the European definition of a film,

---

[177] Militsyna K., "Human Creative Contribution to AI-Based Output - One Just Can('t) Get Enough", *GRUR Int.*, 2023, pp. 939-949.
[178] Article 3(1) (d) of Directive 2006/115/EC, Article 3(3) of Directive 2006/116/EC, and Article 2(d) and Article 3(2)(c) of Directive 2001/29/EC.
[179] Loef R. and Verwehen U., "One more Night – Überlegungen zum abgeleiteten fremdenrechtlichen Filmherstellerschutz", *Zeitschrift für Urheber- und Medienrecht*, 2007, pp. 706-711.
[180] Baumann M., "Presseleistungsschutzrecht: Der Schlüssel zum Schutz KI-generierter Erzeugnisse?", AfP – *Zeitschrift für Medien- und Kommunikationsrecht*, 2024, pp. 193-197; Ebers M. et al., "§ 9 KI und Urheberrecht", *Künstliche Intelligenz und Robotik*, Ebers M. et al. (eds.), C.H.Beck, 2020, marg. no. 70; Becker M., "Das Urheberrecht als Trostpreis für den Menschen? Die überraschende Verteilung von Leistungsschutzrechten für KI-Erzeugnisse", *GRUR*, 2024, pp. 505-514.

© European Audiovisual Observatory (Council of Europe) 2024

even videos generated completely by AI could enjoy protection. It is therefore irrelevant whether or not the film qualifies for protection as a work according to the criteria set out above.[181] The neighbouring right is created in connection with the physical medium on which the film was first fixed, regardless of the content. The holder of the right is the person or entity who provides the organisational and economic effort involved.

The neighbouring right of the broadcasting organisation[182] also protects the broadcast material irrespective of the content. As such, a broadcasting organisation can receive rights in audiovisual AI content that does not reach the threshold for protection as a work.[183]

4.2. Liability for AI output

4.2.1. When is an infringement deemed to have occurred?

As a basic principle, it can be assumed that the traditional and established general rules of copyright law have to be applied when answering the question of whether or not output can be deemed to have caused an infringement.[184] As such, the existing provisions under EU copyright law also apply to audiovisual AI output.

EU copyright law contains explicit provisions for adaptations only for certain types of work, such as software[185] or copyright-protected databases.[186] For other types of works, and in particular for audiovisual works, only the copyright laws of the individual EU member states (for example Belgium,[187] France[188] or Germany[189]) contain explicit provisions on adaptations requiring authorisation. Rights of reproduction are fully harmonised under EU law in Article 2 of the InfoSoc Directive.[190] Harmonisation also covers reproduction in

[181] Hartmann C. et al., *Trends and developments in artificial intelligence*, Publications Office of the European Union, September 2020, p. 91.
[182] See Article 13 of the Rome Convention for the Protection of Performers, Producers of Phonograms and Broadcasting Organisations of 26 October 1961,; Article 3(2)(d) of Directive 2001/29/EC; Article 14(3) of TRIPS.
[183] Becker M., "Das Urheberrecht als Trostpreis für den Menschen? Die überraschende Verteilung von Leistungsschutzrechten für KI-Erzeugnisse", *GRUR*, 2024, pp. 505-514.
[184] Finke M., "Urheberrechtliche Zulässigkeit der Nutzung des Outputs einer Künstlichen Intelligenz", *Zeitschrift für Geistiges Eigentum*, 2023, pp. 414-444, "Generative KI: Eine "Blackbox" urheberrechtlicher Haftungsrisiken? Balanceakt zwischen Innovationsförderung und effektivem Rechtsschutz für Werke Dritter", MMR – Zeitschrift für IT-Recht und Recht der Digitalisierung, 2024, pp. 298-304.
[185] Article 4(1)(b) of Directive 2009/24/EC.
[186] Article 5(b) of Directive 96/9/EC.
[187] Article 1(1), *Loi n° 2006-961 du 1 août 2006 relative au droit d'auteur et aux droits voisins dans la société de l'information* (Law no. 2006-961 of 1 August 2006 on copyright and related rights in the information society).
[188] L122-4, *Code de la propriété intellectuelle* (French Intellectual Property Code).
[189] § 23 (1), *Urheberrechtsgesetz* (German Copyright Act).
[190] *Infopaq*, Judgment of 16 July 2009, C-5/08.

© European Audiovisual Observatory (Council of Europe) 2024

part.[191] When it comes to the reproduction right in relation to the neighbouring right of the phonogram producer, the CJEU has also found that modified reproductions also fall under the concept of a reproduction if the original is still recognisable despite the modification.[192]

According to the Bundesgerichtshof (German Federal Court of Justice), this case law is also applicable to the reproduction right of genuine copyright in accordance with Article 2(a) of the InfoSoc Directive.[193] However, the Svea hovrätt, Patent- och marknadsöverdomstolen (Svea Court of Appeal, Patent and Market Court of Appeal) has referred a question to the CJEU for clarification concerning the applicability of the CJEU's Pelham case law on the neighbouring right of the phonogram producer to exploitation rights of the copyright author.[194] It is not entirely clear how the CJEU will rule on this question. But if one assumes that the CJEU will apply its case law in the Pelham case to the right of reproduction under genuine copyright as per Article 2 of the InfoSoc Directive, the only relevant factor under EU law when assessing whether an infringement has occurred is recognisability. Accordingly, the question would then be: may the copyright-protected elements of an earlier work be recognised in the newly created (later) work?[195]

Applying this to AI output, the question to be asked is whether copyright-protected elements from earlier works are recognisable in the AI output. That said, there is no principle of priority in copyright law, meaning that independent creations are not considered copyright infringements. They are not deemed to be a reproduction of the earlier work because the author of the later work has created their work independently. To date, there is no CJEU case law on the question of when a work constitutes independent creation. Nevertheless, the principle that independent creations are permitted is widely recognised throughout the EU.[196] Historically, the claim that a work is an independent creation has only been used successfully on rare occasions. In Germany, for example, similarities in terms of relevant creative parts generally suffice as *prima facie* evidence that the work in question is a modified reproduction requiring authorisation. This *prima facie* assumption can only be rebutted if it is likely, from the relevant circumstances, that there is another explanation for the similarities than that the creator of the later work drew from the earlier work.[197]

---

[191] *Infopaq*, Judgment of 16 July 2009, C-5/08.
[192] *Pelham*, Judgment of 29 July 2019, C-476/17.
[193] Bundesgerichtshof, Judgment of 7 April 2022, I ZR 222/20.
[194] *Mio and others*, Request for a preliminary ruling, C-580/23.
[195] Bundesgerichtshof, judgment of 7 April 2022, I ZR 222/20; *Mio and others*, request for a preliminary ruling, C-580/23.
[196] See Peukert A., "Copyright in the Artificial Intelligence Act – A Primer", *GRUR Int.*, 2024, pp. 497-509; Iaia V., "To Be or Not to Be...Original Under Copyright Law, That Is (One of) the Main Questions Concerning AI-Produced Works", *GRUR Int.*, 2022, pp. 807-812; Inguanez D., "A Refined Approach to Originality in EU Copyright Law in Light of the ECJ's Recent Copyright/Design Cumulation Case Law", *International Review of Intellectual Property and Competition Law*, 2020, pp. 797-822,; on the legal situation in the United Kingdom: Guadamuz A., "A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs", *GRUR Int.*, 2024, pp. 111-127.
[197] Bundesgerichtshof, judgment of 3 February 1988, I ZR 142/86.

© European Audiovisual Observatory (Council of Europe) 2024

As far as independent creation is concerned, the question is whether this principle can apply in the case of AI output. One might assume that it can.[198] The principle would be: if an AI system has been trained on recognisable work, the output does constitute a copyright infringement but if an AI system produces an output which is purely coincidentally similar to another work, without AI having been trained on that work, this will constitute an acceptable independent creation.[199] Rules for a *prima facie* assumption should be applied also here.

4.2.2. Exceptions and limitations to copyright applicable to AI output

Depending on the specific output, exceptions and limitations to copyright might apply, in particular those under Article 5(3)(k) of the InfoSoc Directive for parody, caricature or pastiche. Other exceptions and limitations, specifically in regard to audiovisual content, also have to be considered. The provision under Article 4 DSM Directive for text and data mining is not applicable to AI output, however.

4.2.3. Responsibility of the user

There are no AI-specific provisions at EU level that target the issue of responsibility. For the question as to who is liable for the use of AI output, there seems to be a compelling case for applying the existing principles, albeit slightly modified.

First of all, the general rules should apply where AI users use the AI output themselves in a manner that has copyright relevance. This would mean that AI users bear responsibility whenever they reproduce AI output (Article 2 of the InfoSoc Directive), distribute it (Article 4 of the InfoSoc Directive) or communicate it to the public (Article 3 of the InfoSoc Directive). While the AI user directly carries out the use and is liable as perpetrator, it is conceivable that he or she does act with negligence (lack of culpability), hence claims for damages are excluded. At the end of the day, anyone using AI tools will have to check whether the output contains elements of third-party works – even if certainty in this regard will ultimately not always be possible. If rights-infringing AI output is communicated to the public via a hosting provider, the CJEU liability model from

[198] Also of this opinion, on UK copyright law: Guadamuz A., "A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs", *GRUR Int.*, 2024, pp. 111-127; on German copyright law: Käde A., "Do You Remember?", *Zeitschrift für Urheber- und Medienrecht*, 2024, pp. 174-183; Nordemann J.B., "Generative AI, copyright infringements and liability – My guess for a hot topic in 2024", *Kluwer Copyright Blog*, 23 January 2024; Baumann M., "Generative KI und Urheberrecht - Urheber und Anwender im Spannungsfeld", *Neue Juristische Wochenschrift*, 2023, pp. 3673-3678.

[199] Baumann M., "Generative KI und Urheberrecht - Urheber und Anwender im Spannungsfeld", *Neue Juristische Wochenschrift*, 2023, pp. 3673-3678.

© European Audiovisual Observatory (Council of Europe) 2024

"YouTube and Cyando"[200] can be applied. That liability model holds video platforms and other hosting providers liable for infringements of the right of communication to the public that they have indirectly caused, provided firstly that they play an indispensable role in the infringement and secondly that they have breached duties of care.

Another conceivable outcome is liability of the AI provider. There is currently no specific provider liability for copyright infringements in the area of generative AI. The European Commission's draft revision of the Product Liability Directive[201] expressly includes software as a product as per the definition in Article 4(1). However, copyright infringements still do not fall within the scope of damage covered by that Directive.[202] The draft AI Liability Directive also does not include any provisions regarding the attribution of liability in the event of copyright infringements by AI output, instead it only facilitates the enforcement of rights.[203]

The existing rules must therefore be applied to AI providers. It should be noted, however, that in most cases this will only concern liability for unauthorised reproduction in the form of AI output (Article 2 of the InfoSoc Directive).

One possible solution is to utilise the CJEU liability model for parties that indirectly cause infringements, as already mentioned above. Up to now, the CJEU liability system has only been applied to the right of communication to the public. It would make sense, however, to extend it also to cover other exploitation rights like the right of reproduction.[204] One particular argument in favour of this is that it could be used to create a well-balanced liability system also for AI providers. Moreover, the content is not purely user-generated as would be the case, for example, for typical hosting platforms. The content is determined to a significant degree by the AI system. AI providers also have the means (filters, blacklists, metaprompts, red teaming, training methods, etc.) to integrate measures to reduce the risk of infringements into the system itself. In this regard, the AI Act stipulates that providers of general-purpose AI models must develop a policy to ensure compliance with EU copyright law (Article 53(1)(c) of the AI Act). Against this background, it makes sense that both users and providers can be liable for the output and to impose duties of care on both.

For the CJEU liability model to apply, it is firstly important that the AI provider plays an indispensable role. In the *Ocilion IPTV Technologies*[205] case, the CJEU found that hardware and software providers do not play an indispensable role as they lack an ability to influence the infringement. In that case, the software and hardware provider supplied

---

[200] *YouTube/Cyando*, Judgment of 22 June 2021, C-682/18 and C-683/18.
[201] COM(2022) 495 final.
[202] Baumann M., "Generative KI und Urheberrecht - Urheber und Anwender im Spannungsfeld", *Neue Juristische Wochenschrift*, 2023, pp. 3673-3678.
[203] Ibid.
[204] Nordemann J.B., "Neu: Täterschaftliche Haftung von Hostprovidern im Urheberrecht bei (Verkehrs) Pflichtverletzungen im Internet", *Zeitschrift für Urheber- und Medienrecht*, 2022, pp. 806-816,; Nordemann J.B. "Generative AI, copyright infringements and liability – My guess for a hot topic in 2024", *Kluwer Copyright Blog*, 23 January 2024; of a different opinion: *Sony Computer Entertainment Europe*, Opinion of Advocate General Szpunar of 25 April 2024, C-159/23.
[205] *Ocilion IPTV Technologies*, Judgment of 13 July 2023, C-426/21.

© European Audiovisual Observatory (Council of Europe) 2024

its services to a third party and only that third party was in contact with the end customers who were recipients of infringements.[206] It should be noted, however, that in the CJEU case, the end users were not the infringers (of the right of communication to the public) either but only received the infringements and did not have any influence on the infringements themselves. Accordingly, it is possible that a software and hardware provider - such as an AI provider - does play the required indispensable role if it makes the infringement (in the form of an unauthorised reproduction) available to the user. An AI provider that outputs an infringement to a user would therefore be deemed to be playing an indispensable role. There would then be certain duties of care incumbent upon the AI provider if it wishes to avoid liability.

A nuanced assessment and thus a nuanced generation of duties of care appears to be called for when it comes to copyright-infringing AI output. The AI system can determine the content of the output to a significant degree. Therefore, providing an AI system involves more than just providing software that allows users to create reproductions at their own discretion. The duties of care must be attributed according to who induced the actual infringing content.

To the extent that AI is merely a technical tool for the user and the key parameters for the determination of the content are set by the AI user (e.g. through his or her prompts), only the AI user may be considered as a perpetrator. Example: the AI user's prompts are designed to generate infringing content. The situation is different, however, if the infringements were induced primarily by the generative AI. In that case, liability could be attributed to the AI provider on the grounds of a breach of duty of care. This would be the case if, for example, an AI user has only input very minor specifications in their prompts and the AI system has ultimately generated the infringement autonomously. The AI provider would have a duty of care at least to prevent clear copyright infringements. That would apply even if the AI provider was previously unaware that its AI system was capable of generating the infringement in question.

If the user is primarily responsible for the AI output, then the duties of care have to be limited. Even then, however, the AI provider could bear some responsibility. After all, AI still generates the content and plays an indispensable role in the infringement (see above). A conceivable way to deal with such cases could be to apply the three duties of care set out in YouTube/Cyando[207] in a slightly modified manner:[208]

- Upon becoming aware of the problem, the AI provider would have to do everything technologically possible that could reasonably be expected of it, to prevent the infringement being output again.
- The AI provider must, if it is aware or ought to be aware that users are reproducing protected content illegally via its system, put in place the appropriate technological measures that can be expected from a reasonably diligent provider

[206] Ibid.
[207] *YouTube/Cyando*, Judgment of 22 June 2021, C-682/18 and C-683/18.
[208] Nordemann, J.B., "Generative Künstliche Intelligenz : Urheberrechtsverletzungen und Haftung", *GRUR*, 2024, pp. 1-2.

© European Audiovisual Observatory (Council of Europe) 2024

in its situation in order to counter copyright infringements in a credible and effective manner.

- The AI provider may not participate in the selection of protected content that is illegally reproduced, nor may it provide tools specifically intended for the illegal reproduction of such content, nor may it knowingly promote such reproduction. One factor which could suggest that an AI provider is knowingly promoting such reproduction would be if the provider has adopted a financial model that encourages its users to have protected content reproduced as AI output.

It will have to be discussed further, if this path could be followed further.

4.2.4. providers' terms of use

The terms of use of most AI providers prohibit users from generating illegal content and postulate that the user alone is responsible for the content generated. The provisions in this regard are often accompanied by indemnity clauses that protect the AI provider.[209] As such, the AI providers are attempting to relieve themselves, as far as possible, of any responsibility for the content.

Such provisions generally have no effect on who bears liability, as far as third parties like rightsholders are concerned. They are very much relevant, however, in the internal contractual relationship between user and provider. If a user (deliberately) generates infringing material, he or she breaches contractual obligations and the AI provider has claims for recourse against the user if the provider themselves is subject to claims for rights infringements.

As such, the fact that some AI providers offer indemnity clauses for copyright-infringing output is especially relevant for users.[210] However, AI providers often incorporate a broad catalogue of exceptions which significantly limit the extent of the indemnification.

4.2.5. Reducing potential liability

No conclusive analysis has yet been performed on the probability that individual AI systems will generate rights-infringing output. It is already becoming apparent, however,

[209] Such indemnity clauses can be found in almost all AI providers' terms of use. By way of example, Section 11.3 of the Terms of Use of Mistral AI.
[210] See, for example, 50.10. of the AWS Service Terms or the Customer Copyright Commitment for Microsoft Azure OpenAI Services.

© European Audiovisual Observatory (Council of Europe) 2024

that this probability will be largely determined by the specific intended purpose of the AI system, the training data or the type of training[211] and the prompts.[212]

There are also other ways to further minimise the risk of producing rights-infringing output. One way is to use an AI tool that has only been trained with one's own or licensed material. Even if the foundation model has been trained with a wide range of works, a specific second training using one's own materials can reduce the system's tendency to produce infringing output.

When wording their prompts, users can avoid referencing protected works. This precautionary measure has the equivalent effect of "blacklisting" by the provider, where the provider blocks certain prompts. In addition, "metaprompts" offer the possibility of writing general instructions for the system. These metaprompts can be used to reduce the probability that the AI system generates rights-infringing content.[213]

AI providers can carry out regular evaluations to determine the systemic risk of rights infringements (red teaming) and filter out reported rights-infringing content from the output.[214]

4.2.6. Transparency

Transparency regarding the fact that the output in question has been artificially generated protects the recipients. This can protect consumers but also end-users acting in a professional capacity. For the latter, it is especially important that the content is eligible for protection so that they can license it. In addition, clients need to be able to assess whether their suppliers are using AI in order to gauge potential liability. The normal guarantees in the area of film, that all rights in the supplied material are held by the supplier, have to be critically scrutinised. If AI has been extensively used, it will not be possible to provide this guarantee with absolute certainty. As a result, the number of users demanding comprehensive disclosure of the use of AI in their contractual agreements increases.

It is worth mentioning that it may be important for the user to disclose the use of AI and specifically to inform contractual partners about it even if for liability reasons alone. This applies particularly if the intention is to grant exclusive exploitation rights in

[211] Militsyna K., "Human Creative Contribution to AI-Based Output - One Just Can('t) Get Enough", *GRUR Int.*, 2023, pp. 939-949; Pesch P.J. and Böhme R., "Artpocalypse now? - Generative KI und die Vervielfältigung von Trainingsbildern", *GRUR*, 2023, pp. 997-1007.
[212] Marcus G. and Southen R., "Generative AI Has a Visual Plagiarism Problem. Experiments with Midjourney and DALL-E 3 show a copyright minefield", *IEEE Spectrum*, 24 June 2024; Carlini N. et al., "Extracting Training Data from Diffusion Models", *arXiv*, 30 January 2023; Henderson P. et al., "Foundation Models and Fair Use", *Journal for Machine Learning Research 24*, 2023, pp. 1-79.
[213] See Microsoft Azure, Customer Copyright Commitment Required Mitigations.
[214] Ibid.

© European Audiovisual Observatory (Council of Europe) 2024

the output. Under the Writers' Guild agreements in the USA, screenwriters even have to obtain their client's (film producers) permission in advance if they want to use AI.[215]

Limited transparency obligations also exist within the AI Act. According to Article 50(1) of the AI Act, providers of AI systems must ensure that users can tell that they are interacting with AI. However, this obligation only applies to *direct* interaction and not every form of output. Providers of general-purpose foundation models, however, have a comprehensive obligation to label all output. Article 50(2) of the AI Act stipulates that such providers must generally mark all AI output, in a machine-readable format, as artificially generated.

Those using AI tools to create deepfakes must disclose this (Article 50(4) of the AI Act). It should be noted in this regard that the definition of 'deep fake', in Article 3, number (60) of the AI Act, is very broad: it covers all image, audio or video content that resembles existing persons, objects or places and that would seem to a person to be authentic. How dangerous the deception is for the individuals concerned is not relevant. In the area of film specifically, AI can be used to generate real places as artificial sets or AI lookalikes of real people. These are then deep fakes.

In conclusion, the use of AI affects a number of different areas within copyright law. The AI Act regulates the issue from the perspective of product safety law and leaves open many questions regarding the traditional and established general rules of copyright law. Despite the autonomy of AI systems, the focus remains centred on the human: copyright protection depends on the human contribution and liability depends on humans meeting duties of care. Transparency obligations regarding the output are aimed primarily at the legitimate interests of human recipients.

[215] Article 72 D, Memorandum of Agreement for the 2023 WGA Theatrical and Television Basic Agreement of 25 September 2023.

© European Audiovisual Observatory (Council of Europe) 2024

5. Personality Rights & Transparency

Kelsey Farish, Media and entertainment business affairs lawyer, Reviewed & Cleared, London

5.1. Setting the Scene

When artificial intelligence company OpenAI released the ChatGPT-4o system in May 2024,[216] people noted the model's impressive text-to-speech capabilities, including its seemingly flawless mimicry of vocal intonations across multiple languages. But Hollywood star Scarlett Johansson noticed something else: she claims the platform's voice sounds "so eerily similar" to hers that even her closest friends could not tell the difference.[217]

Complicating matters is the fact that, according to Johansson, OpenAI approached her multiple times to officially voice the product; she declined. So when a Scarlett-esque voice was somehow used anyway, the actor was "shocked, angered and in disbelief", leading her legal team to demand details of how ChatGPT-4o's voice was developed. "In a time when we are all grappling with deep fakes and the protection of our own likeness, our own work, our own identities, I believe these are questions that deserve absolute clarity," her press statement explained.

Although OpenAI quickly disabled the system's voice, this incident highlights the challenges of protecting one's likeness in the age of generative artificial intelligence (genAI) content like deep fakes. Increasingly called digital doubles, replicas, or clones,[218] deep fakes first gained notoriety as "face swapping" videos or those manipulating someone's lip movements to match altered speech.[219] Voice-only deep fakes are now on the rise, too. For instance, Warner Music partnered with Edith Piaf's estate to create a biopic of the chanteuse, to be narrated using her "own" voice through the power of genAI.[220] As a more controversial example, Drake's now-infamous diss-track *Taylor Made*

---

[216] https://openai.com/index/hello-gpt-4o/

[217] Johansson S., Scarlett Johansson's Statement About Her Interactions With Sam Altman (20 May 2024) *The New York Times*

[218] "Deep fake" is still common parlance, however this term has an entrenched association with intimate image-based abuse (i.e. so-called "deep fake porn"), which may constitute a criminal offence; an important topic which falls beyond the scope of this chapter.

[219] Lees D., (2024). *Deepfakes in documentary film production: images of deception in the representation of the real*. Studies in Documentary Film, 18(2), 108–129.

[220] Keslassy E., *Creators of the Edith Piaf AI-Generated Biopic Speak Out* (22 November 2023). Variety.

© European Audiovisual Observatory (Council of Europe) 2024

Freestyle featured unauthorised cloned vocals of fellow rap icons Snoop Dogg and Tupac "2Pac" Shakur.[221]

The technology presents remarkable creative opportunities. But as acting legend Tom Hanks observed, genAI poses "an artistic challenge, but also a legal one",[222] because hyper-realistic digital doubles can be created without legitimate approval of the cloned individual. And when a person's protected works, voice or image "are used without their knowledge, consent and remuneration to generate content[,] such uses may harm their moral, economic and personality rights".[223] It is the latter issue – personality rights – that this chapter will focus on in particular. Put simply, personality rights empower an individual to (in certain circumstances and to various degrees) protect and control how their likeness or other personal attributes (their "persona" or "personality") are used. In an age where anyone can replicate the appearance and voice of another quickly, convincingly and without consent, personality rights have become subject to increased debate and discussion.

New transparency obligations are set out in the European Union's recently-approved AI Act (AI Act)[224] and the Council of Europe's new Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law (Convention).[225] In the United States, legislation at both state and federal level may overhaul long-standing publicity laws to protect individuals from unauthorised genAI content. The United Kingdom has (as of July 2024) thus far resisted codified regulation to address AI specifically, and instead relies upon its framework of extant common law (judicial precedent), technology-agnostic statute such as consumer protection regulations, and contract law to regulate AI. However, one can expect that at least some form of AI-specific regulation will be introduced in due course. This chapter explores the technological advancements and commercial pressures driving these new laws, and focuses on the theme of transparency to consider how they protect performers' personality rights.

---

[221] Horowitz S., *Drake Removes 'Taylor Made Freestyle,' Featuring AI Tupac Shakur Vocals*, From Social Media After Threat of Lawsuit (28 April 2024) Variety.

[222] See The Adam Buxton Podcast, Ep.201 'Tom Hanks' (12 May 2023) at https://shows.acast.com/adambuxton/episodes/ep201-tom-hanks

[223] https://europeanjournalists.org/blog/2023/11/23/ai-transparency-must-be-put-back-at-the-heart-of-the-ai-act/

[224] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 And (EU) 2019/2144 And Directives 2014/90/EU, (EU) 2016/797 And (EU) 2020/1828

[225] https://www.eeas.europa.eu/delegations/council-europe/text-first-legally-binding-global-instrument-address-risks-posed-artificial-intelligence-finalised_en

© European Audiovisual Observatory (Council of Europe) 2024

5.2. Commercial Drivers

5.2.1. The Evolution of Digital Doubles

Despite considerable advancements in computer-generated images (CGI) and synthetic voice programs over the years, the chances of faithfully reproducing a performer's appearance or voice using this legacy technology remain slim to none. Artists could therefore reasonably expect to maintain at least some control and bargaining power over the use of their inimitable likeness. But after generative adversarial networks (GANs) were invented in 2014, AI's capabilities were soon unleashed to generate realistic human-like performances as never before.[226]

GenAI algorithms are "trained" on extensive datasets comprised of genuine media, such as footage of actors or samples from singers' albums.[227] When a user instructs or "prompts" the algorithm to create new content, the AI system utilises its training to generate the desired text, images, audio, or video (each a type of "output"). Early genAI output often appeared amateurish, with discrepancies and visual glitches known as "artefacts" or "hallucinations" making it easy to identify the media as fake. But today, genAI output can rival authentic performances thanks to increasingly sophisticated GANs, and the subsequent development of diffusion models and generative pre-trained transformers ("GPTs", like OpenAI's ChatGPT). These advancements paved the way for complex content like the human face and voice to be generated realistically, quickly, and at scale. Amongst other things, diffusion models refine poor-quality or "noisy" images to levels of hyperrealism, and GPTs create synthetic texts which are nuanced and contextually responsive.

Since the European Audiovisual Observatory's 2020 report on *AI in the Audiovisual Sector* 2020,[228] still more impressive genAI techniques have emerged. Gaussian splatting improves lighting, shadow effects and textures, and Neural Radiance Fields (NeRF) transform just a handful of selfies into intricate 3D scenes.[229] Importantly for theatrical performances, NeRFs can generate compelling emotional expressions for digital doubles. GenAI can also be integrated with more traditional software, including that for pose estimation, photogrammetry, motion capture and video editing.

---

[226] Cole S., AI-Assisted Fake Porn Is Here and We're All Fucked, Motherboard Tech by VICE

[227] This is of particular importance when considering the transparency obligations set out in Article 53(1)(d), discussed below.

[228] See in particular Farish, K., *Personality Rights: From Hollywood to Deepfakes* in *Artificial Intelligence in the Audiovisual Sector* (2020), IRIS Special 2020-2, European Audiovisual Observatory.

[229] See Mildenhall, B. et al, *NeRF: representing scenes as neural radiance fields for view synthesis* (2021), Communications of the ACM, Volume 65, Issue 1 pp 99–106

© European Audiovisual Observatory (Council of Europe) 2024

5.2.2. Performers' Perspectives: Empowerment or Exploitation?

While on set for her sci-fi movie *The Beast* (the plot for which, coincidentally, involves AI), French actor Léa Seydoux playfully suggested that film crew should clone her voice. "I shouldn't be working. I shouldn't be losing time on [automated dialogue replacement],"[230] she recalls saying. Although intended as a joke, having one's likeness or voice cloned may certainly save a performer time and effort, and even provide new income streams. Canadian singer Grimes, known for her electronic dance music and endorsement of futuristic technologies, announced she would "split 50% royalties" on any "successful AI generated song that uses [her] voice," and artists should "feel free to use [her] voice without penalty".[231] Efficiency and financial gain are just part of the story. Country Music Hall of Famer Randy Travis lost the ability to speak and sing after suffering a stroke in 2012; a decade later, he permitted his record label to generate a new song featuring synthetic vocals trained on his back catalogue.[232] Heralded as an example of how genAI may empower disabled persons, the song had important emotional implications, too. "All I ever wanted since the day of the stroke was to hear that voice again," Randy's wife Mary remarked. "The ability to have it back is a beautiful gift."

In these cases, genAI usage was endorsed by the person whose characteristics were digitised. Unfortunately, this is not always guaranteed. "I don't mind if someone takes a blink out during an edit," action star Keanu Reeves once explained, but he draws the line at "scary" deep fakes which threaten a performer's agency.[233] Hollywood veteran Sean Penn took his criticism further, calling it "insulting" and indicative of a "lack of morality" that studios would use digital doubles without a performer's willing involvement.[234]

Similarly, Shakespearian thespian and Succession star Brian Cox lambasted a studio that "in no uncertain terms" told another actor it would retain rights to their image "and do what the fuck they liked with it", which Cox found "completely unacceptable".[235] Even those who support AI adoption in the entertainment industry urge caution, with Legally Blonde's Reese Witherspoon admitting "AI should be a tool upon which we layer our own creativity, our own humanity and our own ethics. We need to have our say."[236]

---

[230] Lattanzio R., *Léa Seydoux and George MacKay on the Darkness of L.A.* (3 April 2024). IndieWire
[231] https://x.com/Grimezsz/status/1650304051718791170
[232] Carras, C. *Randy Travis releases new music with the help of AI after a stroke* (7 May 2024)., Los Angeles Times
[233] Watercutter A., *Keanu Will Never Surrender to the Machines* (14 February 2023), Wired.
[234] Rodrick S., *Sean Penn's Crusade: Why He's Risking It All for Ukraine, Furious at Will Smith and Ready to Call Bulls*— on Studios' AI Proposals (13 September 2023), Variety.
[235] Parkel I., *Brian Cox Rages against 'Scary' AI at SAG-AFTRA SolidarityRally in London* (21 July 2023).,The Independent
[236] Desborough J., *Reese Witherspoon says artificial intelligence in Hollywood must not be feared amid actor backlash* (15 April 2024), Mirror.

© European Audiovisual Observatory (Council of Europe) 2024

5.2.3. Regulatory Gap

The idea that people should control how their name, appearance, and public image are used is the legal foundation of personality rights. Personality rights evade a strict definition *per se*, so it may be helpful to consider them as a bundle of causes of action rooted in intellectual property, consumer protection, and privacy, as well as economic torts, publicity, defamation, and certain human rights.[237] Data protection is a related principle but serves a different purpose and, in some instances, personality rights can protect intangible assets which are not personal data, like one's "brand magnetism" and reputation.

Subject to factual circumstances and jurisdiction, personality rights can be asserted through a variety of sources, to include contract, litigation, and statute. Taking contracts as a first example, a musician's agreement with their record label might establish the boundaries of how their voice may be digitally enhanced or cloned. But at present many contracts are silent on genAI, meaning there may be no practical limit as to how the label, studios, agencies, or other counterparties can generate and distribute digital doubles. In any event, some performers lack the bargaining power or legal counsel needed to sufficiently protect their contractual position.

As for dispute resolution, substantial resources are typically required to initiate legal proceedings or make public statements to "name and shame" offenders into compliance. Whilst Scarlett Johansson and 2Pac's heirs[238] could afford to instruct lawyers and PR experts, many people lack practical access to such remedies. Furthermore, legal battles fought through the court system can be lengthy, uncertain, and inadequate insofar as outcomes are concerned, meaning an injured party may find litigation to be more trouble than it is worth.

Normally, legislation would establish certain guardrails for contracts and provide statutory rights as a mandatory minimum, thereby reducing the need to turn to potentially protracted and unbalanced negotiations or lawsuits. But as existing statutes have largely failed to address the rapid advancements and complexities of genAI, obligations regarding transparency, consent, and accountability are inadequate or otherwise non-existent. These factors, coupled with the potential insufficiencies of contract law, have led concerned stakeholders to demand new legislation to protect personality rights effectively.

British singer-songwriter FKA Twigs may have put it best during testimony before the United States Congress in April 2024.[239] "I will be engaging [my own digital double] to extend my reach and handle my online social media interactions, whilst I continue to focus on my art from the comfort and solace of my studio," she explained. Notwithstanding her fondness for AI however, she argued that "what is not acceptable is

[237] See Farish, K., *Personality Rights: from Hollywood to Deepfakes*, p. 150

[238] Donahue B. *Tupac Shakur's Estate Threatens to Sue Drake Over Diss Track Featuring AI-Generated Tupac Voice* (24 April 2024), Billboard.

[239] FKA Twigs appeared before the U.S. Senate Judiciary Subcommittee on Intellectual Property on 30 April 2024 to comment on the NO FAKES Act, discussed below.

© European Audiovisual Observatory (Council of Europe) 2024

when my art and my identity can simply be taken by a third party and exploited falsely for their own gain without my consent due to the absence of appropriate legislative control."[240] Fortunately for FKA Twigs and others who share her opinion, new legislation is on the horizon.

5.3. Transparency in European Instruments

5.3.1. European AI Act

Transparency encompasses the practice of being honest, open and clear about a particular activity or decision. In the artificial intelligence context, this involves the disclosure of accessible and straightforward information about a system's training data, functionalities, and outcomes. With this information, talent, audiences, and other stakeholders can make informed decisions about their interactions with AI – to include giving or withholding consent for the creation of digital doubles. Statutory transparency obligations can therefore serve as a useful mechanism for fostering legitimacy and upholding personality rights, because individuals are better equipped to understand and control how their personas are utilised by such systems.

The AI Act's[241] transparency obligations differ depending on the system's particular risk profile, meaning the "type and content" of rules are tailored to "the intensity and scope of the risks that AI systems can generate."[242] This risk-based approach reflects the European Union's general principle of proportionality,[243] but requires a case-by-case analysis to determine which obligations apply, taking into account the probability and severity of potential harm.[244] AI posing "unacceptable" risks, such as those in biometrics, profiling, or behavioural manipulation systems, is banned from the Union outright.[245] "High-risk" systems are those used in products with safety implications (for example aviation or children's toys), or otherwise critical services like education, financial and legal services, and healthcare. Such systems are permitted but will attract robust regulatory oversight, due to the potential harms they may inflict.

Transparency obligations for high-risk systems are consequently substantial, but European Parliament guidance expressly states that "generative AI, like ChatGPT, will not be classified as high-risk".[246] That said, it would be wrong to assume genAI evades

[240] https://www.judiciary.senate.gov/imo/media/doc/2024-04-30_-_testimony_-_twigs.pdf
[241] All "Recitals" and "Articles" referenced below are from the AI Act unless otherwise noted.
[242] Recital 26
[243] Article 5(4) of the Treaty on European Union states that "under the principle of proportionality, the content and form of Union action shall not exceed what is necessary to achieve the objectives of the Treaties".
[244] Article 3 and Recital 52
[245] Article 5
[246] https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

© European Audiovisual Observatory (Council of Europe) 2024

transparency obligations altogether. Instead, genAI is considered a type of "general-purpose AI" (GPAI), a unique risk category with its own set of transparency responsibilities: namely, the deployer's labelling requirements under Article 50, and the provider's documentation and compliance requirements under Article 53.

Taking these in turn, a "deployer" is the natural or legal person using the AI system.[247] This would include talent agencies, production companies, record labels, individual creators and so on, but exempt those using genAI for "personal non-professional purposes". The deployer must label genAI output as "artificially generated or manipulated",[248] and provide this information "in a clear and distinguishable manner" no later than the viewer's "first interaction or exposure" to the content.[249] However, of interest to those in retouching and post-production roles, this obligation does not apply where AI is merely used to "perform an assistive function for standard editing",[250] ostensibly with software like Adobe Premiere Pro and Avid ADA.

Special requirements attach to deep fakes, which the legislation defines as "AI-generated or manipulated image, audio or video content that resembles existing persons [and] falsely appear to a person to be authentic or truthful".[251] Deployers must label deep fakes as artificially generated or manipulated, but here too an interesting carve-out applies. Where the deep fake forms part of an "evidently artistic, creative, satirical, fictional or analogous work or programme", labels can be limited "in an appropriate manner that does not hamper the display or enjoyment of the work".[252] This potentially offers wide discretion insofar as style and substance of labels are concerned.

Article 53 contains the key transparency obligations of the GPAI "provider", being the natural or legal person who develops the AI and then places it on the market.[253] These primarily concern record-keeping and documentation to be made available to regulators, interested third parties, and members of the public. The latter is arguably the most relevant to the protection of personality rights, and requires providers to "make publicly available a sufficiently detailed summary about the content used for training" the GPAI,[254] based on a template from the newly-established European AI Office. This requirement seeks "to increase transparency on [training] data", and is intended "to facilitate parties with legitimate interests, including copyright holders, to exercise and enforce their rights".[255]

GPAIs provided on a "free and opensource licence" normally post their technical documentation and architecture details online. In such cases, the opensource GPAI will be exempt from some of the transparency requirements stipulated by the Act, unless the

[247] Article 3
[248] Article 50(2)
[249] Article 50(4)
[250] Article 50(2)
[251] Article 3
[252] Article 50(4) and Recital 134
[253] Article 3
[254] Article 53(1)(d)
[255] Recital 107

© European Audiovisual Observatory (Council of Europe) 2024

GPAI model is deemed to present "significant risk".[256] However, in all cases, opensource GPAI providers must still publish public summaries of training data and implement internal policies to comply with copyright law.

5.3.2. Transparency in the Framework Convention on AI

The Framework Convention on artificial intelligence marks the first legally binding international treaty of its kind. Led by the Council of Europe's Committee on Artificial Intelligence, representatives from countries including the United States, Australia and Japan also contributed to the Convention's development, and both European and non-European countries can ratify its terms.[257] If ratified, the country's domestic AI regulations must protect human rights and the rule of law, and adhere to the Convention's[258] enumerated principles – including transparency.[259]

The Convention states that the complexity, opacity and varying levels of autonomy of AI systems necessitate "safeguards in the form of transparency".[260] This is described as "openness and clarity", meaning the logic and operational details of algorithms should be "understandable and accessible".[261] What this means in practice however is open to interpretation, as the Convention does not impose the sort of specific, prescriptive obligations seen in the AI Act. Rather, it is "purposefully drafted at a high level of generality" to be "applied flexibly in a variety of rapidly changing contexts".[262]

For genAI content, the Convention addresses the need to avoid deception, and suggests "techniques such as labelling and watermarking", especially for genAI tools which can spread disinformation and misinformation.[263] The focus here largely centres on public trust, consumer protection, and prevention of electoral interference. While certainly important, this does not speak to the risk of harm faced by someone whose digital double appears in such content, nor to how an aggrieved individual may seek redress.

Comfort might instead be found in the Convention's principle of human dignity, which requires "acknowledging the complexity and richness of human identity [and] emotions".[264] A sympathetic interpretation of this provision, together with the Convention's call to respect "the inherent value and worth of each individual", supports normative arguments in favour of strong personality rights generally. The Convention requires "human-centric regulation and governance" that gives due regard to individual

[256] Recital 104
[257] Lamont C., *The Council of Europe's draft AI Treaty: balancing national security, innovation and human rights?* (18 March 2024) Global Governance Institute.
[258] All paragraphs referred to below are paragraphs of the Explanatory Report to the Council of Europe Framework Convention on AI and Human Rights, Democracy and Rule of Law.
[259] Paragraph 49
[260] Paragraph 56
[261] Paragraph 57
[262] Paragraph 49
[263] Paragraph 43
[264] Paragraph 53

© European Audiovisual Observatory (Council of Europe) 2024

autonomy, defined therein "as the capacity for self-determination and free choice".[265] Under this approach, protecting an individual's dignity may theoretically extend to an obligation to mitigate emotional and psychological harm, for example if a digital double is used in a defamatory or otherwise non-consensual manner.

Complementing human dignity and individual autonomy is the Convention's principle of privacy, framed broadly therein to include, *inter alia*, the protection of "personhood (individuality or identity, dignity, individual autonomy)" and "physical, psychological or moral integrity". This is drawn from Article 8 ECHR[266] which offers much by way of case law. Helpfully for our purposes, "privacy" in the sense of dignity and autonomy can be understood as a right to "ensure the development, without outside interference, of the personality of each individual in his relations with other human beings".[267]

When taken together, the Convention's principles of transparency, human dignity, individual autonomy, and privacy suggest safeguards against the misuse of digital doubles. Of course, this will ultimately depend on how member states choose to interpret and implement these provisions in national legislation.

5.3.3. Different angles: The United States and the United Kingdom

Home to Dolly Parton, Miley Cyrus, Justin Timberlake, Elvis Presley and countless other musicians, the U.S. state of Tennessee has a vibrant music industry, especially in its Memphis and Nashville metropolitan areas.[268] It is unsurprising, then, that the state was the first in the country to pass specific legislation to safeguard musicians' voices (and the interests of its prominent recording industry) against unwanted AI cloning. Under the Ensuring Likeness Voice and Image Security (ELVIS) Act of 2024,[269] a person must first provide authorisation before their voice is broadcasted, performed, or otherwise made publicly available.[270] The ELVIS Act also introduces a new offence of supplying "an algorithm, software [or] other technology" designed to capture or clone someone else's likeness or voice without consent".[271]

California is another AI regulatory hotspot, with more than 130 proposals made in the 2023-2024 legislative session alone.[272] Those relevant to personality rights include

[265] Paragraph 55
[266] Article 8 of the European Convention on Human Rights guarantees that "everyone has the right to respect for his private and family life, his home and his correspondence".
[267] *Botta v. Italy*, Appl. No. 21439/93, Eur. Ct. H.R. (1998).
[268] https://tnecd.com/wp-content/uploads/2018/10/Entertainment2015.pdf
[269] https://publications.tnsosfiles.com/acts/113/pub/pc0588.pdf
[270] Section 6(a)(2)
[271] Section 6(a)(3)
[272] https://leginfo.legislature.ca.gov/faces/billSearchClient.xhtml?session_year=20232024&keyword=artificial%20intelligence&house=Both&author=All&lawCode=All

© European Audiovisual Observatory (Council of Europe) 2024

new restrictions on using "digital replicas" of entertainment industry employees,[273] and requirements for "imperceptible and maximally indelible watermarks" on any genAI material.[274] The State's existing publicity statute is also likely to be amended to clarify that a digital double is a protected aspect of one's persona, and oblige genAI systems to come with consumer warning labels that explain misuse could lead to civil or criminal liabilities.[275] However, progress may be slow, as Californian politicians are often caught between the pressures of Hollywood creatives on the one hand, and Silicon Valley innovators – to include OpenAI – on the other.

Work is also underway in Washington D.C. to establish a unified national framework at federal level, with the current 118th Congress actively considering several regulations designed to protect individuals from unauthorised AI cloning and deep fakes. Notably, these include the Nurture Originals, Foster Art, and Keep Entertainment Safe (NO FAKES) Act,[276] the No Artificial Intelligence Fake Replicas and Unauthorized Duplications (No AI FRAUD) Act,[277] and the Defending Each and Every Person from False Appearances by Keeping Exploitation Subject (DEEPFAKES) to Accountability Act.[278] Each act takes a different approach to personality rights – broadly understood in America as the right of publicity – but all three introduce consent requirements and statutory damages for violations.

The NO FAKES Act is a bipartisan proposal to effectively establish a federal right of publicity, which currently exists only within certain states and with varying levels of protection.[279] NO FAKES would protect an individual's "digital replica" through a new property right enabling "certain economic control over their identity",[280] with exceptions for digital doubles appearing in news or public affairs broadcasts, documentaries, biopics, satire, scholarly work, and more. No AI FRAUD's scope is similar, but broader: it would grant every individual an intellectual property right in their own "likeness" and voice,[281] AI generated or not. Whilst the DEEPFAKES Accountability Act primarily focuses on intimate image-based abuse, it also introduces several labelling and disclosure requirements to ensure genAI media is clearly identifiable as such.

Key to shaping these regulations is their compatibility with the U.S. Constitution's First Amendment, which restricts how government may curtail freedom of expression. As currently drafted, the proposals may have an unintended chilling effect on legitimate creative expression, with the No AI FRAUD attracting criticism for potentially "unconstitutionally vague" provisions.[282]

---

[273] Assembly Bill (AB) 2602, https://legiscan.com/CA/text/AB2602/id/2928937
[274] AB 3211 https://legiscan.com/CA/text/AB3211 and AB 3050 https://legiscan.com/CA/text/AB3050
[275] Senate Bill (SB) 970 https://legiscan.com/CA/text/SB970/
[276] https://www.coons.senate.gov/imo/media/doc/no_fakes_act_draft_text.pdf
[277] https://www.congress.gov/bill/118th-congress/house-bill/6943/text
[278] https://www.congress.gov/bill/118th-congress/house-bill/5586/text
[279] Tennessee's Personal Rights Protection Act of 1984, brought about by litigation from Elvis Presley's estate, together with California's Civil Code § 3344, are examples of state publicity laws.
[280] Nair P., *Imitation Is Not Flattery: Introducing the NO FAKES Act* (16 January 2024) ACT | The App Association
[281] No AI FRAUD, Section 3(1) and (2)
[282] Klosek K., *No Frauds, No Fakes… No Fair Use?* (1 March 2024). Association of Research Libraries.

© European Audiovisual Observatory (Council of Europe) 2024

In contrast to the European risk-based approach, the United Kingdom has been deliberately *laissez faire*. Its principal roadmap for AI legislation was set out in the 2023 pro-innovation approach to AI regulation "White Paper",[283] which opens by asserting "heavy-handed and rigid" legislation "can stifle innovation and slow AI adoption". It claims that as the UK is "home to a third of Europe's total AI companies and twice as many as any other European country", the British Government will consult with sector-specific regulators and industry stakeholders to design a "proportionate" and "flexible" framework to "ease the burden on business". It also suggests that extant legislation, for example regarding product safety and consumer protection, may be sufficient to address the harms posed by AI.

Nevertheless, the British parliament does appear to acknowledge that genAI can create material that "deliberately misrepresents someone's behaviour, opinions or character", and that some AI models do not sufficiently disclose or explain technical information.[284] To combat these and other challenges, the White Paper introduced "appropriate transparency and explainability" as one of five principles developers should respect when designing and providing AI solutions. In its February 2024 response[285] to the White Paper, Government stated it was "exploring mechanisms for providing greater transparency, including measures so that rights holders can better understand whether content they produce is used as an input into AI models". However, the focus here is expressly upon copyright rather than personality rights or reputational protections, and as the British government continues to stand firmly by a non-statutory approach, any such transparency mechanisms would be voluntary.

Notwithstanding the above, the UK is unlikely to become a "Wild West" in which AI rides off unregulated into the sunset. Firstly, the White Paper concedes that "AI will ultimately require legislative action" in due course.[286] Secondly, the outcome of the UK's July 2024 elections resulted in a change of national ruling party, to one which has been vocal about the need to introduce AI regulations.[287] It is therefore probable that the approach outlined above will succumb to more stringent transparency obligations and codified protections for individuals under the new government.

5.4. Transparency as a keystone to uphold personality rights

Of course, legislation forms only part of the personality rights saga of the AI era. Many businesses are self-regulating, with groups like the Coalition for Content Provenance and

[283] https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper
[284] https://publications.parliament.uk/pa/cm5803/cmselect/cmsctech/1769/summary.html
[285] https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response
[286] ibid.
[287] See, *inter alia*, Landi, M. *Labour commits to introducing AI regulation for tech giants.* The Independent (13 June 2024) and the Labour Party Manifesto

© European Audiovisual Observatory (Council of Europe) 2024

Authenticity (C2PA)[288] and the National Institute of Standards and Technology (NIST)[289] establishing voluntary technical and governance standards. Public pressure and advocacy groups are likewise making an impact, as evidenced by the resolution of the SAG-AFTRA walkout in December 2023.[290] After five months of industrial action, the entertainment workers' union approved a deal with the Alliance of Motion Picture and Television Producers (AMPTP),[291] the major American trade association for film studios, television networks, streaming services, and production companies. Amongst other things, AMPTP entities must now secure a performer's consent when making digital doubles. As of June 2024, there are also more than 20 active intellectual property lawsuits across both sides of the Atlantic involving genAI companies, the outcome of which will almost certainly influence how genAI is developed and used.[292]

It remains an open question as to how legislation, industry-led initiatives, contractual negotiations, and case law will evolve to address the challenges of genAI. What is clear, however, is that transparency is crucial for the meaningful exercise of personality rights as digital doubles and AI cloning become more commonplace. Consent is crucial for control, and legitimate consent requires honest, accessible information about genAI risks and benefits. Moreover, when providers and deployers are transparent about how genAI is developed and utilised, this helps ensure they may be held to account so that harmed individuals have proper means of redress. This can safeguard performers, creators, and audiences, as well as encourage trust in the proper use of digital doubles and the systems that create them lawfully. Perhaps most importantly, when people are fully informed as to how their likeness, voice, or other personal attributes are digitised, this goes some distance to affirm their rights to self-determination, dignity and autonomy. In the words of actor Talulah Riley: "It is vital that my voice and my image are my own, no matter how easily and cheaply those things can be digitally replicated."[293]

[288] https://c2pa.org/specifications/specifications/2.0/index.html
[289] https://www.nist.gov/itl/ai-risk-management-framework
[290] https://www.sagaftra.org/sag-aftra-members-approve-2023-tvtheatrical-contracts-tentative-agreement
[291] Summary of the updated contract
[292] Lee E., *Status of all 24 copyright lawsuits v. AI companies* (24 May 2024). Chat GPT is Eating the World.
[293] Vallance C., *Actors launch campaign against AI 'show stealers'* (21 April 2022). BBC.

© European Audiovisual Observatory (Council of Europe) 2024

6. Impact of AI on the audiovisual labour market in Europe

Elodie Migliore, PhD at CEIPI, University of Strasbourg

6.1. Introduction

« Toute puissance est faible, à moins que d'être unie. »[294]

Le vieillard et ses enfants – Jean de La Fontaine

The rise of AI technologies is impacting every aspect of our daily life. As depicted in many works of science fiction, such as I, Robot,[295] the common perception of AI technologies is that they will replace humans in many tasks. Whilst it can be perceived as a pipe dream for some, there are already specific sectors where this prediction is becoming a reality, with technologies effectively replacing human labour.

One prominent example is the creative sector. An early study by OpenAI indicates that the exposure risk for writers and authors is at 82.5%.[296] The audiovisual sector is no exception. A report by KPMG[297] indicates that creative occupations have some of the biggest shares of tasks susceptible to automation, with a 43% share of tasks automated for authors, writers, and translators, with humans "fine tuning" machine output.[298] This could lead to many consequences such as squeezing the pay of professional writers further. However, workers have decided not to sit back and stay passive. This has led to major strikes and disruption in the last few months.

This Chapter seeks to analyse the current state of labour law in the audiovisual sector concerning the use of AI, drawing upon the two major strikes that happened in the United States (US).

---

[294] "Any power is weak unless it is united", free translation.
[295] I robot, Alex Proyas, 20th Century Fox, 2004.
[296] Tyna Eloundou and others, "GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models" (arXiv, 21 August 2023)
[297] "Generative AI and the UK Labour Market", KPMG UK
[298] "Writers and AI" (Writers' Guild of Great Britain, 12 July 2023)

© European Audiovisual Observatory (Council of Europe) 2024

6.2. Impacts of AI on labour law in the audiovisual sector in the US

6.2.1. The WGA and SAG- AFTRA strikes

On 2 May 2023, the Writers Guild of America (WGA), a labour union representing 11 500 screenwriters went on strike. The strike order concerned all companies that are signatories to the Minimum Basic Agreement (MBA),[299] a collective agreement that sets out the rules and pay rates applicable to WGA scriptwriters.[300] The writers had not been on strike since the historic 100-day strike in 2007.[301]

Then, on 14 July 2023, the Screen Actors Guild-American Federation of Television and Radio Artists (SAG-AFTRA), a labour union representing 160 000 media professionals worldwide, also went on strike. The strike order concerned all services covered under the Producer SAG-AFTRA Codified Basic Agreement, and SAG-AFTRA Television Agreements and their related agreements.[302] It was the first time actors had engaged in a labour dispute in the United States since the 1980 actors' strike.[303] More importantly, for the first time since 1960, actors and writers were simultaneously on strike.

Both unions fought against the Alliance of Motion Picture and Television Producers (AMPTP), a trade association representing 350 American television and film production companies in collective bargaining with entertainment industry unions.

The WGA strike ended after 146 days, on 27 September 2023, following an agreement reached with the AMPTP, covering the period from 25 September 2023 to 1 May 2026.

The SAG-AFTRA strike ended on 9 November 2023 with an agreement ratified on 5 December, covering the period from 9 November 2023 to 30 June 2026.[304]

Both strikes shared common revendications such as negotiating new residuals linked to the rise of streaming services, but they also shared the common objective to regulate the use of genAI in the course of their employment.

WGA screenwriters feared that AI-generated works could compete with their jobs and that training AI models with professional writers' material could diminish their credit

---

[299] Memorandum of Agreement for the 2023 WGA Theatrical and Television Basic Agreement, 2023

[300] Elodie Migliore, "Fin de la grève des scénaristes américains : quand l'union fait la force", Intelligence artificielle | Dalloz Actualité (2023)

[301] Cal Berry, "Blueprint for a Strike in the Entertainment Industry: Lessons from the 2007 WGA Strike", (*Left Voice*, 5 November 2021). See also, Pencils Down! The 100 Days of the Writers Guild Strike, Brian S. Kalata, 2014.

[302] "SAG-AFTRA Strike Order for TV/Theatrical/Streaming Contracts" (*SAG-AFTRA*)

[303] Cynthia Littleton, "Revisiting the 1980 SAG-AFTRA Strike with 'MASH' Stars, an Emmy Boycott and All-Night Negotiating Sessions: 'We'Re Going to Strike Like Hell'"(*Variety*, 1 September 2023)

[304] 2023 TV/Theatrical Contracts Tentative Agreement

© European Audiovisual Observatory (Council of Europe) 2024

and residuals. SAG-AFTRA was concerned that studios might use AI and digital technologies to replicate performers' faces and voices, reducing actors' rights and work opportunities.

6.2.2. The WGA agreement after the strike

The demands of the WGA were numerous. While some of them were accepted, others were not incorporated into the agreement as they stood.[305]

Firstly, it was decided that the use of genAI is not permitted to write or rewrite literary material/content. Moreover, AI-generated content cannot be considered as source material under the agreement.[306]

In addition, a screenwriter may use an AI system as part of their services if the company agrees, provided that the screenwriter complies with the company's policies. However, a company may not impose on a scriptwriter the use of an AI system to deliver its services. The company may also reject the use of an AI system, particularly if it has doubts about the possibility of benefiting from copyright protection for the content produced, or about its ability to exploit said content. The company must also inform the writer if the documents communicated to them have been generated by an AI system or contain elements generated by an AI system.

Finally, a contentious issue was the training of AI systems, to which the WGA was strongly opposed. This issue was one of the most difficult to settle, and the agreement maintained a clause implying that if the writers retained reserved rights on their material, they could – or the WGA on their behalf – forbid the use of said material for the training of a GAI. On the contrary, this also means that if a studio fully retains the reserved rights on the material, they can exploit it to train GAI systems.[307] This clause is not as victorious as it appears, though, since there is no ban on studios using scripts they own to train AI systems; all will depend on the rights retained.

6.2.3. The SAG-AFTRA agreement

The WGA agreement embodied a step forward concerning the regulation of AI in the audiovisual sector. The SAG-AFTRA agreement similarly integrates interesting

[305] Article 72, page 68, WGA Proposal No. 29, Memorandum of Agreement for the 2023 WGA Theatrical and Television Basic Agreement, 2023.
[306] Source material "means all material upon which the screenplay is based other than story as hereinabove defined, including other material on which the story is based. Credit shall be given on the screen for story authorship of feature-length motion pictures [...]", see the 2020 WGA Minimum Basic Agreement, page 403. AI-generated content cannot be used to infringe a writer's credit or rights.
[307] Article 72, page 68, WGA Proposal No. 29, Memorandum of Agreement for the 2023 WGA Theatrical and Television Basic Agreement, 2023.

© European Audiovisual Observatory (Council of Europe) 2024

provisions.[308] The 2023 SAG-AFTRA memorandum of agreement (MOA) governs theatrical motion pictures and scripted dramatic content produced for television and new media platforms, with a specific focus on the issues surrounding outputs.[309] The agreement underlines consent and compensation as two core notions, consistently present throughout the AI provisions.

6.2.3.1. Training data

Firstly, concerning the issue of training data, it appears that the MOA of the SAG-AFTRA does not provide additional payments for the inclusion of footage or voice recordings of an actor's performance in a training dataset. It does not mean it would be impossible to seek compensation, but it is left to actors or performers to negotiate their own deals. This situation appears possible for famous actors with enough bargaining power, but less realistic for new actors entering the market.

The only provision dealing with training data is Paragraph C of the Title II «Artificial Intelligence» of the Summary of 2023 Tentative Successor Agreement to the 2020 Producer-SAG-AFTRA Codified Basic Agreement and 2020 SAG-AFTRA Television Agreement[310], providing regular meetings to "[…] discuss remuneration, if any, for use of work produced under [the Collective Bargaining Agreement] to train GAI system for creation of Synthetic Performers".[311]

6.2.3.2. Synthetic performers

Secondly, the agreement defines the concept of a synthetic performer as a digitally-created asset which "is intended to create, and does create, the clear impression that the asset is a natural performer who is not recognizable as any identifiable natural performer; is not voiced by a natural person; is not a Digital Replica; and no employment arrangement for the motion picture exists with a natural performer in the role being portrayed by the asset."[312]

It then provides additional requirements for the use of recognizable synthetic performers. This notion refers to synthetic performers including recognizable features of an actor, such as a "principal facial feature (i.e., eyes, nose, ears and/or mouth)" that is requested through a "prompt to a GAI system".[313] In this situation, the producer must bargain and obtain the performer's consent. For example, it means that if one would like

[308] However, there are dissenting opinions, see on this issue Laura Weiss, "SAG-AFTRA's New Contract Falls Short on Protections from AI"(*Prism*, 5 December 2023). See also, "How SAG-AFTRA's AI Provisions Work: A Lawyer's View"
[309] 2023 TV/Theatrical Contracts Tentative Agreement, *op.cit*
[310] https://www.sagaftra.org/files/sa_documents/TV-Theatrical_23_Summary_Agreement_Final.pdf
[311] 2023 TV/Theatrical Contracts Tentative Agreement, Section C, page 3, *op.cit*
[312] 2023 TV/Theatrical Contracts Tentative Agreement, Section C, page 4, *op.cit*
[313] Ibid

© European Audiovisual Observatory (Council of Europe) 2024

to generate a synthetic performer with Emma Stone's eyes, one would need to bargain and obtain her explicit consent.

6.2.3.3. Digital replicas

Thirdly, the agreement defines two types of replicas, namely employment-based replicas and independent replicas.[314]

Employment-based digital replicas are defined as digital reproductions of a performer's voice or likeness that are created in connection with their employment on a motion picture, using digital technology and the performer's physical participation, to portray the performer in photography or soundtracks where they did not actually perform.[315] For instance, creating a replica of Kyle MacLachlan to portray a young Henry MacLean in Fallout.[316]

In this situation, the producer must provide advance notice prior to service creation, and obtain the actor's "clear and conspicuous"[317] consent in a separate document from the employment contract signed by the performer, alongside additional requirements for compensation. Then, the agreement provides a whole section dedicated to the use of these replicas, providing rules on when consent or extra compensation is required.[318] It must also include "a reasonably specific description of the intended use".[319]

Independently created digital replicas are replicas designed to convincingly portray a natural performer by using recognizable features such as their voice and/or likeness. The replica will be used to perform a character rather than the natural performer and there is no employment arrangement with the natural performer for the motion picture in which the replica is used. It is often created by using existing materials to portray the actor in scenes they did not actually shoot,[320] – for example, Paul Walker as Brian O'Conner in Fast and Furious 7. For this type of replica, a producer must negotiate and obtain consent prior to use. It also provides for pension and health contributions.

6.2.3.4. Digital alteration

Finally, the agreement also deals with the concept of digital alteration, a common phenomenon in the movie industry. Digital alteration can be performed for cosmetics purposes for example and might not always involve AI processes. Consent will not be

[314] Ibid
[315] Ibid
[316] For a digital replica which is not to everyone's taste and has been the subject of some discussion amongst fans, some of them speculating that it was AI-generated and mentioning the SAG-AFTRA agreements, see this discussion on Reddit for example
[317] 2023 TV/Theatrical Contracts Tentative Agreement, Section C, *op.cit*
[318] The agreement also contains provisions concerning deceased actors or the use of these replicas for a sequel or a prequel for example, see *ibid*, page 5
[319] 2023 TV/Theatrical Contracts Tentative Agreement, Section C, *op.cit*
[320] Ibid

© European Audiovisual Observatory (Council of Europe) 2024

needed when "the photography or soundtrack of the performer remains substantially as scripted, performed and/or recorded", but consent is needed for more significant alterations. Similar rules concerning consent and compensation are included.

6.3. Impacts of AI on labour law in the audiovisual sector in the EU

6.3.1. European Union policy

The two strikes in the US managed to achieve improved working conditions for workers in the audiovisual sector. It appears that the European Union (EU) is also starting to consider these issues.

Firstly, it must be noted that social policy is primarily the responsibility of EU member states, and it could impact the EU's ability to improve workers' rights in relation to AI. However, certain domains are a shared competence with the EU.[321] Indeed, Title X of the Treaty on the Functioning of the European Union (TFEU)[322] defines social policy in the EU.[323] A horizontal social clause is also introduced by Article 9 of the TFEU. In addition, Article 6 of the Treaty on European Union (TEU)[324] grants binding authority to the social rights outlined in the EU Charter of Fundamental Rights.

The European Parliament and the Council may adopt incentive measures to support and complement the actions of EU countries in specific areas. They may also adopt minimum requirements through directives to enable EU countries to adopt additional stricter provisions. These directives concern limited domains, including but not limited to, health and safety of workers, information and protection of workers or protection of workers in the case of termination of their employment contract.[325]

It means that the European Commission will have limited competence in social matters, notably concerning remuneration, explaining why the European Commission may not move as quickly in these areas.

For now, there is no EU binding legislation specifically focused on the audiovisual sector imposing new terms on member states, but a step forward can be observed at the

[321] Shared competence refers to areas in which legislation and the adoption of binding acts can be carried out both at European level and by each of the Member States, independently of the others. However, Member States can only exercise their competence to the extent that the EU has not exercised, or has decided not to exercise, its competence. Concerning social policy, it only concerns aspects specifically defined in the treaty.
[322] Consolidated versions of the Treaty on European Union and the Treaty on the Functioning of the European Union 2012/C 326/01
[323] Social policy objectives are detailed by Article 151 of the TFEU.
[324] Consolidated version of the Treaty on European Union C 202/1

[325] "Social policy", Glossary of Summaries (EUR-Lex)

© European Audiovisual Observatory (Council of Europe) 2024

European Parliament with its resolution of 21 November 2023, with recommendations to the Commission on an EU framework for the social and professional situation of artists and workers in the cultural and creative sectors.[326]

The resolution is composed of 73 recommendations, with a large variety of measures relating to the status of the artist, social protection, decent working conditions, fair remuneration, education and training, artistic freedom or collective bargaining.

The resolution calls for a strategic approach at the Union level to address the impact on jobs, working methods, worker conditions, and the need for upskilling and reskilling. It also recommends job creation plans and sector-specific financial support to ensure social protection for those affected by digitalisation and AI-related job losses. The resolution highlights challenges like job loss and transformation of work and urges the Commission to address AI's effects on workers' rights and well-being in future legislation.

The European Commission welcomed the resolution,[327] and highlighted some initiatives already conducted in this field.[328] A high-level group is planned to explore the best way forward to address the needs of the sector. The Commission concluded its response by enumerating several initiatives, such as fair remuneration of authors and performers for the exploitation of their artistic work with a focus on the challenges posed by AI.

6.4. Analysis of the different initiatives of selected stakeholders

The following sections will review the role played by different actors, essential to the audiovisual sector, such as collective management organisations, associations or federations, and trade unions, to study their actions and initiatives.

6.4.1. Collective management organisations (CMOs)

Among the first entities that come to mind when talking about improving workers' conditions in the audiovisual sector are the Collective Management Organisations

[326] European Parliament resolution of 21 November 2023 with recommendations to the Commission on an EU framework for the social and professional situation of artists and workers in the cultural and creative sectors (2023/2051(INL))
[327] Answer of the European Commission to the resolution of 21 November 2023.
[328] European Commission, Directorate-General for Education, Youth, Sport and Culture, The status and working conditions of artists and cultural and creative professionals – Report of the OMC (Open Method of Coordination) group of EU Member States' experts, Publications Office of the European Union, 2023

© European Audiovisual Observatory (Council of Europe) 2024

(CMOs).[329] As a consequence, the SACD and Playright will be analysed due to the nature of their actions, as well as the SAA, an association of different CMOs.[330]

To begin, the French *Société des Auteurs et Compositeurs Dramatiques* (SACD) "has adjusted its general contracts with users to include new clauses to protect the works of our authors. These clauses prevent users from licensing the rights of the authors we represent to AI companies."[331] For example, the template contract governing the relationship between the producer and the author contains several provisions similar to the WGA memorandum.[332] It contains clauses preventing authors being forced to use AI systems, a requirement of disclosure of the use of an AI system by the author or by the producer to the other party, or clauses requiring the producer not to use elements generated by AI to create visuals to promote the film without the express prior agreement of the author. Moreover, contrary to the WGA and SAG-AFTRA agreements, the use of the work produced by the author to train an AI system is expressly forbidden. Yet, the reach of these template contracts will be limited compared to what has been achieved in the US, as their use is conditional. They do not constitute a minimum mandatory basis to use, and in practice it is left up to the author and his/her bargaining power to impose these contractual terms on the other party. Finally, a new clause has been submitted to the SACD's general meeting, which would enable the SACD to "intervene unequivocally on behalf of its member authors on artificial intelligence".[333] The 2023 Annual report also evokes several questions that need to be asked, such as how to reach an agreement with AI companies, as they have with all broadcasters and platforms.[334]

PlayRight, a Belgian CMO, shared its opinion on AI through a position paper[335] or news update on AI. It also enacts guidelines for contracting with AI,[336] without providing for a contractual template, but only several recommendations on the possible clauses to include in contractual agreements. Some recommendations are also similar to some provisions of the WGA and SAG-AFTRA agreements.[337]

Finally, the actions of the Society of Audiovisual Authors (SAA) are worth mentioning. The SAA is an association of European collective management organisations representing audiovisual authors. The SAA is actively advocating regulation of the use of

---

[329] Collective management organisations aim to provide authors with an efficient and cost effective way to manage their rights worldwide to ensure that their works are used per governing laws

[330] Even though this part focuses on three actors, a large variety of CMOs exists, such as the BECS, the GVL or the AIE, and they are also acting with regard to AI

[331] SAA, Expert Seminar about audiovisual authors' rights and Artificial Intelligence, 30 January 2024, Expert Seminar about audiovisual authors' rights and Artificial Intelligence. See also 5 takeaways from the SAA Expert Seminar on Artificial Intelligence 30 January 2024

[332] "Modèles de contrats audiovisuels" (*SACD*, 8 February 2017).

[333] Free translation, see 2023 Annual Report, (SACD, 2024).

[334] *Ibid*, page 10.

[335] Playright, POLITICAL MEMORANDUM OF THE PERFORMING ARTIST 2024-2029, MAY 16 2024

[336] Team COMPLUS, "Contractual Guidelines in Relation to AI" (*PlayRight*, 27 February 2024).

[337] For example, they discuss the possibility to include a specific contractual clause requiring the prior authorisation of the performer for any new use, Playright, "Contractual Guidelines in Relation to AI", *op.cit.* See also Team COMPLUS, "End of Strike in Hollywood: SAG-AFTRA Reaches Agreement" (*PlayRight*, 28 November 2023)

© European Audiovisual Observatory (Council of Europe) 2024

AI through joint statements or position papers.[338] They also raise awareness on issues linked to AI through seminars for example.[339] They hence provide a wide range of information regarding AI and its impact on the audiovisual sector.[340] Moreover, they recently published a position paper[341] focusing on intellectual property issues and highlighting core principles for human-centred AI regulation to foster creativity.[342] It establishes transparency, authorisation/licensing and remuneration as fundamental principles while emphasising the role CMOs can play.[343]

6.4.2. Associations and federations

European or international associations and federations are also converging towards common actions to make progress. Firstly, it is important to note that European associations might not have the same bargaining power as their American counterparts due to their geographical fragmentation and structural differences. However, European associations and federations are actively working to improve workers' conditions.[344]

A notable example is the Federation of Screenwriters (FSE), representing together 26 screenwriters' organisations. The actions of this federation will be studied since it represents workers at risk. They aim to enhance workers' conditions through awareness-raising campaigns and representation activities for instance.[345] They released a joint resolution with the International Affiliation of Writers Guilds (IAWG) calling for "ethical use" of AI based on guidelines.[346] Moreover, one of the key priorities of the FSE is to facilitate collective bargaining both at EU level and at national and regional level. Collective bargaining enables the establishment of minimum terms and conditions of employment. It was one of the strengths of the unions studied in the US, which allowed them to reach such agreements. A similar approach could prove useful to reach new deals framing AI.[347]

The FSE is far from being the only association or federation acting to improve workers' conditions regarding the use of AI in the audiovisual sector. For instance, the

---

[338] See EU AI Act: Joint statement from European creators and rightsholders (*SAA*, 13 March 2024)
[339] See for example, Expert Seminar about audiovisual authors' rights and Artificial Intelligence, *op.cit.*
[340] See their different actions concerning AI https://www.saa-authors.eu/en/tags/222-ai#.ZIm3Y3bP06Q.
[341] 'Artificial intelligence must serve society and enhance human creativity' (*SAA*, 4 October 2023)
[342] *Ibid*, page 5.
[343] *Ibid.*
[344] See for example, the Federation of European Screen Directors (FERA), the European Audiovisual Production Association (CEPI), the Association of European Performers' Organisations (AEPO-ARTIS) or the SAA.
[345] See for example the numerous joint letters signed to regulate AI or concerning AI regulations: "Joint letter in response to the dialogue with the Audiovisual Sector on Copyright & AI" (FSE, 30 November 2023), "For an innovation and creator friendly AI Act" (FSE, 24 November 2023) and "Joint Statement on Artificial Intelligence and the Draft EU AI Act" (FSE, 26 September 2023)
[346] Federation of Screenwriters in Europe, "Artificial Intelligence: Global Screenwriters Call for Ethical Use" (*FSE - Federation of Screenwriters in Europe*, 11 April 2024). See also, "Global Screenwriters Call for AI Regulation"
[347] Federation of Screenwriters in Europe, "Collective Bargaining" (*FSE - Federation of Screenwriters in Europe*)

© European Audiovisual Observatory (Council of Europe) 2024

Federation of European Screen Directors (FERA), the European Audiovisual Production Association (CEPI) or the Association of European Performers' Organisations (AEPO-ARTIS) are all participating through joint letters, events or by communicating news on the matter.

The Confederation of Societies of Authors and Composers (CISAC), one of the most important non-profit organisations, is also focused on lobbying through joint letters,[348] communication,[349] education,[350] and news. For example, in their recent annual report,[351] they underlined AI as one of their key priorities. In this report, they refer to three principles. The three principles do not necessarily focus on labour law issues, but rather on authorisation to use the artist's works,[352] remuneration, and transparency.[353]

6.4.3. Trade unions

Trade unions appear to be an important actor to frame the use of AI, as demonstrated in the US.

UNI Europa is a European trade union federation, using collective strength to expand collective bargaining. UNI MEI is their audiovisual section. [354] UNI Europa adopted a resolution on AI in 2021, advocating for fair wages and working conditions through collective bargaining.[355] This resolution lays out a "forward-looking political platform for common frameworks on collective bargaining demands".[356] However, it must be kept in mind that wages and working conditions are primarily determined by different national legal and institutional settings. Collective bargaining is primarily the responsibility of national unions.[357] Moreover, UNI Europa also undertakes lobbying and political campaigns.[358]

---

[348] "Global Creators and Performers Demand Creative Rights in AI Proliferation | CISAC" (20 July 2023)

[349] "CISAC VP Ángeles González-Sinde Calls for Ethical Rules on AI Use in the Film Industry" | CISAC' (20 October 2023)

[350] For example they organise seminars, see "How Will AI Transform the Music Industry? Expert Panel at the IPRS International Music Creators' Seminar Discuss | CISAC" (10 January 2024)

[351] "CISAC Annual Report Highlights Its Work Programme on Behalf of CMOs Worldwide | CISAC" (22 May 2024)

[352] They state that "creators must have the right to license the use of their works by AI tools", CISAC Annual Report, page 20, *op.cit.*

[353] Transparence means that "AI providers must be obliged to inform on the training of copyrighted works", ibid.

[354] UNI-MEI (*International Arts and Entertainment Alliance*)

[355] "Forward through collective bargaining – Resolutions adopted by the 5th UNI Europa Conference Brussels, 27-29 April 2021" (*UNI Europa*)

[356] *Ibid*, page 16.

[357] Ibid.

[358] For example, see "Digital Working in the Media, Arts & Entertainment Sector: Challenges and Opportunities" (*UNI Global Union Media and Entertainment The FIM, FIA, and EFJ*) or also "Solidarity with SAG-AFTRA and FIA" (*UNI-MEI*)

© European Audiovisual Observatory (Council of Europe) 2024

The International Federation of Actors (FIA) is a global federation composed of performers' trade unions, guilds and professional associations.[359] AI is a major focus, as demonstrated by the dedicated page on their website.[360] They released a guide to explain the implications of AI for the performers FIA affiliates represent.[361] It also sets out FIA's core principle regarding the use of AI while offering guidance on "how trade unions can structure their bargaining strategy to enhance protections and compensation for their members."[362] While not all FIA affiliates can negotiate AI safeguards in their collective bargaining, the guide provides advice for unions able to do so, such as clear consent and additional remuneration for digital replicas of performers.[363] Finally, it suggests how FIA affiliates can act with regard to an appropriate policy and regulatory environment while stressing the importance of a mandatory legislative framework.[364]

The Writers' Guild of Great Britain (WGGB) is also an interesting union since it has similarities with the WGA. The WGGB is a trade union representing professional writers such as those who work in audiovisual sectors. They can negotiate better working conditions with major British industry players.[365] These standard agreements with major organisations are beneficial to all writers, but only WGGB members have access to certain benefits.[366] It differs from the US since joining the WGGB is entirely optional while it is mandatory for the WGA in some instances.

For the moment, the WGGB has not struck any new contractual agreements with provisions related to the use of AI but it is actively communicating on this issue.[367] Finally, it has pledged support for the strikes in the US, to show solidarity with the "sister union".[368] The WGGB has the power to negotiate new deals, as in the past, which could integrate provisions pertaining to the use of AI. It is worth following the WGGB's actions to observe if such an outcome will actually happen.

6.5. Concluding remarks: remaining gaps and path forward

For now, discussions in Europe are mostly focused on intellectual property issues such as copyright and training data, remuneration in a copyright perspective, the protection by

[359] "About – FIA" (*FIA)*
[360] "Artificial intelligence – FIA" (*FIA*)
[361] "Guide de la FIA sur l'intelligence artificielle" (*FIA*, 23 November 2023)
[362] "FIA Policy and Practical Guide with Respect to Artificial Intelligence – FIA" (*FIA*, 23 November 2023)
[363] "Guide de la FIA sur l'intelligence artificielle", page 9 and 10, *op.cit*.
[364] *Ibid*, page 12 and 13.
[365] "About" (*Writers' Guild of Great Britain*)
[366] Jonny Walfisz, "WGA Strikes: What Is the State of Writers' Unions in Europe?" (*euronews*, 31 May 2023)
[367] See for example their policy statement, "Writers and AI" (n 7).
[368] Sarah Woodley, "WGA Strike" (*Writers' Guild of Great Britain*, 2 May 2023). See also Sarah Woodley, 'WGA Strike Ends' (*Writers' Guild of Great Britain*, 27 September 2023). (*Writers' Guild of Great Britain*, 27 September 2023).

© European Audiovisual Observatory (Council of Europe) 2024

copyright of content generated by AI systems, and the notion of transparency or licensing.[369]

The discussions and initiatives dealing with labour law per se remain quite general and are not directly targeted at specific challenges faced by the audiovisual sector.[370] There is also increasing note taken of the need for training and awareness of creators, as highlighted by reports and in the resolution of the European Parliament.[371]

To conclude, one promising lead is the resolution from the European Parliament, which indicates a similar philosophy to the US but is not yet detailed enough to be indicative on its substance. The different actions of the various stakeholders also demonstrate a growing need for, and interest in, regulation of the use of AI in the audiovisual sector. All these initiatives need to be closely followed for future developments.

[369] See on this, Chapters 3 and 4 of this study. See also Article 53 of the AI Act and Articles 3 and 4 of the Directive (EU) 2019/790

[370] For example, see the AI Act. See also, "Un an après l'arrivée de ChatGPT: Réflexions de l'Obvia sur les enjeux et pistes d'action possibles face à l'IA générative" (*Observatoire international sur les impacts sociétaux de l'IA et du numérique (OBVIA)*, 15 January 2024)

[371] See for example, European Parliament (2023), recommendation 61, *op.cit*. See also the legal obligation to ensure "AI literacy" of Article 4 of the AI Act. See Commission de l'Intelligence artificielle, IA : notre ambition pour la France, March 2024, recommending to train people on AI.

© European Audiovisual Observatory (Council of Europe) 2024

7. Disinformation and AI in the AV Sector

Judit Bayer, University of Münster

7.1. Defining Disinformation

The transformation of the information environment in the past decades erased previous structures of public information, such as the dominance of traditional news sources, along which the trustworthiness of information was evaluated by the audience. Entry barriers to the public realm have vanished; publishing has hyperdemocratised; quality control over the published information has become the exception. This enormous potential has been exploited not only by honest participants of the public discourse, but also by malicious actors. Such individuals or networks have been forging and disseminating manipulative or false information with the intention to advance their financial or political goals even at the price of harming individuals or societies. In the participatory communicative space, sharing, liking and other amplification actions further promote disinformation which is often more sensational, more attractive than real information. When genuine believers share or publish false or manipulated information, or when the original intent of the author remains unknown, we talk about misinformation.[372] Similarly, when established providers of edited content who normally take efforts to ensure the quality of their products make occasional mistakes, we can also call those instances misinformation. The following piece will generally focus on intentional disinformation created and disseminated with the help of AI, briefly touching upon unintentional instances of misinformation related to genAI.

[372] Wardle, C. and Derakhshan, H., "Information Disorder", *Council of Europe report* DGI(2017)09

© European Audiovisual Observatory (Council of Europe) 2024

7.2. . AI applications in the disinformation industry

7.2.1. Generative AI

Since 2022, genAI models have become openly accessible to anyone for free, or for a low price. These models are a type of general purpose AI systems that are capable of generating text, audio, images or video, on the basis of text prompts.[373] The easy accessibility of genAI technology is likely to increase the volume of all content, including disinformation.[374] However, so far, we have not seen the expected flood of AI-based disinformation.[375] The 2024 European Parliamentary elections have also occurred in relative tranquillity, without major AI-generated disinformation campaigns, and previous elections also appear to have been relatively immune to the influence of AI-generated disinformation.[376]

7.2.2. Text and images

Ascertaining the difference between enhancing and manipulating existing content or generating new content can be elusive. Hence, the AI Act handles generated and manipulated content together.[377] If AI is only used to edit existing content or enhance its quality but does not substantially alter the input data, then it is exempted from the obligation of transparency. Also, if the publisher undertakes editorial responsibility and the AI-generated content undergoes human review, then the artificial generation or manipulation need not be revealed.[378]

Nonetheless, it is good journalistic practice to be transparent about the use of AI, and the European Media Freedom Act (EMFA) also mentions this as an element of qualifying as a "media service provider" (declaring that one does not provide content

[373] Recital (105): Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (AI Act)
[374] Cook, L. and Chan, K., "AI could supercharge disinformation and disrupt EU elections, experts warn", *AP,* 5 June, 2024
[375] Zöldi, B., "Not only images of real suffering, but also AI creations of the Hamas-Israeli war are spreading", *Lakmusz,* 27 November, 2023. EEAS, "2nd Report on FIMI and Interference Threats", January 2024
[376] Güttel, L., and al. "The European Elections 2024: Between Digital Policies and Radical Right Success Online."
OII. 4 June 2024.; Canetta, T., "EU Elections 2024: the battle against disinformation was won, but the attrition war is far from over", *EDMO Blog,* June 11, 2024; Łabuz, M. and Nehring, C. "On the way to deep fake democracy? Deep fakes in election campaigns in 2023". *Eur Polit Sci* (2024). https://doi.org/10.1057/s41304-024-00482-9
[377] Article 52 1a. AI Act
[378] Article 52 c. AI Act

© European Audiovisual Observatory (Council of Europe) 2024

generated by artificial intelligence systems without subjecting it to human review or editorial control).[379]

However, as the thin line between enhancing and generating images blurs, so do the boundaries between real and "generated" and "false" images vanish. Real news may be illustrated by AI-generated photos,[380] whereas strategic disinformation is often illustrated with real, but decontextualised photo or video.[381]

Among generative AI models, the linguistic models are currently the most developed and widely used and these are the most used for generating disinformation as well. Paradoxically, machine-generated text can appear more credible than human-written text.[382] This phenomenon is attributed to the "fluency bias" and the "truthiness" factor, when users conclude from unrelated features, such as the quality of the text, that the text is trustworthy. These allow malevolent actors to exploit users' intuitive conclusions.[383] These same actors may also deliberately introduce grammatical errors and typos to make the text seem more "organic".

7.2.3. Deepfakes

AI-generated or manipulated videos are generally called deepfakes. The same disclosure obligation applies as for other generative AI products. Even where the product serves artistic, creative, satirical or fictional analogous work or programmes, the AI use should be revealed but merely in a way that does not hamper the display or enjoyment of the work. Videos are often manipulated either by face swapping or face alteration. Swapping the original face in a video to the face of a public figure has often been used in sarcastic political opinion videos, similarly to analogue caricatures.[384] The imperfection of the technology is used as an advantage to achieve the "surprise effect" that is necessary for humour to function. For a fully fictitiously generated fake video, many items in the picture should be generated by AI, such as the background, lights, and movements. Such fakes are not easily created: we see them in movies, but not in free online disinformation (yet). Another popular technique is face alteration, e.g. lip syncing. Simple online tools allow existing videos of public figures to be altered so that they represent the person as saying something different from that which they originally said, such as the deepfake of Maria

[379] Article 18 1e. EMFA
[380] See: Ghost Archive, Stock images from the Gaza war. https://ghostarchive.org/archive/aFK4n; Fingas, J. "Adobe Accepts AI-Generated Stock Art, with Limits: The Company Thinks It Can Minimize the Risk of Copyright Disputes." *Engadget*, December 5, 2022
[381] https://rtl.hu/belfold/2023/11/23/orosz-propaganda-dezinformacio-magyar-kormanykozeli-sajto
[382] Zellers, R., et al. "Defending against neural fake news." *Advances in neural information processing systems* 32, 2019
[383] Hanson, R., Grissom, A.R., and Mouton, C.A., "The Future of Indo-Pacific Information Warfare: Challenges and Prospects from the Rise of AI" *RAND Corporation*, 2024. p. 4
[384] RoW, "A political speech meets a cinematic betrayal", *Rest of World* 2024, AI Elections Tracker, 2024.

© European Audiovisual Observatory (Council of Europe) 2024

Ressa,[385] of the German daily news show,[386] and several others for criminal uses in the private interest.

Such deepfakes are not as widespread as to have a meaningful impact on the public discourse. However, by interfering in high-stakes discourses, they carry a specific risk of inducing important individual or political decisions based on false impression, for example, the deepfake depicting Ukrainian Commander-in-Chief General Valery Zaluzhny seemingly announcing that President Zelensky had assassinated his assistant and was surrendering, and calling upon Ukrainian citizens to launch a coup.[387] Such deepfakes also represent a risk in private and business relationships. For example, when a zoom meeting is interrupted by "Elon Musk" entering the call,[388] this may be considered by some to constitute a funny hack, but it signals a potential leak of business secrets and misleading of high-level government officials in real-time. This, similarly to the Doppelgänger disinformation campaign, raises the issue of cybersecurity.[389] The highly coordinated Russian disinformation campaign dubbed Doppelgänger by the EU Disinfo Lab cloned at least 17 authentic European media providers' websites, to mislead users into believing they were seeing the original website.[390]

7.2.4. Audio

According to EDMO, AI-generated audio currently presents the biggest concern regardingdisinformation.[391] This technology is approximately as advanced as text generation, and in combination, they can mutually enhance each other's potential, as well as overall impact. They are "cheap and easy"[392] and tests have generated convincing election lies.[393] It seems a miracle that we have not had more known incidents. Merely a few political disinformation actions by voice imitation have been reported, from the US, UK,[394] Slovakia, and Sudan. In Slovakia, it was strategically used in the last days of the 2023 parliamentary election which prevented mainstream media from discussing and

[385] Vera Files, "VERA FILES FACT CHECK: Maria Ressa NOT endorsing 'bitcoin platform'", VERA Files, 28 Feb. 2024
[386] Reveland, C., and Siggelkow, P., "Falsche tagesschau-Audiodateien im Umlauf", *Tagesschau,* 13 Nov. 2023
[387] EEAS, "2nd Report on FIMI and Interference Threats", January 2024
[388] Fanatical Futurist by 311 Institute "DeepFake Elon Must bombs a Zoom call", *Youtube,* 2021.
[389] Alaphilippe, A., et al, "DoppelGanger: Media Clones Serving Russian Propaganda", *EU Disinfo Lab*, 27 Sep. 2022. Goujard, C., "Big, bold and unchecked: Russian influence operation thrives on Facebook", *Politico,* April 2024., Cook, L. and Chan, K., "AI could supercharge disinformation and disrupt EU elections, experts warn", *AP,* 5 June, 2024
[390] Council of the EU, "Information manipulation in Russia's war of aggression against Ukraine: EU lists seven individuals and five entities/", Press Release, 28 July, 2023.; Bouchaud, P., Faddoul, M., and Buse Cetin, R., "No Embargo in sight", AI Forensics, 2024
[391] EDMO, "Prebunking AI-generated disinformation ahead of EU elections", March 24, 2024.
[392] Jingnan, H., "It's quick and easy to clone famous politicians' voices, despite safeguards", *NPR, Untangling Disinformation,* 31 May, 2024
[393] Swenson, A., "Tests find AI tools readily create election lies from the voices of well-known political leaders", *AP News,* 31 May, 2024
[394] Meaker, M., "Deepfake Audio is a Political Nightmare", *Wired*, 9 Oct, 2023.

© European Audiovisual Observatory (Council of Europe) 2024

debunking the online "leak" audio of the candidate allegedly planning election fraud.[395] In the Sudanese case, the focus is not on the content of the audio, which was sourced from an innocent blogger, but rather the apparent public reappearance of Sudan's former leader, an accused war criminal, who had avoided public appearances for a year and was believed to be seriously ill.[396]

7.2.5. Bots

While genAI is the latest hype, other types of bots are still extremely active and functional in disseminating disinformation. They automatise dissemination, fake participation and support. Combined with genAI, they can automatically design and carry out an online disinformation campaign, by automatically distributing personalised messages based on users' behavioural profile, without the involvement of human "trolls".

As an interesting example, X's chatbot "Grok", which is supposed to give context and headline to trending topics, misrepresented a trending video that shows a missile attack in the Ukraine war. It amplified the video with the headline referring to an Iranian attack on Tel Aviv..[397] While this piece was not a coordinated disinformation campaign, dissemination by bots can have a similar effect. Bots are also able to deploy anthropomorphising strategies such as posting at irregular intervals and with less consistency, or purposely making spelling errors or using trendy words and phrases. In addition, empathetic chatbots are being intensively developed and tested for healthcare, educational use and potentially customer service. While the academic discussion is diverse on this,[398] the threshold might be lower than we think: tests showed the Google chatbot to be more empathetic than real doctors.[399] As a side-effect of such experimenting, extremist groups are also using chatbots to recruit and convince followers. Researchers found that extremist bots were more successful at persuading prospective members to join than human head-hunters. According to their findings, this was due to the bot's upbeat tone, quick response, and non-judgmental attitude towards applicants.[400]

---

[395] Kőváry, K., "Slovakia: Deepfake audio of Denník N journalist offers worrying example of AI abuse", *IPI NEWSROOM,* 31 Oct, 2023
[396] Goodman, J., and Hashim, M., "AI: Voice cloning tech emerges in Sudan civil war", *BBC*, 5 October 2023.
[397] Thomas, J., "No, Iran has not started attacking Tel Aviv" *Euronews*, April 11, 2024
[398] Seitz, L., "Artificial empathy in healthcare chatbots: Does it feel authentic?" Computers in Human Behavior: Artificial Humans, Vol.2., Issue 1, 2024
[399] Quach, K., "Google AI chatbot more empathetic than real doctors in tests", *The Register,* 16 Jan 2024
[400] Mantello, P., and Ho, M. T., "Losing the information war to adversarial AI", *AI & Society*, 28 April 2023

© European Audiovisual Observatory (Council of Europe) 2024

7.2.6. AI and misinformation

Training genAI with unreliable data, including disinformation, leads to the generation of misinformation. The models themselves, despite their apparent accuracy, are not capable of fact-checking themselves.

The objective of AI training is not to convey truth, nor is it to subject the content discussed to critical evaluation. It merely recompiles and reproduces content that has already been written. This leads to occasional confabulations (popularly called hallucinations).[401] Both misinformation and confabulations by LLMs could be then mistakenly used and disseminated as part of a journalistic product, and thus further pollute the sea of information.

Telling what truth from falsity is presents a profound philosophical, semantic, legal and epistemological challenge. For the same reason, this task cannot be expected from AI chatbots even in the future. Considering how widespread they might become as conversational agents, this might be a cause for concern as well as for regulation. Voice-enabled chatbots may give the perfect illusion of an intelligent discussion partner and are already embedded in smartphones.

7.3. The fight against disinformation

7.3.1. Regulation

7.3.1.1. Legislative rules

As the quality of content cannot be the basis for prohibition, legal tools focus on the techniques that increase the impact of falsity and manipulation. As described above, the AI Act requires that the use of deepfakes and generative AI be made transparent. Specifically, providers of genAI systems must ensure that outputs of their systems are watermarked and thus detectable.[402] Purposefully manipulative or deceptive techniques are also prohibited but only with narrow conditions. Elements of the legal prohibition are that the use of such techniques takes place with the purpose or with the effect of material

[401] The term "hallucination" is inaccurate as hallucination means false sensory perception, which is not the case with genAI. Scientists have proposed the word "confabulation" which means "mistaken reconstructions of information which are influenced by existing knowledge, experiences, expectations, and context", Smith, A. L., Greaves, F., and Panch, T., "Hallucination or confabulation? Neuroanatomy as metaphor in large language models." *PLOS Digital Health* 2.11, 2023; Maleki, N., Padmanabhan, B., and Dutta, K., "AI Hallucinations: A Misnomer Worth Clarifying." *arXiv preprint arXiv:2401.06796* (2024); "Emsley, R. ChatGPT: these are not hallucinations – they're fabrications and falsifications." *Schizophr* 9, 52, 2023

[402] Article 52 1a AI Act

© European Audiovisual Observatory (Council of Europe) 2024

distortion of behaviour; impairing the manipulated or deceived person's ability to make an informed decision; and that the decision ultimately taken caused harm or is likely to cause significant harm to individuals.[403] One typical scenario under this definition would be fake voice scams when a phone caller imitating the voice of a family member presses the victim to transfer a sum of money.[404]

The DSA lists among very large online platforms' and search engines' (hereafter together called: VLOPSE's) systemic risks, among others, dissemination of illegal content and "any actual or foreseeable negative effects on civic discourse and electoral processes, and public security". The fight against illegal content is thus duplicated: it must be removed when the provider acquires knowledge about it, and is listed as a systemic risk as well. The latter underscores the basic obligation of making efforts to maintain a safe platform environment. When conducting such risk assessments, VLOPSEs must consider their algorithmic and recommendation systems and their content moderation systems, among other elements.[405] In case a service provider notices the intentional manipulation of its service, or deceptive high-volume commercial content, such as inauthentic use by bots, fake accounts or large disinformation schemes, it should remove the content or restrict its visibility. The latter includes techniques such as downranking the content, blocking the user without their being aware (shadow banning), or demonetising the service. Reasoning is not required when the cause is such inauthentic use, nevertheless, the user is still entitled to effective remedy before national courts.[406]

7.3.1.2. The Code of Practice on Disinformation

The DSA is completed with codes of practice, among them the Strengthened Code of Practice on Disinformation (2022) [407] as the most relevant in this context.[408] Rather than directly tackling disinformation content, the Code addresses the circumstances of online communication, in particular two clusters: advertising and manipulative practices. The signatories to the Code commit to separating advertising from disinformation, with regard to the entire value chain of advertising, including online e-payment services, e-commerce platforms, crowdfunding or donation systems. The Code also incorporates the draft EU

[403] Article 5 AI Act
[404] Bethea, C., "The Terrifying A.I. Scam That Uses Your Loved One's Voice", *The New Yorker,* 7 March 2024.
[405] Article 34 DSA
[406] Recital 55 DSA
[407] Strengthened Code of Practice on Disinformation, 16 June 2022. Other codes are the Code of Conduct countering illegal hate speech online, the codes of conduct for online advertising (Article 46 DSA), and the codes of conduct for accessibility (Article 47 DSA)
[408] More information on the Code of Practice on Disinformation in the European Audiovisual Observatory's report: Cabrera Blázquez F.J., Cappello M., Talavera Milla J., and Valais S., "User empowerment against disinformation online", IRIS Plus 2022-3, European Audiovisual Observatory, 2022

© European Audiovisual Observatory (Council of Europe) 2024

Regulation on Political Advertising,[409] to provide transparency already before it comes into force.[410]

Regarding AI-assisted manipulative behaviour, the Code specifically names fake accounts, account takeovers (such as the Doppelgänger scheme), bot-driven amplification, hack-and-leak, impersonation, and malicious deepfakes.[411] The Code invites developers or operators of AI systems that are capable of creating or disseminating deepfakes and other AI-generated content, to commit, by respecting the recommended transparency commitments and the list of prohibited manipulative practices listed in the AI Act. Furthermore, the Code sets out commitments to empower users through fostering media literacy, safety, and accountability in the platform design, and to empower the fact-checking community, expanding the effect of the AI Act's requirement for AI literacy.[412]

7.3.1.3. Enforcement

The giant platforms have were providing transparency reports for the third time at the time of writing.[413] Most of them made appreciable efforts to combat disinformation and safeguard against generative AI systems on their services.[414] Still, despite their, and the Code's, efforts to design measurable commitments, recent analysis has indicated that the compliance rate falls short of what was desired.[415] Besides the suspicion that their policies and practices to combat deceptive advertising and ensure transparency of political content do not satisfy the DSA requirements, the procedure also concerns the lack of effective AI-monitoring tools ahead of the EP elections, and the elimination of the

[409] This Regulation was published on 20 March 2024, however, it will become effective only 18 months after its publication in the Official Journal. Griera M., "EU agrees on political advertising rulebook effective after Parliament elections", *Euractive,* 7 Nov. 2023

[410] Commitments 4-9. Strengthened Code of Practice on Disinformation.

[411] The actions "the purchase of fake engagement", and "artificially amplifying the reach or perceived public support for disinformation" would be included in "bot-driven amplification", in my view, therefore they are not separately mentioned in the list. See Commitment 14.

[412] EU AI Act, "Article 4: AI literacy"

[413] European Commission, "Online platforms put special focus on elections in the third batch of reports under the Code of Practice on Disinformation", *Pub Affairs Bruxelles*, 26 March 2024

[414] EDMO, "Prebunking AI-generated disinformation ahead of EU elections", March 24, 2024

[415] EDMO News, "EFCSN review of the fulfillment of the Code of Practice on Disinformation by the very large online platforms and search engines", 11 January 2024.; Hernández-Echevarría, C., "Major Tech Platforms Fail to Deliver on EU Fact-Checking Commitments, Risking DSA Compliance", 11 Jan, 2024.; Lai, S., and Yadav, K., "Operational Reporting in Practice: The EU's Code of Practice on Disinformation", *Carnegie Endowment for International Peace*, 21 Nov. 2023

© European Audiovisual Observatory (Council of Europe) 2024

CrowdTangle public monitoring tool[416] without adequate replacement.[417] The researcher and activist community is protesting against its shutdown.[418]

The German Parliament drafted a bill that would criminalise deepfakes that depict a person in violation of that person's rights (without their consent).[419] This amendment to the criminal law would render such deepfakes illegal in Germany, and therefore subject to the removal obligation under DSA. As DSA refers the definition of "illegal" to Member States competence, other Member States may also follow this route.

7.3.2. Factchecking with the help of AI

Fact-checking is helpful in getting a clear picture of what is happening in the information landscape, and in debunking meaningful pieces of disinformation. Even if reacting to each and every disinformation piece might seem like a whack-a-mole game and remain ineffective, knowing the trends helps design policy strategies.

7.3.2.1. Identifying disinformation with traditional methods

Even though genAI may increase the credibility of the generated disinformation, the distribution patterns of inauthentic information remain the same, therefore, the tools for fact-checking have also remained effective.[420] However, the qualitative evaluation of content still requires human work.[421]

Guo et al. identified three steps for identifying disinformation: 1) claim detection, 2) evidence retrieval and 3) claim verification.[422] Formulating the suspicion based on content requires intensive training on a vast amount of data, and is currently not the most promising method.[423] Rather, a combination of recognising suspicious patterns of

[416] CrowdTangle is a content analysis tool provided by Facebook to analyse all public content by Meta, including engagement data, such as shares, views, comments, likes and other reactions. It was applied by publishers, journalists, researchers and fact-checkers. See Tess, "What data is CrowdTangle tracking?" *Crowdtangle*, 2023-2024

[417] Albert, J., "Facebook's gutting of CrowdTangle: a step backward for platform transparency", *Algorithmwatch*, 3 August, 2022

[418] Letter Urging Meta to Maintain CrowdTangle Tool Through Upcoming Elections, May 15, 2024. See also: Open Letter to Meta, Mozilla Foundation, 2024

[419] Bundestag: Gesetzesantrag: Entwurf eines Gesetzes zum strafrechtlichen Schutz von Persönlichkeitsrechten vor Deepfakes. 14.05.2024

[420] EEAS, "2nd Report on FIMI and Interference Threats", January 2024

[421] EDMO: Fact-Checking Overview

[422] Guo, Z., Schlichtkrull, M., and Vlachos, A., "A Survey on Automated Fact-Checking", *Transactions of the Association for Computational Linguistic* 2022; 10 178–206

[423] See more in: Grimme, Ch., and others (eds), "Disinformation in Open Online Media: First Multidisciplinary International Symposium", *MISDOOM 2019,* Hamburg, Germany, February 27 – March 1, 2019, Revised Selected Papers, vol 12021 (Springer International Publishing 2020)

© European Audiovisual Observatory (Council of Europe) 2024

dissemination, propagation and speaker type is applied with the best success rates.[424] AI can also be helpful in retrieving evidence, such as finding further contextual information beyond the original claim. Claim verification is the most complex task which includes interpreting and assessing the information, and comparing it with the retrieved set of evidence.[425] Current artificial fact-checking (AFC) technologies are not well suited to this task, but there are experiments already with LLMs.[426] However, the fundamental problem is the reliability of the information retrieved for evidence. There must be human staff manual supervision,[427] or reliance on a human-prepared database of previously fact-checked information. Especially in cases where information is not readily available, for example during a crisis situation, or due to deliberate restrictions by authoritarian states, maintaining an up-to-date and comprehensive database is a big challenge, or even impossible.[428] Human fact-checkers, however, can be inventive in uncovering hidden clues to refute claims, for example in the case of decontextualised videos.

Once data is available, it can be structured by AI to make it easily accessible for other AI tools as trustworthy databases for automated fact-checking (e.g., ClaimReview),[429] thereby narrowing the topics that are vulnerable for disinformation. Further, AI is also useful in tagging already fact-checked materials, disseminating them and streamlining international collaboration. AI is further used to simplify the publication process (Chequeado) or to share real-time fact-checking results through a mobile app even live during political speeches (Factstream).[430]

Recognising propaganda narratives with the help of AI analysis has also been successful. In a study, the ML system was trained to capture the frequency of whole sets of topics in the mainstream media of a nation, rather than individual words, to create a baseline. Then the app associated each word with a multidimensional vector, to establish a relationship between the terms, based on the angles and distances of the vectors, using hundreds of dimensions. The model trained this way was then given opposing words (good-bad), to which each article was assigned with a score, based on the matrix of vector-analysed text, with the ability to span several sentences, rather than words only. The AI scores demonstrated that in September 2021, the pro-Russian propaganda in the Hungarian mainstream (i.e. government-dominated) press significantly grew. Human researchers came independently to the same results. The model can be constantly

---

[424] McGovern, A., "Artificial Intelligence System Could Help Counter the Spread of Disinformation", *MIT News | Massachusetts Institute of Technology*. See also: Smith, S.T., et al., "Automatic detection of influential actors in disinformation networks", Computer Sciences, Jan. 7, 2021; See also: Cassauwers, T., "Can Artificial Intelligence Help End Fake News?" *Research and Innovation* 15 April 2019.

[425] Sittmann, J., Tompkins, A., "The strengths and weaknesses of automated fact-checking tools" *Deutsche Welle Akademie,* 17.07.2020

[426] Guo, Z., Schlichtkrull, M., and Vlachos, A., "A Survey on Automated Fact-Checking", *Transactions of the Association for Computational Linguistics* 2022; 10 178–206

[427] Borel, B., "*The Chicago guide to fact-checking*". University of Chicago Press, 2023

[428] Graves, L., "Understanding the Promise and Limits of Automated Fact-Checking", Reuters Institute, Oxford, February 2018

[429] See: https://www.claimreviewproject.com/

[430] Adair, B., "FactStream app now shows latest fact-checks from Post, FactCheck.org and PolitiFact", *Reporterslab*, 7 October, 2018

© European Audiovisual Observatory (Council of Europe) 2024

updated and responds to similar prompts within minutes.[431] This model follows the old path of semi-supervised learning, which combines reasonable amounts of human labour with reliable outcomes, and thereby reaches the best scores.[432] Developing the model required manual labour and remained context-bound, but it was capable of recognising more complex meaning than earlier models.

7.3.2.2. Recognition of Gen-AI through AI

Occasionally, AI text appears more trustworthy than human-written text, especially as efforts are made to reduce synthetic perfection.[433] Still, developers are working on identifying AI-generated content. For example, synthetic image detection tools focused on properties of the generative architectures.[434] Other researchers found that a person's representation on manipulated videos or audios is inconsistent, in terms of tiny details, with the real identity, which becomes detectable through AI, provided a real talking-face video or real audio record is at hand.[435] The same logic is not applicable with text though, as text in itself is constructed; in addition, many genuine texts deploy AI enhancement. Even though the AI Act prescribes the obligation of watermarking, this is reportedly easily to remove. At the same time, it is easy to add a fake watermark to discredit authentic content, and also to add a falsified watermark that is supposed to prove that the image comes from a real camera.[436] To sum up, AI tools are capable of supporting savvy humans in detecting disinformation, and extrapolating the findings to reach wider audiences, but the human in the loop cannot and should not be omitted.[437] In addition, they can be useful in generating and disseminating easily accessible, truthful information, with a "watermark" of trustworthiness.

---

[431] Ries, U., "AI Helps Uncover Russian State-Sponsored Disinformation in Hungary", *DarkReading*, 20 November, 2023. Note the cited research of Novak and Weddingensen

[432] Xin, L. and others, "A Novel Self-Learning Semi-Supervised Deep Learning Network to Detect Fake News on Social Media", *Multimedia Tools and Applications,* 2021.

[433] Zellers, R., et al. "Defending against neural fake news." *Advances in neural information processing systems* 32 (2019)

[434] R. Corvi, and al., "On The Detection of Synthetic Images Generated by Diffusion Models," *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023, pp. 1-5

[435] Cozzolino, D., et al., Audio-visual person-of-interest deep fake detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,* 2023 (pp. 943-952)

[436] Hoffman-Andrews, J., "AI Watermarking Won't Curb Disinformation." *EFF*, 5 January, 2024

[437] Montoro-Montarroso et al. (., "Fighting disinformation with artificial intelligence: fundamentals, advances and challenges." *Network activisms*, Vol. 32. N. 3. 2023

© European Audiovisual Observatory (Council of Europe) 2024

7.4. Conclusion

One of the greatest challenges posed by genAI lies in the sheer volume and velocity of content that can be created, and the endless capacity to tailor messages to diverse audiences, increasing the credibility of the disinformation through personalisation.[438]

And the biggest threat presented by the potential of AI-generated content is, again, that we lose another set of our traditional anchors of trust: high quality, good structure, word choice or the evidential value of picture and video. Our epistemic realities are becoming even more confused. New anchors of trust and evidence systems must be developed.

We also need to prepare to deal with the other angle of the same problem: discrediting authentic content as AI-generated, for instance, when claiming that the official photo of the Bulgarian prime minister giving a speech at the European Parliament was a fake.[439] False positive results of openly accessible fact-checking systems may feed such distrust. Exploiting the hype around the threat posed by AI[440] is another weapon that can sow general distrust in the democratic system, such as the claim that any content may be manipulated, and hence, objective truth is non-existent.[441]

[438] Wach, K., et al., "The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT," *Entrepreneurial Business and Economics Review* 11, no. 2, 2023, 7-24. See also: Dobber, T., et al., "Do (Microtargeted) Deep fakes Have Real Effects on Political Attitudes?," *The* International Journal of Press/Politics 26, no. 1, 2021, 69–91

[439] Bontcheva, K., "Generative AI and Disinformation: Recent Advances, Challenges, and Opportunities", White Paper, Vera.ai project, February 2024.

[440] EEAS, "2nd Report on FIMI and Interference Threats", January 2024.

[441] Political Capital, "A significant part of Hungarians already doubt the existence of facts and reality", *PCBlog.Átlátszó*, 30 November, 2023.

© European Audiovisual Observatory (Council of Europe) 2024

8. Diversity and Pluralism

Mira Burri, University of Lucerne

This chapter examines the affordances of AI and how and to what extent these have impacted the conditions of freedom of expression, as well as media pluralism and cultural diversity. While acknowledging some of the positive implications, such as potentially improved accessibility and efficiency in content creation, the chapter, in particular, sheds light on the negative implications of AI, such as content personalisation, bias and representation, curation and gatekeeping, as well as the broader disruptive impact on traditional media models that may degrade trustworthiness of the available information, distort public discourse and reduce cultural diversity.

8.1. Setting the scene: AI as a disruptive technology

The digital transformation epitomised by the advent and spread of the Internet continues to change the ways content is produced, distributed, accessed, and consumed, and has modified the patterns of user experience and participation with critical implications for pluralism and diversity.[442] AI as a set of new technological developments marks a new stage of a, very likely, disruptive transformation that comes with distinct challenges. It is the purpose of this chapter to explore the affordances, i.e. its current and potential capabilities, of AI and its effects on freedom of expression, media pluralism and cultural diversity. The chapter highlights distinct concerns that have been raised in this context that may call for measures, of a regulatory and non-regulatory nature, to address the new environment, so that core societal values are safeguarded, and the benefits of AI reaped for a healthy and sustainable pluralistic cultural environment.

For starters, it should be noted that there is no commonly agreed upon definition of AI – not only because technology evolves so rapidly but also because perceptions, societal contexts, and ethical sensitivities may vary. For the purpose of this chapter, we use the (recently updated) definition of the Organisation for Economic Co-operation and Development (OECD), which captures in a neutral and succinct way the different

[442] See e.g. See Y. Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (Yale University Press, 2006); C. R. Sunstein, *Republic.com 2.0* (Princeton University Press, 2007).

© European Audiovisual Observatory (Council of Europe) 2024

affordances of AI.[443] While the impact on the cultural sector of conventional AI systems that are predominantly used to analyse data and make predictions, has been discussed, although not conclusively,[444] the more recent genAI poses new challenges. In particular, since these AI systems, such as ChatGPT, can generate *new* content and outputs that resemble but also expand upon the patterns and structures present in their training data.[445]

Overall, AI has become recognised as a general purpose technology (GPT)[446] that has an enormous transformative potential with spillover effects across diverse sectors and industries that go beyond its initial applications.[447] In addition and perhaps in contrast to other GPTs, such as the printing press, electricity or indeed the internet, AI is of iterative nature – in the sense that AI systems continuously learn and adapt their performance through exposure to new data, experiences and feedback loops.[448] GPTs in general and AI in particular, do pose a major regulatory challenge, as their evolution and distribution is neither linear nor predictable.[449]

[443] The OECD definition and the EU's AI Act definition are presented in chapter 1.

[444] See e.g. M. Burri, "Cultural Diversity Policies in the Age of AI", *AI in the Audiovisual Sector, IRIS Special Report* (European Audiovisual Observatory, 2021), 69–84

[445] See e.g. C. Gros and D. Gros, *Generative AI: A Concise Primer for Non-Experts* (University of Bocconi, 2023)

[446] See e.g. B. Jovanovic and P. L. Rousseau, "*General Purpose Technologies"*, in P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth* (Elsevier, 2005), 1182–1224; J. Manyika et al., *Disruptive Technologies: Advances That Will Transform Life, Business, and the Global Economy* (McKinsey Global Institute, 2013).

[447] Another example of GPT from history is the printing press. It was first used as a way to make the Bible accessible but became instrumental for the leaders of the Reformation, who adopted the technology to print the pamphlets that spread the movement. The printing press also helped spark the scientific revolution and the Enlightenment by disseminating research and discoveries. Indirect effects included accelerated city growth. Some historians attribute Europe's rapid growth and global influence and the eclipse of Islamic nations after the 15th century to the quick adoption of printing in Europe and its slow adoption in Islamic economies. See Manyika et al. (2013), ibid., at 25; J. E. Dittmar, "Information Technology and Economic Change: The Impact of the Printing Press", *The Quarterly Journal of Economics* 126 (2011), 1133–1172

[448] On GPTs in general, see e.g. E. Brynjolfsson, D. Li and L. R. Raymond, "Generative AI at Work", *NBER Working Paper* 31161 (2023). On AI as a GPT, see e.g. EU AI Act, at Preamble

[449] See e.g. Y. Benkler, "Growth-Oriented Law for the Networked Information Economy: Emphasizing Freedom to Operate over Power to Appropriate", in Kauffman Taskforce on Law, Innovation and Growth (ed), *Rules for Growth: Promoting Innovation and Growth through Legal Reform* (Kauffman Foundation, 2011), 313–342; N. Helberger, J. Pierson and T. Poell, "Governing Online Platforms: From Contested to Cooperative Responsibility", *The Information Society* (2017); B. Judge, M. Nitzberg, S. Russell, "When Code Isn't Law: Rethinking Regulation for Artificial Intelligence", *Policy and Society* 43 (2024) [advance article].

© European Audiovisual Observatory (Council of Europe) 2024

8.2. AI's impact on freedom of expression, media pluralism and cultural diversity

8.2.1. Introductory remarks: what's different?

It is important to understand that the implications of conventional AI and generative AI are both important for the availability, distribution and consumption of content but may also differ, and have become intertwined. With regard to AI-driven recommender systems, what was key was the changed dynamics of the media space and the emergence of these systems as critical intermediaries with distinct editorial functions. This was in stark contrast to the offline/analogue world, where editorial roles were concentrated under the roof of a single media institution (be it a newspaper or a broadcaster), which was also given a certain regulatory mandate – among others, to feature local and national content; to ensure the quality and trustworthiness of information. This legacy model also supported the assumption, which underlies almost all national media and cultural policies, that diversity in supply will be reflected in diversity of consumption – with the follow-up assumption of beneficial effects on opinion building, political participation and cultural engagement.

AI-enabled algorithmic-driven platforms, such as notably social media sites, completely changed this picture and triggered "fundamental shifts in the composition and consumption of media products".[450] These "new editors" are multiple, disintegrated and distributed[451] and it is ultimately their embedded algorithms[452] that define the new media space and condition the exercise of freedom of speech, both in its active and passive dimensions.[453] And again, this brings about broader implications for pluralism and diversity.

[450] R. Kleis Nielsen, R. Gorwa and M. de Cock Buning, *What Can Be Done? Digital Media Policy Options for Strengthening European Democracy* (Reuters Institute Report, 2019).
[451] M. Latzer, K. Hollnbuchner, N. Just and F. Saurwein, "The Economics of Algorithmic Selection on the Internet", in J. M. Bauer and M. Latzer (eds), *Handbook on the Economics of the Internet* (Edward Elgar, 2016), pp. 395–425
[452] For a comprehensive definition of algorithms, see M. Latzer and N. Just, "Governance by and of Algorithms on the Internet: Impact and Consequences", in *Oxford Research Encyclopedia, Communication* (Oxford University Press, 2020).
[453] J. Balkin, "Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation", *UC Davies Law Review* 51 (2018), 1149–1210.

© European Audiovisual Observatory (Council of Europe) 2024

8.2.2. Implications for content distribution and consumption

In this context, it is first necessary to acknowledge the possible interferences with users' individual autonomy and freedom of choice. While AI-driven curation and recommender systems reduce search and information costs and may facilitate social orientation,[454] this can be compromised by the production of social risks, such as "threats to basic rights and liberties as well as impacts on the mediation of realities and people's future development".[455] Important in this context is also the increasing personalisation of the media diet, which has become based on (and biased by) users' social media profiles and previous experiences (e.g. "likes"; "friends"; browsing history, including location), and ultimately "promotes content that is geographically close as well as socially and conceptually familiar".[456] This keeps users within certain familiar boundaries, feeding their curiosity with more of the same, which reinforces existing opinions and removes conflicting ones.[457] One can of course state that this has been the case with legacy media as well, where people were naturally drawn to content that they have liked in the past – the key difference in the current space is that users see *only* this content, and their active choice is so lessened or manipulated. Hoffman et al. argue that social media only exacerbate this effect by combining two dimensions of "homophily" – that is, similarity of peers and of content.[458]

One should keep in mind in this context that despite a slight reduction in the use of social networking sites as an entry point to content and variations across countries,[459] they still are important gatekeepers. This reinforces the effect of homophily and potential bias, as well as clearly illustrates the power of a few players and the deep impact of their decisions – decisions that are notably taken with primarily economic motivation[460] and under no specific mandate to safeguard certain rights and/or values.

The radical increase in commercially or politically driven "fake news" with serious repercussion for democracies[461] but also for other fundamental values, such as public

---

[454] Latzer et al. (2020).
[455] Ibid., at 29–30.
[456] C. P. Hoffman, C. Lutz, M. Meckel and G. Ranzini, "Diversity by Choice: Applying a Social Cognitive Perspective to the Role of Public Service Media in the Digital Age", *International Journal of Communication* 9 (2015), 1360–1381.
[457] Ibid. For earlier literature, see e.g. C. R. Sunstein, *Going to Extremes: How Like Minds Unite and Divide* (Oxford University Press, 2009); E. Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (Viking, 2011).
[458] Ibid.; also J. A. Tucker et al., *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature* (Hewlett Foundation, 2018).
[459] For country analyses, see Reuters Institute, *Digital News Report 2024* (Oxford, 2024).
[460] For instance, Meta has been trying to reduce the role of news across Facebook, Instagram and Threads, and has restricted the algorithmic promotion of political content. It has also reduced support for the news industry, including by not renewing deals and removing its news tab in a number of countries. See Reuters Institute (2024).
[461] See e.g. L. W. Bennett and S. Livingston, "The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions", *European Journal of Communication* 33 (2018), 122–139. V. Dan et al., "Visual Mis- and Disinformation", *Journalism and Mass Communication Quarterly* 98 (2021), 641–664.

© European Audiovisual Observatory (Council of Europe) 2024

health, should also be underscored.[462] As disinformation becomes widespread on platforms, algorithmic amplification may also augment its negative impact, "legitimizing deceptive narratives, conspiracies, and untruths. In that sense, the algorithmic recommendation of deceptive content may impede learning from the other side, create an illusion of social support, and herewith reproduce (dis)information biases".[463] The new affordances of AI only enable more sophisticated versions of misrepresentation, especially in images and videos; there are also dangers linked to mistakes and hallucinations that AI systems are often prone to.[464]

In the context of both mis- and disinformation and algorithmic bias, it should be noted that despite a slight shift towards reader payment models for news,[465] it is worth remembering that the vast majority of online consumption still happens through few platforms that are largely supported by advertising and thus dependent on the battle for viewers' attention.[466] While some of the aggregated content stems from legacy media,[467] which may disperse some of the conventional criticism that aggregators amplify the impact of unreliable non-traditional sources,[468] it is still true that content is not made more abundant but has merely become more distributed. There is also a valid question to be asked as to how diverse the available content is, since some of it may be simple re-used entertainment or commercially driven productions (such as influencers' vlogs) that also have low value in enriching cultural and/or political engagement. Still, it is fair to note that legacy media have responded to the technologically enabled aggregation and offer much more content online than in their print or broadcast versions. With specific regard to news, the Reuters Institute found that legacy news organisations are making major investments in social media and report receiving significant amounts of traffic, off-site reach, and/or additional digital subscribers.[469]

While this may enable access to a variety of content over more platforms, also such that may entice young people, two possible drawbacks need to be highlighted: the

---

[462] See e.g. M. Burri, "Fake News in Times of Pandemic and Beyond: Exploring of the Rationales for Regulating Information Platforms", K. Mathis and A. Tor (eds), *Law and Economics of the Coronavirus Crisis* (Springer, 2022), 31–58.

[463] D. Shin et al., "Countering Algorithmic Bias and Disinformation and Effectively Harnessing the Power of AI in Media", *Journalism and Mass Communication Quarterly* 99 (2022), 887–907, at 890.

[464] Reuters Institute (2024).

[465] Across markets, only around a fifth of respondents (22%) identify news websites or apps as their main source of online news (10% down from 2018). Publishers in a few Northern European markets have managed to counter this trend, but younger groups everywhere are showing a weaker connection with news brands than they did in the past. See Reuters Institute (2024).

[466] Reuters Institute (2024).

[467] Ibid.

[468] While legacy media are important for news (especially in social networks such as Facebook and X), the Reuters Institute (2024) finds an increasing focus on partisan commentators, influencers, and young news creators, especially on YouTube and TikTok. This is also linked to the trend that video is becoming a more important source of online news, especially with younger groups.

[469] They identify three main strategic aims shaping the different ways in which news organisations approach social media: (1) driving on-site traffic through referrals, (2) driving off-site reach through native formats and distributed content, and (3) driving digital subscription sales, often in part through advertising content on Facebook.

© European Audiovisual Observatory (Council of Europe) 2024

first relates to the "platform risk" that comes with it, as legacy companies are highly reliant on few platforms (which next to the traditional players of Google and Meta, now also include TikTok[470]). The second is that private sector legacy news organisations' approaches to social media are strongly shaped by path-dependent business models oriented towards advertising, subscriptions, or a mix thereof.[471]

8.2.3. Implications for content creation

AI has also permitted highly sophisticated metrics as to the distribution and consumption of content, which may lead to a bias towards more mainstream reporting (especially with the new rise of short-video formats[472]) rather than one that is critical and challenging – what some authors have referred to as "click-based versus editorial goals".[473] Thereby, algorithms drive decision-making in media organisations by predicting audiences' consumption patterns and preferences.[474] While in some areas this may be viewed as beneficial in giving the audiences what they want, in other areas, such as for news, this may be highly problematic, as local news and current affairs become tailored to the demographic, social and political variables of specific communities.[475] Overall, this may do little to contribute to a sustainable offer of diverse local, regional and national content[476] and there may be real difficulties in finding it, because it is or becomes marginalised on online platforms.[477]

As mentioned, AI content production is a relatively new phenomenon. There are of course some earlier "unfortunate" examples with the early generation of the so-called "content farms", which, based on search-engine data (such as popular search terms, ad word sales and the actual available content), produced content rapidly and cheaply in

[470] Reuters Institute (2024).
[471] A. Cornia, A. Sehl, D. A. Levy and R. K. Nielsen, *Private Sector News, Social Media Distribution, and Algorithm Change* (Reuters Institute, 2018).
[472] Reuters Institute (2024).
[473] See e.g. T Dodds et al., "Popularity-Driven Metrics: Audience Analytics and Shifting Opinion Power to Digital Platforms", *Journalism Studies* 23 (2023), 403–421; see also N. Helberger, "FutureNewsCorp, or How the AI Act Changed the Future of News", *Computer Law and Security Review* 52 (2024) 105915.
[474] P. M. Napoli, "On Automation in Media Industries: Integrating Algorithmic Media Production into Media Industries Scholarship", *Media Industries Journal* 1 (2014), 33–38; also F. Saurwein, N. Just and M. Latzer, "Governance of Algorithms: Options and Limitations", *info* 17 (2015), 35–49.
[475] Napoli, ibid., at 34. Reuters Institute (2024) found that publishers may be focusing too much on updating people on top news stories and not spending enough time providing different perspectives on issues or reporting stories that can provide a basis for occasional optimism. In terms of topics, they found that audiences feel mostly well served by political and sports news but there are gaps around local news, as well as health and education news.
[476] See e.g. P. M. Napoli, M. Weber, K. McCollough and Q. Wang, *Assessing Local Journalism: News Deserts, Journalism Divides, and the Determinants of the Robustness of Local News* (News Measures Research Project, August, 2018).
[477] M. Burri, "Discoverability of Local, Regional and National Content", A Thought Leadership Paper for the Canadian Commission for UNESCO and the Canadian Heritage, February 2019.

© European Audiovisual Observatory (Council of Europe) 2024

order to meet that demand.[478] Such creation of content is evidently commodified and possibly harmful to any public interest function of the media one may think of, including in the cultural and political spheres. With the new affordances of AI systems, the transformation with regard to content production is much further-reaching. For instance, a white paper of the University of Amsterdam shows how AI is changing journalism and the media – which ranges from new ways of supporting journalists' research, to assistance in writing and new forms of engaging and interacting with the audience, such as through personalised distribution of content or use of virtual agents.[479] On the positive side, this increases efficiency of content production (e.g. with regard to more mundane activities, such as transcription, copy-editing, and layout) and can also be applied to tackle some of the possible negative implications noted earlier, such as for fact-checking and detecting deep fakes.[480] It also has the potential to completely transform content production and distribution to an extent that is at this point barely imaginable.[481]

Yet, there are downsides too. A critical one is the increased dependence on AI-driven solutions and the linked influence of technology companies over content production that plays an important role in all societal facets. As has long been acknowledged, technology companies also come with their own objectives, such as efficiency, scalability and speed.[482] But also values that are linked to their jurisdiction, which at this point is the United States – that reflect deeply rooted perceptions of economic freedom, free speech or privacy and become embedded in the technology itself.[483] In the context of AI dependence and technology companies' power , there may be a mismatch with the paradigm of public service media and the role of journalism in a democratic society that pursues pluralism and diversity.[484] With regard to the primarily US dominance in the AI area, there may be a mismatch between the societal values shared by different countries that is somewhat implicit in the underlying technology but potentially with far-reaching effects.[485]

---

[478] Napoli (2014), at 35.

[479] A. Schjøtt Hansen et al., *Initial White Paper on the Social, Economic, and Political Impact of Media AI Technologies* (AI4Media, 2023); also C. Beckett and M. Yaseen, *Generating Change. A Global Survey of What News Organisations Are Doing with AI* (JournalismAI, 2023).

[480] Schjøtt Hansen et al. (2023), ibid.; see also N. Newman, *Journalism, Media, and Technology Trends and Predictions 2023* (Reuters Institute, 2024).

[481] For insightful future scenario thinking, see Helberger (2024).

[482] Ibid.

[483] See e.g. A. Chander, "How Law Made Silicon Valley", *Emory Law Journal* 63 (2014), 630–694; K. Klonick, "The New Governors: The People, Rules, and Processes Governing Online Speech", *Harvard Law Review* 131 (2018), 1598–1670.

[484] F. M. Simon, "Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy", *Digital Journalism* 10 (2022), 1832–1854; M. Moore and D. Tambini (eds), *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple* (Oxford University Press, 2018); R. Kleis Nielsen and S. A. Ganter (eds), *The Power of Platforms Shaping Media and Society* (Oxford University Press, 2022); J. van Dijck, T, Poell and Ma. de Waal, *The Platform Society: Public Values in a Connective World* (Oxford University Press, 2018).

[485] A lesson learnt in this regard comes from the domain of privacy/personal data protection, as the EU and the US share very different perceptions and have accordingly adopted very different regulatory frameworks.

© European Audiovisual Observatory (Council of Europe) 2024

8.2.4. Additional aspects to consider

Other problematic aspects can be identified. Some are linked to the intrinsic opacity and lack of awareness of AI applications. As AI systems cannot understand the societal context, their filtering may often be inaccurate. Thereby, false positives can lead to unjustified limitations on speech, while false negatives may cause a chilling effect, leading to self-censorship.[486] It should also be acknowledged that AI can facilitate surveillance and censorship with serious implications for the right to seek and receive information, as well as to media pluralism.[487] Inequalities in terms of access to and use of AI systems pose further challenges – this is true across regions and countries but also across sectors and companies in these sectors. In the first sense, this is an amplification of the "original" digital divide between developed and less developed countries with serious implications also for language representation, as training data may be largely in English or other dominant languages within a country and minority languages experience performance would be worsened.[488] Such discriminatory effects may also be true with regard to distinct sectors, and cultural industries can be particularly vulnerable in this regard, as still few artists and entrepreneurs know how to use AI tools and if they do, become dependent on few commodified systems.[489] Finally, there are inequalities within the media sector too, where local and regional media with smaller budgets lag behind and struggle to make good use of AI affordances.[490] This again presents certain dangers of reduced diversity, fragmentation and potential distortion of the public and cultural discourse.

Overall, one can certainly maintain that we are in the midst of a deeply transformational phase of the media space with distinct challenges for pluralism and diversity. There are, however, benefits to be reaped from AI too. Accessibility of information is key but also providing information that is more relevant and potentially more responsive to the interests of a heterogeneous audience.[491] This includes not only contemporary content but also content that comes from archives, including from cultural heritage and memory institutions.[492] As noted earlier, AI tools may also enable more

---

[486] Julia Haas, Freedom of the Media and Freedom of the Media and Artificial Intelligence Artificial Intelligence (Office of the OSCE Representative on Freedom of the Media Representative on Freedom of the Media, 2020); see also N. Helberger et al., "Artificial Intelligence – Intelligent Politics: Challenges and Opportunities for Media and Democracy", Background Paper, CoE Ministerial Conference, Cyprus, 28–29 May 2020.

[487] Haas, ibid., referring also to United Nations High Commissioner for Human Rights, A/HRC/39/29, 2018; S. Feldstein, "The Road to Digital Unfreedom: How Artificial Intelligence Is Reshaping Repression", *Journal of Democracy* 30 (2019), 40–52. A. Ünver, "Artificial Intelligence, Authoritarianism and the Future of Political Systems", *Cyber Governance and Digital Democracy* 9 (2018).

[488] Schjøtt Hansen et al. (2023), at 90.

[489] See e.g. O. Kulesz, Culture, Platforms and Machines: The Impact of Artificial Intelligence on the Diversity of Cultural Expressions, report for UNESCO, DCE/18/12.IGC/INF.4 (2018); B. Caramiaux, "The Use of Artificial Intelligence in the Cultural and Creative Sectors", EU Parliament Briefing PE 629.220 (2020).

[490] Schjøtt Hansen et al. (2023), at 90.

[491] Ibid.

[492] Caramiaux (2020). This comes with its own challenges, however. See B. Caramiaux, "AI with Museums and Cultural Heritage" in *AI in Museums* (De Gruyter, 2023), 117–130; Magdalena Pasikowska-Schnass with Young-

© European Audiovisual Observatory (Council of Europe) 2024

efficient journalistic and editorial processes, which can free up resources for high quality journalism and critical investigation. As to the latter, AI also enables fact-checking and testing the trustworthiness of sources and information. AI tools can also help media companies offer new services, including personalised content and editorial guidance, which can also be linked to new financial models that can assist smaller players.[493] It should also be noted that there is a long way to go until AI systems become fully human-like.[494] Media organisations can invest in "delaying" this process by building trust and showing that their journalism is built on accuracy, fairness, and transparency and that humans remain in control, especially as audiences are likely to respond positively, since they seek diverse and trustworthy content in a highly populated and messy informational space.[495]

8.3. Addressing the real and potential effects of AI systems on pluralism and diversity

In comparison to only three or four years ago, there is a wave of initiatives that regulate platforms because of their critical role in society, including with regard to speech mediation, mis- and disinformation.[496] There is in addition a new complementary wave of regulation that specifically targets AI systems and seeks to prevent and reduce their risks.[497] While these initiatives do not (yet) address specifically the concerns related to pluralism and diversity flagged in this chapter, the implications of these new regulations – e.g. with regard to transparency, accountability and standard-setting – can be significant and trigger impact on the ground.[498] This development will hopefully not merely constrain the employment of AI but support responsible AI practices.[499] There is also a clearly increased awareness among the international community that action needs

Shin Lim, "Artificial Intelligence in the Context of Cultural Heritage and Museums: Complex Challenges and New Opportunities", European Parliament Briefing PE 747.120 (2023).

[493] Schjøtt Hansen et al. (2023); Shin et al. (2022).

[494] Shin et al. (2022), at 902.

[495] Reuters Institute (2024).

[496] See e.g. Burri (2022). The European Union (EU) has been the champion of such initiatives with generic platform regulation coming from the Digital Services Act (DSA), which introduces wide-ranging transparency measures around content moderation and advertising, and binding and enforceable legal obligations in particular for very large online platforms to assess and address systemic risks for fundamental rights or presented by the intentional manipulation of their services, including with regard to fighting misinformation online. In addition, there are a number of more specific instruments, such as the Media Freedom Act or the Code of Practice on Disinformation. See e.g. Burri (2022), ibid.

[497] Here again the EU is at the forefront with the AI Act but other states, although with differences, have adopted AI legislation too. See e.g. T. Giardini and J. Fritz, *The Anatomy of AI Rules: A systematic comparative analysis of AI Rules across the Globe* (Digital Policy Alert, 2024).

[498] See e.g. Helberger (2024); N. Helberger and N. Diakopoulos, 'The European AI Act and How It Matters for Research into AI in Media and Journalism', *Digital Journalism* 11 (2022), 1751–1760.

[499] Helberger (2024).

© European Audiovisual Observatory (Council of Europe) 2024

to be taken and the recent Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law[500] is proof of this.

In order to address the new, AI-dependent, environment, so that core societal values are safeguarded, and the benefits of AI reaped for a healthy and sustainable pluralistic cultural environment, there have been discussions of more targeted action. Schjøtt Hansen et al. provide, for instance, helpful guidance not only on regulatory but also non-regulatory measures that could be taken.[501] They highlight among other things the need to develop alternative data and content infrastructures that move away from the current commodified model and can accommodate specificities – e.g. of the media sector, of certain societal values intrinsic to the community and of diverse and minority languages. This could be linked to the need to have domain-specific, open-source and non-commercial datasets for training AI systems, as well as ones that specifically target embedded "European values". Further, there seems to be a positive need for best practices with regard to disclosure and the responsible use of AI systems for the media sector, as well as "diversity by design".[502] One should of course not forget that any of the measures are unlikely to come about if there are no appropriate funding schemes and incentives. At the same time, when adopting any of these measures, a balance must be struck, so that innovation is not impeded – a discussion that has received some attention in the wake of the EU AI Act.[503]

8.4. Concluding remarks

This chapter showcased mere snippets of the deep transformation of the media space that has been fueled by rapid technological advances, most recently by the advent and wide spread of genAI. This development is radical and has major implications for content creation, distribution, use and re-use, and consumption that we are yet to fully comprehend. This comes with spillover effects on pluralism and diversity as critical pillars of our democratic societies. To safeguard our fundamental rights and societal values, there may be a need for action but the path of taking such action is neither clear nor easy, as the environment is highly fluid and unintended consequences are possible. Awareness is the first step in the right direction – not only in terms of understanding the technological affordances but also of our fundamental values that need to be safeguarded.

---

[500] CM(2024)52-final, 17 May 2024.
[501] Schjøtt Hansen et al. (2023).
[502] Ibid. On the 'diversity by design', see N. Helberger, 'Diversity Label: Exploring the Potential and Limits of a Transparency Approach to Media Diversity', *Journal of Information Policy* 1 (2011), 337–369; N. Helberger, 'Diversity by Design', *Journal of Information Policy* 1 (2011), 441–469.
[503] Some reports maintain, for instance, that there is no need for intervention and that the EU AI Act may stifle genAI innovation in Europe. See e.g. P. R. Abecasis et al., *Generative Artificial Intelligence: The Competitive Landscape* (Copenhagen Economics, 2024)

© European Audiovisual Observatory (Council of Europe) 2024

PART IV – GenAI in our lives: the need for education and awareness

GenAI is now part of our lives, and we must learn to maximise its potential by understanding how to use it effectively. Setting aside the regulatory challenges it presents, and looking forward, there is a need to educate people about AI and increase AI literacy in daily life. This raises questions about the evolving roles of public institutions.

Should school teach small children how to use and evaluate AI-generated content, and if so, what tools should they use? Could the machine eventually assist teachers by helping them personalise content towards their audience, could they possess a human-like sensitivity?

Public service media could also play a role in raising AI awareness. For instance, a future voice assistant may deliver daily news selectively, potentially using information that is not fact-checked or accurate. Should public service media and their news departments collaborate with genAI companies to ensure news accuracy, and a human touch in journalism and media pluralism? What are the implications if news and journalism are not created or mediated by humans?

These questions remain open, and the future will provide answers as we navigate this rapidly changing landscape.

9. The world of tomorrow: are the texts AI-proof and ready for the AV challenges?

Prof. Dr. Mark Cole, Director for Academic Affairs, Institute of European Media Law & Professor for Media and Telecommunication Law, University of Luxembourg / Dr. Sandra Schmitz-Berndt, LAIWYERS project, University of Luxembourg

9.1. Recap: Existing and forthcoming regulatory approaches

With AI capabilities evolving swiftly, regulation has emerged – actually in parts of the world quite rapidly – in a bid to address the risks and challenges arising from AI use. Early initiatives have mainly focused on an ethical use of AI and sought to introduce principles that deployers should adhere to.[504] These principles concentrated on addressing the underlying technology and included fairness, accuracy, and transparency, in more general terms.

9.1.1. Starting with Recommendations: Early approaches by the OECD and UNESCO

The OECD Recommendation on Artificial Intelligence[505] of May 2019 can be considered as the first "global reference point for trustworthy AI".[506] The Recommendation aims to foster innovation and trust in AI by promoting the responsible stewardship of trustworthy AI while at the same time ensuring respect for human rights and democratic values. Notably,

[504] On the evolution of AI regulation see Mark D. Cole, "AI Regulation and Governance on a Global Scale: An Overview of International, Regional and National Instruments"",(2024) 1(1) AIRe 126.
[505] OECD, "Recommendation of the Council on Artificial Intelligence"",(2019) OECD/LEGAL/0449.
[506] Lucia Russo and Noah Oder, "How Countries Are Implementing the OECD Principles for Trustworthy AI" (oecd.ai, 31.10.2023).

© European Audiovisual Observatory (Council of Europe) 2024

the Recommendation also contains a definition of AI which was already revised in November 2023 to encompass genAI.[507] Complementing existing OECD standards in areas such as privacy, digital security risk management, and responsible business conduct, the Recommendation introduces core value-based principles to lead the trustworthy deployment, development, and use of AI; namely, (1) inclusive growth, sustainable development and well-being; (2) human-centred values and fairness; (3) transparency and explainability; (4) robustness, security and safety; and (5) accountability for the proper functioning of AI systems and respect of the principles. These principles have become an influential reference point in subsequent policy frameworks in the form of political declarations such as the G7 and G20 AI Principles[508] as well as the Bletchley Declaration[509] issued by states attending the AI Safety Summit in November 2023 in the UK.

The UNESCO Recommendation on the Ethics of Artificial Intelligence[510] as the first global normative framework addresses those features of AI systems that are of central ethical relevance. Governance aims of the Recommendation consider the usage of the output in terms of, *inter alia*, sustainability, gender equality and within the employment context in order to ensure a healthy development of AI. As already addressed in the foregoing chapters of this publication, AI raises profound and multiple ethical concerns,[511] the most pertinent issues being bias, fairness, threat to human rights, allocation of services and goods, and economic displacement, which thus risk reinforcing existing inequalities. UNESCO has sought to provide an "ethical compass" via said Recommendation in response to these ethical challenges. Again, this Recommendation is designed as a reference point for further concretisation in future legal obligations.

9.1.2. Moving towards binding law: Developments in the Council of Europe and the EU

In contrast to these soft law instruments of the OECD and UNESCO, going even further and quite notable in terms of the relatively swift agreement, ambition when it comes to territorial reach, and substantive content, the first "hard law" texts of relevance in Europe have recently emerged: a Framework Convention on AI by the Council of Europe and the AI Act by the European Union.

---

[507] According to the updated Recommendation an AI system is defined as "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment".

[508] G20, "G20 AI Principles'",Annex to the G20 Ministerial Statement on Trade and Digital Economy (June 2019); G7, "Hiroshima AI Process G7 Digital & Tech Ministers' Statement", (1 December 2023)

[509] The Bletchley Declaration by countries attending the AI Safety Summit, 1-2 November 2023, (1 November 2023) .

[510] UNESCO, "Recommendation on the Ethics of Artificial Intelligence'",(2022) SHS/BIO/PI/2021/1.

[511] See for an early discussion also Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter and Luciano Floridi, "The Ethics of Algorithms: Mapping the Debate", (2016) 3(2) Big Data Soc.

© European Audiovisual Observatory (Council of Europe) 2024

9.1.2.1. The Council of Europe 2024 AI Framework Convention

Taking into account the aforementioned soft law instruments and political declarations, on 17 May 2024, the Committee of Ministers of the Council of Europe formally adopted the first public international law treaty on AI, the Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.[512] While the Council of Europe brings together its 46 Member States, further states including the United States of America (hereinafter: U.S.) and Japan as well as representatives from civil society, academia and industry were actively involved in the drafting process. Consequently, when the Convention was signed on 5 September 2024, Parties that expressed their intention to be legally bound by the Convention in future by signing the Convention were not only Member States of the Council of Europe,[513] but also Israel, the EU and the U.S..[514] The Convention will enter into force shortly after the first five ratifications or comparable instruments of signatories, three of which need to be Council of Europe Member States, have been deposited. It can be expected that many more States will now add their signatures to the Convention. Moreover, the Council of Europe considers the area of artificial intelligence as a "cross-cutting priority"[515] and the Convention in its preamble speaks of "the framework character of this Convention, which may be supplemented by further instruments to address specific issues relating to the activities within the lifecycle of artificial intelligence systems", which means that possibly further sector-specific texts will follow also from the Council of Europe.

Putting its emphasis on the commitment to protect human rights, the Framework Convention follows a risk-based approach and seeks to complement existing human rights frameworks to ensure that state parties' existing protection levels of human rights, democracy and the rule of law also apply to current and future challenges raised by AI.[516] Accordingly, the Framework Convention formulates fundamental principles aimed at filling in any legal gaps that emerge as a result of technological advancement; the Convention requires the Parties to take stock of whether their existing human rights framework sufficiently and effectively responds to the AI risks and challenges and, if not, how they have to adapt to uphold the existing protection levels. By way of the legal instrument of a framework convention rather than a comprehensive regulation in the form of a convention creating rights and obligations, the instrument sets out broad principles, core values and areas for action, seeking to put existing standards on human rights, democracy and the rule of law in the context of AI. Broad principles allow an interpretation that can be adapted to the challenges of a changing world, whereas, in particular in a technology context, detailed legal provisions often may impede an

[512] Council of Europe, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, ETS No. 225.
[513] Andorra, Georgia, Iceland, Norway, Republic of Moldova, San Marino, and the United Kingdom.
[514] See here for list of signatures and ratifications (no ratifications so far as the Treaty has only just been opened for signature).
[515] Council of Europe, "The Council of Europe & artificial intelligence", (March 2023), p. 3.
[516] Cf. the Chair of the CAI's address following the adoption, "Word of the Chair".

© European Audiovisual Observatory (Council of Europe) 2024

application when conditions change.[517] Nonetheless, the Framework Convention does not limit itself to giving substantive elements to consider by the ratifying parties, but in addition also demands the establishment of efficient supervision and institutional structures for this.

The fundamental principles enshrined in the Framework Convention slightly depart from the OECD principles in that they also include human dignity and individual autonomy, privacy and data protection, equality and non-discrimination, reliability, and safe innovation. Most of these additional principles had already been included in the values and principles enshrined in the UNESCO Recommendation. While the reference to human dignity and fundamental rights is more in the focus of the Council of Europe work, the inclusion of allowing safe innovation at the same time may be more surprising in the context of a document by this organisation. The principles are complemented by the obligation upon states to adopt a risk and impact management framework and the obligations with regard to the implementation of the Convention. It must be noted that although the Convention requires the parties to lay down enforcement and supervisory mechanisms, much of its effectiveness will in the future rely on national implementation of the broadly formulated requirements. In order to enhance international cooperation and facilitate efforts to harmonise governance of AI at a global level, the Framework Convention replicates the definition of an artificial intelligence system from the latest revised definition adopted by the OECD. By not only being open for ratification by non-Council of Europe members but by already having integrated them, in view of the location of many of the innovation-driver companies in this field (most notably the U.S.) in the drafting process, the Framework Convention has the potential to be the "hard law global standard'" with a kind of "Strasbourg effect" resembling the "Brussels effect" of some legislation of the European Union concerning the digital sphere.[518]

While the Framework Convention covers the public sector and private sectors acting on behalf of public authorities, the treaty gives parties much leeway in applying the provisions to the private sector. This "pick and choose approach"[519] allows parties to choose not to apply the principles and the obligations of the Framework Convention directly in relation to private sector activities.[520] Accordingly, the relevance of the

---

[517] Cf. Council of Europe Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108), "Report on Artificial Intelligence'", (2019) T-PD(2018)09Rev.

[518] On the "Brussels effect" see Mark D. Cole and Christina Etteldorf, "The Implementation of the GDPR in Member States' Law and Issues of Coherence and Consistency", in: Inge Graef and Bart van der Sloot (eds.), *The Legal Consistency of Technology Regulation in Europe* (Hart Publishing, 2024), 131-156. Consider also for an earlier standard-setting instrument by the Council of Europe concerning the online environment the Cybercrime Convention of 2001, ETS No. 185, which has been signed and ratified by 76 States among which are the U.S., Japan, Australia, Brazil and Nigeria.

[519] See Christopher Lamont, "The Council of Europe's draft AI Treaty: Balancing National Security, Innovation and Human Rights?'",(18.03.2024) Global Governance Institute.

[520] In this case, however, already in the drafting, the expectation was expressed that the approaches of those parties would develop over time as their approaches to regulate the private sector evolve. See Explanatory Report, para. 30.

© European Audiovisual Observatory (Council of Europe) 2024

(horizontally applicable) Framework Convention in the audiovisual media sector very much depends on the national implementation of the treaty. While it is not obligatory for ratifying States to extend the application in this sense, it does provide for an obligation of the States to anyway take "appropriate measures to fulfil the obligation" of addressing possible "risks and impacts arising from activities within the lifecycle of artificial intelligence systems by private actors",[521] meaning that they cannot ignore the relevance of the sector when complying with the requirements stemming from the Framework Convention.

9.1.2.2. The European Union 2024 AI Act Regulation

In contrast to the principles-based Framework Convention of the Council of Europe, the EU AI Act[522] follows a market-driven approach establishing common rules to harmonise the EU internal market. It was passed in the form of a Regulation, making it a directly applicable and binding instrument for all 27 EU Member States. It entered into force on 1 August 2024 and – for the most part – will become applicable from 2 August 2026 after a transition period.[523] The much more detailed AI Act aims to foster the development of safe and trustworthy AI systems across the EU's single market by both private and public actors and to stimulate investment and innovation related to AI in the EU. At the same time, the Act seeks to ensure respect for fundamental rights of EU citizens. Similar to the Framework Convention, the AI Act follows a risk-based approach whereby different types of AI systems are categorised according to the risks that emanate from them, and consequently prohibits certain AI practices for which the risks are regarded as unacceptable. At the core of the AI Act are product safety requirements. In that regard, the Regulation lays down specific requirements for high-risk AI systems and obligations for operators of such systems. As it applies to both the private and the public sector the impact is broader. It differentiates in the strictness of the applicable norms corresponding with the risk level. For example, there are stricter obligations where high-risk AI systems are put into service by public authorities. Certain requirements are foreseen for high-risk AI systems including a risk management system, data governance rules, technical documentation, record-keeping, transparency and information obligations, human oversight, accuracy, robustness and cybersecurity.

[521] Art. 3(1) Framework Convention.

[522] Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L 2024/1689, 12.07.2024, http://data.europa.eu/eli/reg/2024/1689/oj.

[523] The rules concerning prohibited AI practices will, e.g., already become applicable on 2 August 2025.

© European Audiovisual Observatory (Council of Europe) 2024

With these requirements there is a clear relation – and in future after applicability of the AI Act also overlap – with the existing acquis of the EU for the digital sector.[524] Transparency requirements derive from *inter alia* the GDPR,[525] the P2B Regulation[526] or the DSA;[527] cybersecurity rules have been put in place amongst others with the GDPR, the NIS Directives,[528] and the forthcoming Cyber Resilience Act;[529] data law beyond the GDPR includes the Data Act,[530] and the Data Governance Act.[531] Where these legislative acts address algorithms and automatic decision-making processes, they are precedences also for the AI Act. This includes even the Audiovisual Media Services Directive which after its revision in 2018[532] and extended application to video-sharing platforms also referred to algorithms.[533]

Further progress in AI policies and regulation was certainly triggered by the new challenges posed by general purpose AI (GPAI) systems, some of which had only become evident after the drafting process for the above-mentioned texts had progressed. In fact, the AI Safety Summit and the resulting Bletchley Declaration are a direct response to the increased capabilities of GPAI; the same applies to the U.S. Presidential Executive

---

[524] For an overview cf. Christina Etteldorf, The European Union Regulatory Framework, in: Cappello M. (ed.), *Algorithmic transparency and accountability of digital services*, IRIS *Special*, European Audiovisual Observatory, Strasbourg, 2023, p. 14 et seq.

[525] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119/1, 4.5.2016.

[526] Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services, OJ L 186/57, 11.7.2019.

[527] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277/1, 27.10.2022.

[528] Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union (NIS Directive), OJ L 194/1, 19.7.2016, which will be replaced by Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive), OJ L 333/80, 27.12.2022.

[529] European Commission, "Proposal for a Regulation of the European Parliament and of the Council on horizontal cybersecurity requirements for products with digital elements and amending Regulation (EU) 2019/1020" COM/2022/454 final. For an overview see Mark D. Cole and Sandra Schmitz-Berndt, "Towards an Efficient and Coherent Regulatory Framework on Cybersecurity in the EU: The Proposals for a NIS 2.0 Directive and a Cyber Resilience Act"",(2022) 1(1) Applied Cybersecurity & Internet Governance.

[530] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), OJ L 2023/2854, 22.12.2023.

[531] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), OJ L 152/1, 3.6.2022.

[532] Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities, OJ L 303/69, 28.11.2018.

[533] Art. 28b AVMSD when it comes to possible measures that the Member States can expect VSP providers to take, cf. on this Mark D. Cole and Christina Etteldorf, Future Regulation of Cross-Border Audiovisual Content Dissemination, Schriftenreihe Medienforschung Bd. 84 (Nomos, 2024), 108 et seq., 125 et seq., 206 et seq.

© European Audiovisual Observatory (Council of Europe) 2024

Order[534]of October 2023 pushing for national AI standards through the executive branch to respond to increasing AI capabilities and their implications. However, both the Framework Convention and the AI Act address genAI. Important in a media context, the Framework Convention as well as the AI Act address the necessity to avoid the risk of manipulation and deception when requiring transparency with regard to the identification of content generated by AI systems.[535]

9.2. Reality bites!? Applicability and limitations of regulatory approaches to the specifics of the AV sector

In view of the variety of initiatives introduced above, questions emerge as to whether these frameworks are capable of responding to the specifics of the AV context. Chapters 2 to 8 have already highlighted a number of legal challenges for the use of AI in the AV sector while making clear that no legally directly binding sector-specific rules have been put in place so far, which makes all the more pertinent the question of the extent to which existing legislation indirectly impacts also AI systems or whether the newly created frameworks have been designed in a way allowing them to respond to the specific challenges and risks.[536]

9.2.1. Data protection aspects

Being data-driven, the training of an AI system is unavoidably dependent on large datasets which can lead to data protection and privacy law infringements. As outlined in Chapter 2 in detail, AV training materials most often contain personal data, necessitating a legal basis under EU law for activities like scraping and copying content.[537]

Compliance with data protection requirements reaches further and also concerns data transfers from the EU to third states, for which limitations are in place that ensure that personal data protection is not undermined by the transfer of such data to third states that do not have an adequate level of protection compared to the one guaranteed in the EU or for which no other safeguards are in place. As datasets used to train, validate and test AI systems, the interaction of individuals with AI systems and the content generated by an AI system have a cross-border dimension and namely take place between persons in the EU and processing companies located e.g. in the U.S. or Asian countries,

[534] Executive Order (E.O.) 14110 on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 30.10.2023.
[535] Art. 6 Framework Convention; see also Explanatory Memorandum, paras. 59 and 104.
[536] Cf. also the contributions in Cappello M. (ed.), *Artificial intelligence in the audiovisual sector*, IRIS *Special*, European Audiovisual Observatory, Strasbourg, 2020.
[537] See Chapter 2.3.1. For the EU, see Art. 6 GDPR.

© European Audiovisual Observatory (Council of Europe) 2024

the potential limitations by rules prohibiting data transfers without safeguards are impactful. Global initiatives agree that further collaboration is needed on these issues in particular in the context of genAI,[538] not least with the goal of ensuring comparative levels of protection that could facilitate data flows. In that sense, for instance, the G7 DPA Roundtable[539] is currently exploring how best to protect privacy and urges AI developers to ensure that personal information generated by genAI is kept accurate, complete and up-to-date[540] - another obligation under EU data protection law. Recital 9 EU AI Act clarifies that the harmonised rules laid down in the Act complement existing EU law in particular on fundamental rights and data protection. As a fundamental rights-based framework, the Framework Convention, too, underlines the importance of personal data protection in safeguarding privacy rights in the digital word,[541] as does the OECD Recommendation. The Framework Convention complements existing human rights frameworks; as regards data protection, a binding legal instrument exists within the Council of Europe with the modernised Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data.[542] The modernisation of the Convention in 2019 ensured synchronisation with the 2016 EU reform and passing of the GDPR and also revised the regime for transborder flows of personal data. The latter seeks to ensure that personal data are protected with appropriate safeguards even if they fall within the jurisdiction of a non-Party.[543] Guaranteeing such protection is obviously challenging considering that the U.S., where many tech companies are based, is a non-Party and has a very different approach to limitations on processing of personal data. This highlights the importance of the work of the G7 in that field and the data transfer agreement between the U.S. and the EU and for obvious reasons is very relevant for the content industry for which until today large amounts of the AV production come from the U.S.

9.2.2. Intellectual property rights aspects

As outlined in Chapter 3 in more detail, training data often includes copyright-protected material. Since genAI models require vast amounts of data, text and data mining (TDM) techniques are used extensively for retrieval and analysis of the data. The use of copyrighted content typically requires authorisation from rightsholders. To prevent unauthorised use, press and media entities are implementing anti-scraping tools, and

[538] See sect. 2.6 and Roundtable of G7 Data Protection and Privacy Authorities, "Statement on Generative AI", (2023).
[539] The G7 DPA Roundtable brings together representatives from the data protection supervisory authorities of the G7 states to intensify regulatory cooperation and discuss questions of secure and trustworthy processing of personal data across borders in an increasingly digitised and global society.
[540] See ibid.
[541] Cf. Art. 11 Framework Convention.
[542] Council of Europe, "Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data'" ETS 108+ (ETS 223 which after entry into force will modernise ETS 108, hence the common reference to it as "108+".
[543] Art. 14 Convention 108+.

© European Audiovisual Observatory (Council of Europe) 2024

copyright holders are filing lawsuits against tech companies for infringement, arguing that AI outputs rely heavily on training datasets which may have included their copyrighted material. This indicates that the regulatory approach so far does not extend specifically to the IP-related issues and there is a perception of lack of protection on the side of rightsholders.[544]

The question of whether the unauthorised use of copyrighted material to train AI models is legitimate relates also to the AI Act. According to the AI Act, providers of genAI models are obliged to put in place a policy to comply with EU law on copyright and related rights and a detailed summary of the content that is used to train the genAI model.[545] Currently, providers are in a situation in which compliance with copyright law is a challenge unless they have acquired a license and thereby authorisation of use by the rightsholder. It will be challenging to determine what has to be considered a sufficiently detailed summary of the content used for training the genAI model until the newly-established EU AI Office[546] has provided a template for this type of summary. According to Recital 107 of the AI Act, this summary shall be generally comprehensive in its scope to facilitate the exercising and enforcement, by parties with legitimate interests, including copyright holders, of their rights under EU law. Besides the copyright aspects, the disclosure obligation is one of the means to satisfy the principle of transparency enshrined in all of the AI-related instruments mentioned above.

AI-generated content also raises questions about copyright protection for the output. As outlined in chapter 4.1.1., in most jurisdictions including the U.S. and the EU, only a human creation can enjoy copyright protection. Where computer-generated content is protected, this raises the question of whether it is merely authorship that is shifted to the "human behind the machine" or whether the so-far demanding requirements for protection as a "work" are reduced.[547] A work usually necessitates an element of creativity. This leads to the question of whether the output of AI can be classified as a new creative work. It has to be taken into consideration that the output produced is not merely based on statistical patterns in the training dataset; there is also significant human influence in the process during creation and configuration of the AI system, which includes the selection of training data, but also in the way a genAI system is prompted and potentially in the subsequent editing of the system's output.[548] As can be concluded from the analysis of case law in chapter 4.1.2. the degree of human creative contribution that must also be reflected in the AI output varies in national law and has to be assessed on a case-by-case

[544] As regards the perceived lack of protection, see European Guild for Artificial Intelligence Regulation, "Our Manifesto for AI Companies Regulation in Europe'" (4.11.2023).
[545] Art. 53(1)(c) and (d) AI Act.
[546] The EU AI Office was established in May 2024, European Commission, Press Release, "Commission Establishes AI Office to Strengthen EU Leadership in Safe and Trustworthy Artificial Intelligence", (29 May 2024), https://ec.europa.eu/commission/presscorner/detail/en/IP_24_2982. See also https://digital-strategy.ec.europa.eu/en/policies/ai-office.
[547] Cf. Chapter 4.1.1.
[548] These stages have been compared by Baumann and Nordemann with regard to the different steps taken in portrait photography, cf. Chapter 4.1.2.

© European Audiovisual Observatory (Council of Europe) 2024

basis. If the threshold is reached, authorship is then attributed to the user of the AI system.[549]

IP protection is not restricted to copyright. There are neighbouring rights which do not require creativity but protect investment or economic or organisational efforts in a work; an example in the AV sector would be the neighbouring right of the film producers. Since these rights do not depend on a creative element, these neighbouring rights can be much more easily claimed for AI content below the threshold for protection as a (copyrighted) work.[550]

Obviously, questions of copyright and neighbouring rights are posing new challenges in the AV sector, when for instance the supplier of special effects to a film can no longer per se be presumed to have all rights to the material supplied. This makes it even more important to have full transparency on the usage of AI. In that regard, the Framework Convention and under certain conditions the AI Act require transparency – however, it remains to be seen whether this also means that a content supplier has to disclose the use of AI to other parts of the supply chain in AV production. One risk that arises, in particular with regard to AI-generated AV content, is that elements of an earlier work contained in the training data may be recognisable in the output. The determination and consequences of recognisability of an earlier work constitute an original question of copyright law and, again, are not AI-specific. Problems arise where the user of AI output does not know that they are using third-party works, and thus, they may not be, at least under EU law, liable for damages. As regards the AI provider, taking again the example of EU law, neither the Product Liability Directive[551] nor the proposed AI Liability Directive,[552] which is to complement the AI Act, cover the attribution of liability in the event of copyright infringements of AI output.[553] This legal uncertainty will certainly lead to legal disputes in the future.

9.2.3. Personality rights aspects

Besides the risk of copyright infringements, AI systems and, once more, in particular genAI models facilitate infringements of personality rights when content impersonates individuals or their attributes such as through the generation of deep fakes including also voice-only deep fakes. Personality rights allow an individual to protect and control the

[549] See Chapter 4.1.3.
[550] See ibid.
[551] Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 210/29, 7.8.1985.
[552] European Commission, "Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)", COM/2022/496 final.
[553] See Chapter 4.2.3.

© European Audiovisual Observatory (Council of Europe) 2024

use of their likeness or other personal attributes such as their voice.[554] This allows the individual to also exploit the use of their identity economically. Chapter 5 lists a variety of examples for unauthorised uses of imagery and voice ranging from AI generated non-consensual intimate imagery to the AI assisted enhancement of a singer's voice and voice cloning. It follows from the aspects discussed on copyright, that transparent information about a system's training data, functionalities, and output is also crucial for the enforcement of personality rights in order to identify infringements and to hold an infringer accountable. Transparency is one of the core principles enshrined in policy documents concerning AI from an early stage onwards and that is also included in the AI Act and the Framework Convention.[555]

The transparency obligation within the AI Act is however limited to direct interactions with an AI system unless the interaction with it "is obvious from the point of view of the natural person".[556] Providers of genAI models, where it may be less clear for an individual that they are exposed to AI-generated content, are under a more comprehensive obligation and need to ensure that genAI-generated output is marked in a machine-readable format and detectable as artificially generated or manipulated.[557] As mentioned before, this raises questions as to whether along the supply chain of an AV production, the use of AI has to be disclosed in order to allow the deployer to comply with their transparency obligation set out in Article 50 AI Act. Where an AI tool is used to create deep fakes, disclosure is always mandatory.[558] Deep fakes forming part of an "evidently artistic, creative, satirical, fictional or analogous work or programme" only have to disclose the AI involvement "in an appropriate manner that does not hamper the display or enjoyment of the work'" Overall, the transparency requirements and issues pertaining to the use of AI in the production phase are one of the core challenges for the AV sector.

There is no guidance as to how a deployer can satisfy these transparency obligations in relation to AV works, i.e. where, how and when the information needs to be placed and its exact wording. To date, it is unclear whether it will for instance be sufficient to have a disclaimer in the movie credits. The Explanatory Memorandum to the Framework Convention suggests as a means for disclosure labelling or watermarking to allow the identification of content as AI-generated. However, the focus of transparency in the Framework Convention lies more on safeguarding public trust, consumer protection and prevention of electoral interference, whereas protection against unauthorised use of personality rights is provided under the interplay of the principle of human dignity, individual autonomy and privacy.[559] In contrast to the emerging legislation in the U.S. that

[554] See Chapter 5.1.
[555] See generally on Transparency Cappello M. (ed.), *Transparency of media ownership*, IRIS *Special*, European Audiovisual Observatory, Strasbourg, 2021 and Cappello M. (ed.), *Algorithmic transparency and accountability of digital services*, IRIS *Special*, European Audiovisual Observatory, Strasbourg, 2023.
[556] Art. 50(1) AI Act.
[557] Art. 50(2) AI Act.
[558] Art. 50(4) AI Act.
[559] Cf. Chapter 5.3.2.

© European Audiovisual Observatory (Council of Europe) 2024

seeks to protect individuals from unauthorised AI cloning and deep fakes,[560] neither the AI Act nor the Framework Convention addresses this issue in particular due to their complementary nature to existing human rights frameworks. First agreements between a film and TV actors' union which also address the personality rights dimension were found in the SAG AFTRA agreement outlined in Chapter 6.

9.2.4. Disinformation as an important challenge

In relation to deep fakes it is also necessary to point out the effects these may have on the recipients. First of all, where audiences cannot distinguish between real and fake content, the credibility of the AV provider can be undermined. At the same time, high-quality deep fakes can have a meaningful impact on public discourse and influence public opinion.[561] The evidential value of pictures and video material gets lost, when real content can no longer be distinguished from synthetic content.[562] Furthermore, AI-empowered bots are functional in disseminating disinformation. It is also noteworthy that AI systems do not necessarily produce content that conveys truth but can also generate false information. Where such information enters into a journalistic product without verification, this can have the same negative consequences mentioned before: undermining trust in and credibility of the medium as well as influencing public discourse and opinion. As mentioned above, under the AI Act, deployers of an AI system that produces deep fakes always have to be transparent, but guidance is lacking as to how this transparency can be achieved in the AV sector. In contrast, providers of an AI system intended to interact directly with natural persons only have to disclose the involvement of an AI system when this is not obvious. Similarly, deployers of AI systems that generate or manipulate text which is published with the purpose of informing the public on matters of public interest shall disclose that the text has been artificially generated or manipulated unless the content has undergone a process of human review or editorial control and where a natural person holds editorial responsibility for the publication of the content.[563] Accordingly, separate rules apply for text and AV media with AV media providers always under an obligation to disclose the use of an AI system.

Addressing disinformation, one must consider that malicious deployers may intentionally disseminate manipulated materials. While obligations to remove illegal materials exist for hosting services under the DSA[564] and the eCommerce Directive[565]

[560] For an overview of the emerging legislation and bills in the U.S. see Chapter 5.4.
[561] See Chapter 7.2. Cf. further Angelica Fernandez, "Deep fakes": disentangling terms in the proposed EU Artificial Intelligence Act, 85 (2) UFITA 2021, p. 392-433
[562] Ruth Meyer and Rupprecht Podszun, „Künstliche Intelligenz und die Medienpolitik", (2024) ZRP 41.
[563] Art. 50(4) AI Act, Recital 134 AI Act.
[564] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (2022) OJ L 277/1.

© European Audiovisual Observatory (Council of Europe) 2024

respectively, the voluntary Code of Practice on Disinformation (2022)[566] considers AI-assisted manipulative behaviour (e.g. impersonation and malicious deep fakes) as impermissible manipulative behaviour. Considering the risks associated with AI-generated disinformation campaigns and their algorithmic amplification, fact-checking becomes essential as well as reinforcing trust in the media as a public watchdog. This will be a difficult task considering that AI-driven curation and recommender systems succeed in personalising content with the effect that individual users are more or less locked in a bubble where existing opinions are reinforced and conflicting ones are removed.[567]

With the steep rise in deployment of AI systems, a transformation of work takes place that is not exclusive to the media sector. Accordingly, initiatives addressing labour issues are commonly of a rather general nature. Exemplary for the AV sector may be that workers are particularly likely to lose control over their work, and challenges arise in relation to the authorship and ownership of creations. As has been outlined in previous chapters, AI technologies have been used to create "younger versions" of actors or to relive the voices of deceased singers.[568] Furthermore, genAI models are capable of writing scripts and generating video clips. While concerns with regard to these issues have culminated in strike actions in the U.S. and led to industry agreements with trade unions, the European Parliament is seeking to improve the working conditions for workers in the AV sector in general and has adopted a Resolution for an EU Framework for the social and professional situation of artists and workers in the cultural and creative sectors.[569] In the meantime, however, we may also see collective management organisations or European or international associations and federations as well as trade unions getting active.[570]

9.2.5. The reach of the AI Act: geographical scope

As has become evident, the AI Act is the main reference point when it comes do detailed AI systems-related rules, especially within the EU. Therefore, it has been prominently placed throughout the previous chapters of this publication and in this concluding overview. Nonetheless, besides illustrating the novelties coming with the newly introduced rules, it is also necessary to briefly mention the broad scope and thereby territorial reach of the AI Act. Like other product safety regulations, the AI Act has effects

---

[565] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services in particular electronic commerce, in the Internal Market (2000) OJ L 178/1.
[566] Strengthened Code of Practice on Disinformation (16 June 2022).
[567] See Chapter 8.2. See further Christina Etteldorf, Standard-Setting of the Council of Europe, in: Cappello M. (ed.), *Algorithmic transparency and accountability of digital services*, IRIS *Special*, European Audiovisual Observatory, Strasbourg, 2023, p. 4, 5 et seq.
[568] See Chapters 5.1., 6.2.3.3. and 10.4.
[569] European Parliament, "EU Framework for the social and professional situation of artists and workers in the cultural and creative sectors", (2023/2051(INL)).
[570] See Chapter 6.4.

© European Audiovisual Observatory (Council of Europe) 2024

beyond the geographical boundaries of the EU. It applies to "providers placing on the market or putting into service AI systems or placing on the market general-purpose AI models in the Union, irrespective of whether those providers are established or located within the Union or in a third country".[571] Further, it applies *inter alia* to deployers that have their place of establishment or are located in the EU;[572] where the output produced by an AI system is used in the Union but the provider and deployer are located in a third country;[573] and importers and distributors of AI systems.[574]

This shows the ambition to provide a secure and trustworthy use of AI systems for Union citizens and companies alike, irrespective of the actual location of the systems and their providers. The EU as a marketplace is the connecting factor, similarly to the case with the GDPR, for applicability of EU rules.

9.3. Looking ahead: On "future-proofness" and global standards

9.3.1. Towards global and flexible risk-based standards in specific legislation

At EU level, the AI Act as a product safety regulation seeks to reduce the risks for humans when using the product "AI systems", but it is uncertain to what extent it provides for legal certainty in this new market. Some requirements are drafted in a very general form leaving room for interpretation if not even the need for further detailing. Accordingly, several rather generic provisions require clarification through implementing acts by the European Commission. For instance, the Commission may adopt implementing acts establishing common specifications for the requirements for high-risk AI systems.[575] This regulatory methodology was chosen, as less abstraction and more detail in the AI Act itself may have come at the risk that the rules may become outdated soon. Therefore, "merely" outlining requirements at the normatively highest level of EU secondary law is characteristic for modern technology regulation, at least where technology evolves fast. Guidance on the interpretation of certain obligations can be filled with technical standards clarifying expectations and for instance in the case of the AI Act guiding providers through the risk assessment process. While providers benefit from a presumption of conformity when adhering to such standards, the downfall of standards is

[571] Art. 2(1)(a) AI Act.
[572] Art. 2(1)(b) AI Act.
[573] Art. 2(1)(c) AI Act.
[574] Art. 2(1)(d) AI Act.
[575] Art. 41(1) AI Act.

© European Audiovisual Observatory (Council of Europe) 2024

that they are de facto norm-setting by private actors and there is a problematic risk that the development of standards does not lead to increased transparency.[576]

While much remains unclear in terms of concrete implementation of the flexible, yet binding legislation, it is worth taking a step back and appreciating what has already been achieved, considering that the urge to regulate AI only became pressing in recent years due to the large-scale rollout of products and services to general users. At macro level, we can see that principles and values have evolved on which a consensus has been reached as they are frequently taken as the main (and minimum) expected basis for dealing with AI systems. These are now also enshrined in the first legally binding international treaty on AI that has been opened for ratification. Furthermore, legal definitions of AI systems are aligning increasingly.[577]

The OECD has been engaging in empirical and policy activities on AI to support emerging policy debates for almost a decade with its activities culminating with the first intergovernmental standard on AI policy. Soft law instruments like said OECD Recommendation and the UNESCO Recommendation have reflected a minimal consensus in the form of common principles and values which at the same time can positively be acknowledged as already being ambitious for the establishment of common ground. This consensus is now also enshrined in legally binding instruments. Namely, the possibility for a more global approach has been established with the Council of Europe Framework Convention. Not only is the Framework Convention open for signature for non-Members, it is also the product of drafting by a group of international stakeholders and experts and reflects their consensual work. Moreover, the Convention introduces a common terminology which can serve as a basis for further regulation, also where necessary sector-specific regulation. Limitations, for example when it comes to applicability to the private sector, result from the instrument being a compromise text between states that may not all have the same idea of the level of obligations that should be imposed on the private sector. Given these unsurprising divergences between states' approaches, it is a significant step forward that such agreement was reached, opening the door for a more global minimum level of binding rules.

Besides the AI Act of the EU which has already been passed, several national jurisdictions around the globe are working on domestic AI regulation, partly in response to commitments following the above-mentioned recommendations. The examples provided in the previous chapters indicate however that most of the regulation is restricted to a certain (AI system) application, for instance the Ensuring Likeness Voice and Image Security (ELVIS) Act of Tennessee concerning unwanted voice cloning, or the currently discussed bills at federal level in the U.S. such as the Nurture Originals, Foster

[576] Cf. Michael Veale and Frederik Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act—Analysing the good, the bad, and the unclear elements of the proposed approach", (2021) 22 Computer Law Review International 97. At EU level, CEN and CENELEC are working on standards.

[577] For definitions of AI systems see chapter 1.1. See also Sandra Schmitz-Berndt, AI Regulation and Governance on a Global Scale: Overview of Scope, Definitions and Key Elements, Annex to Mark D. Cole, "AI Regulation and Governance on a Global Scale: An Overview of International, Regional and National Instruments", (2024) 1(1) AIRe 126, p. 141-142.

© European Audiovisual Observatory (Council of Europe) 2024

Art, and Keep Entertainment Safe (NO FAKES) Act attempting to deal with giving individuals a property right to their "digital replicas" or the No Artificial Intelligence Fake Replicas and Unauthorized Duplications (No AI FRAUD) Act that goes even further in the proposed way of protecting a person's own "likeness".[578]

9.3.2. Considering sector-specific aspects for (AV) media

What makes the AI Act stand out is that it is a comprehensive, product-focused framework. With its general horizontal approach, it does not address the particularities of certain sectors, but attempts at finding all-encompassing basic rules. The need for differentiating among those for certain risks and those for providing more technical guidance led to 13 annexes being joined to the AI Act.

In general, the global and regional initiatives avoid addressing specificities of certain sectors.[579] Again this has to do with the fact that general principles are outlined that can and should be applied universally. Accordingly, questions pertinent to the AV sector, in particular those relating to risks to media pluralism and new forms of market concentration due to scale effects, have to be addressed in more specific legislation. While media pluralism is addressed in the recently adopted European Media Freedom Act (EMFA) of the EU,[580] market concentration, which is also part of the EMFA, is now on the agenda of the European Commission in view of AI systems. The Commission is currently examining how tech giants might limit the development of competing models of generative AI or favour the integration of their AI applications into other products and ecosystems.[581]

Even though the regulatory approaches so far do not specifically address the media or AV sector, the relevance of AI systems especially for this sector is apparent. Therefore, media providers have started to publicly declare how they position themselves in view of using AI in their content production and dissemination. These commitments partly respond to emerging legal obligations, but already different types of media providers, either individually or via associations of journalists or publishers, have committed themselves in a self-regulatory sense on how they apply – or do not – AI systems in their processes. Recent examples are the German Press Council, which has already amended the "Pressekodex" (the self-regulatory deontological code),[582] the

[578] Cf. on these Chapter 5.4.
[579] Cf. Jo Pierson, Aphra Kerr et al., "Governing artificial intelligence in the media and communications sector", (2023) 12(1) Internet Policy Review.
[580] Regulation (EU) 2024/1083 of the European Parliament and of the Council of 11 April 2024 establishing a common framework for media services in the internal market and amending Directive 2010/13/EU (European Media Freedom Act) (2024) OJ L 1083/1.
[581] European Commission, "Competition in Virtual Worlds and Generative AI – Call for Contributions", (9 January 2024).
[582] The preamble now states that anyone who commits to adhere to the Pressekodex bears the ethical responsibility for all editorial contributions, regardless of how they are created, and that this responsibility

© European Audiovisual Observatory (Council of Europe) 2024

German public broadcaster Bayerischer Rundfunk (BR), which has already updated its 2020 AI ethics guidelines in view of GPAI,[583] Sveriges Radio in Sweden, which has introduced a policy for GPAI,[584] the Swiss broadcaster SRG SSR, which has implemented "National AI Principles",[585] or the British public service broadcaster BBC, which has set out principles that shape the BBC's approach to working with GPAI.[586]

In order to better deal with the challenges that AI systems pose in the media – and more concretely the AV sector – the route to more transparency as demanded by the rules put in place so far is the first important step. For example, labelling obligations empower readers, listeners and viewers to at least be aware of the new dimension of content production and/or content exposure. Considering the potential (negative) impact on the functioning of public-opinion-making procedures which democracies are built on, the next step will have to be to further evaluate whether the "integrity of democratic processes and respect for the rule of law" will need to be given more attention. Art. 5 of the Framework Convention emphasises the need to "adopt or maintain measures that seek to ensure that artificial intelligence systems are not used to undermine the integrity, independence and effectiveness of democratic institutions and processes". Said article further highlights the importance of protecting democratic processes throughout the AI lifecycle, ensuring *inter alia* fair access to public debate, and allowing individuals to freely form opinions. This underscores the significance of considering "the effects of AI systems and their use on media and their role for a pluralistic debate".[587]

extends to artificially generated content. See also Presserat, Press Release „Redaktionen auch für KI-generierte Inhalte ethisch verantwortlich", 18.9.2024.

[583] Katharina Brunner, Rebecca, Ciesielski, Philipp Gawlik et al, „Unsere KI-Richtlinien im Bayerischen Rundfunk", (BR, 12.7.2024).

[584] Olle Zachrison, Press Release, "Därfor publicerar Sveriges Radio en policy för generativ AI", 7.7.2023.

[585] SRF SSR, « SRG's National AI Principles »,(2023).

[586] BBC, Press Release, "Generative AI at the BBC'",5.10.2023.

[587] Cf. also Framework Convention, Explanatory Memorandum para. 46.

© European Audiovisual Observatory (Council of Europe) 2024

10. Ethical Dilemmas and Societal Challenges Raised by Generative AI

Bart van der Sloot, Tilburg Institute for Law, Technology, and Society (TILT) Tilburg University, Netherlands[588]

10.1. Introduction

The previous chapters in this volume have provided the reader with insight into AI, its inner workings and data technologies that power it. In particular, they described the use of AI in the AV sector and the questions that this raises in terms of privacy and data protection, copyright and intellectual property, liability and accountability, personality rights, labour law, disinformation and diversity. In addition, Chapter 9 offered a peak into the world of tomorrow and provided examples of the many consequences for the AV sector. This chapter will close off and delve into the ethical fundaments that form the essential pillars not only of legal concepts, but of our society as a whole. It will discuss how these ethical fundaments may be challenged by the unfolding technological reality, in particular the advances in genAI.

Section 2 of this chapter will briefly touch upon three essential pillars of open societies: trust, truth and friction. Section 3 will show how AI affects those values, in particular pointing to the loss of authenticity, the effects the spread of AI-manipulated and -generated content has on trust and the capacity of AI to relieve people of many arduous aspects of life. Section 4 will turn to the ethical question for the AV sector and discuss the challenges raised by genAI for journalism and news media, questions over the use of AI in the movie industry and finally move to evaluate the contrast between AI vs human creativity. Section 5 will map several societal challenges and in particular point to the legitimacy of the democratic process, the effects of genAI on legal procedures and on questions over granting legal persona to AI-driven entities. Finally, section 6 offers a small recap.

[588] Bart van der Sloot is associate professor at the Tilburg Institute for Law, Technology and Society, Tilburg University. He has written extensively about genAI in an open access book: Van der Sloot, B. (2024). Regulating the Synthetic Society: Generative AI, Legal Questions, and Societal Challenges, Bloomsbury.

© European Audiovisual Observatory (Council of Europe) 2024

10.2. Ethical foundations

This section will explain the relevance of three ethical foundations to our contemporary society: truth, trust and friction.

First, it is difficult to overestimate how essential truth is for personal, social and societal interest. "Know thyself" is perhaps the archetypical epigraph of Western civilisation, in which life is understood at least in part as a continuous quest to understand oneself. Being wrong about oneself or acting in a way that feels untrue to yourself can be one of the most eerie experiences of life. In social relationships, having a shared conception of truth is necessary to form a bond and many of our societal institutions depend on a baseline agreement on a basic set of facts, such as, not in last place, democracy itself. It should consequently come as no surprise that in societies where there is considerable disagreement on basic facts, such as the United States, the democratic process is cracking and social relationships across groups are formed less and less frequently.[589]

Second, truth is deeply connected to the notion of trust. Trust is foundational for people's perception of reality. People trust that the sun will rise the next morning, that their friend will still be their friend tomorrow and that they will still be able to walk next day. Although changes do occur, they usually do so gradually. This epistemological stability provides people with a basis for their perception of the world, others and themselves. For example, if a person is unsure whether they are going to live or die the next day, this has huge repercussions on their ability to engage with others and find meaning.[590]

Third, both trust and truth are consequentially essential to form social relationships, reflect on and develop oneself as well as for societal institutions.[591] But although they are essential, life confronts us with many instances where we find we were wrong about ourselves, others or the world around us, where no agreement can be found on a shared conception of truth, where we are betrayed and violated in our trust or where our probabilistic prediction of the (near) future proved to be wrong. Although these moments are stressful and sometimes even traumatic, and we consequently have a deep urge to avoid, prevent and remedy those situations, they are essential for our personal development. Although we would rather not be confronted with world views that radically conflict with our own, we would rather not face the reality of a loved one being terminally ill, and we would rather avoid a friend who is telling us the hard truth about ourselves, it

---

[589] Baumer, D. C., & Gold, H. J. (2015). *Parties, polarization and Democracy in the United States*. Routledge. Pausch, M. (2021). The future of polarisation in Europe: relative cosmopolitanism and democracy. *European Journal of Futures Research*, *9*(1), 12. Horonziak, S. (2022). Dysfunctional democracy and political polarisation: The case of Poland. *Zeitschrift für Vergleichende Politikwissenschaft*, *16*(2), 265-289.

[590] Keymolen, E (2016). *Trust on the line: a philosophical exploration of trust in the networked era*. Wolf Legal Publishers.

[591] Hegel, G. W. F. (1967). Phänomenologie des Geistes. Suhrkamp, Frankfurt 1969. Especially in the interpretation of Hegel by Kojève: Kojève, A. (1980). *Introduction to the Reading of Hegel*. Cornell University Press.

© European Audiovisual Observatory (Council of Europe) 2024

is through friction, through uneasy moments, that we learn about ourselves, our loved ones and the world around us. Without these moments, we would become static and one-dimensional beings.

10.3. How AI-generated content affects these concepts

This section will point to some of the potential effects of Generative AI on truth, trust and friction.

The first challenge is that with the advance of chatbots, humanoid robots, augmented reality, virtual reality, deep fakes and other applications powered by genAI, an increasing amount of online material is produced or manipulated by AI. GenAI is not only used to deliberately mislead people, but for generating beautiful atmospheric images of artificial forest cabins on Facebook or photos of non-existent children who appear to be doing something spectacular, which cannot be distinguished from authentic material. Photo cameras increasingly operate on the basis of blueprints: a burning forest might consequently still look green in the photo because the AI "knows" that a forest is green and the moon may appear crystal clear, even on a foggy evening.[592] With video calling, as a standard, high audio registers are filtered out and the skin tones of people are softened. Because deep fake technology is freely accessible to anyone, citizens are increasingly generating deep fakes for homely and satirical purposes. All these trends (and many more) taken together may mean that in a few years' time, more than 90% of all digital material is AI-generated or AI-manipulated.[593] This would result in a push towards a post-truth society.

Second, a more structural impact of the increased production of AI-generated content is that people will be unsure what they can believe. People who have mistakenly believed in a falsehood before are known to be more cautious when seeing shocking or sensational news. Insecurity about the veracity of communication in a world that is almost fully digitised and mediated can be significant, both on a personal and a societal level. The already declining trust in "mainstream media" may accelerate, and the trend according to which people choose the media that reaffirms their pre-established worldview may deepen. This may mean that groups become increasingly transfixed by their own perceptions of reality, which may lead to polarisation and societal discontent. A synthetic society may consequently cause epistemological insecurity: is what I am seeing actually true? Am I actually speaking to the person I think I am? Is this a human or a humanoid robot? Is the avatar resembling a friend really controlled by that friend? Is the avatar a "truer version" of the natural person controlling it or is it a deceptive version of them, or are the physical and the virtual avatar both equally important representations of one person? Is the avatar human or AI-generated and/or controlled? Combined with a

[592] https://www.theverge.com/2023/3/13/23637401/samsung-fake-moon-photos-ai-galaxy-s21-s23-ultra.
[593] Schick, N. (2020). Deepfakes and the infocalypse: what you urgently need to know. Hachette UK.

© European Audiovisual Observatory (Council of Europe) 2024

form of hyperpersonalisation, where everyone lives in their own reality and the "sequestration of experience",[594] it may be difficult to verify whether what one has experienced is true: "Hyperpersonalisation can thus lead to a loss of confidence in (objective) perception. What value can we still assign to a story or an eyewitness account if we cannot judge whether what the person has actually experienced is "real'?"[595]

Third, when work in the household, in factories and even in creative sectors is taken over by AI, humans may become increasingly dependent on AI-driven entities for physical, intellectual and creative activities. Although maybe not to all, there will be an intuitive appeal to many to take the easy road. If a robot can lift a shopping bag, why would a person do it; if a virtual avatar can take a person on a tour in a foreign country and Google lens can translate texts, why try and learn the basics of that language on vacation; if a humanoid love robot is available, why go out dating and go to lengths to buy smart clothes and nice perfume; if a care robot can help a parent to go to the toilet, why would their child bother; when ChatGPT can write an essay, why would a student do so themselves? Additionally, through deep fake technology, it is possible to create news bulletins that fit people's established world views; it is also possible to bring back to life one's deceased partner, so that the surviving partner does not have to be confronted with the loss and grief; it is possible for teenagers who are overweight and bullied at school to flee into a virtual reality with an avatar that mirrors their body-ideal, so they do not have to be confronted with the painful truth in physical reality. Thus, in many ways, the effect of genAI is that friction is removed from the lives of people.

10.4. Ethical dilemmas within the audiovisual sector

This section will home in on three ethical dilemmas for the AV sector specifically, namely the impact of genAI on journalism, its use by the movie industry and the more general point of human vs. AI-driven creativity.

First, the rise of synthetic media will have an impact on the functioning of the press. It will be used for positive use cases. For example, synthetic technologies can be utilised to visualise situations where no camera footage can be shot, such as in warzones. Also, it is possible to put out AI-newsreaders that can speak any language of the world. Or, media can enhance user experience by allowing for 3D live participatory reporting in virtual reality. But it might also have negative effects. The media is already struggling to properly check all user-generated content and online material for accuracy and authenticity. A small but telling example is a football player who was carried off the pitch during a European championship match because of a medical condition, shortly after which a picture appeared on Twitter (now known as X) that would prove that he was still

[594] Giddens, A (1991). *Modernity and self-identity: self and society in the late modern age*. Stanford University Press.
[595] Schermer, BW, & Ham, JV (2021). Regulering van immersieve technologieën.

© European Audiovisual Observatory (Council of Europe) 2024

alive when he was taken to the hospital. However, it took quite a while for many traditional media to mention the photo, as it might have been a fabrication or a postdated image. In a world where every citizen has access to synthetic technologies and can create and distribute fake videos, photos or audio files within minutes, the question is how media can ensure in practice that their reporting remains accurate. Quality media that invests in such procedures not only runs the risk of making less profit because of the cost involved, but also of becoming "obsolete", because other media, with less due diligence, would be quicker with coverage and post sensational stories, even those that would later turn out to be false. In addition, the use of genAI in journalism raises questions such as: Generative AI tools are trained on data containing public service media (PSM) content. If PSM agree to it, it ensures citizens access reliable content.[596] However, some PSM may not want to deal with AI companies. Should this be seen as part of the PSM remit/mission? It could change the view of what journalism is. The debate may grow bigger and more philosophical: How can we preserve the human touch in journalism? How do we preserve media pluralism? Reality should remain mediated by humans and less by AI. What happens if culture overall is not created and mediated by humans?

Second, genAI is already used in the movie-industry, and will most likely be used even more as the technology advances. AI is used, among things, to make actors artificially younger or older, to bring back deceased actors to finish a movie or to have AI-models of an actor perform dangerous stunts, while in the adult industry, actresses can have their AI double perform the most perverse scenes or remarkable requests by their users. This raises the question: if cinema becomes AI-generated, will the audience keep believing in stories? Wouldn't it be too perfect, so the audience wouldn't recognise itself in the AV content? In the future, AI "risks" rendering content perfect, so would the use of human creativity eventually make content imperfect? But genAI may also have other consequences, three of which stand out. First, it might lead to replacement. Script-writers, for example, fear that their profession might suffer from genAI's capacity to produce not only new ideas, but entire story lines, character descriptions and episodes for series. Second, actors might be replaced by AI-driven versions of themselves or of non-existent people, which would not only have the advantage for the industry of reduced costs, but also of flexibility. It will be cheap and fast, for example, to produce a movie and show it to a test audience, and then to redo some scenes or the outline of the movie altogether. Third, what is new with genAI is that it is democratised. It gives any citizen the tools to make a movie of their liking. Not only does this feed into the fear of job replacement, more importantly, citizens are bound by less professional standards than established industries. To give just one of many examples, according to some reports, more than 95% of all deep fakes concern non-consensual deepfake porn of women.[597] Female celebrities and politicians are regularly targeted, undermining their credibility and causing reputational damage, which has resulted in politicians resigning in order to protect themselves and in particular their family members from harmful content. Perhaps even

[596] More from EBU: https://www.ebu.ch/guides/loginonly/report/ai-regulation-and-its-importance-for-public-service-media-a-look-ahead

[597] https://regmedia.co.uk/2019/10/08/deepfake_report.pdf

© European Audiovisual Observatory (Council of Europe) 2024

more problematic, with deepfake technology democratised and available to anyone, the female body is further sexualized, unrealistic beauty ideals are reinforced, and women are stigmatised and every teenage boy can generate a fake porn video of a girl in class and distribute it privately, on social networks or specialised porn sites. This can have catastrophic effects on girls' social status, perception of self and personal development.

Third, genAI raises complicated questions over creativity and intellectual property. For example, most AI models are trained by simply scraping as much data as possible from the internet. The question is whether this falls under the fair use exception, whether that information should be considered free for use or whether this practice is in clear violation of intellectual property rights and of the possibility of creatives to receive remuneration for their productions. In addition, there are no clear rules on intellectual property rights over AI-produced content. Suppose a person asks a Large Language Model (LLM): can you watch all movies by the Coen brothers and produce a new Coen-like movie for me that is in their style. Who has the intellectual property rights over that new movie: the Coen brothers, the user who came up with the prompt, the AI company, none of them, or all of them? AI-generated creativity also poses the question: what is actually creativity? What is a new work, what should be understood to be the fruit of intellectual processes and can AI ever be creative like humans can?[598] None of these questions can be answered correctly, but at some point the legal paradigm needs to provide clarity and the choices made can and will have considerable effects on the future of creativity.

10.5. Societal challenges raised by AI

This section discusses three societal challenges raised by genAI: election integrity, evidentiary problems in court proceedings and the question of sentience and whether AI-driven entities, at some point in time, will need to be provided protection.

The first challenge relates to the use of AI-driven content in politics. Such technologies are used by politicians, political parties, and the entities and people that support them, as well as by their adversaries. For example, some politicians have their hologram tour the country,[599] create deep fakes of themselves to reach their voters while in jail,[600] have AI tweak their facial features to appear more relatable,[601] or use deep fake technology to give a speech in a minority language the candidate does not actually speak.[602] Although these uses mostly go against the terms and conditions of the various products and services, there are no legal provisions explicitly establishing whether they amount to voter deception or not. Deep fake technology can also be used against

[598] See Chapter 4 on the protection of the output.
[599] https://www.theverge.com/2014/5/7/5691714/indian-politician-uses-holograms-to-reach-voters
[600] https://www.forbes.com/sites/siladityaray/2023/12/18/imran-khan-pakistans-jailed-ex-leader-uses-ai-deepfake-to-address-online-election-rally/
[601] https://www.cfr.org/blog/ai-context-indonesian-elections-challenge-genai-policies
[602] https://www.theverge.com/2020/2/18/21142782/india-politician-deepfakes-ai-elections

© European Audiovisual Observatory (Council of Europe) 2024

politicians, such as by using voice cloning to have a candidate say something compromising or damaging[603] or creating entire fake media environments: a deep fake video of a political candidate doing something outrageous, fake news websites that seem to report on it, fake X accounts that discuss the fake video, fake Instagram accounts that generate memes using frames from the video, etc. Harmful use of AI by political adversaries comes both from domestic rival parties and from foreign powers. Countries like Russia, China and Iran are also targeting the Global South with fake and manipulated news for a variety of reasons, such as influencing concrete decision making (e.g., so that a Russian state-owned company gets a contract), influencing local elections (e.g., to bring a China-friendly regime to power) or influencing decisions at the international level (e.g., by getting a country to vote in favour of lifting sanctions against Iran).[604] Consequently, one threat of deep fake technology is that it is utilised to undermine or influence the political process.

The second challenge is that with the advance of chatbots, humanoid robots, augmented reality, virtual reality, deep fakes and other applications powered by genAI, an increasing amount of online material is produced or manipulated by AI. Currently, under most jurisdictions, evidence introduced in a legal procedure is deemed to be authentic unless there are contraindications. That assumption may need to be reversed: evidence has probably come into contact with AI and the question should be who made which manipulations for what purpose and whether they are relevant to the case at hand.[605] AI-detection programs can only filter out half of the AI-generated or -manipulated material and often only give an "authenticity percentage", eg "the likelihood that this material is authentic is 67%". It is consequently clear that the rise of genAI and in particular the democratisation of deep fake-technology will have significant repercussions for legal procedures. The process might take longer as parties may claim that evidence is inauthentic, there might be an even bigger reliance on expert witnesses and there will be uncertainty about which standards or bars should be met by which party at which point in the process. With the rise of synthetic content two arguments will be increasingly made in the court room, namely the "at the time, I thought it was true" argument (although later it may be established that that belief was based on a deep fake) and the "at the time, I thought it was not true" argument (ibid). Neither of them can be *prima facie* excluded. In addition, because there are no adequate safeguards in place, judges are increasingly likely to make a mistake and incorrectly assume evidence to be authentic or inauthentic.

---

[603] https://www.nbcnews.com/politics/2024-election/democratic-operative-admits-commissioning-fake-biden-robocall-used-ai-rcna140402

[604] Dias, J. A., Doca, H. H., & da Silva, F. F. (2021). Bots, fake news, fake faces and deep fakes: automation, under the bias of dromology, as a sophisticated form of biopower to influence the democratic election process. *Revista de Ciências Jurídicas*, *26*(3), 1-14. Barari, S., Lucas, C., & Munger, K. (2021). Political deep fake videos misinform the public, but no more than other fake media. *OSF Preprints*, *13*. Wilkerson, L. (2021). Still waters run deep (fakes): the rising concerns of "deepfake" technology and its influence on democracy and the first amendment. *Mo. L. Rev.*, *86*, 407. Whyte, C. (2020). Deep fake news: AI-enabled disinformation as a multi-level public policy challenge. *Journal of cyber policy*, *5*(2), 199-217.

[605] https://cyfor.co.uk/deepfake-audio-evidence-used-in-uk-court-to-discredit-father.

© European Audiovisual Observatory (Council of Europe) 2024

Third, a legal question that is looming on the horizon, and if ever, will become relevant only in a decade or even a century, is the move of AI towards sentience and Artificial General Intelligence (AGI). AGI is the moment when computer intelligence is similar to and indistinguishable from human intelligence. It is clear that on some points, AI has already superseded human intelligence, while on other points, it is still lagging behind. The capabilities of AI have advanced exponentially over the last three years, and if they continue to do so at the same speed for another decade, AGI might indeed be attained.[606] Whether AI has self-awareness, or sentience, is even more difficult to establish. However, both points are not black and white issues. As the technology advances, AI will be more and more capable of performing a diversity of tasks and become more and more intelligent. Even if AI does not supersede human intelligence on all aspects of life, it is likely that on many fronts AI will become more potent than humans. The question of sentience is ultimately dependent on a subjective experience, making it difficult to objectively verify what an entity is "really" feeling. Even apart from what is 'really' happening inside a machine, it is clear that humans easily get attached to objects, and attribute to them anthropomorphic qualities. The Tamagotchi, an egg-shaped computer that should represent a digital animal, for example, showed how quickly people become emotionally attached. Now, with the advances in soft robots, LLMs and AI, that may be truer than ever.

Thus, the question has been posed to what extent and when AI-driven entities should be attributed both moral and legal protection or even "human" rights. For example, having intelligent machines work in long and degrading factory jobs has been likened to modern forms of slavery, just like having intelligent entities serve as a sex robot has been likened to forced prostitution. There are also moral and legal questions revolving around the status of deep fake images of non-existent people. Highly realistic images and videos can be produced of people that you would swear exist, but don't. What, if any, moral limits should be imposed on using images of non-existing people? In many jurisdictions, but certainly not all, generating child pornography of non-existent people is prohibited, but what about extremely violent and dehumanizing snuff porn of non-existing people? Finally, there is the question around deep fakes of deceased persons. There is currently no clear legal protection against such usage, as rights are typically only awarded to living natural persons.[607] The question is whether it falls onto the relatives to decide what to do with the data and likeness of the deceased person, whether there should be a legal default setting that such is not allowed, unless the person has indicated in their will that they are okay with such use or, the other way around, that such is allowed, unless people have explicitly made clear that they do not condone such conduct.

[606] See however: Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2023). The curse of recursion: training on generated data makes models forget.
[607] See Chapter 5 on personality rights.

© European Audiovisual Observatory (Council of Europe) 2024

10.6. Conclusion

It is clear that there are many ethical dilemmas and societal challenges raised by genAI. This contribution has homed in on the three potentially most disruptive ones, namely the potential move towards a post truth society, the effects that has on trust and the removal of friction in life and thus the moments, however difficult, through which we learn and develop. In particular, it has discussed the effects of genAI on the AV-sector specifically and discussed the use of genAI by media, but also the challenges that the technology raises in terms of content verification; it has discussed the use of AI by the movie industry, to relieve actors and writers from arduous work, but also the potential effect of job replacement and the fact that users having access to the technology are not bound by the same ethical standards as professional industries; and it has touched upon the question of intellectual property and ownership over AI-produced content as well as the question of what actually is creativity. Finally, it has touched upon broader societal challenges and put the lens on three aspects in particular: the threat of genAI for democratic elections, its effects on court proceedings and the rule of law and the question of whether at some point in time AI-driven entities should be afforded protection and/or whether moral as well as legal limits should be set on their creation and exploitation.

Although it is clear that important personal and societal interests are at stake, good regulatory options are few and far between, as they often entail a choice between Scylla and Charybdis. Regulating deep fakes and other AI technologies for political purposes, for example, raises the question of where the boundary between unproblematic and problematic manipulations should be drawn. Is faking a massive crowd at a rally so bad that it should be prohibited; is altering one's facial features to come across as more approachable so much different from altering one's micro-expressions in photoshop? Even for deep fakes used against a political candidate, it is not always clear where the line should be drawn between an innocent satirical video, and election interference. Similarly, what is to be considered "fake" and what is "real" is often one of the crosslines between voters of different political parties. Prohibiting "fake" news thus easily becomes a power tool in the hands of the incumbent, allowing them to unduly limit freedom of speech; yet doing nothing opens the risk to election interference. Many of the same questions apply to new media.

It is difficult to see how the rise of synthetic content and with it a post-truth era can be prevented, as so many AI-driven tools, products, and services have been democratised and others have AI embedded in their functional design. Although there are several ways of addressing this issue, each raises its own dilemmas. For example, that the AI stimulates hosting providers, platforms and content services to run AI detection programs to filter out AI-generated and -manipulated content. However, this would block an enormous amount of legitimate and unharmful content as well, as most content will be (mostly marginally) manipulated by AI and even substantially AI-modified and AI-generated content is often legitimate. Another way would be to only rely on watermarked content, at least when used by the media or in courtrooms; a watermark is a logbook attached to a photo, video or other material that shows what has been altered or manipulated, when and by whom. However, requiring a watermark could have the effect

© European Audiovisual Observatory (Council of Europe) 2024

that potentially authentic material that could exculpate a person is declared inadmissible by a judge or important and valuable content is missed by media.

As to the removal of friction, this does not undermine the autonomy of people, on the contrary. It feeds into deep human instincts to avoid difficult or even excruciating experiences. A regulator that would force friction upon the lives of citizens because that is what is supposedly best for them runs the risk of becoming overly paternalistic. It is possible to set limits on technologies in specific settings, such as prohibiting students from using ChatGPT to write essays, but even here there are many complicated questions, such as, but not limited to: is it possible to enforce such a prohibition? Can students ask ChaptGPT to find relevant information or academic sources? What about non-native speakers who want their linguistic disadvantage removed by having an LLM check their language, etc? In addition, there is no clear line between instances in which it is valuable to have humans be replaced by AI, for example when actors do not have to perform dangerous stunts, and the many instances leading to the fear of mass job replacement becoming reality. Also, there are often no clear lines regarding originality and creativity when humans use AI for human-AI co-creativity.

Finally, the current legal regime is not designed to protect non-human entities. There are various types of questions raised by genAI on this point. First, most legal regimes do not adequately protect avatars created by people, which often leaves them unprotected against theft and against virtual violence or rape. Second, the rise of deep fakes of deceased persons has raised discussion as to their protection. Many people, for example, would not like their surviving partner to communicate with their deep fake on a daily basis or have them give a speech at their own funeral, but there is no legal protection against such use. It is, however, not evident where the boundary should be drawn. A partner can, for example, communicate on a daily basis with a photo or play a video over and over; AI-empowered daily communication seems just the next step on a sliding scale. Third, there is the question of the extent to which there are moral limits to what can be done with highly realistic representations of non-existent people. Although it is clear that leaving this aspect unregulated raises moral complexities and might have negative real-life consequences, the question is whether it is not too stifling of freedom of speech, creativity and what people do in the privacy of their home to sanction certain immoral acts. Fourth, a question may emerge as to whether AI-driven entities should receive moral and/or legal protection. If so, it should be established at what point machines are so intelligent and have so much self-awareness that they should be considered AGIs and sentient beings, for which there is currently no generally accepted test.

Deepporn is already prohibited through criminal law in most legal systems. Given the enormous number of Deepporn images and videos, it is, however, difficult for law enforcement authorities to adequately tackle the issues, a problem that is aggravated by the fact that it is not always clear who made the video and that the services through which they are published are often based in foreign jurisdictions.

© European Audiovisual Observatory (Council of Europe) 2024

A ban on technologies, apps and services, at least those that explicitly advertise the production of sexual content,[608] could address this issue, but is a radical measure, as it also disallows legitimate and positive use cases, denies innovation of technologies through serendipitous experimentation and prohibitions are often easily circumvented, especially in the online environment.

[608] https://www.vice.com/en/article/deepnude-app-creates-fake-nudes-of-any-woman/

© European Audiovisual Observatory (Council of Europe) 2024

11. Concluding remarks

Time is of the essence. We undoubtedly stand at a pivotal moment in terms of technological and regulatory evolution. The integration of AI into the audiovisual sector offers opportunities and new assisting tools to further develop innovation, creativity, and efficiency. But it also comes with challenges that may require regulatory action.

The current regulatory framework comprises various legislations, and some are still in their nascent stage. The effectiveness of the legislative framework in ensuring the beneficial and sustainable integration of AI into the audiovisual sector remains to be seen.

With the expansion of AI usage by citizens, from recreational purposes at home to content restoration and creative assistance, freedom of expression (Art. 10 ECHR) should remain a key concern when regulating AI. The work has already begun. The Council of Europe's creation of a committee of experts on the impacts of generative AI for freedom of expression in April 2024 is an important step.[609] The committee is tasked with drafting a non-binding Guidance Note on the implications of generative AI for freedom of expression by the end of 2025. According to the draft meeting report from the committee's meeting in April 2024,[610] the Guidance Note should be framed around benefits and systemic risks. The benefits include: expanding access to information at a larger scale, adapting the information format to the individual, for example making the language simpler and communicating visually, or enabling a better understanding and use of information; increasing the visibility of diverse voices, and providing a suitable platform also for groups and individuals in vulnerable situations.[611] The risks encompass the spread of disinformation, de-skilling of people, digital exclusion, manipulation, cheating, deep fakes, and environmental aspects of foundation models.[612]

Furthermore, following the adoption by the Council of Europe's Committee of Ministers of the Framework Convention on AI and Human Rights, Democracy and the Rule of Law in May 2024, the committee on AI is now developing the first legally non-binding methodology for the risk and impact assessment of AI systems from the point of view of

[609] https://www.coe.int/en/web/freedom-expression/msi-ai-committee-of-experts-on-the-impacts-of-generative-artificial-intelligence-for-freedom-of-expression

[610] Committee of experts on Generative AI implications for Freedom of Expression (MSI-AI), 1st meeting 23-24 April 2024, Draft meeting report, MSI-AI(2024)04.

[611] Draft meeting report, point 14.

[612] Ibid, point 15.

© European Audiovisual Observatory (Council of Europe) 2024

human rights, democracy and the rule of law (HUDERIA) to support the implementation of the Framework Convention.[613]

These initiatives show the importance of freedom of expression and the need to balance citizens' interests with the rapid evolutions of AI technologies.

[613] Committee on AI's terms of reference, 1 January 2024 – 31 December 2025

© European Audiovisual Observatory (Council of Europe) 2024

A publication
of the European Audiovisual Observatory