AI4media
ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

Deliverable 2.4

Pilot Policy Recommendations for the use of AI in the Media Sector

Project Title	Artificial Intelligence for the Media and Society
Contract No.	951911
Instrument	Research and Innovation Action
Thematic Priority	H2020-EU.2.1.1. - INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Information and Communication Technologies (ICT) / ICT-48-2020 - Towards a vibrant European network of AI excellence centres
Start of Project	1 September 2020
Duration	48 months

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu

Deliverable title	Pilot Policy Recommendations for the use of AI in the Media Sector
Deliverable number	D2.4
Deliverable version	1.0
Previous version(s)	-
Contractual date of delivery	31 August 2022
Actual date of delivery	15 September 2022
Deliverable filename	AI4Media_D2.4_PolicyRecommendations_final
Nature of deliverable	Report
Dissemination level	Public
Number of pages	104
Work Package	WP2
Task(s)	T2.2
Partner responsible	KUL
Author(s)	Noémie Krack (KUL), Lidia Dutkiewicz (KUL), Emine Ozge Yildirim (KUL), Symeon Papadopoulos (CERTH), Filareti Tsalakanidou (CERTH), Ioannis Kompatsiaris (CERTH), Anna Schjøtt Hansen (UvA), Lucile Sassatelli (UCA), Atia Cortés (BSC), Dario Garcia-Gasulla (BSC), Patrick Aichroth (FhG), Andreas Steenpass (FhG), Luca Cuccovillo (FhG), Birgit Gray (DW), Ipek Baris Schlicht (DW), Rasa Bocyte (NISV)
Editor	Noémie Krack (KUL), Lidia Dutkiewicz (KUL), Emine Ozge Yildirim (KUL)
EC Project Officer	Evangelia Markidou

Abstract	Deliverable 2.4 “Pilot Policy Recommendation for the use of AI in the media sector” reflects on the lack of specific policy recommendations regarding the use of AI applications and tools in the media sector. This absence of guidance creates important challenges for accomplishing the EU's vision on trustworthy AI. The deliverable investigates how this could be fixed. It is divided into three main parts. Section 3 identifies challenges for the use of AI applications in the media sector. It also provides the initial policy recommendations addressing these challenges. Section 4 assesses the feasibility of adopting the European Digital Media Code of Conduct to mitigate some of the challenges faced by the media sector. Section 5 offers concluding remarks and explains the next steps planned by the AI4Media consortium towards the final policy recommendations due in month 48 of the project.
Keywords	AI, Media, Trustworthy AI, EU Values, Policy Recommendations, Challenges.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

Copyright

Contributors

NAME	ORGANISATION
Noémie Krack	KUL
Emine Ozge Yildirim	KUL
Lidia Dutkiewicz	KUL
Symeon (Akis) Papadopoulos	CERTH
Filareti Tsalakanidou	CERTH
Georgia Pantalona	CERTH
Stavroula Rizou	CERTH
Ioannis Kompatsiaris	CERTH
Anna Schjøtt Hansen	UvA
Lucile Sassatelli	3IA-UCA
Atia Cortés	BSC
Dario Garcia-Gasulla	BSC
Patrick Aichroth	FhG
Andreas Steenpass	FhG
Luca Cuccovillo	FhG
Birgit Gray	DW
Ipek Baris Schlicht	DW
Rasa Bocyte	NISV

Peer Reviews

NAME	ORGANISATION
Daniel Gatica-Perez	IDIAP
Georgi Kostadinov & Chris Georgiev	IMAGGA

Revision History

VERSION	DATE	REVIEWER	MODIFICATIONS
0.1	25/05/2022	Noémie Krack, Emine Ozge Yildirim	Table of content, description of the different sections.
0.2	27/06/2022	Noémie Krack, Emine Ozge Yildirim	New version with harmonised layout for partners to upload their contribution.

0.3	15/07/2022	ALL contributors to the deliverable	Partners finished uploading their contributions in the relevant sections.
0.4	17/08/2022	Noémie Krack, Lidia Dutkiewicz, Emine Ozge Yildirim	Version ready for internal review.
0.5	31/08/2022	Daniel Gatica-Perez Georgi Kostadinov Chris Georgiev	Version with reviewers' comments.
0.6	13/09/2022	Noémie Krack, Lidia Dutkiewicz, Emine Ozge Yildirim	Updated version after internal review comments addressed.
1.0	15/09/2022	Noémie Krack, Lidia Dutkiewicz, Emine Ozge Yildirim	Final version to be submitted to EC.

Table of Abbreviations and Acronyms

Abbreviation	Meaning
AI	Artificial Intelligence
AIA	AI Act
AIDA	International AI Doctoral Academy
API	Application programming interface
AVMS Directive	Audiovisual and Media Services Directive
BR	German Bayerischer Rundfunk
CDJ	Le Conseil de déontologie journalistique
CEDEFOP	European Centre for the Development of Vocational Training
CJEU	Court of Justice of the European Union
CoE	Council of Europe
DGA	Data Governance Act
DSA	Digital Services Act
DMA	Digital Markets Act
DL	Deep learning
EBU	European Broadcasting Union
ECHR	European Convention of Human Rights
ECPMF	European Centre for Press and Media Freedom
ECtHR	European Court of European Rights
EDMO	The European Digital Media Observatory
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EFJ	European Federation of Journalists
eg	Example
EGDF	The European Games Developer Federation
EU	European Union

Abbreviation	Meaning
FTC	Federal Trade Commission
GDPR	General Data Protection Regulation
HCI	Human-computer interaction
HLEG	High-Level Expert Group
HRIA	Human Rights Impact Assessment
IFJ	International Federation of Journalists
IGDA	International Game Developers Association
ISFE	The Europe's Video Games Industry
IoT	Internet of Things
KPI	Key Performance Indicator
LEA	Law enforcement agency
ML	Machine Learning
MLEP	The BBC's Machine Learning Engine Principles
NGO	Non-Governmental Organisation
NPO	Nederlandse Publieke Omroep
ÒECD	The Organisation for Economic Co-operation and Development
OSCE	Organization for Security and Co-operation in Europe
PEGI	Pan-European Game Information
PET	Privacy-enhancing technologies
PSM	Public service media
Q	Question
RFoM	The OSCE Representative on Freedom of the Media
R&D	Research and Development
US	United States
UN	United Nations
WIPO	World Intellectual Property Organization

Index of Contents

Index of Figures

Index of Tables

Index of Appendixes

1. Executive Summary

Deliverable 2.4 “Pilot Policy Recommendation for the use of AI in the media sector” reflects on the lack of specific policy recommendations regarding the use of AI applications and tools in the media sector. This absence of guidance creates important challenges for accomplishing the EU's vision on trustworthy AI. The deliverable investigates how this could be fixed. It is divided into three main parts.

Section 3 identifies challenges for the use of AI applications in the media sector divided into three categories. Primarily, the challenges for media companies developing and applying AI applications (sub-section 3.2), which include: staff and knowledge gap within companies, the limited resources available and data issues, the power imbalance vis-à-vis third parties and big tech online platforms who act as providers of AI services, tools, infrastructure, and the lack of guidance and standards to assess and audit the trustworthiness and ethicality of the AI used in media applications. Secondly, the challenges for academia and researchers (sub-section 3.3) that predominantly relate to data: the lack of real-world, quality, and GDPR-compliant data sets to develop AI research. Researchers are left to balance the need for innovation and competitiveness and the need to ensure trustworthy AI without appropriate guidance on how to achieve these aims. Thirdly, legal and societal challenges which apply to both of the abovementioned groups (sub-section 3.4). These include: the complex legal landscape and plethora of initiatives that indirectly apply to media, the lack of certainty on whether and how the AI Act applies to the media sector, and societal risks of bias and discrimination in AI media applications.

In the same section, we provide the initial policy recommendations to the European Union policymakers, as well as European policymakers at large, addressing the challenges identified above. Those include:

- promoting EU-level programs for training media professionals, leveraging on existing schemes and instruments such as the European Center for the Development of the Vocational Training (CEDEFOP) or International AI Doctoral Academy (AIDA);
- promoting and funding the development of national or European clusters of media companies and AI research labs that will focus on specific topics of wider societal impact;
- adopting initiatives such as the Media Data Space, which would enable pooling together AI solutions, and applications in the media sector;
- providing room for early-stage AI innovation and “regulatory sandboxes”, promoting the development of European clusters of media companies and AI research labs that will focus on specific topics of wider societal impact;
- Sharing best practices and developing guidance on how to practically implement the principles of the High-Level Expert Group Guidelines on AI (AI HLEG) in a media context;
- supporting the development of public datasets for AI research, cleared and GDPR-compliant (a go-to place for sharing AI datasets);
- issuing a practical guidance on addressing disinformation, as well as using and publishing datasets with social media data for AI Researchers;

- providing formal guidelines on AI and the GDPR which will address practical questions faced by media sector;
- facilitating a process of establishing standardised processes to audit AI systems for bias/discrimination;
- promoting the development of AI fairness audit reports and certificates for the AI media applications;
- providing a legal certainty on what is media and on the relationship between legacy media and 'new media' (i.e. social media, online intermediaries);
- clarifying what is the place of media in the AI Act proposal;
- ensuring the coordination with other standard setting organisations (CoE, UN, OECD,...) to create a one stop shop easy access to all the different guidelines and instruments.

Section 4 assesses the feasibility of adopting the European Digital Media Code of Conduct to mitigate some of the challenges faced by the media sector. It recalls the main findings of the AI4Media survey performed in D2.3, which demonstrate a lack of awareness or interest in AI ethics management among the AI research community and media professionals. Only 13% of both communities follow or are members of a Code on AI ethics. Then, it investigates on the existing and forthcoming codes of conduct/guidelines around the use of AI in the media sector (sub-section 4.1): press and journalism, broadcasting and public service media, and video games. As our analysis of the existing media codes and guidelines showed, many have overlapping principles such as: transparency, explainability, and keeping humans in the loop. What is challenging, however, is how to apply these principles in daily media and/or research activities. In the light of that, this section concludes that having a generic 'European Digital Media Code of conduct' risks being too general and not addressing the needs and challenges faced by the media companies. The reasons for that are explained in sub-section 4.2: the lack of definition of the 'media', the varying challenges faced by different media sub-sectors, and different stages of AI training/application when the unethical use of AI may appear. Instead, an alternative approach is suggested.

Finally, section 5 offers concluding remarks and explains the next steps planned by the AI4Media consortium towards the final policy recommendations due in month 48 of the project. Section 6, the appendix, provides the overall view of initial policy recommendations addressing the challenges identified in Section 3.

2. Introduction

One of the goals of WP2 is to assess under which circumstances AI-driven tools can be used to improve the media value chain while at the same time enhance the democratic role of media and respect European values, freedom, and rights. This is challenging to achieve, given the lack of specific policy recommendations regarding the use of AI applications and tools in the media sector, which creates important challenges for accomplishing the EU's vision on trustworthy AI.

As part of WP2, AI4Media partners are reflecting on the latest EU policy developments on AI research from different angles to be able to propose novel recommendations for the use of AI in the media sector. The overall goal is to ensure that the European values of ethical and trustworthy AI will be embedded in future AI deployments in the media sector. Through this work for developing a set of initial policy recommendations, the AI4Media consortium aims to raise awareness about the specific challenges that AI media applications trigger and search for possible mitigation measures, especially given the crucial role of media for society and the daily lives of citizens. Being key to shaping citizens' opinions, participating in a democratic and balanced debate, enabling freedom of expression, freedom of the arts, and right to information, the media sector deserves the greatest attention in light of its powerful impact on society. The diverse expertise of AI4Media partners (technical, legal, societal) should inform and improve regulatory and policy initiatives on the topic.

To accomplish the WP2 objectives (providing policy recommendations, a common research agenda on AI and media, a detailed assessment of the political, economic, and social risks, the development of a Media AI Observatory), in the first two years of the AI4Media project WP2 partners worked on three deliverables. Each provides key information that helps to understand the opportunities and challenges for embedding European Values of ethical and trustworthy AI in the media sector. The deliverables are the following (Table 1):

Table 1: An overview of WP2 deliverables

- **D2.1 - Overview & Analysis of the AI Policy Initiatives in EU level (led by KUL)**
  It provides an analysis of a selection of the international and EU policy initiatives on AI. It also analyses the recent EU legislative efforts and proposals on AI regulation which could impact the media sector. The aim is to provide a clear overview of existing and upcoming policy frameworks and an analysis of the ensuing principles and requirements.

- **D2.2 - Initial White Paper on the social, economic, and political impact of media AI Technologies (led by UvA)**
  It provides an overview of some of the core discussions of AI for media from a media studies/social science perspective, identifying the main potentials and challenges connected with AI applications across the media cycle. These concrete challenges are then discussed more widely in terms of how they might impact society (socially, economically, or politically) and what mitigating measures will be important to ensure that the use of AI in the media sector remains responsible and that it positively affects society. The whitepaper is based on a thorough literature review of academic journals

published by scholars within the field of humanities, social science, media, and legal studies, as well as reports developed either with a specific focus on AI in the media sector or with a broader outlook on AI in society. Furthermore, a range of examples of concrete AI applications are described to provide context for the reader and some of the mediated responses to the applications.

- **D2.3 - AI technologies and applications in media: State of Play, Foresight, and Research Directions (led by CERTH)**
  It provides a detailed overview of the complex landscape of AI for the media industry. It analyses the current status of AI technologies and applications for the media industry, highlights existing and future opportunities for AI to transform media workflows, assist media professionals and enhance the user experience in different industry sectors, and offers useful examples of how AI technologies are expected to benefit the industry in the future, and discusses facilitators, challenges and risks for the wide adoption of AI by the media.

These deliverables provide key insights for Task 2.2 "Policy recommendations in the field of AI and Media" as they focus on elements grounded from various perspectives and domains of expertise namely: legal, technical, societal & economical. The purpose of D2.4 is not to provide a summary of the main findings of these deliverables. Instead, it consolidates their findings, and building on them, it provides a follow-up with the aim of formulating initial policy recommendations.

Two other important work efforts which fuel this deliverable and related task 2.2 are: the Online survey on AI for the Media Industry and the survey on Media AI in the service of Society & Democracy (that took place in December 2021 and January 2022). These surveys (Table 2) required the collaboration of various AI4Media partners to design the questions in order to collect the most useful data in line with the project and WP2 goals.

Table 2: An overview of AI4Media WP2 surveys

- **AI4Media Online survey on AI for the Media Industry**

  This online survey (Figure 1) was addressed to both AI researchers working on multimedia AI but also to people working in the media industry or whose work is closely related to this industry (e.g. researchers studying the media, media regulators, people working in relevant NGOs, etc.). The survey aimed to collect their opinions on the benefits, risks, technological trends, and challenges of AI use in the media industry, as well as their experience with AI strategies and AI skills in media organisations, their insights on the most promising ways to facilitate AI adoption and knowledge transfer and, finally, their perceptions about the ethical use of AI. It gathered 150 responses from AI researchers and media professionals from 26 countries in Europe and beyond.

- **AI4Media survey on Media AI in the service of Society & Democracy**

  This short survey was internal, addressed only to AI4Media partners, and aimed to collect their opinions on the benefits and risks of media AI for the society and democracy as well as to record their views with regard to potential policies for the ethical use of media AI, aiming to safeguard fundamental human rights. 31 responses to this survey have been received from media professionals and AI researchers that are part of the project consortium.

For more information on these two surveys, we invite you to read pp. 84-123 of D2.3, which includes extensive analysis and visuals of the survey responses.

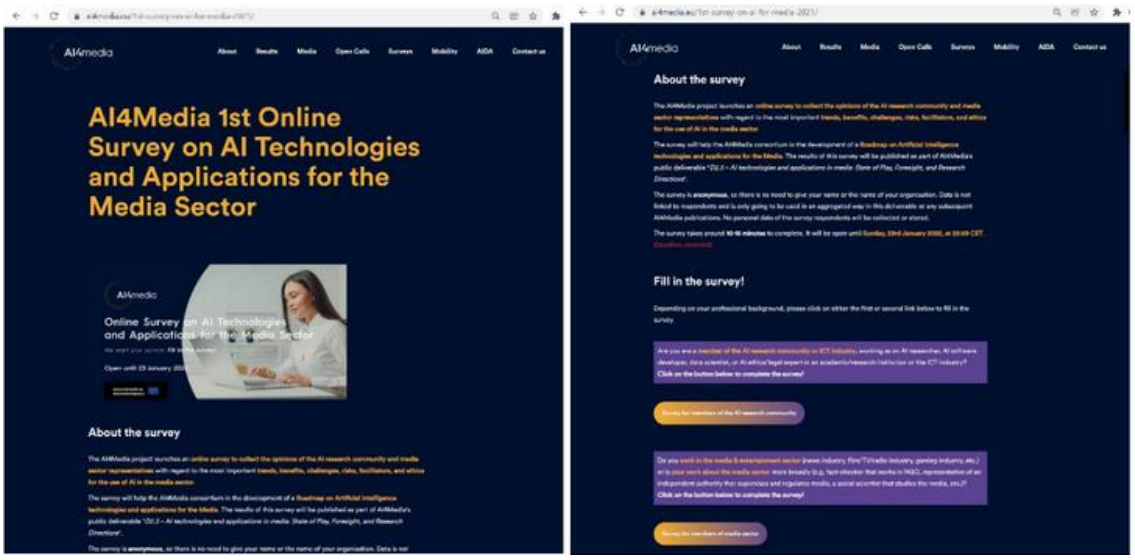


Figure 1: A screenshot of AI4Media Online Survey on AI Technologies and Applications for the Media Sector

Based on the knowledge gathered through desk research and survey analysis, WP2 partners also recently conducted a stakeholder consultation entitled “Towards policy recommendations in the field of AI and media” (on 29 June 2022). The purpose was to hear and discuss directly with three groups of stakeholders about the use of AI in the media sector and identify several challenges encountered in their effort to achieve trustworthy AI (Table 3).

Table 3: An overview of AI4Media stakeholder consultation: "Towards policy recommendations in the field of AI and media"

AI4Media stakeholder consultation: "Towards policy recommendations in the field of AI and media".

The purpose of this event was to present AI4Media and disseminate the main findings of D2.1, D2.2, and D2.3. Partners wanted to hear directly from the various stakeholders their perspective on opportunities and challenges in relation to AI and media, including their ideas on how to tackle those challenges. The event was structured in three different sessions, one for each relevant stakeholder group:

1. Media stakeholders (media companies, media associations)
2. Civil society, NGOs and academia
3. Policymakers (international institutions, civil servants, national media authorities)

The event was held under the Chatham House Rule∗ in order to incentivise a free and open debate. Each session started with a presentation of the project, the deliverables, and the lessons learned. Then the floor was open for discussion and input.

During the event, the smaller nature of groups enabled a fruitful and participatory discussion on the topic. The event gathered around 55 participants, all sessions combined. First, the AI4Media partners briefly presented the project and WP2 outcomes (analysis of policy landscape, whitepaper on social/economic/political impact of media AI, and roadmap) but also the opportunities, challenges and risks in relation to AI and media, which were identified as a result of WP2 activities. Then, based on a few selected points we initiated a discussion inviting participants to share the challenges they face as part for their work and to offer insights and suggestions on how these challenges can be addressed at different levels (media companies, academia & research, EU policy makers). The discussion allowed us to take a deep dive on the actual/practical problems and challenges faced by the different groups of stakeholders and identify a lot of pain points but also potential solutions that may have been missed before. The audiences showed a substantial interest in the AI4Media presentations, which confirms and reiterates the importance of the project in raising awareness about the challenges related to the use of AI in the media sector. Therefore, the event also contributed

to the dissemination of information about the project and the results of the work carried out so far.

The inputs gathered during this consultation are directly incorporated into this deliverable (in sections focusing on challenges and recommendations).

* When a meeting, or part thereof, is held under the Chatham House Rule, participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed. The purpose of the rule is to encourage open discussion since anything said is "off the record".

Thanks to all the work carried out through desk-research, data analysis from surveys and consultation, AI4Media partners now have a better picture of the different kinds of challenges that AI media applications raise for the embodiment of EU values and the respect of fundamental rights. Section 3 will firstly present the various challenges identified, organised around distinct categories: challenges for media companies, challenges for researchers and academia, legal and societal challenges. AI4Media partners worked to group challenges together to provide the best picture possible and include nuances.

Section 3 will then provide the initial/pilot recommendations for the use of AI in the media sector, which aim to address the challenges identified. The recommendations are based on the collaborative and interdisciplinary efforts of the AI4Media consortium. These recommendations are also briefly summarised in the Appendix.

Another aspect of the work for T2.2 is the potential development of a European Digital Media Code of Conduct. The development of a European Digital Media Code of Conduct was initially thought as a possible way to tackle the challenges related to the use of AI in media. Section 4 will firstly analyse and map initiatives in relation to codes of conduct for the use of AI in media. The purpose is to assess the need and feasibility of drafting such a code. The section will analyse whether the subject is already well covered by self-regulation, and if yes at which level, and whether there is the need and the willingness from the media sector to have and comply with such a code.

Finally, Section 5 will summarise the main findings of this deliverable and provide directions for future work in D2.6, which will provide an EU Research Agenda and the final policy recommendations for AI in the Media Sector. It will also contain the final version of the State of Play, Foresight, and Research Directions in relation to AI technologies and applications in media. D2.6 is due in August 2024.

3. Challenges on the road for ethical and trustworthy AI in the media sector and initial policy recommendations

3.1. Methodology

This section presents **identified challenges for the use of AI applications in the media sector** divided into three categories: challenges for media companies developing and applying AI applications (sub-section 3.2); challenges for academia and researchers (sub-section 3.3); legal and societal challenges which apply to both abovementioned groups (sub-section 3.4). The schematic overview of this categorization can be found in the figure below (Figure 2). Furthermore, this section also presents **initial policy recommendations addressing the challenges** identified under each sub-category.

The methodology for the delivery of the policy recommendations included a mix of methods: interdisciplinary research by legal, technical, and societal experts within the AI4Media consortium, as well as an analysis of the 150 responses from AI researchers and media professionals from 26 countries in Europe and beyond which were collected as part of the AI4Media survey (see D2.3).

First, each WP2 partner was requested to provide a selection of the challenges identified from their field of expertise and research. They were also requested to provide initial recommendations on how they think these challenges would be best addressed.

Once the challenges and recommendations were collected, KUL made a categorisation effort by grouping similar challenges together in order to form a comprehensive text and avoid any redundancies. This summarisation and categorisation effort led to the following three types of challenges for the use of AI in media: challenges for media companies, challenges for academia and researchers, and legal and societal challenges (Figure 2).

When the section was ready, it was offered for review to not only the internal reviewers but also the entire AI4Media consortium. This enabled them to validate the categorisation adopted but also provide additional input in relation to the identified challenges and elaborated recommendations.

Challenges for Media Companies

Challenges related to staff and knowledge gap

- Challenges related to the internal organisation of media organisations
- Challenges related to AI innovation
- Difficulty in handling legal and ethical aspects in early-stage AI innovation
- Lack of AI talent in the media industry
- Knowledge gap
- Lack of information related to Trustworthy AI in the innovation context

Challenges related to limited resources and the bargaining power

- Media concentration and journalistic autonomy
- Limited bargaining power
- Licensing tensions between data sets free to re-use for research and for commercial applications
- Lack of coordination between media partners to seek collective solutions

Compliance with legal and ethical frameworks

- Challenges for legal and regulatory compliance
- A need for accessible ethics and legal advices for the media staff
- Lack of information on how AI systems address trustworthy AI challenges
- Lack of instruments for media organisations to audit, assess the ethicality and trustworthiness of AI systems
- Business needs not aligned with ethical concerns
- A need for standardised data documentation
- Transparency of AI models and workflows documentation
- Lack of combination of legal and technical "templates" to simplify data protection compliance
- A need for space for experimentation to support policy

Challenges for Academia and Researchers

Challenges related to the lack of data and data access

- Lack of real-world data to train AI systems for the media sector
- Lack of quality data
- Lack of data for developing synthesis detection and Privacy Enhancing Technologies
- Lack of common understanding that (AI) systems and tools can address privacy aspects without the need to sacrifice utility or performance

Challenges for AI and disinformation research for media

- Lack of common best practices and standards for disinformation analysis
- A need for sustainable R&D for disinformation analysis
- Lack of sufficient API/data access to tackle disinformation

Challenges related to competitiveness and resources

- Tension between aiming at AI Excellence and complying with Trustworthy AI desiderata and requirements
- Open platforms for AI research and evaluation

Legal and Societal Challenges

Complexity of Legal Landscape

- Plethora of policy initiatives and a complex regulatory landscape
- Compliance with the General Data Protection Regulation
- Monolithic policy regulations

Conceptual and definitional challenges

- The role of 'media'
- The so-called 'media exemption' in content moderation
- The place of media in the AI Act

Fundamental rights and societal challenges

- AI-driven Manipulation and Propaganda
- AI bias and discrimination against underrepresented or vulnerable groups
- Filter-bubbles in recommender systems
- Transparent communication

IP challenges

- Copyright challenges of AI use in media

Figure 2: Challenges for the use of AI in media

3.2. Challenges for Media Companies

3.2.1. Challenges related to staff and knowledge gap

Challenges related to the internal organisation of media organisations

While innovation work in media organisations can be conducted by general staff members or teams, in many cases, there are also specialised divisions that are referred to as R&D or content/innovation labs. The latter are often deliberately somewhat removed from the operational part of the media organisation, but in close contact with the organisation's staff, processes or media users. These innovation related divisions differ significantly in terms of size, activity focus, and available budget. While some are fully funded by the media company they belong to, others also acquire additional co-funding from national, European or international research and innovation programmes.

Most of these units are engaged in diverse media innovation topics. Ranging from the application of (newer) technologies such as IoT, AI, Robotics, or Blockchain to media/journalism processes, to the development of new content formats or products (with or without underlying technologies) and the research of associated innovative business and operational issues. In case of success, the resulting tested prototypes, beta-level products, or new knowledge might be transferred into the operative side of the media business as appropriate, including the launch of internal pilot projects for the operative realisation of new solutions, products or services. This process can take weeks or years, e.g., starting with initial ideas, research, and proof-of-concepts, moving towards tested/evaluated mock-ups or prototypes, and then possibly evolving into operative tools, services, formats, or products. Innovation staff and units are often also active in sharing knowledge and open-source results with similar organisations across Europe, with a view to benefit media innovation. This can also include collaborative projects at national or European level.

Challenges related to AI innovation

Recently, the topic of AI and its potential use in media/journalism has become an important aspect of media and journalism innovation, with a number of pioneering examples for media solutions that have come out of or have been shaped by AI innovation processes, to support:

- media processes (such as automatic translation or content verification),
- journalistic approaches (such as the re-use of archive content or investigative projects),
- new AI-powered services (such as content personalisation/recommendation).

Related to what has been mentioned above, there are a number of challenges and concerns, which are more relevant for lower-scale R&D and innovation activities than larger corporate AI implementation projects that are taken up to support the day-to-day core business.

In comparison to general AI technology or tool implementation projects at the operative level, innovation activities in this field are likely to happen in a less defined and ad-hoc work environment. They often have a limited budget with more difficult access to general corporate

resources and an unclear position in the company's strategic priorities. These characteristics are typical for the innovation domain and also a result of the nature of the work, especially in early-stage experiments or proof-of-concept developments with usually uncertain outcomes.

Difficulty in handling legal and ethical aspects in early-stage AI innovation

Members of innovation teams who work on AI-related experiments are increasingly likely to be subject to corporate AI guidelines that their media organisation may have published. Like other AI guidelines from various organisations (including the EU), such corporate guidelines describe principles and "what should be". In practice, however, it can be difficult to implement those principles for small innovation teams and early-stage development projects, unless specialist Trustworthy AI expert advice and support is at hand.

Lack of AI talent in the media industry

One of the reasons hindering the adoption of AI in the media industry is the lack of relevant skills by media professionals and difficulties in recruiting AI experts. This challenge does not only affect media organisations; academic/research institutes are affected as well. This is mainly because they cannot afford to compete with big companies in order to attract AI talent. To overcome this obstacle, AI training and education are necessary for media professionals, but also funding and being able to offer attractive conditions to retain talent. Moreover, collaboration of the media industry with academia/research but also with other media organisations or industries on AI topics of common interest would also be beneficial.

Knowledge gap

Despite the important developments in the use of AI in media, there is still a knowledge gap on what AI is and how it is being used, including in the media sector. This affects media professionals but also end-users, and is due to a low competency of AI, often mixed with high expectations of what the technology can do.

The survey results show that most organisations (58% of respondents) do not have a clear AI strategy in place; furthermore, 17% of respondents do not have ethical frameworks to manage relevant risks or are unaware of the existence of such frameworks (19% of respondents).

Lack of information related to Trustworthy AI in the innovation context

Innovation-related staff are likely to be more aware of the concept and the need for Trustworthy AI. They can be instrumental in bringing related state-of-the-art knowledge into a media company. Accordingly, there are many circumstances where they would seek information related to Trustworthy AI aspects, such as Fairness, Privacy, and Robustness, as well as Explainability, Accountability, or Transparency. For example, when running experiments or developing proof-of-concept services that collect data from public content platforms, when using open datasets from other providers, or connecting third-party AI services via Application Programming Interfaces (APIs). At present, due to being largely subject to research and an emerging field of AI, this type of information is either very difficult to obtain, or not at all provided. In this case – if Trustworthy AI information is not available – media innovation staff cannot ensure that what they are newly developing meets a required "trusted" standard. For

example, that a result/prediction is without bias and/or they can play a part in mitigating bias as part of their development efforts (by-design). It can also be an issue that more explainable AI models need more time for processing than those without explanations, which prevents real-time performance that is sometimes required in a media context.

Initial policy recommendations addressing the challenges related to staff and knowledge gap

In the following, we present the initial policy recommendations addressing the challenges related to staff and knowledge gap, which were discussed in the previous paragraphs.

Table 4: Initial Policy Recommendations Addressing the Challenges Related to Staff and Knowledge Gap

Initial Policy Recommendations Addressing the Challenges Related to Staff and Knowledge Gap	
Challenges related to the internal organisation *Challenges related to AI innovation* *Difficulty in handling legal and ethical aspects in early-stage AI innovation* *Lack of AI talent in the media industry*	• Allow access to ethical guidance provided by specialised public committees for ethical problems that can arise during machine learning model development and AI service development. • Issue practical, easy-to-use guidance and solutions on how to practically implement the responsible, ethical and trustworthy principles listed in corporate as well as other AI guidelines in an innovation context in the media sector. • Establish AI curricula at all education levels. • Start EU-level programs for training media professionals, leveraging on existing schemes and instruments such as CEDEFOP or AIDA. • Start mobility programs for internships or secondments of media professionals in AI research labs or of AI researchers in media companies. • Promote the development of national or European clusters of media companies and AI research labs that will focus on specific AI topics of wider societal impact. These clusters can among other things offer training to media professionals or retraining of technical personnel.
Knowledge gap *Lack of information related to*	• Promote strategies to raise awareness and engage the society into the process of creating a culture for Trustworthy AI. People need to be trained to adopt ethical values and understand capabilities and limitations of AI. Strategies include, but are not limited to, (i) reinforcing education at school on STEM (and in

Trustworthy AI in the innovation context	particular new technologies such as AI) but also on social sciences and humanities (SSH) to adopt European values from early stages; (ii) strengthen collaborations among AI researchers and media professionals to improve communication skills for a general audience, with a common, simple and rigorous language that can inform the society avoiding misconception or overhype of the capabilities of the technology; (iii) involve the society with participatory methods such as open consultations or debates to make them feel part of the technology progress and gain trust towards it; (iv) awareness campaigns such as those launched with the release of the GDPR by national authorities and the EU.

3.2.2.Challenges related to limited resources and the bargaining power

Media concentration and journalistic autonomy

With the current AI Act proposal[1] there is a risk of incentivising further media concentration or outsourcing as the compliance is based on the developer. As the burden of compliance is high, only large newsrooms can potentially be developers, but in most cases, newsrooms would have to outsource their AI needs to large tech companies, minimising the incentive to make media-specific systems and minimise media diversity. The gap between large (national) newsrooms and local newsrooms' access to AI and the resources to build AI should also be considered, as this is already widening the competitive divide and producing more media concentration.

Additionally, as shown by Simon,[2] large technology companies and platform companies[3] play a significant role in news organisations' processes - they act as providers of AI services, tools, and infrastructure, making news organisations dependent on them. This 'infrastructure capture' allows gatekeepers to control the channels of communication and distribution, which pushes the news towards values and logics of online platforms encoded in their algorithms. In turn, these risks undermine news's autonomy.

Limited bargaining power

Organisations that do not develop their own AI tools or do not have the technical capacity to implement open-source solutions need to rely on technology providers and off-the-shelf solutions. In such a context, it is much more difficult, especially for smaller organisations, to raise

[1] European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts COM/2021/206 final.

[2] Simon Felix M., "Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy", [2022] Digital Journalism, DOI: 10.1080/21670811.2022.2063150.

[3] The term "platform companies" is used as an umbrella term encompassing general services such as hardware (cloud computing, cloud storage), software (analysis tools e.g. Google News Consumer Insights, advertising e.g. Google Marketing Platform), AI models, APIs).

ethical and legal concerns as the technology providers are often not economically motivated to address them as there are no market incentives for them to do so. To be commercially successful, AI tools are designed to be scalable, and making custom product changes or updating the product for all customers requires substantial additional resources. Investors, funding bodies and legislators have not put in place framework conditions under which commercial companies would be stimulated to prioritise ethical IA concerns.

Licensing tensions between data sets free to re-use for research and for commercial applications

It is typical in an innovation environment that AI models and services are first developed in a research context and as experiments or proof-of-concepts. This enables staff members to use open-source or publicly available data sets, which is in many cases also essential due to limited research/innovation budgets. Although such datasets carry licences, they are free to use for research activities. If and when experimental AI models/services move to become a pilot project within the company or even operational products, they are therefore subject to licence fees. For innovation teams, it can be challenging to obtain professional business/legal advice on a) which AI resources they can use freely during research and b) at what point in the concept-to-product transition phase a licensing fee will apply, if it is at all possible to obtain a proper commercial licence for the dataset (e.g., due to the way the dataset has been harvested, the provider of the dataset may not be in position to grant certain usage rights). There are also mixed situations, where a machine learning model might be trained on a research dataset, and then used to collect further labels from users of the AI service in a more real-life context.

Lack of coordination between media partners to seek collective solutions

European media organisations in both the private and public sector are losing their market share to big tech platforms. To remain relevant and competitive, they need economies of scale. However, this is not possible due to lack of cooperation between existing players (e.g. in many EU countries television broadcasters do not see themselves as a single market but rather as direct competitors).[4] It is even more prominent because of collusion prevention and regulation. In a more cooperative environment, such organisations could exchange know-how, collectively promote certain standards, adopt codes of conduct, share the burden of creating datasets adhering to European legal standards, and demand changes, which would in turn strengthen their collective market share.

Initial policy recommendations related to limited resources and the bargaining power

In the following, we present the initial policy recommendations addressing the challenges related to limited resources and the bargaining power, which were discussed in the previous paragraphs.

[4] EUI Centre for Media Pluralism and Media Freedom, ' The Media Pluralism Monitor 2022 (MPM2022)', available at : https://cmpf.eui.eu/mpm2022-results/

Table 5: Initial Policy Recommendations Addressing the Challenges Related to Limited Resources and the Bargaining Power

Initial Policy Recommendations Addressing the Challenges Related to Limited Resources and the Bargaining Power	
Media concentration and journalistic autonomy	• Consider where the burden of compliance with the proposed AI Act and similar regulations lies and ensure this will not be harmful to media diversity or to producing responsible AI solutions for the sector. • Consider a solution to the need of levelling up between news organisations and platforms regarding the information asymmetries and resource redistribution.
Limited bargaining power	• Invest in and, importantly, sustained funding for platforms and networks that enable media partners to work in coordinated action, including instruments to continuously gather data on the challenges and needs, and a forum to communicate this to relevant stakeholders (policy and decision makers, industry representatives).
Licensing tensions between data sets free to re-use for research and for commercial applications	• Issue practical, easy-to-use guidance and solutions on how to practically implement the responsible, ethical and trustworthy principles listed in AI guidelines in a media-innovation context. • Issue guidance from business and legal perspectives on which AI resources and datasets can be freely used and certainty about the legal status of the datasets and applicable licence fees.
Lack of coordination between media partners to seek collective solutions	• Provide support on a national level to incentivise coordinated action. • Facilitate more cooperation on the level of pooling together AI solutions and applications in the media sector, apart from initiatives such as the Media Data Space.

3.2.3.Compliance with legal and ethical frameworks

Challenges for legal and regulatory compliance

Research and innovation projects in the area of AI may involve using and testing AI functions from third-party providers via Application Programming Interfaces (API). It can be difficult and time-consuming for media innovation staff to obtain information from the provider on how the AI service they wish to use in experimental projects meets the requirements of existing legislation, such as GDPR[5]. While some more established providers of AI services publish such information on their websites (especially related to GDPR), others do not provide any information related to legal compliance issues. Since the publication of the EU's proposal for AI regulation (the AI Act), there are also concerns that some experimental and early-stage AI developments for the purpose of testing or proof-of-concept may be subject to one of the risk-based categories that are prohibitive or require compliance with substantial legal and transparency conditions (e.g., sentiment analysis). Although the AI Act proposal foresees a "regulatory sand-box" provision, this remains a legal environment focused on "market ready" solutions that require final public testing in an environment under special legal conditions (but where legal conditions still apply)[6]. This instrument, which is designed to support "innovation" in SMEs and Start-ups[7], may not match small innovation teams in the media sector and early-stage experiments. There is also some uncertainty, whether innovation related research activities in commercial or public service media organisations are classified as "academic" research and therefore eligible for (possible) regulatory exceptions and provisions. For instance, if it would qualify as freedom of expression exemptions in data protection laws (eg: art. 85 of the GDPR).

A need for accessible ethics and legal advice for the media staff

Staff working on innovative AI service development in a media context may come across ethical issues with potentially negative consequences on staff, teams or companies involved. For example, they might have to use controversial labels for training a machine learning model with publicly available (transparent) datasets, or inaccuracies in a model could impact negatively on a person/company or their social media account. To avoid such problems, innovation staff would have to undertake significant effort, which may include the seeking of ethical guidance or the human control of the outputs of controversial AI models.

Lack of information on how AI systems address trustworthy AI challenges

AI systems are usually black boxes: we know the input and output, but we do not know how they internally work (how/why they make a decision, what kind of data was used to train them, whether they exhibit some kind of bias, whether they are vulnerable to attacks, etc.). While

[5] 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, 4.5.2016, p. 1–88, ELI: http://data.europa.eu/eli/reg/2016/679/oj

[6] CLAIRE project: What is regulatory sandboxing for AI?, https://claire-ai.org/brainfoods4/.

[7] AI4Media project: D2.1 Overview & Analysis of the AI Policy Initiatives on EU level.

trustworthy AI is a widely used term, it is not clear if/how such considerations have been taken into account when developing AI systems. For AI systems to be widely adopted by media professionals as part of everyday media workflows, it is important for media professionals to know that they can trust the AI systems they use. This means that they should know how trustworthy AI requirements have been considered or to what extent they have been addressed. For instance in the Netherlands, there is the plan to develop a label of trustworthy AI when AI systems are used in public spaces in order to increase transparency around their use.

Lack of instruments for media organisations to audit, assess the ethicality and trustworthiness of AI systems

Whether using in-house built or vendor solutions, it is hard for organisations to practically assess the ethicality, performance and trustability of AI tools. This will only become more difficult as models become more complex and trained on increasingly larger datasets (for instance, the recent example of a Google employee claiming they believed their chatbot was sentient)[8]. While there are benchmarking instruments that researchers can use to assess AI systems, they are not accessible to non-experts.

Business needs not aligned with ethical concerns

Decisions on technical solutions that an organisation will be using are often taken in isolation by CTOs and ICT departments that only consider technical and business aspects (efficiency, costs, scalability, etc.). However, these business needs might not be aligned with ethical considerations (e.g. efficiency often comes at a cost to environmental impact). The challenge here is to ensure that decisions about AI are equally informed by other priorities, such as societal and environmental commitments expressed in the mission and vision of an organisation, and its core values. Involvement of a broader range of internal (as well as external) stakeholders is important for this.

A need for standardised data documentation

With the growing interest and needs for AI technologies, media companies become data consumers, data providers, or both. By incorporating third-party AI models in their production processes, media companies indirectly consume data used to (pre-)train the models. When fine-tuning, transferring, or designing new models to better approach their specific needs with new learning tasks and/or their own data, media companies become data providers (and possibly AI model providers) , be it to their internal varied workflows, or to external partners as well.

While data fuels AI models, the domain of data has been paradoxically under-valued as a scientific domain and under-resourced as a critical process in a ML workflow.[9] Beyond the

[8] Luscombe R., "Google Engineer Put on Leave after Saying AI Chatbot Has Become Sentient", [2022], The Guardian, available at: https://www.theguardian.com/technology/2022/jun/12/google-engineer-ai-bot-sentient-blake-lemoine&gt.

[9] Sambasivan N. and others, "Everyone Wants to Do the Model Work, Not the Data Work': Data Cascades in High-Stakes AI" [2021] Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.

examination of biases in the datasets[10,11,12] and unfairness produced by AI systems relying on these datasets[13], recent works have produced a critical examination of the genealogy of major datasets[14] by analysing the practices and discourses at the time of their creation, and the situatedness of the values they embed. These are characterized by the invisibilisation of the human annotator crowd workers, by the lack of rules for interpretation to document "the process of subjective meaning, bias, and discriminatory classifications" in order to produce accountable datasets.

It is, therefore, crucial to enable media staff to be data-educated and implement rigorous and well-documented dataset creation processes.[15] Such a rigorous process must allow preventing the above lack of transparency and unaccountability issues in the dataset creation, but also prevent so-called data-cascades, defined as "compounding events causing negative, downstream effects from data issues."[16]

Transparency of AI models and workflows documentation

Media companies design their own AI workflows, which may be as simple as applying a third-party tool to their own data, or as complex as building new AI models from existing (possibly pre-trained on third-party data) models for new learning tasks, and possibly chaining different processes or changing parts of the processes (see, e.g., in the law enforcement domain, or face recognition being one building brick of more complex workflows).[17] In such case,, the risk for unintended use and often harm to prejudiced groups of people arises. A stark example is that of hate speech detection systems censoring the very groups they are supposed to protect from harm.[18] This has led to recent studies to understand the demographics of annotator crowd workers and how their identities may impact their rating of subjective interpretive concepts.[19]

It is crucial to enable the prevention or mitigation of such unforeseen impact of AI models, by documenting every step of the model creation in a transparent way (target task, intended use,

[10] Goyal P. and others, "Fairness Indicators for Systematic Assessments of Visual Feature Extractors" [2022] 2022 ACM Conference on Fairness, Accountability, and Transparency.
[11] Wang A, Narayanan A and Russakovsky O, "Revise: A Tool for Measuring and Mitigating Bias in Visual Datasets" [2020] Computer Vision – ECCV 2020 733.
[12] Fabbrizzi S., Papadopoulos S., Ntoutsi E., and Kompatsiaris I., "A Survey on Bias in Visual Datasets," [2021], available: http://arxiv.org/abs/2107.07919.
[13] Buolamwini J. and Gebru T., "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," [2018], Proceedings of the 1st Conference on Fairness, Accountability and Transparency, PMLR 81:77-91, 2018.
[14] Denton E. and others, "On the Genealogy of Machine Learning Datasets: A Critical History of ImageNet" (2021) 8 Big Data & Society 205395172110359.
[15] Sambasivan N. and others, op.cit.
[16] Ibid.
[17] Garvie C., "Garbage in, garbage out: face recognition on flawed data", Georgetown Law, Center on Privacy and Technology, 2019, available at: https://www.flawedfacedata.com/.
[18] Sap M. and others, "The Risk of Racial Bias in Hate Speech Detection" [2019] Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics.
[19] Goyal N., Kivlichan I., Rosen R., and Vasserman L., "Is Your Toxicity My Toxicity? Exploring the Impact of Rater Identity on Toxicity Annotation" [2022], available at: http://arxiv.org/abs/2205.00501

foreseen negative impact, training data, detailed performance analysis disaggregated over diverse groups of data - political sides, demographics, etc.).

Lack of combination of legal and technical "templates" to simplify data protection compliance

Dealing with data protection tends to require individual preparation and use of legal information and technologies, but there are recurring patterns of uses that could be greatly simplified by reusing easy-to-understand "templates" for legal communication and technology, not unlike "Creative Commons" being developed and applied for common types of copyright uses.

A need for space for experimentation to support policy

There are no clear-cut answers to many questions regarding trustability, ethical and legal use of AI systems (e.g. many recommend creating more diverse datasets for training AI, but bias is unavoidable, and a perfectly diverse dataset is an impossible task). Answers to these questions are often multifaceted and require probing from different angles and perspectives. Experimentation environments that allow for cutting-edge experimentation, trial and error (especially error!) are extremely beneficial. An example of such an environment could be the S+T+ARTS project - an initiative of the European Commission to foster alliances of science, technology, and the arts.[20]

Initial policy recommendations addressing challenges concerning compliance with legal and ethical frameworks

In the following, we present the initial policy recommendations addressing the challenges concerning compliance with legal and ethical frameworks, which were discussed in the previous paragraphs.

Table 6: Initial Policy Recommendations Addressing Challenges Concerning Compliance with Legal and Ethical Frameworks

Initial Policy Recommendations Addressing Challenges Concerning Compliance with Legal and Ethical Frameworks	
Challenges for legal and regulatory compliance	• Facilitate access to legal information and issue guidance related to early-stage AI innovation and "regulatory sandbox" for AI development in the context of existing and planned legislation. • Facilitate access to easy-to-integrate, affordable Trustworthy AI enhancement tools, understandable transparency information and trustworthy datasets, both for third-party AI functions that are used in experimental tools/services as well as own early-stage AI development.
A need for accessible ethics advice for the media staff	

[20] S+T+ARTS website, (last accessed 4th August 2022), https://starts.eu/

Lack of information on how AI systems address trustworthy AI challenges	• Ensure that AI systems come with trustworthy AI certificates, ensuring that AI systems have been audited to address issues such as explainability, robustness, fairness, privacy, etc. • Ensure that AI providers and big only platforms i) apply "trustworthiness by design" principles when developing AI systems, ii) provide periodic public reports on how they address trustworthy requirements and particularly explain where they come short, iii) make public any instances in which their systems demonstrably failed to comply with trustworthy AI requirements, leading to negative impacts on media companies or media users using such systems, iv) share information about their data and algorithms with independent researchers and independent authorities that could act as auditors/testers. • Issue guidance on development of trustworthy and explainable AI solutions and relevant certifications by independent authorities that can act both as advisors as well as enforcers. • Ensure that AI providers and big platforms enable independent research on their services and products to analyse potential impact and risks. • Invest in practical solutions that allow media practitioners that do not have a background in AI to critically engage with and assess the ethicality and trustworthiness of AI. • Ensure continued investment in training. Organisational structures in the media industry are needed to create space for continuous learning and keeping up with state-of-the-art research.
Lack of instruments for media organisations to audit, assess the ethicality and trustworthiness of AI systems	
Business needs not aligned with ethical concerns	• Incentivise organisations to prioritise ethical, societal and environmental considerations, for instance, by setting up specific KPIs.
A need for standardised data documentation	• Ensure that the media companies adopt transparent and accountable practices in their dataset creation, enabling auditing of the data they produce to fuel AI models. The AI research community, in a multi-disciplinary approach involving

	interpretive methodologies in semiotics and information science, human computer interaction (HCI) and software engineering, has started to conceptualise the questions around dataset creation for ML and how to design transparent and accountable dataset creation.[21,22]
Transparency of AI models and workflows documentation	• Ensure that the media companies document in a standardised way their model creation process by using existing tools[23]already partly adopted by the AI community.[24] Companies must be enjoined to follow such processes and produce the said documentation when acting as a public media stakeholder.
Lack of combination of legal and technical "templates" to simplify dealing with data protection	• Incentivise and fund the development of "templates" for dealing with common data protection issues, using legal and technical means, tailor-made for the needs of media organisations applying AI.
A need for space for experimentation to support policy	• Invest into creating experimentation and validation environments, such as sandboxes, that would bring interdisciplinary, cross-sector actors and, importantly, provide a direct link to and inform policymakers and regulators.

[21] B. Hutchinson et al., "Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure," (2021), Available: http://arxiv.org/abs/2010.13561

[22] M. Pushkarna, A. Zaldivar, and O. Kjartansson, "Data Cards: Purposeful and Transparent Dataset Documentation for Responsible AI.", (2022), Available: http://arxiv.org/abs/2204.01075

[23] M. Mitchell et al., "Model Cards for Model Reporting," in Proceedings of the Conference on Fairness, Accountability, and Transparency, (2019), pp. 220–229. doi: 10.1145/3287560.3287596.

[24] Model Cards, 'The value of a shared understanding of AI models', (last accessed on 25th July.), https://modelcards.withgoogle.com

3.3. Challenges for Academia and Researchers

3.3.1.Challenges related to the lack of data and data access

Lack of real-world data to train AI systems for the media sector

AI researchers need large volumes of real and high-quality data to train AI models for the media industry. However, access to such data is limited since most media companies, e.g. news organisations or social media companies, are reluctant to share their data, which is of course a source of monetisation for them. It is important to create an environment of collaboration between AI researchers and media companies that will allow sharing of data and development of better and more trustworthy AI systems.

Lack of quality data

Over the last years, there have been several examples of how AI techniques based on Deep Learning (DL) have led to biased outcomes, discriminating against underrepresented communities. DL needs vast amounts of data to train its models and this process can be expensive and time-consuming. Nowadays public large datasets are available and commonly used for this task. However, there is also a problem with the lack of quality in the datasets used to train these neural networks, due to a lack of documentation on the collection, label and training process[25]. Data annotation has been shown to perpetuate discrimination against underrepresented communities and reproduce socially constructed biases.

There is also a problem related to the characteristics of the data used, often classified as personal or sensitive data. The use of synthetically generated data allows to work on a privacy-preserving approach complying with the regulatory aspects of disclosing personal data. However, the faithful or deceiving use of the technology is an emerging risk, since media stakeholders might not be able to recognise and deal with this kind of content.

Lack of data for developing synthesis detection and Privacy Enhancing Technologies

The development of tools for synthesis detection or Privacy Enhancing Technologies often requires person-related data for training and testing. However, there is a lack of sufficient data that is available to use from a data protection perspective. Older published datasets do not explicitly authorise this type of application, and the creation of newer datasets require substantial financial resources.

Lack of common understanding that (AI) systems and tools can address privacy aspects without the need to sacrifice utility or performance

Public discussions about AI applications often suffer from two alternative and somewhat extreme positions. The first extreme considers data protection as an absolute goal, tends to see legal regulation as the one and only means to make progress, and is often critical towards technical innovation. The second extreme considers technical innovation and utility as an absolute goal, which should not be bothered by any regulations related to data protection, since

[25] Excavating AI website, (last accessed 7th July 2022), https://excavating.ai

regulations automatically cripple utility. Both extremes share the (false) view that it is only possible to have either privacy and data protection, or functionality and performance, but never both. However, in many if not most cases, this is wrong: It is often possible to achieve both, at least to some extent, by applying suitable organisational and technical means (e. g. Privacy Enhancing Technologies), resulting in business opportunities. Regulation and innovation are not contradictions, but can complement each other; in other words, the widespread misconception of privacy and utility / performance is an obstacle to many opportunities, especially in Europe, which would be in an advantageous position to combine both aspects.

Initial policy recommendations addressing the challenges related to the lack of data and data Access

In the following, we present the initial policy recommendations addressing the challenges related to the lack of data and data access, which were discussed in the previous paragraphs.

Table 7: Initial Policy Recommendations Addressing the Challenges Related to the Lack of Data and Data Access

Initial Policy Recommendations Addressing the Challenges Related to the Lack of Data and Data Access	
Lack of real-world data to train AI systems for the media sector	• Promote the development of national or European clusters of media companies and AI research labs that will focus on specific topics of wider societal impact. In the context of such initiatives, the clusters will pursue the development of public datasets for AI research, the development of standard and transparent mechanisms for the formation of bilateral agreements for data sharing between media industries and AI researchers, the establishment of benchmarking datasets and competitions for testing AI algorithms, etc. • Promote the European AI on demand platform and its marketplace as the go-to place for sharing AI datasets while also providing incentives for developing datasets or sharing data.
Lack of quality data	• Ensure that AI researchers and practitioners adopt best practices of data management that guarantee the highest quality of datasets, enabling reusability and accountability. • Ensure that the use of synthetic data be clearly disclaimed on applications that can be considered as high risk, such as deep fakes. There should be a clear definition of what is considered a dangerous application of synthetic data and when it is

	necessary to inform media stakeholders.
Lack of data for developing synthesis detection and Privacy Enhancing Technologies	• Incentivize / fund activities that check and "clear" older datasets with respect to data protection by getting consent, to make them usable for synthesis detection and privacy enhancing technologies (PET) development. • Incentivize / fund activities for creating new datasets for synthesis detection and PET development.
Lack of common understanding that (AI) systems and tools can address privacy aspects without the need to sacrifice utility or performance	• Incentivize the development and use of privacy enhancing tools, especially considering usability aspects, for R&D projects and services. • Incentivize the communication of cases in which alleged conflicts between privacy and utility could be resolved using technology.

3.3.2.Challenges for AI and disinformation research for media

Lack of common best practices and standards for disinformation analysis

Disinformation analysis within media companies currently suffers from the lack of common guidelines and standards as to which AI tools to use, how to interpret results from AI tools, how to document the overall verification process including such tools, how to ensure replicability of verification steps by others using respective tools, and how to minimise human bias (especially confirmation bias) within content verification processes. For some of the mentioned aspects, it would be possible to draw from results from research projects, industry initiatives, or from extensive activities from adjacent sectors such as law enforcement (see e. g. "Best Practice Manual for Digital Image Authentication"[26]), but there are white spots, and the different aspects and "pieces" need to be combined into one coherent document and tool set.

[26] Best Practice Manual for Digital Image Authentication, ENFSI-BPM-DI-03, available at: https://enfsi.eu/wp-content/uploads/2021/10/BPM_Image-Authentication_ENFSI-BPM-DI-003-1.pdf

A need for sustainable R&D for disinformation analysis

Due to the lack of private / industrial investment and potential conflicts of interest, sustainable and flexible public R&D funding for media disinformation information analysis (detection of decontextualization, manipulation, and synthesis) is extremely important to be able to address the related challenges. The continuous technological improvements of production tools (which can be used by attackers) can only be addressed with sustainable and flexible funding to develop of defensive tools ("cat-and-mouse-game"). Technologies for speech and video synthesis, for instance, have seen an impressive improvement over the last years, but media-related calls on national and EU levels that addressed such topics have been scarce, delayed (especially considering that research needs to address respective topics years in advance to be able to have tools that can be applied in practice) and tend to lack flexibility to quickly react to the fast technological developments.

As for other sources of funding for disinformation-related R&D, the LEA domain comes to mind, and indeed, many of the challenges are similar and there is great (albeit until now, largely unused) potential for synergies between the media and the LEA domain related to disinformation analysis. However, LEA-related calls related to media forensics and disinformation, which could serve as an alternative funding source, are also scarce: While being considered an important topic and e.g. being mentioned in the SOCTA report,[27] the topic was not included in the current EMPACT priorities[28] and therefore, will not be part of forensics R&D calls until 2025(in contrast to the US, which supports DARPA programs such as the MediFor[29] and SemaFor[30]).

To address the aforementioned problems, it would be great if R&D for disinformation analysis and detection could be supported with more frequent project calls (accepting the fact that this is not a one-time issue, but will require R&D for many years to come), possibly favouring shorter project durations and smaller budgets while allowing more flexibility. Moreover, it would be good to incentivize a stronger cooperation between the media and LEA in this research domain, to exploit the synergies among them.

Lack of sufficient API/data access to tackle disinformation

The ability of independent researchers to access online platforms' data is a precondition for effective platform governance, independent oversight and to understand how these platforms work. For instance, access to online platforms' data allows researchers to carry our public interest research into platforms' takedown decisions, recommender systems and advertising. It

[27] Europol, "European Union Serious and Organised Crime Threat Assessment (SOCTA) 2021. A corrupting influence: the infiltration and undermining of Europe's economy and society by organised crime", available at: https://www.europol.europa.eu/publications-events/main-reports/socta-report

[28] Europol, "EU Policy Cycle - EMPACT 2022+ Fighting crime together" available at: https://www.europol.europa.eu/crime-areas-and-statistics/empact

[29] DARPA website, (last accessed 10th August 2022), https://www.darpa.mil/program/media-forensics

[30] DARPA website, (last accessed 10th August 2022), https://www.darpa.mil/program/semantic-forensics

also allows researchers to identify organised disinformation campaigns, to understand users' engagement and to identify how platforms enable, facilitate, or amplify disinformation (e.g. via their recommender systems, engagement-based business model). This evidence-based work critically depends on access to platform data. Yet, whereas the number of platforms' data is constantly growing, it has become increasingly difficult for researchers to access that data.[31]

First, it is a common knowledge by now that access to platforms' data for researchers is currently mainly governed by contractual agreements, platforms' own terms of service and public application programming interfaces (APIs). APIs access can be restricted or eliminated at any time and for any reason. Moreover, platforms' APIs vary significantly: Twitter and YouTube are considered to have relatively generous data access through APIs vs Facebook or Instagram, for example. The over-reliance on a single platform, raises justified questions of representativeness and validity of research results.[32] The fact that the APIs are not sufficient and impose restrictions on further use of data which they select, made researchers coined the term 'the post APIcalypse'[33] and a 'post-API age'.[34] The UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, stressed a lack of transparency and access to data as 'the major failings of companies across almost all the concerns in relation to disinformation and misinformation'.[35]

Further, many researchers experience problems with understanding and interpreting platforms' APIs terms of use. To give an example, Twitter permits the sharing of Tweet IDs and User IDs in a dataset for others to use (for academic research the number is currently unlimited). The recipient must then 're-hydrate' those IDs into Tweets using Twitter's API. A limited number of 'hydrated' Tweets can be shared, but only privately, and the dataset creator who is sharing their dataset must ensure that the recipient has agreed to the Twitter terms before doing so (ensuring that the recipient must then also respect the synchronisation and other requirements). Twitter also places particular restrictions on the form in which Tweets may be published, requiring certain items of data to be retained in the published form. In addition, its synchronisation requirements apply to published material, meaning that should a user delete or protect a Tweet that has been quoted in a paper, that paper would need to be modified to remove it. This is a

[31] Dutkiewicz L., "From the DSA to Media Data Space: the possible solutions for the access to platforms' data to tackle disinformation", 2022, European law blog, available at: https://europeanlawblog.eu/2021/10/19/from-the-dsa-to-media-data-space-the-possible-solutions-for-the-access-to-platforms-data-to-tackle-disinformation/

[32] Tufekci Z., "Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls" [2014], Proceedings of the 8th International AAAI Conference on Weblogs and Social Media, available at: https://arxiv.org/abs/1403.7400.

[33] Bruns A., "After the 'Apicalypse': Social Media Platforms and Their Fight against Critical Scholarly Research" [2021] Disinformation and Data Lockdown on Social Platforms 14.

[34] Freelon D, "Computational Research in the Post-API Age" [2018] 35 Political Communication 665.

[35] Khan. I., "Disinformation and freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression", A/HRC/47/25 13 April 2021, available at: https://undocs.org/A/HRC/47/25.

well-recognised issue in the ethics literature and there is general guidance to seek individual informed consent from the user whose Tweet is intended to be published so that even if they withdraw the Tweet from Twitter, the researcher has clear ethical consent to publish it.[36]

It is also worth recalling that in January 2022, in a cross-border case, the Belgian Data Protection Authority imposed a fine on EU DisinfoLab[37] – an NGO that fights disinformation – and on one of its researchers for violating the GDPR. The penalties relate to i.a. the publication of raw data from Twitter accounts. This decision is notable as it holds one of the researchers personally liable (and fined with €1,200).[38] This is to illustrate that media organisations and fact-checkers communities struggle with understanding and complying with different APIs conditions of use and GDPR obligations.

Initial policy recommendations addressing the challenges for AI and disinformation research

In the following, we present the initial policy recommendations addressing the challenges for AI and disinformation research for the media, which were discussed in the previous paragraphs.

Table 8: Initial Policy Recommendations Addressing the Challenges for AI and Disinformation Research

Initial Policy Recommendations Addressing the Challenges for AI and Disinformation Research	
Lack of common best practices and standards for disinformation analysis	• Establish R&D calls to develop a coherent best practice document and tool set for content verification in the media sector, using and combining existing information.
A need for sustainable R&D for disinformation analysis	• Establish *recurring* R&D calls that include media forensics and media disinformation analysis development to ensure that technologies can cope with the overall technology development that is available for attackers.
Lack of sufficient API/data access	• Address the clear need for a legally binding data access framework at the EU level that provides researchers with access

[36] Williams, M. L., Burnap, P., & Sloan, L. (2017). Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views, Online Context and Algorithmic Estimation. Sociology, 51(6), 1149–1168. https://doi.org/10.1177/0038038517708140.

[37] Belgian DPA, Decision nr 13/2022 of 27 January 2022, DOS-2018-04433.

[38] Vanleeuw R., "Processing of public Twitter data: as free as a bird?" (2022), CiTiP blog, available at: https://www.law.kuleuven.be/citip/blog/processing-of-public-twitter-data-as-free-as-a-bird/.

to tackle disinformation	to a range of different types of platform data. • Further promote and expand initiatives such as Media Data Space, or the EDMO's draft Code of Conduct on how platforms can share data with independent researchers while protecting users' rights.

3.3.3.Challenges related to competitiveness and resources

Tension between aiming at AI Excellence and complying with Trustworthy AI desiderata and requirements

The AI realm has grown into a hyper-competitive research space, where cutting edge developments are pushed forward at a remarkable pace, mainly by big AI labs in the US, Canada and China which have massive resources at their disposal and typically operate under relatively lax ethical standards. At the same time, EU-based AI researchers typically operate under more limited resources (funding, compute resources, access to relevant datasets) and additionally are under intense pressure to comply with a wide range of regulations and ethical standards. This disparity/disadvantage is even more pronounced for the AI for the media sector, where industrial funding is extremely limited or non-existent (since media organisations are themselves under-funded), the regulatory landscape is highly complex and evolving (the Digital Service Act (DSA)[39], Digital Markets Act (DMA)[40], AI Act (AIA)[41]) and access to datasets is controlled by non-EU companies (big digital platforms). This puts media AI researchers at a great disadvantage and considerably compromises their capacity to deliver world-class advances in this area. This echoes the work of a cross-sector, multi-stakeholder Working Group established by the European Digital Media Observatory (EDMO) in May 2021. The goal was to start drafting a Code of Conduct under Article 40 of the GDPR on platform-to-researcher data access in order to find guidance and solutions to the legal tensions between platform's users' privacy and the need for data access for research.[42]

Open platforms for AI research and evaluation

AI platforms tend to substantially benefit from "positive network effects", i. e. the value that an individual (business) user derives from a platform grows with the overall number of (business)

[39] European Commission, Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC COM/2020/825 final.

[40] European Commission, Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act) COM/2020/842 final.

[41] AI Act op.cit.

[42] EDMO & the Institute for Data Democracy & Politics (The George Washington University), 'Report of the European Digital Media Observatory's Working Group on Platform-to-Researcher Data Access', (2022), available at : https://www.google.com/url?q=https://edmo.eu/wp-content/uploads/2022/02/Report-of-the-European-Digital-Media-Observatorys-Working-Group-on-Platform-to-Researcher-Data-Access-2022.pdf&sa=D&source=docs&ust=1662130030233908&usg=AOvVaw0hQHiaBjhPj6xIxiIfS8Ty

users: the more (business) users use an AI platform and provide their services to it, the greater the overall value of all services, because (a) various complementary functionalities can be combined into more powerful services, (b) data from more sources and end-users can be leveraged, and (c) maintenance costs for tool development by users can be kept to a minimum. Current or future platforms for AI research tend to fall into this category: The more AI tools are integrated into such a platform, the more powerful workflows can be supported, and the more users are using it, the greater the incentive for other users to also use it. As a consequence, it is often the case that organisations that have the resources and the willingness to take the risk of investing considerable resources to set up and maintain a platform early on, can create quasi-monopolies ("winner-takes-all"): Net effects can be so strong that entry barriers for new offerings are very high, except if they are willing to make even larger investments and take bigger risks, and that can lead to platform users having to pay a high price (e.g. in terms of price, lock-ins, etc.) for usage.

In some cases, it can be of public interest to establish (not necessarily maintain!) open platforms, to promote the broad development of competences and know-how. Platforms to integrate, combine and evaluate AI components for research purposes fall exactly in this category. On the technical level, such platforms can be implemented in a similar way as AI4EU Experiments, which is also supported by AI4Media's WP7: Platforms define open APIs, which are used to implement and integrate individual tools, and provide a common model to ensure interoperability among the integrated tools. This does not necessarily mean that developed tools must be open source. Indeed, one possible approach is to provide only the core platform functionalities as business-friendly, non-copyleft open source, while the tools using the platform can be both open source or closed source, allowing for maximum openness regarding the related business approaches. The problem is, such platforms do not emerge by themselves, due to the need for investment outlined above.

Initial policy recommendations

In the following, we present the initial policy recommendations addressing the challenges related to competitiveness and resources, which were discussed in the previous paragraphs.

Table 9: Initial Policy Recommendations Addressing the Challenges Related to Competitiveness and Resources

Initial Policy Recommendations Addressing the Challenges Related to Competitiveness and Resources	
Tension between aiming at AI Excellence and complying with Trustworthy AI	• Ensure sufficient public funding to research projects on AI for the media. • Ensure fair and straightforward access of European researchers to large compute infrastructure and relevant datasets. • Create a "safe regulatory space" for researchers that provides

desiderata and requirements	clear guidance on the limits of ethical AI research without over-constraining freedom of thought and innovation.
Open platforms for AI research and evaluation	• Incentivise the use / implementation of open APIs and public ontologies by public funding to improve interoperability and evaluation of tools developed within funded projects. • Prepare (regular) R&D calls for the development and maintenance of application-specific open platforms and ontologies, especially in domains with large impact for the public (e. g. disinformation analysis).

3.4. Legal and societal challenges

3.4.1.Complexity of Legal Landscape

Plethora of policy initiatives and a complex regulatory landscape

As already observed in our deliverable D2.1 "Overview & Analysis of the AI Policy Initiatives at EU level", AI systems are the heart of international, European and national and sectoral policy initiatives. We also noted that regulatory initiatives impacting the use of AI media applications are also rapidly developing. This mass of content and obligations constitutes a complex policy and regulatory framework to be aware of, understand and implement. This constitutes a specific burden for smaller organisations which must employ more forces and more staff to digest and apply all this content to their practical AI applications.

On the policy aspect, we have seen that there has been a plethora of policy initiatives. Coming from different levels of power and different institutions, it's not always easy for a non-lawyer or policy expert to find its way through the variety of content and determine which one should be applicable when developing the use of AI systems in the media sector. Especially as most of the AI initiatives are very high level and hard to implement specifically for the media sector.

In the recent years we have nevertheless seen some specific policy initiatives specifically in relation to AI, media and freedom of expression with different stakeholders as addressee. For instance, the Organisation for Security and Co-operation (OSCE) has a Representative on Freedom of the Media (RFoM). In December 2020, it released a Policy Paper on freedom of the media and AI[43] and in April 2021, it published a policy paper on AI and freedom of expression in

[43] OSCE Representative on Freedom of the Media, 'Policy paper on freedom of the media and artificial intelligence' (2020), https://www.osce.org/files/f/documents/4/5/472488.pdf

political competition and elections[44]. It has released on 20 January 2022 a manual containing policy recommendations on the most effective ways to safeguard freedom of expression and media pluralism, when deploying advanced machine-learning technologies such as AI.[45] 14 recommendations are addressed to the OSCE Participating States. The manual contains the findings of experts about the main challenges that AI tools pose to human rights, in particular, the right to freedom of expression and opinion, and media freedom and pluralism. It is organised around AI in content moderation and curation. The analysis focuses on challenges relating to the use of AI in these domains such as security threats, hate speech, media pluralism, and surveillance-based advertising. The Council of Europe (CoE) hosted a conference[46] on artificial intelligence and the challenges and opportunities for media and democracy, where they published a background paper[47] about the impacts of AI-powered technologies on freedom of expression and they adopted a Declaration and Resolution on AI tools used for the creation, moderation and distribution of online content.[48] We can also mention CoE guidance note on AI and freedom of expression, with clear recommendations to member states, platforms and civil society on content moderation, as well as best practices.[49] Finally, the CoE also adopted recommendations on the impacts of digital technologies on freedom of expression[50] and on the human rights impacts of algorithmic systems[51].

[44] OSCE Representative on Freedom of the Media, ' Policy paper on AI and freedom of expression in political competition and elections', (2021), HYPERLINK "https://www.osce.org/representative-on-freedom-of-media/483638"https://www.osce.org/representative-on-freedom-of-media/483638

[45] OSCE Representative on Freedom of the Media, ' Spotlight on Artificial Intelligence and Freedom of Expression: A Policy Manual', 20 January 2022, https://www.osce.org/representative-on-freedom-of-media/510332

[46] Council of Europe, 'Conference of Ministers responsible for Media and Information Society Artificial intelligence – Intelligent politics Challenges and opportunities for media and democracy' (2021), https://www.coe.int/en/web/freedom-expression/media2021nicosia

[47] Helberger N., Eskens S., van Drunen M., Bastian M., Moeller J. for the Council of Europe, 'Implications of AI-driven tools in the media for freedom of expression', (2019), https://rm.coe.int/coe-ai-report-final/168094ce8f https://rm.coe.int/cyprus-2020-ai-and-freedom-of-expression/168097fa82

[48] Council of Europe, 'Final Declaration of the Conference of Ministers responsible for Media and Information Society and resolutions on freedom of expression and digital technologies, on the safety of journalists, on the changing media and information environment, on the impacts of the COVID-19 pandemic on freedom of expression 10-11 June 2021', https://rm.coe.int/final-declaration-and-resolutions

[49] Council of Europe, 'Guidance note on content moderation. Best practices towards effective legal and procedural frameworks for self-regulatory and co-regulatory mechanisms of content moderation', (2021) accessible at : https://rm.coe.int/content-moderation-en/1680a2cc18

[50] Council of Europe, 'Recommendation CM/Rec(2022)13 of the Committee of Ministers to member States on the impacts of digital technologies on freedom of expression', (2022), available at : https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680a61729

[51] Council of Europe, 'Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States
on the human rights impacts of algorithmic systems', (2020), available at :
https://search.coe.int/cm/pages/result_details.aspx?ObjectId=09000016809e1154

On the regulatory aspect, there are still only very few sectors specific legislative and binding initiatives in force. However, as outlined by the HLEG guidelines on trustworthy AI, AI systems must comply with existing legislations which are already applicable to them even though they are not explicitly targeting the media sector or AI technology. This includes, fundamental rights corpus of law, EU primary law[52], relevant EU secondary law (the GDPR and the e-Privacy Directive[53], the Product Liability and Safety Directives[54], the Regulation on the Free Flow of Non-Personal Data[55], Anti-discrimination Directives[56], Consumer law and Safety and Health at Work

[52] Treaty on European Union (TEU), OJ C 326, 26.10.2012, p. 13–390,ELI: http://data.europa.eu/eli/treaty/teu_2012/oj ; Treaty on the functioning of the European Union (TFEU), OJ C326, consolidated version on 26.10.2012, p. 47–390, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012E%2FTXT and the Charter of Fundamental Rights of the European Union, OJ 326, 26 December 2012, p.391-407, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT

[53] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), OJ L201, 31 July 2022, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02002L0058-20091219

[54] 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, 7 August 1985, p. 29–33, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A31985L0374 and Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety (Text with EEA relevance), OJ L 11, 15 January 2002, p. 4–17, https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32001L0095

[55] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union (Text with EEA relevance.), OJ L 303, 28 November 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32018R1807

[56] Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation, OJ L303/16, 2 December 2000, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2000:303:0016:0022:EN:PDF and Council Directive 000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin, OJ L180/22, 19 July 2000, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2000:180:0022:0026:EN:PDF

Directives[57], the E-commerce Directive[58], AVMS Directive[59], Copyright in the Digital Single Market[60],...) and so forth.

This means that AI media applications are already falling in the scope of many EU and national legislations and AI developers/users must comply with them. This patchwork of provisions scattered in various instruments makes it a complex topic to be in compliance with, as many links between provisions must be made. Some existing legislations have been deemed insufficient to address the specific challenges AI systems may bring.[61] Whether not adapted or with too high-level and vague provisions, we have witnessed how the EU policymakers decided to tackle the issue and solve the gap by a wide regulatory effort: AI-specific regulation (the AI Act), other new instruments (the DSA, DMA, Data Governance Act (DGA)[62], Data Act[63],...) or by the revisions of already existing legislations. In some of these initiatives, the AI media applications are sometimes taken into account (although not explicitly) but otherwise are often overlooked despite the crucial impact that media have on EU citizens' lives and on shaping societies and democracies. This complex regulatory landscape is a considerable challenge for media stakeholders, researchers in AI media applications and other interested parties to have a comprehensive view of the relevant obligations and how to make it work altogether. Indeed, some legislations borrow terminology which needs to be consistent to ensure a proper harmonisation.

Compliance with the General Data Protection Regulation

As already described in our deliverable D4.3, entitled "Initial analysis of the legal and ethical framework of trusted AI", even if the GDPR does not refer to 'artificial intelligence', many

[57] All legal and policy initiatives related to safety and health legislation can be found here : European Agency for Safety and Health at work, 'Safety and Health legislation', https://osha.europa.eu/en/safety-and-health-legislation

[58] 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 ('Directive on electronic commerce'), OJ L 178, 17 July 2000, https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031

[59] 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, 28 November 2018, https://eur-lex.europa.eu/eli/dir/2018/1808/oj*

[60] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, OJ L 130, 17.5.2019, p. 92–125

[61] See for example: Eskens S., "A right to reset your user profile and more: GDPR-rights for personalized news consumers", International Data Privacy Law, Volume 9, Issue 3, August 2019, Pages 153–172, https://doi-org.kuleuven.e-bronnen.be/10.1093/idpl/ipz007.

[62] European Commission, Proposal for a regulation of the European Parliament and of the CouncIL on European data governance (Data Governance Act) COM/2020/767 final

[63] European Commission, Proposal for a regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act), COM/2022/68 final, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN

provisions of the legal text prove to be relevant for AI systems. However, the extent to which the GDPR applies to AI is being (re)defined. The guidance from official institutions such as the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) focuses only in part on the AI systems. There is a lack of sufficient clarity, and many uncertainties and diverging opinions between scholars and interpretative guidelines. The academic literature showed sometimes converging and in other cases conflicting opinions among the research community on the scope of some GDPR provisions applied to AI systems.

Many challenges exist in relation to the interpretation and the respect of the GDPR in an AI context. The challenges vary from complexities related to the different stages of AI processing to the transparency problems. It also includes a lack of friendly interfaces and technical tools to enforce data subjects' rights. The issues with GDPR enforcement also contribute to trade-offs strategies by companies. This harms considerably the enforcement of data subject's request and the protection of the rights. The absence of explicit reference and further explanations on how the GDPR concepts could be applied to an AI system environment, especially those applied in the media sector, creates a need for further guidance on the topic. Indeed, compliance with data subjects' rights is a growing challenge.

Monolithic policy regulations

At the moment, policies regulating the use of AI tend to focus on addressing the negative impacts that were largely caused by commercial big-tech companies. Regulation should not only address and mitigate the risks caused by the use of AI systems but also incentivize and protect their positive use. More on the latter is needed to rebalance the approach and ensure to maximise the potential of AI systems.

Initial policy recommendations

In the following, we present the initial policy recommendations addressing the complexity of the legal landscape for AI and media discussed in the previous paragraphs.

Table 10: Initial Policy Recommendations Addressing the Complexity of Legal Landscape

Initial Policy Recommendations Addressing the Complexity of Legal Landscape	
Plethora of policy initiatives and a complex regulatory landscape	• Ensure a practical guidance document for the compliance with relevant regulations when applying AI in the media sector. • Consider an instrument focusing specifically on the use of AI in the media to ensure legal certainty and answer the calls from society and professionals. • Have a place where a mapping of relevant legislation for the development, the use of AI systems in media applications is easily accessible with different categories of media actors/sectors represented, so each one knows their main obligations and rights

	according to the different frameworks. • Ensure during the drafting process of regulatory initiatives that the different existing instruments, along with ongoing proposals, would function together, allowing to have a global and systemic approach to this complex regulatory landscape. • Address the need to ensure consistency to improve harmonisation and legal certainty. Ensuring more coordination and better communication on the interplay between the different instruments seems necessary.
Compliance with the General Data Protection Regulation	• Solve the terminologies inconsistencies when it comes to AI and GDPR, especially for interpretable, explainable and transparent. • Define a formal framework for reliability, transparency and fairness in AI in the media applications. • Develop a multidisciplinary definition of interpretable AI which can be adopted in both the social and the computer sciences. • Resolve the issue of the power asymmetries for AI development with limited private companies' monopoly over data and develop AI inclusive policies and regulations. This can be achieved by putting citizen-centred innovation with class action, activism, and whistle-blowers schemes. • Provide more guidance on how to produce, find, use and re-use clean data sets used to develop AI applications in the media sector. • Provide best practices/guidelines on AI and GDPR (in a comprehensive manner) which will address practical questions and doubts faced by media staff.
Monolithic policy regulations	• Ensure that regulations do not create barriers for organisations to innovate areas that are seen as high-risk. • Address the need for mechanisms to prevent and effectively address misuse of AI technologies in media (e.g. deepfakes) but also mechanisms to incentivise their use for social good.

3.4.2.Conceptual and definitional challenges

The role of 'media'

When one talks about the 'AI for media', it is crucial to understand what is meant by both 'AI' and 'media'. While the legal definition of AI has been subject to many discussions[64], the term media, less so. As observed by Coe, 'the media landscape is undergoing profound change, on an unprecedented scale and at an exponential pace, at the forefront of which is new media'.[65] We now live in the age of 'new media' which 'has become an increasingly important source of news; new media platforms are now a vital, and often the preferred, method of imparting and receiving news'.[66] The question then raises: what is media? What is the role and the position of these 'new media' such as social media vis-à-vis legacy media? The so-called 'internet information gatekeepers' often argue that they do not produce any content themselves but merely facilitate the access to content created by their users. Consequently, they consider themselves tech - and not media - companies. This makes them fall under far narrower and less stringent regulation for tech companies and not those applicable to media.[67] However, some argue that considering their main activities, large online gatekeepers should no longer be considered tech companies.[68] In particular, they make certain editorial-like decisions and engage the algorithms that make such decisions automatically (i.e. through selection and prioritisation of content). The concept of 'editorial activities' is a key part in determining the status of 'media'. As explained by Koltay, 'if the activities of gatekeepers are similar to such editorial activities, the gatekeepers themselves may be subject to media regulation; otherwise, they may be considered technology companies'.[69]

However, under the current legal approach, information gatekeepers are not considered as 'media services'. This means that they are not bound by the various legal requirements concerning the right of individuals to access the news and they are not subject to obligations and standards that are otherwise applicable to the media. The question raises whether this should remain a status quo under the legal doctrine, for example in relation to distributing news. As Tambini asks: 'Given that Google, Facebook, Twitter and others were increasingly taking positions previously adopted by journalists and news organisations (selecting and framing news,

[64] See: AI4Media project: D2.1 Overview & Analysis of the AI Policy Initiatives on EU level.
[65] Coe P., "Redefining 'Media' Using a 'Media-as-a-Constitutional-Component' Concept: an Evaluation of the Need for the European Court of Human Rights to Alter Its Understanding of 'Media' within a New Media Landscape" [2017] 37 Legal Studies 25.
[66] ibid.
[67] Koltay A., "New media and freedom of expression: Rethinking the constitutional foundations of the Public Sphere", [2019] Oxford, UK: Hart Publishing.
[68] Napoli Philip M. and Caplan R., "When Media Companies Insist They're Not Media Companies and Why It Matters for Communications Policy", [2016], available at SSRN: https://ssrn.com/abstract=2750148.
[69] Koltay A., op. cit.

for example), should they not be subject to the same checks, constraints and professional ethics as journalism?'[70]

The so-called 'media exemption' in content moderation

The so-called 'media exemption' in the Digital Services Act (amendments 511 and 513 to article 12(1) and recital 38) caused heated discussions around an important issue: once the media content is already subject to editorial responsibility should it be subject to additional scrutiny by online platforms? Some media organisations pointed out that the 'non-interference principle' would ensure that platforms do not undermine the independence of media publishers by arbitrarily deleting legal, public-interest content.[71] Under the current internet intermediary rules, platforms can delete legal (editorial) content on the basis of their terms & conditions. Allegedly, public broadcasters across the EU have experienced removals and restrictions of their accounts or content.[72] Social media platforms have also applied visibility restrictions on media organisations' content. Some argued that since 'broadcast media are already heavily regulated, both at European and national level'[73] and subject to oversight by national regulators, then "there is no justification for imposing on legitimate news publishers a second, parallel system of regulation".[74]

On the other hand, the fact-checkers community and experts argued that "a media exemption would reverse years of progress in the fight against hate speech and disinformation online, preventing very large online platforms from down ranking, deleting, or even labelling any content coming from a press publication — regardless of whether a given post is actively peddling hateful, patently false or otherwise harmful content".[75] It was also unclear which stakeholders would qualify under the so-called media exemption: what constitutes a "press publication" and who are "editorial content providers" lacks clear definitions.

The place of media in the AI Act

The AIA does not provide clarification as to whether some AI journalistic media applications are falling in the scope of one of the risk-based categories. As provided by Helberger et al. (2022), 'under the current proposal most applications of AI in media and journalism would not be subject to the regulation once it has been adopted.'

[70] Tambini D., "Media Freedom", [2021], Polity Press 2021.
[71] EBU, "Position Paper: The Digital Services Act must safeguard freedom of expression online", available at: https://www.ebu.ch/files/live/sites/ebu/files/News/Position_Papers/open/2022/220118-DSA-media-statment-final.pdf
[72] EBU, "Protecting Media Content Online: A Decisive Moment", [2021], available at: https://www.ebu.ch/news/2021/10/protecting-media-content-online-a-decisive-moment
[73] Ibid.
[74] EBU, Position Paper, op.cit.
[75] EU Disinfo lab, "Fact-checkers and experts call on MEPS to reject a media exemption in the DSA" [2021], available at: https://www.disinfo.eu/advocacy/fact-checkers-and-experts-call-on-meps-to-reject-a-media-exemption-in-the-dsa/

In relation to the high risks category, subject to the main corpus of obligations contained in this proposal, here AI media applications are not listed and part of the Annex II, it's far from clear which AI media applications would fall under this risk category and attached obligations. The question of whether AI in the media should be labelled high-risk, is both legal, but maybe more a societal question around the impact of the use of AI in the media on the exercise of fundamental rights, such as freedom of expression and the right to privacy.[76]

Doubts arise whether the AI Act applies to some social media applications. Regarding the prohibited practices category: the notion of 'subliminal techniques' is not defined, which makes the scope of application of this provision far from clear. One may wonder whether and to which extent the online social media practices such as deceptive design patterns (known also as "dark patterns") fall within the scope of this provision. Would a recommender system be considered as materially distorting a person's behaviour or exploit the vulnerabilities of a specific group such as Instagram with teenagers in light of the psychological harm caused to self-image of this specific group of people?

Moreover, deepfakes and chatbots are specifically targeted by the text under the category of limited risk. However, the scope of this provision is also not so clear for the media and social media sector, i.e. does the 'AI systems intended to interact with natural persons' encompass recommender systems or robot journalism? Furthermore, the paragraph targeting deepfakes seems to have a very narrow scope considering the freedom of exceptions assigned to it. The transparency obligation is not applicable in case of personal non-professional activity, in case of law enforcement purposes, in case necessary to safeguard freedom of expression and freedom of the art. Some may wonder what the exact impact of this provision is and how it will improve the transparency about deepfakes in practice.

Initial policy recommendations addressing the conceptual and definitional challenges

In the following, we present the initial policy recommendations addressing the conceptual and definitional challenges discussed in the previous paragraphs.

Table 11: Initial Policy Recommendations Addressing the Conceptual and Definitional Challenges

Initial Policy Recommendations Addressing the Conceptual and Definitional Challenges	
The role of 'media'	• Reflect on the role of media and 'new media' in the digital era, e.g. whether online intermediaries should be considered media and should they benefit from the legal privileges and protections that are enjoyed by the press and broadcasting. • Consider adopting a general European definition of 'media' under

[76] Helberger N., Diakopoulos N., (2022) "The European AI Act and How It Matters for Research into AI in Media and Journalism", Digital Journalism, DOI: 10.1080/21670811.2022.2082505.

	the upcoming European Media Freedom Act.
The so-called 'media exemption' in content moderation	• Reflect on the treatment of traditional media in the online environment, which should be reflected in legal norms. • Consider clarifying the DSA platforms' responsibilities with regard to content moderation towards media organisations, i.e. how platforms should deal with lawful content under the editorial control and legal liability of the publisher (or broadcaster).
The place of media in the AI Act	• Clarify the scope of the AI Act vis-a-vis media sector. • Consider extending the scope of Art. 52(3) of the AI Act to contain an obligation to make the deep fake identification information undeletable in case of transfer or further modification of the material in order not to lose track of the deepfake's original information. The transparency requirements could include more precisions on what should be communicated (the type of information), when (at which stage this should be revealed) and how.

3.4.3.Fundamental rights and societal challenges

AI-driven Manipulation and Propaganda

A functioning democracy relies on open social and political discourse, free from improper influence and manipulation on individuals. While propaganda and manipulation are not new, the data driven economy, as well as the proliferation in use of AI technologies, has enabled private and public actors to exploit the potential of influence and control on citizens for monetary or political gain.[77] While big tech companies, namely online platforms, started out as simple channels of communication, their role in the public sphere had changed over time. They have been playing an increasingly prominent role in shaping political and otherwise behaviour online, with the aid of microtargeting based on profiles, persuasive technologies, exploitation of vulnerabilities, data driven surveillance, rapid dissemination of disinformation, misinformation, and malinformation, and alike.[78] Therefore, contemporary AI-driven manipulation techniques, coupled with abundance of personal and non-personal data, is not only utilised to predict

[77] Council of Europe, 'CAHAI Feasibility Study' (17 December 2020), available at: https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da.

[78] Council of Europe, 'Study on the impact of digital transformation on democracy and good governance' (26 July 2021), CDDG(2021)4 Final, available at: https://rm.coe.int/study-on-the-impact-of-digital-transformation-on-democracy-and-good-go/1680a3b9f9

choices, but also to influence emotions and thoughts, on subliminal and even supraliminal levels.[79] As a result, apart from the risk they pose to democracies and public interest at large, they also pose the risk of weakening the exercise and enjoyment of individual fundamental rights such as the right to mental and physical integrity, the right to human dignity, the right to freedom of expression, the right to receive and impart information, the right to human autonomy and self-determination, the right to free and fair elections, and more conclusively, the right to freedom of thought.[80]

On the political speech level, political advertising is not a uniformly defined term, as shaping behaviours by political campaigns had been an underestimated concern up until recently. As noted by the Council of Europe Committee of Ministers, "the effects of targeted use of constantly expanding volumes of aggregated data on the exercise of human rights in a broader sense, significantly beyond the current notions of personal data protection and privacy, remain understudied and require serious consideration."[81] Thus, the General Data Protection Regulation (GDPR), while arming data subjects with many significant rights, comes short when it comes to targeted political advertising. As a result, recently, new legislative proposals have been brought to the table by policy makers to overcome the specific challenges addressed above that were not covered or adequately covered by former instruments. For instance, the Proposal on the Regulation on the Transparency and Targeting of Political Advertising has been introduced recently by the European Commission. This proposal aims to regulate the fragmented approach regarding political advertising across Member States, without contradicting the provisions of the GDPR and the proposed Digital Services Act that concerns modernising the e-Commerce Directive.

As we mentioned in D2.1, another example of legislative proposals targeting manipulation both on political and private speech level is Art. 5 of the proposed AI Act. However, with this proposal, it is not clear whether the provision which prohibits the use of subliminal techniques could cover some AI systems used in practice such as the recommender systems or systems used for targeted advertising. The requirements imposed on manipulative AI, such as the use of subliminal techniques or the exploitation of a specific vulnerability of a specific group of persons, as well as the requirement of intent, can result in these provisions having a limited scope. More incidental manipulative systems (such as targeted advertising) are therefore not likely to be covered. Though the explanatory memorandum suggests that other existing instruments still cover manipulative or exploitative practices, apart from practices prohibited under Art. 5, it fails to address that none of this legislation explicitly contains provisions on manipulation. As Bublitz

[79] Council of Europe, 'Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes' (2019), available at: https://search.coe.int/cm/pages/result_details.aspx?ObjectId=090000168092dd4b.
[80] Council of Europe, 'CAHAI Feasibility Study', op. cit.
[81] Council of Europe, 'Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes', op. cit.

and Douglas[82] emphasise, AI systems can powerfully influence or weaken control over individuals' thoughts and behaviours, by bypassing or weakening rational control. This includes micro-targeted advertisement, as well as abuse of trust in recommender systems and their influence on decision-making. Thus, these practices should also be deemed to be manipulative, and they must be fairly addressed in the AI Act, instead of simply referring to other legislation. Perhaps, they could be classified as high-risk AI if they substantially influence thought or behaviour in ways that bypass or weaken rational control.

Lastly, regarding the Human Rights Law framework, while some individual fundamental rights provide some level of protection to citizens, the framework still falls behind the protection level required for illegitimate interference to the minds of individuals. For instance, the right to freedom of thought, enshrined in Art. 9 of the European Convention on Human Rights, has not enjoyed much attention from the ECtHR, despite being an absolute right, not allowing any interference or limitations. As noted by Vermeulen[83], this right encapsulates not to be manipulated, neither by private actors, nor by governments. Neither the ECHR, nor the CJEU (Art. 10 of the Charter of Fundamental Rights) defined the scope of such a right, as well as what is considered illegitimate interference to the mind, as mental autonomy and cognitive liberty has been seen as something that was almost impossible to intervene with. In the political speech and campaign perspective, these rights need to be considered first, while also seeking a balancing act with other fundamental, but restrictable rights. Unfortunately, the current human rights law framework, its interpretation, and its practical application fall short in this regard.

AI bias and discrimination against underrepresented or vulnerable groups

In our survey, respondents highlighted AI bias and discrimination as perhaps the major risk stemming from the adoption of AI in the media and other sectors. AI systems often exhibit bias against specific groups of people, including racial bias, gender bias, etc. due, for example, to prejudiced hypotheses made when designing the models or due to problems of diversity and representation in training data. A prominent example for the media sector is bias that may be embedded in large language models. Such models are trained with swaths of Internet data, which are by definition produced in the biggest or richest countries, in languages with higher linguistic footprint, and by communities with large representation, or mainly by men, thus resulting in models that fail to capture changing social norms or the culture of minorities and underrepresented groups and which will eventually discriminate against such groups or produce language that is not attuned to changing social norms. The gigantic volume of data also makes

[82] Bublitz J. C., Douglas T., "Manipulative Influence via AI Systems and the EU Proposal for Regulation of Artificial Intelligence", [2021], available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/F2665640_en.

[83] Vermeulen, B. P., "Freedom of Thought, Conscience and Religion (Article 9)" [2006], In P. van Dijk, F. van Hoof, A. van Rijn, & L. Zwaak (Eds.), Theory and Practice of the European Convention on Human Rights (pp. 751-771). (4th edition).

it hard to audit such models for embedded bias.[84] Another crucial factor is that the access to platform data varies significantly. Research shows that there is an over-emphasis of a single platform, Twitter, as a data source which raises questions of representativeness and validity of research.[85] Beyond language, similar biases exist in other types of models, such as visual classification models that learn how a certain concept is visually represented, based on the samples found in the dataset, which are likely to be collected in North America or Asia.

Research has shown that AI may discriminate against underrepresented or vulnerable groups sometimes with considerable impact and cost on the lives of individuals.[86] Of special interest is bias in relation to visual datasets, which is often much more challenging to define and quantify.[87] A common challenge faced by research institutions and academics in relation to big AI models is that they have no access to data and compute resources of sufficient scale to help them with replication and more in-depth investigations and auditing of big AI models.

Filter-bubbles in recommender systems

Current content recommenders tend to promote filter bubbles, in that they are often optimised for the one evaluation metric: "utility" measures by the immediate short-term clicks / likes provided by users. By doing that, they tend to emphasise a human bias (confirmation bias, i. e. the tendency to prefer information that is consistent with one's existing worldview), filtering information so that it gets ever more adapted to a user's perspective, effectively creating a "bubble" that users are often unaware of. In a similar way, content providers tend to also rely on clicks / likes for evaluation, and optimise content production for that, using tools which can result in easier-to-consume but also more unified and less diverse information.[88] Such challenges could be avoided by using other / more evaluation metrics for recommenders such as Novelty, Diversity, Unexpectedness or Serendipity, up to the point that recommenders can help avoid filter bubbles by confronting users with more diverse content and opinions and encouraging the use of more diverse formats and content for production.[89] However, such possibilities have been underused until now, probably because it is easier to stick with existing patterns and business models. However, the development and application of recommenders that consider and avoid the dangers of focusing only on (immediate) utility and thereby

[84] Hao K., "We read the paper that forced Timnit Gebru out of Google. Here's what it says." (December 4, 2020), MIT Technology Review, available at: https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/.

[85] Tufekci Z., op. cit.

[86] Heikkilä M., "Dutch scandal serves as a warning for Europe over risks of using algorithms", Politico, (March 29, 2022), available at: https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/.

[87] Fabbrizzi S., op. cit.

[88] https://de.wordpress.org/plugins/semrush-seo-writing-assistant/

[89] Mouzhi Ge, Delgado-Battenfeld C., and Jannach D. [2010], "Beyond accuracy: evaluating recommender systems by coverage and serendipity", in Proceedings of the fourth ACM conference on Recommender systems (RecSys '10). Association for Computing Machinery, New York, NY, USA, 257–260. https://doi-org.kuleuven.e-bronnen.be/10.1145/1864708.1864761.

amplifying confirmation bias can not only provide great benefits for users, but they are also crucial to ensure that users / citizens retrieve the broad and diverse set of news and information needed to keep democratic discourse alive.

Transparent communication

Nowadays, most of the AI-based systems developed by big tech companies rely on deep neural networks. These are considered as black boxes, since the interpretation of their outcomes is almost impossible. Deep Learning has proved to perform with high rates of accuracy, but it is hard to know what happens when we obtain a false negative prediction and to identify which part of the process (data, algorithm, interpretation) might have some bias. In addition, to comply with the principles of Trustworthy AI, there should be a clear communication of the capabilities and limitations of the AI-based system but also a clear communication to end-users that are interacting with an AI-based system.

Initial policy recommendations addressing the fundamental rights and societal challenges

In the following, we present the initial policy recommendations addressing the fundamental rights and societal challenges discussed in the previous paragraphs.

Table 12: Initial Policy Recommendations Addressing the Fundamental Rights and Societal Challenges

Initial Policy Recommendations Addressing the Fundamental Rights and Societal Challenges	
AI-driven Manipulation and Propaganda	• Revise the wording of Art. 5 of the AI Act. • Ban political microtargeting, regardless of whether it utilises personal or non-personal data. "Second-hand-manipulation" conducted by non-political figures, which is not obvious on its face, should also be considered in such a prohibition. • Revise the Human Rights Law framework concerning freedom of thought to make the rights application clear and tangible in practice. This could involve drafting a secondary legislation to explain what this right entails and how it can be enforced to create preventive (ex-ante) measures against manipulation, regardless of its frontiers and technology used.
AI bias and discrimination against underrepresented or vulnerable groups	• Ensure that AI systems should come with AI fairness audit reports certificates, ensuring that the system has been extensively tested/audited to minimise the risk of bias. • Design transparent processes to audit AI systems for bias/discrimination. This can be done by independent authorities on EU level.

	• Ensure that AI providers and big platforms provide clear info/public reports on the data used for training AI systems and how they have considered/addressed AI fairness requirements. • Ensure that AI providers and big platforms enable independent research on their services and products to analyse potential impact and risks. • Consider regulating AI-based systems such as machine translation systems (for language translation) or decision systems. • Support easy and fair access of academic/research institutes to large-scale compute and data management infrastructures.
Filter-bubbles in recommender systems	• Incentivize / fund the development, use and evaluation of recommenders tailored to avoid filter-bubble problems.
Transparent communication	• Promote research on interpretable and explainable AI. • Research on transfer learning discusses the importance of using other metrics (footprint, human cost…) to assess models, beyond the task performance metrics.

3.4.4.Intellectual Property Challenges

Copyright challenges of AI use in media

As is the case with every emerging technology and advancements, AI use in media created many ambiguities in several folds. However, even if the technology is novel, the question of how the law should respond to technological change is not at all new. This stems from the fact that copyright laws should be drafted independent of any technology and continue to apply equal across technologies as they emerge, without favoring or discriminating between new and old. In other words, copyright laws should be based on the notion of technological neutrality. [90]

[90] Carys C., "The AI-Copyright Challenge: Tech-Neutrality, Authorship, and the Public Interest", (December 14, 2021). Ryan Abbott (ed.) Research Handbook on Intellectual Property and Artificial Intelligence (Edward Elgar Press, 2022 Forthcoming), Osgoode Legal Studies Research Paper, available at SSRN: https://ssrn.com/abstract=4014811.

Despite the notion of neutrality, AI use in the media industry still triggered (and has been triggering) some of the following questions:

- "Would extending copyright to AI generated works reflect an appropriate reward for the original creative efforts and investments of the people responsible for AI?
- Would it encourage the kind of authorial creative practices that advances the progress at which copyright is aimed?"[91]
- Would the usage of copyright protected works as training data be deemed infringement?

To adequately answer these questions, one should first investigate the purpose of having copyright laws in the first place. As the Canadian scholar Myra Tawfik explains, copyright law has long been an "integrated system that encouraged creators to generate knowledge, industry to disseminate it and users to acquire it and, hopefully, reshape it into new knowledge."[92] Therefore, "copyright is a system of state-granted entitlements to encourage creative expression and learning; authors' or creators' rights must therefore be balanced with users' rights and the public interest to support a flourishing public domain."[93]

Now, AI-generated works are not copyrightable in most jurisdictions in the world; thus, they are in the public domain. However, the current framework does not stop the debate on the authorship and copyright status of such AI-generated works. This includes the EU, the US, Australia, and Canada, as the originality requirement of copyrightability, is based on the notion that works of authorship must be originating from a human author, who possesses some intellectual involvement in the resulting expression.[94] The World Intellectual Property Organization (WIPO) also seems to agree with this notion with the following statement, "if AI-generated works were excluded from eligibility for copyright protection, the copyright system would be seen as an instrument for encouraging and favouring the dignity of human creativity over machine creativity. If copyright protection were accorded to AI-generated works, the copyright system would tend to be seen as an instrument favouring the availability for the consumer of the largest number of creative works and of placing an equal value on human and machine creativity."[95] In line with WIPO's point of view, it is important to note that the robots are not capable of expressive agency, emotion, or intentionality which makes the work one of protected intellectual expression. Despite the current romanticist trend of associating robot

[91] ibid.
[92] Tawfik, M., "History in the Balance: Copyright and Access to Knowledge. Radical Extremism' to 'Balanced Copyright'" [2010], in Canadian Copyright and the Digital Agenda, available at: https://scholar.uwindsor.ca/lawpub/17.
[93] Carys C., "Copyright, Communication and Culture: Towards a Relational Theory of Copyright Law" [2011]. Books. 53, available at: https://digitalcommons.osgoode.yorku.ca/faculty_books/53.
[94] Carys C., op. cit.
[95] WIPO, 'Conversation on Intellectual Property (IP) and Artificial Intelligence (AI) – Revised issues paper on Intellectual Property and Artificial Intelligence', 21 May 2020, https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=499504

behaviour with anthropomorphic traits, they lack the communicative intentionality and qualities they simply cannot possess by their nature.[96] Hence, as Ryan Calo stated, "the box is 'gorged on data but with no taste for meaning'"[97] Consequently, if the purpose of copyright is to encourage creative practices and dissemination of works of intellectual expression, granting copyright protection to AI-generated-works would not be able to advance such a purpose. On the contrary, leaving such AI-generated works in the public domain would allow the public domain "to flourish as others are able to produce new works by building on the ideas and information contained."[98]

The second crucial debate concerning copyright challenges of AI use in media is whether the usage of copyrighted works for AI training purposes would be infringing. In such a controversial debate, once again, one should look into the purpose and protections/exclusive rights afforded by copyright laws. Within its general framework, copyright consists of the exclusive right to make copies, also known as the right of reproduction. However, during the training process, a massive amount of data needs to be digitally reproduced to train a sophisticated AI. Therefore, creating a hurdle of requiring AI researchers and developers to ask permission from the right-holders for each and any single work would be a considerably huge disincentive, on top of resulting in forcing AI researchers and developers to turn into low-quality datasets to offset such a challenge. Thus, the copyright owner's reproduction right should not be deemed infringed by non-expressive copies, as the main reason for granting such an exclusive right is the human appreciation of the expressive qualities of work in the first place.[99]

Initial policy recommendations addressing the intellectual property challenges

In the following, we present the initial policy recommendations addressing the Intellectual Property challenges discussed in the previous paragraphs.

Table 13: Initial Policy Recommendations Addressing the Intellectual Property Challenges

Initial Policy Recommendations Addressing the Intellectual Property Challenges	
Copyright challenges of	• Ensure that AI systems are not granted authorship or any similar status for AI-generated-works, nor should they be considered copyrightable, to be able to encourage human creativity, flourish and expand the public domain, and incentivize AI research and

[96] C. J., Craig, 'The AI-Copyright Challenge: Tech-Neutrality, Authorship, and the Public Interest', (2021) in Ryan Abbott (ed), Research Handbook on Intellectual Property and Artificial Intelligence (Edward Elgar Press, forthcoming 2022), available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4014811

[97] R. Calo et all, 'Telling Stories on Culturally Responsive Artificial Intelligence', (2020), University of Washington Tech Policy Lab, accessible at http://techpolicylab.uw.edu/wp-content/uploads/2022/04/Telling_Stories_Pages_4-4-22.pdf

[98] CCH Canadian Ltd v Law Society of Upper Canada, [2004] 1 SCR 339, 2004 SCC 13 https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/2125/index.do

[99] M. Sag, 'The New Legal Landscape for Text Mining and Machine Learning', (2019), in Journal of the Copyright Society of the USA, Vol. 66 p.291, accessible at : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3331606

AI use in media	development. • Avoid creating a separate type of protection such as sui generis rights for AI-generated works. • Ensure that the usage of copyright protected works for AI training purposes is not deemed infringing, in order to incentivize research and innovation.

4. European Digital Media Code of Conduct

4.1. The feasibility of the European Digital Media Code of Conduct

4.1.1.Introduction

The present section focuses on the Code of Conduct on AI and the media sector. From surveys, discussions and events held as part of WP2, a remark about the high level and general characteristics of the available ethical instruments (codes, guidelines) emerged. The currently available instruments do not seem to meet the specific needs of the media sector. In this section, we will therefore not focus on overarching policy or ethics initiatives (more information can be consulted in D2.1 for this purpose) but we will have a closer look at the practices of the media sector. Are there any codes of conduct already available focusing on the use of AI in the media context? Are there sectoral, internal company codes or codes developed by research institutions, public entities or NGOs which would tackle the ethical use of AI in media?

After having mapped and identified such codes, we will have a look at their common and distinctive features. What are the lessons we can draw from this analysis? The main question however, is: is there a need for a new Code? If yes, in what format? Does the media sector need such an instrument, and is it willing to implement and use it in its activities? Or, on the contrary, is such a code redundant given the heterogeneity of the 'media' and rather, a specific, tailor-made guidance is needed? Would a unique code for the entire media sector meet the needs of the sector and the staff when it comes to AI? Or is this solution utopic and therefore useless? Should we only focus on the journalistic code of ethics while the media means so much more than the news sector?

We will attempt to answer these questions in this feasibility study for a European Digital Media Code of Conduct.

4.1.2.AI4Media Survey Analysis in relation to ethics and codes of conduct

As part of the Deliverable 2.3 "AI technologies and applications in media: State of Play, Foresight, and Research Directions", AI4Media partners designed two surveys previously presented in the Introduction of this deliverable. These surveys provided key insights in relation to the needs of the sector and highlighted some interesting aspects in relation to codes of conduct and legislation. Indeed, the surveys showed for which areas of AI research and application respondents needed more guidance (see Section 4 of D2.3 and more precisely Section 4.1.8 p. 108 of D2.3). In addition, the survey pointed out that the majority of respondents requested not just guidelines but clear-cut regulations for most of the main issues pointed out.

Another aspect tackled by the survey was the use of ethical tools to mitigate risks and challenges arising from the use of AI in the media sector.

The survey results demonstrated that *"19% of the responders from the AI research community and 17% of media professionals do not even know whether their organisation has any processes or measures related to ethics management. In addition, 17% of AI researchers and 25% of media*

professionals state that they have no measures in place. ***These percentages show that lack of awareness or interest in ethics management*** *(not only related to AI but in general) is a real problem both in the AI research community and also in the media industry. Among those who responded positively in the question whether they have any ethics related measures or processes in place the most common answer is ethical AI principles, followed by ethical board committees for AI researchers (mainly from academia and research centres) and ethical AI checklists for the media community.* ***Only 13% of both communities are following or are members of a Code on AI ethics.*** *Finally, ethics by design processes are followed by 23 % of the AI community and 19% of the media community. It is clear that much more work is required in order to raise awareness about the importance of ethics in both communities, especially in light of the AI breakthroughs that we expect to see in the next couple of decades."* (see Figure 3).[100]

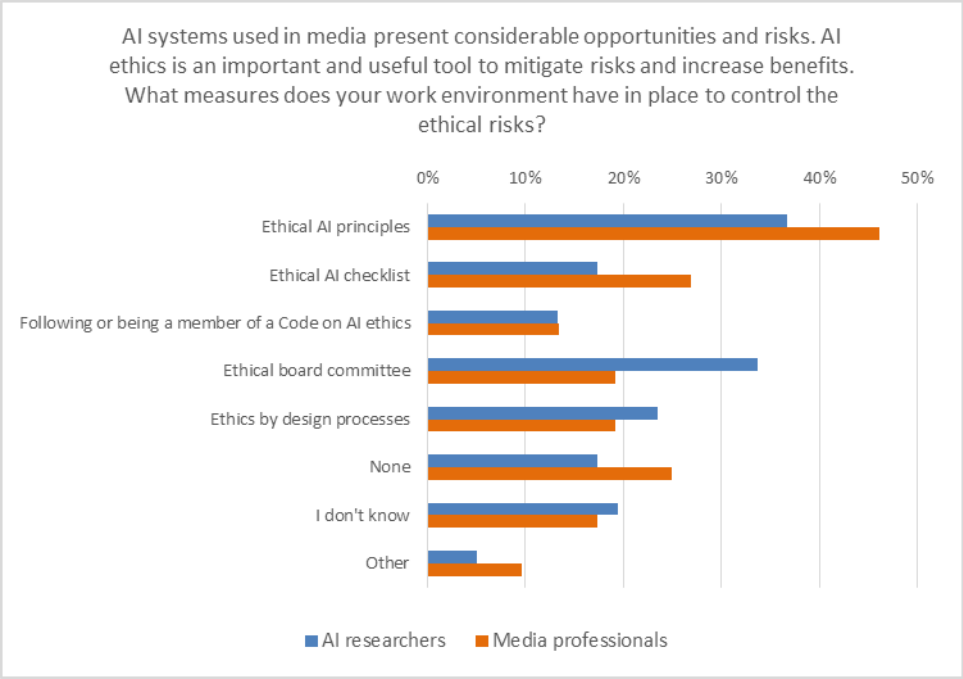


Figure 3: Survey results on a question related to workplace measures to control the ethical risks of AI (check D2.3 for more information).

4.1.3.Overview and analysis of existing Code of conduct on AI and media

This section maps the existing and forthcoming codes of conduct/guidelines around the use of AI in the media sector. We focused on three sub-sectors which correspond to the use cases of the AI4Media project. That is: press and journalists, broadcasting and public service media, video-games. The aim of this section is not to provide an extensive, in-depth analysis of all the existing codes. Instead, it focuses on identifying trends, directions and gaps.

[100] AI4Media project, D2.3"AI technologies and applications in media: State of Play, Foresight, and Research Directions", (2022), p.110 (available at https://www.ai4media.eu/reports/roadmap-ai-technologies-and-applications-in-media/).

4.1.3.1. Press and Journalists

One well-established sector within the media industry is the information sector where journalists and media companies have been operating for several decades. Even though their methods, tools and communication means have evolved, journalism is a long-standing profession, which benefits from a large corpus of self-regulation. These texts contain not only the fundamental principles which an accredited journalist or an alleged journalist must follow to be considered a journalist. They also contain guidance as to how to behave in a certain situation and how to use information sources and journalistic tools. For the public, adherence to such self-regulatory ethics codes constitutes also a guarantee of quality information and news following high standards. The report of Natali Helberger et al. for the Council of Europe demonstrated how AI-driven tools play an increasingly important role in the media in various forms (smart tools, robot journalism, audience analytics, content recommendation). A UNESCO report also points out the opportunities that AI can offer to improve journalism while also investigating relevant threats.[101] The section below will analyse whether the sector's initiatives were adapted to include AI use and provide guidelines on how to use it ethically.

International initiatives

The International Federation of Journalists (IFJ) is "the world's largest organisation of journalists, representing 600,000 media professionals from 187 trade unions and associations in more than 140 countries."[102] In 2019, IFJ adopted the Global Charter of Ethics for Journalists which completes the IFJ Declaration of Principles on the Conduct of Journalists dating back in 1954.[103] Some of the main principles of the Declaration are applicable in an AI context such as searching for the truth, safeguarding media freedom and independence but also the need to minimise harm. Despite being written recently, the Charter does not mention digital challenges nor the use of artificial intelligence in media. The text sticks to overarching principles which should govern the journalistic work.

The Ethical Journalism Network is a coalition of more than 70 groups of journalists, editors, press owners and media support groups from across the globe.[104] Their 5 core principles of ethical journalism could be applicable to an AI media context but stay high level. They include: truth and accuracy, independence, fairness and impartiality, humanity and accountability.[105]

[101] N. Helberger, S. Eskens, M. van Drunen, M. Bastian, J. Moeller for the Council of Europe, 'Implications of AI-driven tools in the media for freedom of expression', (2019), p.5, https://rm.coe.int/coe-ai-report-final/168094ce8f

[102] International Federation of Journalists (IFJ), 'About', (last accessed on 4th August 2022), https://www.ifj.org/who/about-ifj.html

[103] International Federation of Journalists (IFJ), 'IFJ Global Charter of Ethics for journalists', (2019) https://www.ifj.org/who/rules-and-policy/global-charter-of-ethics-for-journalists.html

[104] Ethical Journalism Network, 'About the EJN', (last accessed on 4th August 2022), https://ethicaljournalismnetwork.org/who-we-are

[105] Ethical Journalism Network, 'About the EJN', (last accessed on 4th August 2022), https://ethicaljournalismnetwork.org/who-we-are

European initiatives

The European Federation of Journalists (EFJ) is the largest organisation of journalists in Europe, representing over 320,000 journalists in 73 journalists' organisations across 45 countries.[106] It follows the IFJ Declaration of Principles on the Conduct of Journalists and is part of the ethical journalism network.[107] Even if they do not have a self-regulatory instrument including specifically AI, it is worth mentioning that the EFJ President has participated in 2021 in the Online Conference "Artificial intelligence and the future of journalism: will artificial intelligence take hold of the fourth estate?" organised by the Portuguese Presidency of the Council of the European Union. On this occasion, the EFJ President underlined pressing challenges in relation to the use of AI in newsrooms and how to promote its ethical use. These points include:

- Addressing the gap between big and small media when implementing AI; there is a risk to leave the small-scale media companies behind given the imbalance of financial and human resources to keep up with tech giants. Addressing this risk is of crucial importance as ethics and journalism's role should prevail over the algorithms.
- Data Literacy training (improving the wide accessibility of these trainings, especially thinking about including freelance journalists to have access to these trainings).
- Ethical considerations in relation to the use of AI and algorithms such as bias.

The European Centre for Press and Media Freedom (ECPMF) is a non-profit organisation that promotes, protects and defends the right to a free media and freedom of expression throughout Europe. The Centre adopted in 2009 a European Charter on Freedom of the Press. The charter focuses on principles for the freedom of the press to stay away from government interference and does not include any specific reference to AI or digital technologies.[108]

National initiatives

Some research institutes such as the Reynolds Journalism Institute[109] and Media Wise[110] focus and study the initiatives happening worldwide. Media Wise has assembled a large collection of journalistic Codes of Conduct from around the world.

Several initiatives map Codes of journalism ethics in Europe. We can mention Ethicnet Journalism ethics from Tampere University[111] and Media Council in the Digital Age.

---

[106] European Federation of Journalists (EFJ), 'About EFJ', (last accessed on 4th August 2022), https://europeanjournalists.org/about-efj/

[107] European Federation of Journalists (EFJ), 'Ethics', (last accessed on 4th August 2022), https://europeanjournalists.org/policy/ethics/

[108] European Centre for Press & Media Freedom (ECPMF), 'Governance', (last accessed on 4th August 2022), https://www.ecpmf.eu/about/governance/

[109] Donald W. Reynold Journalism Institute, (last accessed on 4th August 2022), https://rjionline.org/

[110] Media Wise website, (last accessed on 4th August 2022), http://www.mediawise.org.uk/

[111] Tampere University, 'Ethicnet, Journalism Ethics', (last accessed on 4th August 2022), https://research.tuni.fi/ethicnet/

a) Press and Media Council Codes

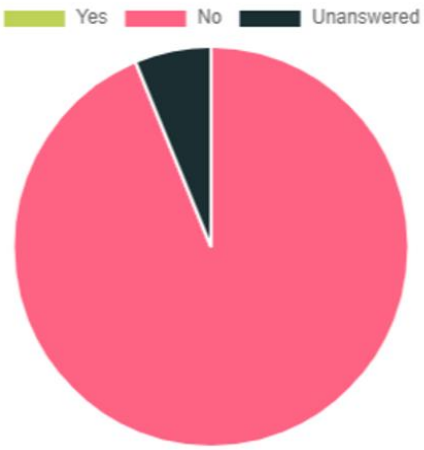
Press or Media Councils are "*self-regulatory bodies set up by the media themselves – although they are normally given a high degree of operational independence in order to maintain public confidence. They oversee Codes of Practice which set out both professional standards for journalists, and a set of rules under which people featured in the news media can complain if something is inaccurate or intrusive. (...) Press Councils represent a form of corporate responsibility which allows people to complain for free and without legal representation, and can help generate trust in the quality of news.*" [112]

The "Media Councils in the digital age" project financed by UNESCO and the European Commission is extremely active when it comes to research and survey among the journalistic community. It has developed a database which analyses different themes in relation to national press councils. As a part of this database, codes of conduct tailored to the industry were one of the abovementioned main research themes. In this database, 55 codes of ethics from 49 countries were analysed. It is important to note that the database only focused on the codes themselves and not their additional guidance. Consequently, the following questions[113] prepared by the project are extremely relevant for the purposes of this Deliverable as well (Figure 4, Figure 5):

[112] Definition from the Alliance of Independent Press Councils of Europe which can be found here: https://presscouncils.eu/about

[113] Alliance of Independent Press Councils of Europe (AIPCE) & Media Council in the Digital Age project, 'Code of Ethics', (last accessed 4th August 2022), https://presscouncils.eu/Code-of-Ethics

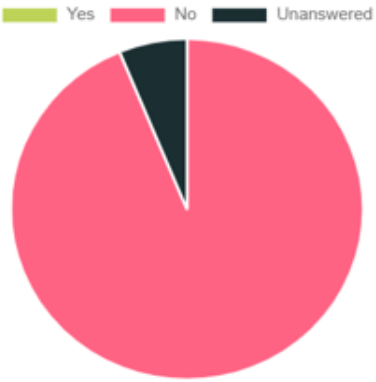
Has the Code of ethics been amended to take responsibility for data journalism and use of algorithms into account?



	Yes	No	Unanswered
%	0	94	6
#	0	30	2
Countries		Albania Austria Belgium - Dutch Belgium - French & German Bosnia and Herzegovina Bulgaria Canada - Québec Croatia Cyprus Denmark Estonia Finland Georgia Germany Hungary Iceland Ireland Kosovo Luxembourg Netherlands North Macedonia Norway Serbia Slovakia Slovenia Spain - Catalonia Sweden Switzerland Turkey UK Impress	Canada France

Figure 4: The survey results concerning data journalisms and the use of algorithms in the Codes of ethics

Has the code of ethics been amended to take responsibility for use of AI or Robot in Journalism into account?



	Yes	No	Unanswered
%	0	94	6
#	0	30	2
Countries		Albania Austria Belgium - Dutch Belgium - French & German Bosnia and Herzegovina Bulgaria Canada - Québec Croatia Cyprus Denmark Estonia Finland Georgia Germany Hungary Iceland Ireland Kosovo Luxembourg Netherlands North Macedonia Norway Serbia Slovakia Slovenia Spain - Catalonia Sweden Switzerland Turkey UK Impress	Canada France

Figure 5: The survey results concerning the use of AI or Robots in journalisms in the Codes of ethics

To go deeper into the impact of technology and AI on journalists and the press, the Media Councils in the Digital Age project conducted a Europe-wide study on the ethical implications of

news automation for the work of media councils. The study showed that even if news automation is still mainly at an experimental level, media councils should stay alert to the possible need for self-regulation and critically revisit their complaints procedure "so that audiences have a genuine opportunity to bring up their grievances with automation". [114]

Since then, the Digital Age project continued research on this topic and delivered a new study on Digital challenges to ethical standards of journalism.[115] The study reported that "several Councils have taken steps to cover AI in their code. As of February 2022, at least the Belgian Raad, the German Presserat and the Impress in the UK had concrete plans to add to their code a clause or guideline on the use of AI in journalism, and the Catalonian council had just published supplementary guidance on the issue".[116] This was discussed during interviews with several countries' media councils. [117]

- Current initiatives on AI and media by local Media Councils
  - The UK

Impress, the UK media council, is currently undertaking a comprehensive review of its Standards Code and Guidance.[118] In the revision plans, a subclause 10.5 should be added to the Code where the following information should be contained[119]:

- **Editorial Responsibility.** Publishers must take responsibility for the use of AI;
- **Human oversight.** Publishers should exercise human editorial oversight (covering the use of AI in generating, publishing and disseminating news);

[114] L. Haapanen, 'Media Councils and Self-Regulation in the Emerging Era of News Automation', (2020), Published by the Council for Mass Media in Finland, https://presscouncils.eu/userfiles/files/Emerging%20era%20of%20news%20automation%20Haapanen.pdf

[115] L. Juntunen for the Council for Mass Media in Finland, 'Digital Challenges to ethical standards of journalism. Responses and needs of European Media Councils', (2022), https://presscouncils.eu/New-report-explores-digital-challenges-to-ethical-standards-of-journalism

[116] L. Juntunen for the Council for Mass Media in Finland, 'Digital Challenges to ethical standards of journalism. Responses and needs of European Media Councils', (2022), https://presscouncils.eu/New-report-explores-digital-challenges-to-ethical-standards-of-journalism

[117] The study from L. Juntunen is based on 16 semi-structured qualitative interviews with representatives of different European media councils. The interviews were conducted between July and October 2021. Where necessary, parts of the information were later updated to reflect the situation at the end of February 2022. It covered the following countries: (both Flanders and Wallonia in Belgium, Bulgaria, Estonia, Germany, Ireland, the Netherlands, Slovakia, Slovenia, Media Councils' views on revising the Codes of Ethics 10 Spain (Catalonia) and Sweden]; (2) EU candidate countries (Republic of North Macedonia, Montenegro and Turkey); and (3) non-EU member countries (Norway and the United Kingdom).

[118] Impress, 'Code Review 2020-2022', (last accessed on 4th August 2022), https://www.impress.press/standards/code-review.html

[119] L. Juntunen for the Council for Mass Media in Finland, 'Digital Challenges to ethical standards of journalism. Responses and needs of European Media Councils', (2022), https://presscouncils.eu/New-report-explores-digital-challenges-to-ethical-standards-of-journalism

- **Transparency.** Publishers should ensure that prominent and clear labels make it explicitly clear that AI has been used (also label for all content that has been generated wholly or partly by AI;
- **Redress.** Ensure that the public has the same ability to complain about AI-generated content as they do with human-generated content.

The revised code was supposed to come during summer 2022, but at the time of finalising this deliverable, the revision was not yet out.

- Belgium

Belgium has a complex media landscape given the presence of three national languages and different cultures associated with these linguistic communities. Hence, there are different press/media councils depending on the language spoken.

Dutch speaking Media Council/Flanders: the Raad voor de Journalistiek[120]

The Dutch Belgian Raad (Council) is preparing two guidelines on AI to add to their existing code. This set of guidelines should be approved in fall 2022. Following the interview with these stakeholders, L. Juntunen reported that according to the draft text the following addition will be brought:[121]

- **Editorial Responsibility:** *"The editorial choices regarding AI must comply with the principles of the Council and that the final responsibility for the produced content rests with the editor-in-chief at all times."* Hence *"the editor-in-chief must ensure that the principles of the ethical code are met in the development of AI-driven systems as well as to monitor the application and implementation of these principles."*
- **Transparency***: "Editors should communicate transparently about automated news production and personalisation of content, so that it would be clear to the public when content has been created or selected by AI. In particular, the editors would be expected to indicate when "the essence of a news item" has been produced on the basis of automated processes and to refer as far as possible to the sources on which the item is based."*

French speaking Media Council: le conseil de déontologie journalistique

In Belgium, le conseil de déontologie journalistique (CDJ) had studied the issue of AI among its members *"but issuing guidance on AI was not yet considered necessary. This is mainly because AI is still hardly used by the media in Wallonia but also because the current provisions of CDJ's*

[120] Raad voor Journalistiek, 'About us', (last accessed on 4th August 2022), https://www.rvdj.be/pagina/english-version

[121] L. Juntunen for the Council for Mass Media in Finland, 'Digital Challenges to ethical standards of journalism. Responses and needs of European Media Councils', (2022), https://presscouncils.eu/New-report-explores-digital-challenges-to-ethical-standards-of-journalism

code are seen as providing sufficient tools for addressing the ethical issues related to news automation."[122]

- Spain, Catalonia

The Catalan Press Council commissioned a study on AI. Following this study, recommendations and requirements were then drafted: [123]

- Data strategy
  - Safeguarding the source and diversity of data
  - Ensuring a constant monitoring of data representativeness
  - Ensuring the technical quality of data processing to minimise risks and mitigate errors
  - Responsible management of data and privacy
- Transparency
  - Making users aware of the existence of algorithms and the basic features of their operation, as much as possible
- Impact assessment and mitigation
  - Preventing the use of tailor-made algorithms to undermine pluralism or cause damage to vulnerable communities
- Human oversight
  - Enhancing the human factor
- Interdisciplinary teams
  - Training and promotion of interdisciplinary teams
- Research for innovation in line with values
  - Promotion of research addressed at exploring convergence between technical efficiency of systems and the values of ethical journalism

- Other respondents

The other respondents pointed out that it was not yet the time for them to create new guidelines on the topic of AI and media. They underlined that the journalistic use of algorithms in their country is not so common and stays at an experimental stage. "Especially in small market areas,

[122] L. Juntunen for the Council for Mass Media in Finland, 'Digital Challenges to ethical standards of journalism. Responses and needs of European Media Councils', (2022), https://presscouncils.eu/New-report-explores-digital-challenges-to-ethical-standards-of-journalism

[123] P. Ventura, 'Algorithms in the Newsrooms. Challenges and Recommendations for Artificial Intelligence with the Ethical Values of Journalism.', (2021), Published by the Catalan Press Council, https://fcic.periodistes.cat/wp-content/uploads/2022/03/venglishDIGITAL_ALGORITMES-A-LES-REDACCIONS_ENG-1.pdf

the opportunities for the media to invest in new technological innovations were said to be scarce."[124]

In addition, several media councils also made the point that it is really unlikely that they would "ever receive complaints about news automation or personalisation even if they were covered by their code".[125]

- Interim Conclusion

The analysis of the different features of the codes addressing media shows the following similarities between them (Table 14). Indeed, while these initiatives are still ongoing or fresh, they share some similar concerns and try to address the use of AI by press with similar strategies and provisions. For all of them, transparency towards the audience about the use of AI is key. Catalonia and the UK have also brought on board the human component while UK and Flanders have focused on clarifying the editorial responsibility for the use of AI. Catalonia has also foreseen several additional components to make sure the use of AI is well framed. This includes, redress mechanisms, data strategy (diversity, quality, and alike), inclusive and diverse team interacting with AI and so forth. Their approach is quite thorough as it addresses explicitly several important aspects of AI use and development.

Table 14: Common principles of the Media Councils codes of ethics

Common features	**UK**	**Belgium, Flanders**	**Spain, Catalonia**
Editorial Responsibility	X	X	
Human Oversight	X		X
Transparency	X	X	X
Redress	X		
Data strategy			X
Impact assessment and mitigation			X

[124] L. Juntunen for the Council for Mass Media in Finland, 'Digital Challenges to ethical standards of journalism. Responses and needs of European Media Councils', (2022), https://presscouncils.eu/New-report-explores-digital-challenges-to-ethical-standards-of-journalism

[125] L. Juntunen for the Council for Mass Media in Finland, 'Digital Challenges to ethical standards of journalism. Responses and needs of European Media Councils', (2022), https://presscouncils.eu/New-report-explores-digital-challenges-to-ethical-standards-of-journalism

Interdisciplinarity			X
Research			X

- Implementation

Another important parameter for the analysis of the efficiency of these codes of conduct is the implementation aspect. The Blanquerna School of Communication and International Relations, Ramon Llull University, has conducted a survey among European journalists and media/press council members during August-November 2020.[126] Journalists are from: Austria, Belgium (Wallonia and Flanders), Estonia, Finland, Hungary, Ireland, Italy, Poland, and Spain.

The following lessons are interesting for our feasibility study and the implementation aspect of Codes of conduct.

The survey showed that most of the respondents have a medium or high knowledge of the Code of Ethics. Especially young journalists are the ones who have the best level of knowledge regarding the existence of the codes of ethics however and surprisingly this group is the one with the least knowledge of the rules (Figure 6).

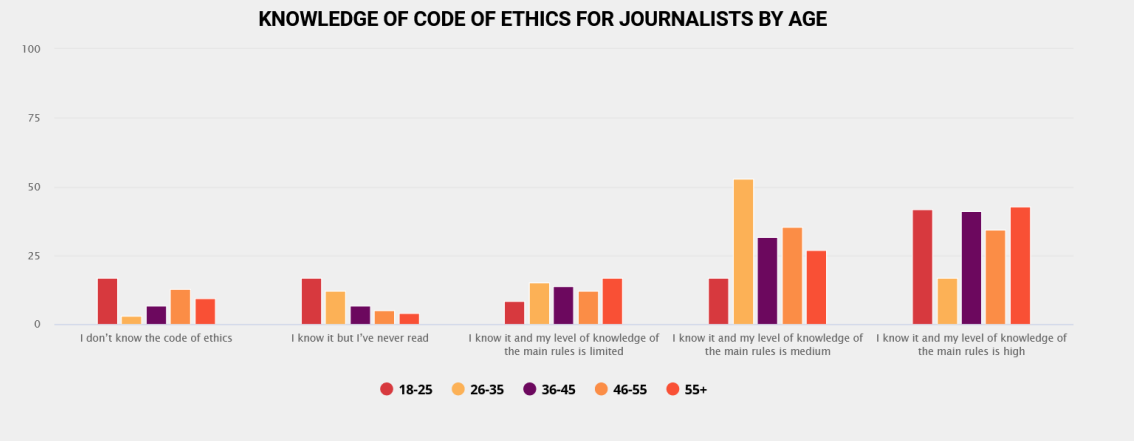


Figure 6: Image about Knowledge of the Code of Ethics coming from the Blanquerna's survey, https://presscouncils.eu/Survey-self-regulation-bodies-challenges-of-digital-age

[126] Alliance of Independent Press Councils of Europe (AIPCE) and Medi Council in the Digital Age project, 'Survey about self-regulation bodies and challenges of digital age', (2020), (last accessed on 4th August 2022), https://presscouncils.eu/Survey-of-professional-journalists-about-self-regulation-bodies-and-challenges-of-digital-age

The survey then showed that only 1 out of 3 respondents thinks that the codes of ethics are actually adapted to answer to the new ethical challenges created by digitalisation and the internet (Figure 7).

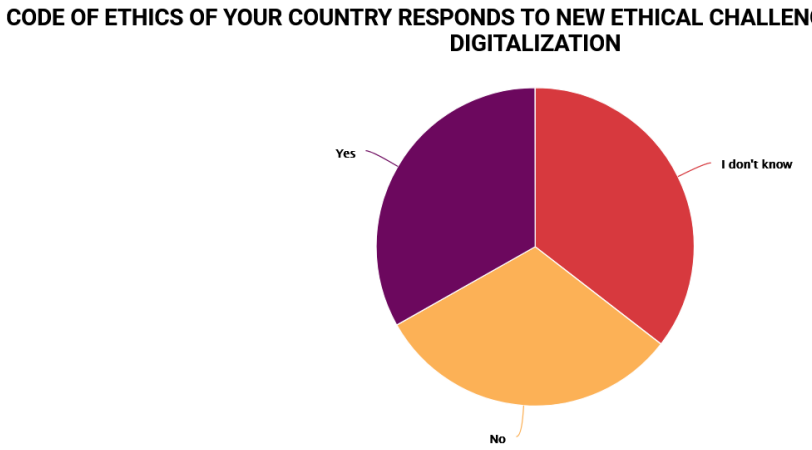


Figure 7: Image about code of ethics and the digital in Blanquerna's survey https://presscouncils.eu/Survey-self-regulation-bodies-challenges-of-digital-age

The survey also pointed out that most of the respondents think that digitalisation does require the creation of new ethical principles (Figure 8).

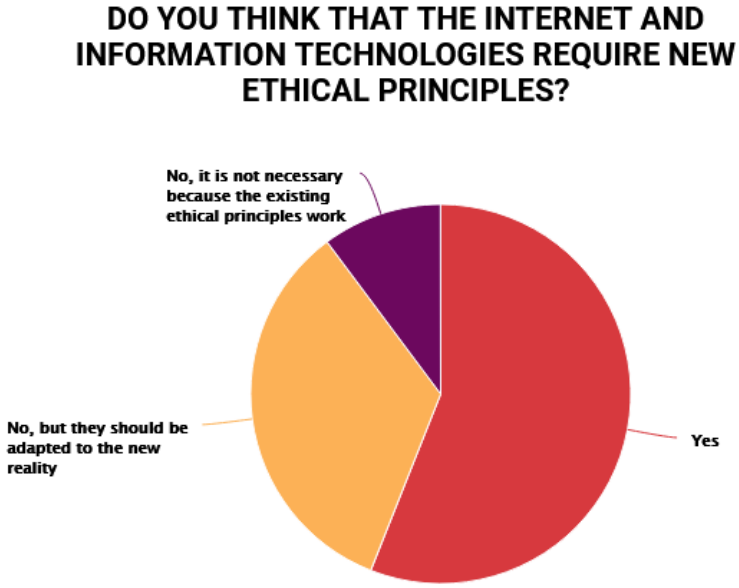


Figure 8: Image about ethics principles and the digital in Blanquerna's survey https://presscouncils.eu/Survey-self-regulation-bodies-challenges-of-digital-age

More specifically, focusing on AI and media, the survey showed that 96,9% of the respondents think that the information produced by robots (such as automated journalism tools) should be identified and labelled as such. Furthermore, 91,1% of the respondents expressed the opinion that content produced by robots should be governed by the same ethical standards applicable to journalists (Figure 9).

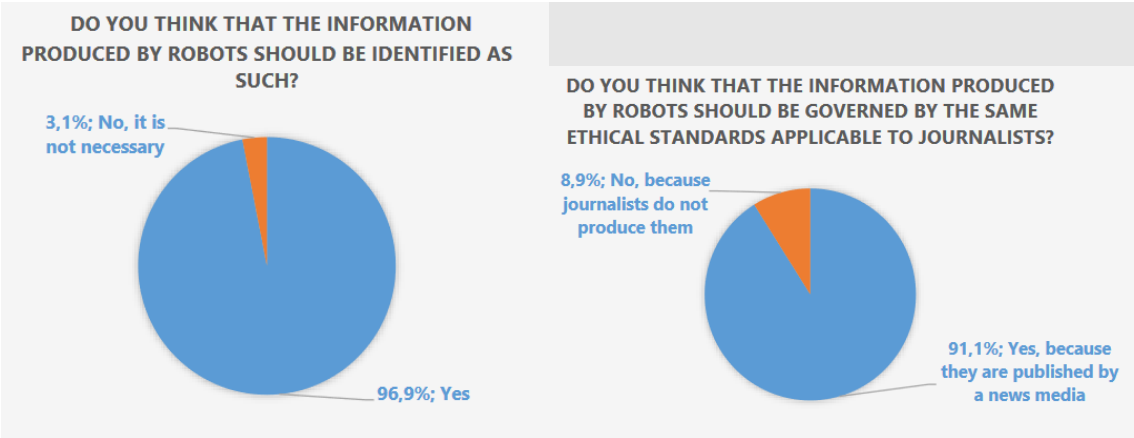


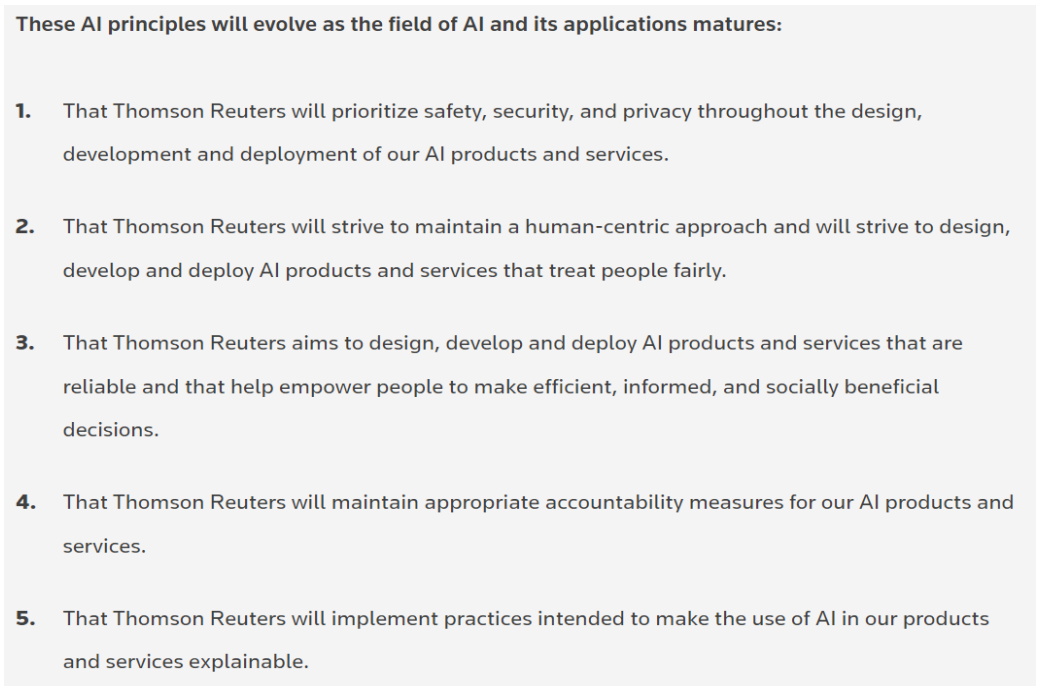
Figure 9: Image about Ethics principles and the robot journalism in Blanquerna's survey https://presscouncils.eu/Survey-self-regulation-bodies-challenges-of-digital-age

This survey showed that the knowledge of the Code's principles and therefore application of those is way lower compared to the knowledge about the existence of Codes of conduct. Several efforts need to be conducted to improve the situation. Firstly, more awareness campaigns about the codes to disseminate them further are needed. Secondly, more efforts on educating about the codes are also imperative. If code principles are not well understood or adapted to today's challenges for journalists, they will not be enforced. Furthermore, the results showed that there is a real need by the professionals to set principles considering the current digital challenges. There is also almost unanimity in the sector in relation to information produced by robots, which is an easy starting point for a press council to integrate in codes of conduct. A revision and an update of the existing codes of conduct appear necessary and would meet the request of the professionals.

b) Private initiatives

Some private initiatives on the use of AI by the media sector have also emerged. This is the case for Thomson Reuters, a Canadian multinational media conglomerate, which has drafted the following AI principles to promote trustworthiness in design, development, and deployment of AI (Figure 10).[127]

[127] Thomson Reuters, 'Our AI Principles', (last accessed on 8th August 2022), https://www.thomsonreuters.com/en/artificial-intelligence/ai-principles.html



These AI principles will evolve as the field of AI and its applications matures:

1. That Thomson Reuters will prioritize safety, security, and privacy throughout the design, development and deployment of our AI products and services.
2. That Thomson Reuters will strive to maintain a human-centric approach and will strive to design, develop and deploy AI products and services that treat people fairly.
3. That Thomson Reuters aims to design, develop and deploy AI products and services that are reliable and that help empower people to make efficient, informed, and socially beneficial decisions.
4. That Thomson Reuters will maintain appropriate accountability measures for our AI products and services.
5. That Thomson Reuters will implement practices intended to make the use of AI in our products and services explainable.

Figure 10: Thomson Reuters AI Ethics principles

4.1.3.2. Broadcasting and public service media (PSM)

Some examples of codes of conducts/AI guidelines can be found in the broadcasting and audiovisual media sector.

Declaration of Intent for responsible use of AI in the media

For example, in the Netherlands a letter of intent for responsible use of AI in media was drafted by media organizations (RTL, Sound & Vision and NPO) which brought in their topical knowledge Participating media companies include, among others: the Netherlands Institute for Sound & Vision, Media Perspectives, NPO, RTL, Omroep Gelderland/Regio Groei, FD Mediagroep. It's worth mentioning that declaration was an outcome of close collaboration by media partners which work closely together in an ecosystem. This translates into an acceptance of the agreed rules within the participating stakeholders.

The statement is voluntary and does not create any legally binding framework.

The document recognizes both positive and potentially negative impacts of the use of AI in media organisations, and therefore advocates for the responsible use of AI in the media industry. The declaration endorses the AI HLEG Guidelines and translates the elements of these guidelines into practice for participants from the media sector.

Chapter I provides that the signatories subscribe to the ethical principles for trustworthy AI: respect for human autonomy, prevention of harm, justice and accountability. In Chapter II, the

signatories intend to comply with the AI HLEG seven requirements for trustworthy AI, acknowledge the tensions between these requirements and commit to develop the algorithms, training data, methods and models as transparently as possible for the end user. Chapter III considers the importance of controlling that the trustworthy AI is applied in the Dutch media sector throughout the life cycle of AI solutions. Moreover, each chapter provides some examples of the application of these principles in the recent past by media companies.

BBC Machine Learning Engine Principles framework

In 2019, the BBC drafted its approach to responsible AI in a set of Machine Learning Engine Principles (MLEP). These guiding principles commit to:

- Reflecting the BBC's values of trust, diversity, quality, value for money and creativity
- Using ML to improve audience's experience of the BBC
- Carrying out regular review, ensuring data is handled securely and that algorithms serve the audiences equally and fairly
- Incorporating the BBC's editorial values and seeking to broaden, rather than narrow horizons
- Continued innovation and human in the loop oversight.

In 2021, the BBC adopted the [BBC Machine Learning Engine Principles framework.](#) The framework comprises six guiding principles and a self-audit checklist for ML teams (engineers, data scientists, product managers etc.) (Figure 11).

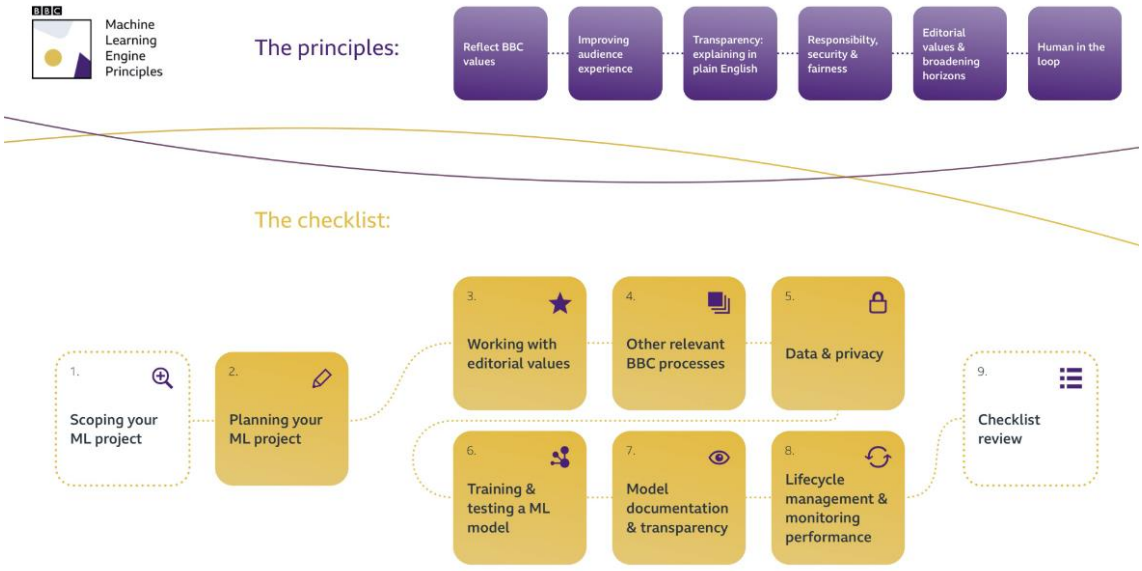


Figure 11: BBC Machine Learning Engine Principles framework. Source: https://www.bbc.co.uk/rd/publications/responsible-ai-at-the-bbc-our-machine-learning-engine-principles

Bavarian Broadcaster AI ethics guidelines

German Bayerischer Rundfunk (BR), a Bavarian public service broadcaster, has issued a new set

of AI ethics guidelines.[128] It offers ten core guidelines for day-to-day use of AI and automation.

The key points are:

- 1) user benefit: AI must add value to users experience and not just be an end in itself;
- 2) transparency & discourse: AI use must be transparent and users should know what technologies are used, what data are processed and which editorial teams or partners are responsible for it;
- 3) diversity & regional focus: BR strives towards dialect models in speech-to-text applications and bias-free training data (algorithmic accountability);
- 4) conscious data culture: BR requires solid information about the data sources from its vendors, i.e. what data was used to train the model. It also recognizes the importance of well-kept metadata, and upholds high standards when it comes to processing personal data;
- 5) responsible personalization: the use data-driven analytics as assistive tools for editorial decision-making;
- 6) editorial control: the editorial responsibility remains with the editorial units;
- 7) agile learning: the AI ethics guidelines offer general orientation in experiments, prototypes and pilot projects up until and including the beta phase. In the final release candidate phase, they are fully binding;
- 8) partnerships: collaboration with research institutions and ethics experts;
- 9) talent & skill acquisition: BR ensures it has employees with the skills to implement AI technologies;
- 10) interdisciplinary reflection: BR integrates the interdisciplinary reflection with journalists, developers and management from the beginning of the development pipeline.

4.1.3.3. Video Games

AI is used in a variety of forms in the game industry. It can be used non-exclusively in their creative process to develop new content, in the player experience, in the digital safety, to fight cheating and abuses and also in player support.[129]

AI used in video games also raises various challenges, which include: “the lack of data, lack of models that capture individual differences and context, and lack of transparency as underlying issues causing several ethical concerns and potential dangers to players such as predatory monetization, marginalisation, and misrepresentation”.[130]

[128] BR, ’Our AI Ethics Guidelines’, 2020, (last accessed on 8th August 2022), https://www.br.de/extra/ai-automation-lab-english/ai-ethics100.html

[129] EGDF & ISFE, ‘EGDF and ISFE Position paper, European Commission Consultation on Digital Principles‘, (2021), https://www.isfe.eu/wp-content/uploads/2021/09/EGDF-ISFE-Position-Paper-Digital-Principles-September-2021.pdf

[130] Magy Seif El-Nasr and Erica Kleinman. 2018. Data-Driven Game Development: Ethical Considerations. In The Fifteenth International Conference on the Foundations of Digital Games (FDG ’20), September 15–18, 2020, Bugibba, Malta. ACM, New York, NY, USA, https://doi.org/10.1145/1122445.11224

The International Game Developers Association (IGDA) is the world's largest non-profit membership organisation representing game developers.[131] IGDA, being aware of the importance of the effect of ideas conveyed through video games, has set up a Code of Ethics for game developers. The code contains provisions about how developers should behave in their workplace, in the leadership position they have, etc.; it also includes fundamental principles which should guide their work. One of them is written as follows: "Promote proper, responsible, and legal use of computing technology at our disposal".[132] Without mentioning specifically AI systems, this provision is broadly defined and enables one to consider that AI development needs to be used in a responsible and legal manner through video games production.

In the US, the Federal Trade Commission (FTC) provided some guidance in relation to the use of AI, especially in order to avoid unfair and deceptive trade practices.[133] Here are the recommendations and key considerations:

- *"Accuracy*. AI components of a game or service should be tested prior to implementation to confirm it works as-intended.

- *Accountability*. Companies should think about how the use of AI will impact the end-user. Outside experts may be used to help confirm that data being used is bias-free.

- *Transparency*. End-users should be made aware that the company may use AI, it should not be used secretively. Individuals should know what is being collected and how it will be used.

- *Fairness*. To further concepts of fairness, the FTC recommends giving people the ability to access and correct information."[134]

In Europe, the Pan-European Game Information (PEGI) has developed a Code of Conduct.[135] This code constitutes a set of rules to which every publisher using the PEGI system is contractually committed. The Code focuses on age labelling, marketing but also on providing transparency to the public in a responsible manner. The Code has detailed compliance, enforcement, and oversight sections. The code contains some provisions about transparency but does not address AI used in the video games. Perhaps the development of the code and more labelling about AI

[131] International Game Developers Association (IGDA), 'About US', (last accessed on 8th August 2022), https://igda.org/about-us/

[132] International Game Developers Association (IGDA), 'Core Values &Code of Ethics', (last accessed on 8th August 2022), https://igda.org/about-us/core-values-and-code-of-ethics/

[133] Federal Trade Commission, 'Using Artificial Intelligence and Algorithms', (2020), https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-and-algorithms

[134] The National Law Review, 'Video Games, AI and the … law? ', (2022), https://www.natlawreview.com/article/video-games-ai-and-law

[135] Pan European Game Information, 'The PEGI Code of Conduct', (last accessed on 8th August 2022), https://pegi.info/pegi-code-of-conduct

systems could be operated via this PEGI Code in the future. This will increase awareness from the audience and the players and strive toward more impactful transparency.

The Europe's Video Games Industry (ISFE) and the European Games Developer Federation (EGDF) published in September 2021, a position paper for the European Commission on Digital Principles.[136] They agreed that AI systems can support the European Commission's efforts to establish a trustworthy, ethical and responsible AI. They called for the EU to take a global approach "to ensure that there is transparency between regulatory approaches to facilitate legal compliance for Europe's businesses and SMEs in particular".[137]

When it comes to video-games companies, we had a look at the 2022 top companies in terms of financial revenue[138] and checked whether they advertised being part of a Code of Conduct or whether they had an internal code or charter of ethics in relation to AI. Many companies, if not all, have Codes of conduct/community standards in relation to the use of their products and services by gamers. They also have codes of conduct applicable to their employees such as the famous Tencent Sunshine Code of Conduct [139]with a zero-tolerance policy for corruption, fraudulent activities or misconduct. However, when it comes to the way AI systems are used by these companies in their products and services, not much can be found. Microsoft, one of the biggest companies active in the video games landscape, has on a cross-sectoral level an approach for responsible AI in line with the HLEG AI ethical principles for trustworthy AI.[140] They also developed a Responsible AI Impact Assessment Guide and Template[141] and responsible AI principles applicable to all the AI systems developed by the company.[142] Tencent, another important actor of the sector coming from China, has set its Vision and Mission as: "Value for User, Tech for Good" and advocates for focusing on principles to establish correct technological values as technologies evolve rapidly.[143] Codes of conduct or Ethical charter for the use of AI systems in their videogames services and product was not present on their official accessible

[136] EGDF & ISFE, 'EGDF and ISFE Position paper, European Commission Consultation on Digital Principles', (2021), https://www.isfe.eu/wp-content/uploads/2021/09/EGDF-ISFE-Position-Paper-Digital-Principles-September-2021.pdf

[137] EGDF & ISFE, 'EGDF and ISFE Position paper, European Commission Consultation on Digital Principles', (2021), https://www.isfe.eu/wp-content/uploads/2021/09/EGDF-ISFE-Position-Paper-Digital-Principles-September-2021.pdf

[138] All Top Everything, 'Top 10 Biggest Video Game Companies in the World', (last accessed on 8th August 2022), https://www.alltopeverything.com/top-10-biggest-video-game-companies/

[139] Tencent, ' Integrity Policy', available at https://www.tencent.com/en-us/integrity-policy.html

[140] Microsoft, 'Empowering impactful responsible AI practices', (last accessed on 8th August 2022), https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6

[141] Microsoft, 'Microsoft Responsible AI Impact Assessment Guide', June 2022, (last accessed on 8th August 2022), https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4ZzOI

[142] Microsoft, 'Microsoft responsible AI principles', (last accessed on 8th August 2022), https://www.microsoft.com/en-us/ai/our-approach?activetab=pivot1%3aprimaryr5

[143] Tencent website, section employees (last accessed 10th August 2022); https://www.tencent.com/en-us/employees.html ; Warc, 'Tencent proposes an ethical framework for AI', 2019, (last accessed on 10th August 2022), https://www.warc.com/newsandopinion/news/tencent-proposes-an-ethical-framework-for-ai/en -gb/42036

documents. When you consider that AI systems are a core part of the design and the functioning of video games, the absence of ethics charters or codes of conduct is surprising.

4.2. Opportunity for a European Digital Media Code of conduct

4.2.1.Conclusion

This section showed that no cross sectoral AI and media codes of conduct could be found. Furthermore, even if the work is in progress, we found little attention to AI in existing sectoral codes of conduct. The reason for this is perhaps because there is no shortage of AI ethics codes[144], and especially because the EU has its own: the HLEG Guidelines on trustworthy AI. As our analysis of the existing media codes and guidelines showed, many have overlapping principles such as: transparency, explainability, keeping humans in the loop etc. As noted by Tambini: "it is relatively easy to come up with these appealing words. The hard work is applying them in practice".[145] In the light of that, having a generic 'European Digital Media Code of conduct' risks being too general and will not address the needs and challenges faced by the media companies. This is because of the following reasons.

First, the lack of agreed definition of media. The term media is widely used in public discourse, encompassing all forms of mass communication across print news, broadcasters, and the internet (e.g., social media) and sometimes even entertainment (e.g., Netflix) or TV more in general. The definition of media at the EU level is, perhaps, in the making, in the upcoming European Media Freedom Act.[146] At this stage however, there is no societal or legal consensus of which entities fall under the EU definition of media: online media (print, broadcasting, online newspapers, websites, news portals), social media actors e.g. bloggers, but also social media influencers; and/or intermediary services who disseminate the news (e.g. social media platforms, search engines, hosting services)? It is therefore unclear who should be the addressee of such a European Code.

Second, different media sectors have different needs, and are using AI at different pace and at different stages of news/content production/distribution. To illustrate, a press agency faces other ethical risks when using AI than public service media.

Third, when it comes to the ethics of using AI in the media, a differentiation needs to be made between the different use cases. The potential and challenges of AI for media across the various stages of the media cycles: ideation and content gathering (AI in story discovery, AI driven

[144] A. Jobin, M. Ienca & E. Vayena, 'The global landscape of AI ethics guidelines', [2019], Nature Machine Intelligence.

[145] Tambini D., "Public service media should be thinking long term when it comes to AI", (12 May 2021), LSE blog, available at: https://blogs.lse.ac.uk/medialse/2021/05/12/public-service-media-should-be-thinking-long-term-when-it-comes-to-ai/. Similar findings can be found in T. Hagendorff, 'The Ethics of AI Ethics: An Evaluation of Guidelines', (2020). *Minds & Machines,* **30**, 99–120

[146] European Commission, 'Press release: European Media Freedom Act: Commission launches public consultation', (2022), https://ec.europa.eu/commission/presscorner/detail/en/IP_22_85

audience measurement), production, curation and distribution (AI in personalisation, AI in advertising), deliberation (AI in content moderation, AI in fact-checking, verification), audio-visual archiving. In other words, there are various levels and places when the unethical use of AI may appear and different mitigation responses are required. It may be at the level of the biased data, but it may as well concern how AI is being applied to end users.

4.2.2.The ways forward

To this end, instead, we propose the following recommendations:

1. First, to raise awareness to the media sector about the importance of AI ethics.
2. This process will be facilitated by the initiatives such as the Media AI Observatory soon available on the AI4Media website. It will serve as a knowledge hub for the media sector and a key place to find resources on AI ethics. The content will be written in an accessible way to benefit a wide range of actors and to draw their attention to the importance of ethical use of AI. Second, to focus on practical guidance on ethical AI applications in the media sector.

   Instead of a high-level list of principles to follow, the media sector needs a practical guide to ethical compliance. This could take the form of a theme-by-theme analysis with selected examples for real life use cases when AI is already used. Such work can entail some input from collaborative, interdisciplinary activities, involving AI developers, media practitioners and ethicists. Preparatory activities towards such practical guidance can be explored in the last two years of the AI4Media project.
3. Third, focus on legal certainty and certifications.

   To raise the level of legal certainty, one can consider a 'stamp of approval', in a form of certification of AI used in media. In order to 'seize the AI opportunity for the media sector and avoiding its potential risks', one of the recommendations of the "Entering the new paradigm of artificial intelligence and series" study commissioned by the Council of Europe and Eurimages was establishing a body of certification to ensure a fair use of AI in media.[147] It was also recommended to consider creating a governing body for a media AI certification. This approach is also in line with the AI Act proposal.

[147] T. Baujard, R. Tereszkiewicz, A. de Swarte and T. Tuovinen, 'Entering the new paradigm of artificial intelligence and series A Study commissioned by the Council of Europe and Eurimages', *2019.* *https://rm.coe.int/eurimages-entering-the-new-paradigm-051219/1680995331*

5. Conclusions

This deliverable provided our initial policy recommendations for the use of AI in the media sector. This initial version of the recommendations built on the work carried out during the first two years of the project. Thanks to the diversity of the AI4Media consortium and the contacts with the different media stakeholders, the overview has been designed as complete as possible. Section 3 of this deliverable presented the various challenges for the use of AI in the media sector encountered by different stakeholders and from different angles. In addition, it laid down our initial policy recommendations in relation to the challenges identified. In what follows, we provide the overview of the most recurring challenges and corresponding recommendations (Figure 12, Figure 13, and Figure 14).

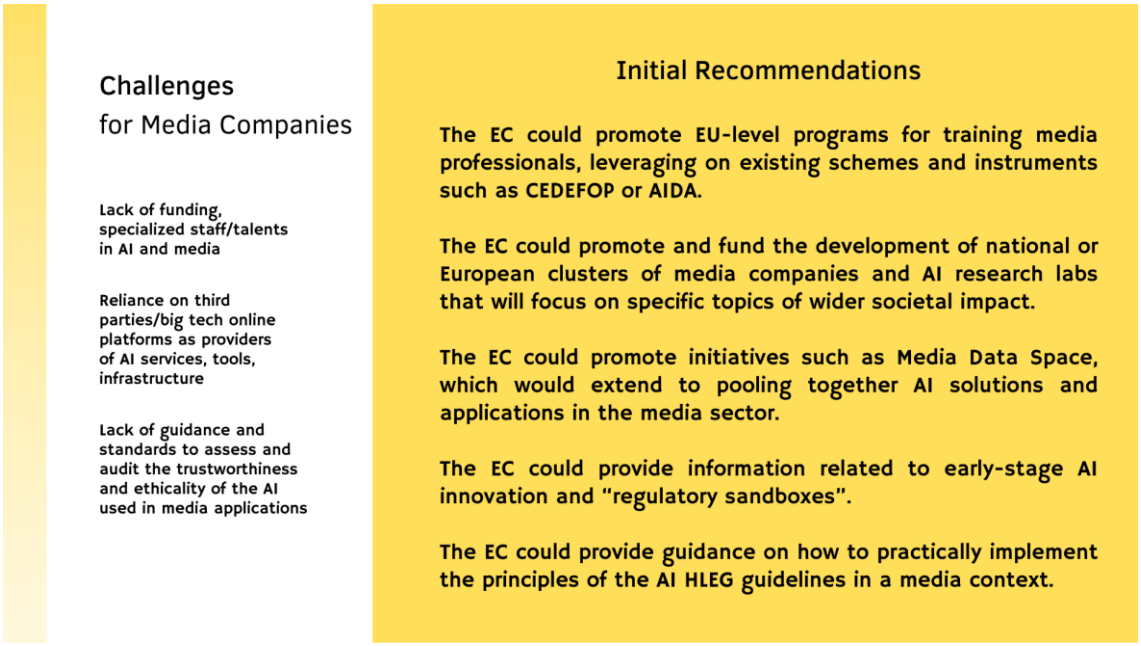


Figure 12: Challenges for media companies and initial recommendations

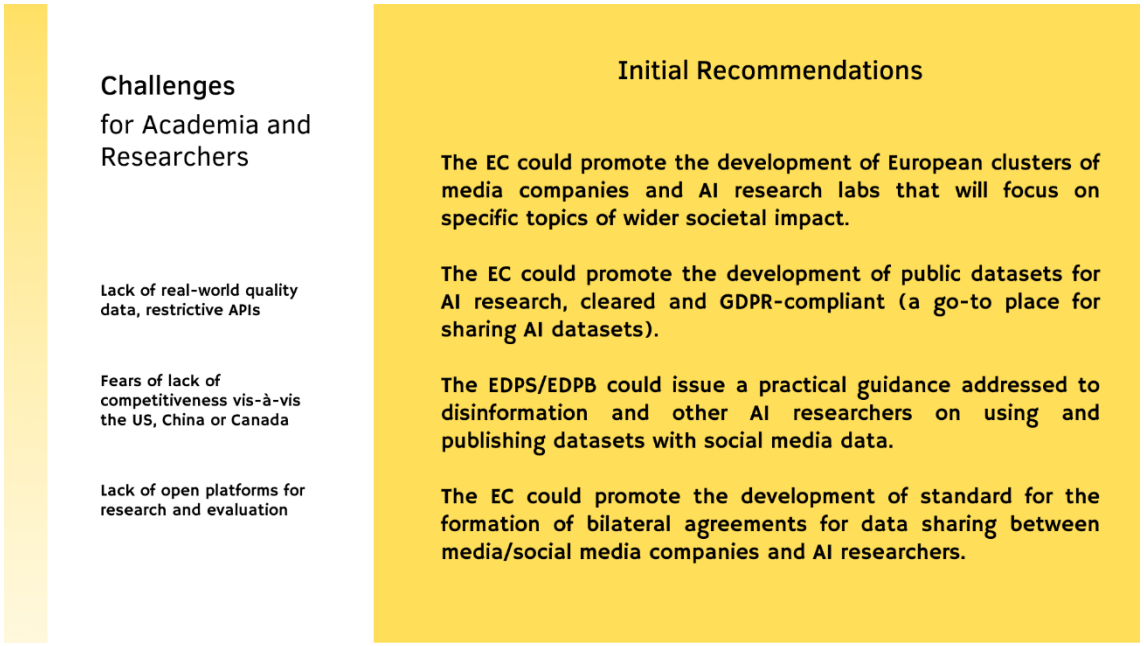


Figure 13: Challenges for Academia and Researchers and initial recommendations

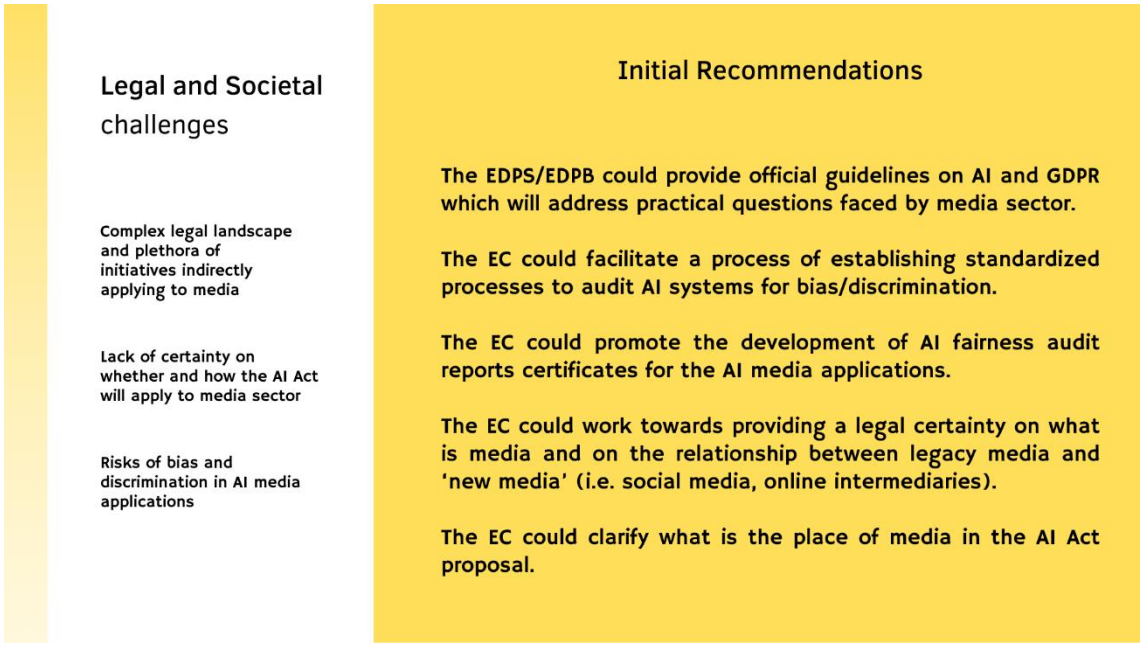


Figure 14: Legal and Societal challenges and initial recommendations

Until the final version of the recommendations, another overarching recommendation will be investigated. It appears necessary to have coordination with other standards setting organisations working on the AI topic. Coordination and collaboration could be necessary to provide clarity, coherence and easy access to guidelines and instruments and to have by topic specific recommendations.

Section 4 reflects on the opportunity to have a European Digital Media Code of Conduct. Through internal consortium discussions in AI4Media it was made clear that partners were not sure whether such a Code would be the best tool possible to ensure a common approach for the use of AI in Media based on European democratic values. Through the various research conducted and discussions held with partners and media stakeholders, it appeared that the media sector is broad and diverse and each sub-sector faces particular difficulties, which cannot be addressed by general criteria and requirements but rather require a tailor-made approach. The study of various codes of conduct showed that they all take from the principles and requirements of the EU HELG guidelines for trustworthy AI. Therefore, the EU seems to already have its EU Code of Conduct on AI and the wheel does not need to be reinvented. However, what is needed by the media sector is specific guidance on how these principles would materialise in a specific media sector and the AI4Media consortium can help in this regard.

While drafting this deliverable, the AI4Media consortium kept monitoring the changes in relation to the EU policy and legal landscape through the various ongoing proposals (for more information see D2.1). Some are closer to be officially adopted such as the Digital Services Act but others remain to be negotiated and can perhaps take into account the series of challenges identified and the preliminary recommendations addressed to it. In the next two years of the project, AI4Media will analyse the evolution of the challenges, policy and regulatory instruments and will proceed to updates, edits and additions for the final version of this deliverable.

On the road towards these final recommendations (due in month 48 of the project), the alternative approaches to a European Digital Media Code of Conduct will be evaluated. Partners led by KUL will consult with the different use cases involved in the project to assess the feasibility of this.

To conclude, AI systems used in the media sector are constantly evolving and bring a series of challenges which need to be addressed in order to promote EU values and fundamental rights. The AI4Media consortium tried to bring keys to bring solutions to these challenges. These initial efforts have set the path for the follow-up efforts of the consortium.

6. Appendix: Initial recommendations for the use of AI in the Media Sector

6.1. Introduction

This appendix provides the overall view of initial policy recommendations addressing the challenges identified in Section 3: initial policy recommendations addressing the identified challenges for media companies (sub-section 6.2., Table 15), initial policy recommendations addressing the identified challenges for academia and researchers (sub-section 6.3, Table 16), initial policy recommendations for legal and societal challenges identified (sub-section 6.4, Table 17).

6.2. Initial Policy Recommendations Addressing the Identified Challenges for Media Companies

In this subsection, we present the policy recommendations addressing the challenges for media companies identified in subsection 3.2.

Table 15: Initial Policy Recommendations Addressing the Identified Challenges for Media Companies

Challenge	*Initial Policy Recommendations*
Challenges related to staff and knowledge gap	
Challenges related to the internal organisation *Challenges related to AI innovation* *Difficulty in handling legal and ethical aspects in early-stage AI innovation*	• Allow access to ethical guidance provided by specialised public committees for ethical problems that can arise during machine learning model development and AI service development. • Issue practical, easy-to-use guidance and solutions on how to practically implement the responsible, ethical and trustworthy principles listed in corporate as well as other AI guidelines in an innovation context in the media sector. • Establish AI curricula at all education levels. • Start EU-level programs for training media professionals, leveraging on existing schemes and instruments such as CEDEFOP or AIDA. • Start mobility programs for internships or secondments of media professionals in AI research labs or of AI researchers in media companies. • Promote the development of national or European clusters of media companies and AI research labs that will focus on specific AI topics of wider societal impact. These clusters can among other things offer training to media professionals or retraining of technical personnel.

Lack of AI talent in the media industry	
Knowledge gap *Lack of information related to Trustworthy AI in the innovation context*	• Promote strategies to raise awareness and engage the society into the process of creating a culture for Trustworthy AI. People need to be trained to adopt ethical values and understand capabilities and limitations of AI. Strategies include, but are not limited to, (i) reinforcing education at school on STEM (and in particular new technologies such as AI) but also on social sciences and humanities (SSH) to adopt European values from early stages; (ii) strengthen collaborations among AI researchers and media professionals to improve communication skills for a general audience, with a common, simple and rigorous language that can inform the society avoiding misconception or overhype of the capabilities of the technology; (iii) involve the society with participatory methods such as open consultations or debates to make them feel part of the technology progress and gain trust towards it; (iv) awareness campaigns such as those launched with the release of the GDPR by national authorities and the EU.
Challenges related to limited resources and the bargaining power	
Media concentration and journalistic autonomy	• Consider where the burden of compliance with the proposed AI Act and similar regulations lies and ensure this will not be harmful to media diversity or to producing responsible AI solutions for the sector. • Consider a solution to the need of levelling up between news organisations and platforms regarding the information asymmetries and resource redistribution
Limited bargaining power	• Invest in and, importantly, sustained funding for platforms and networks that enable media partners to work in coordinated action, including instruments to continuously gather data on the challenges and needs, and a forum to communicate this to relevant stakeholders (policy and decision makers, industry representatives).
Licensing tensions between data sets free to re-use for research and	• Issue practical, easy-to-use guidance and solutions on how to practically implement the responsible, ethical and trustworthy principles listed in AI guidelines in a media-innovation context. • Make guidance from business and legal perspectives on which AI resources and datasets can be freely used and certainty about the

for commercial applications	legal status of the datasets and applicable licence fees.
Lack of coordination between media partners to seek collective solutions	• Provide support on a national level to incentivise coordinated action. • Facilitate more cooperation on the level of pooling together AI solutions and applications in the media sector, apart from initiatives such as the Media Data Space.
Compliance with legal and ethical frameworks	
Challenges for legal and regulatory compliance	• Facilitate access to legal information and issue guidance related to early-stage AI innovation and "regulatory sandbox" for AI development in the context of existing and planned legislation. • Facilitate access to easy-to-integrate, affordable Trustworthy AI enhancement tools, understandable transparency information and trustworthy datasets, both for third-party AI functions that are used in experimental tools/services as well as own early-stage AI development.
A need for accessible ethics advice for the media staff	
Lack of information on how AI systems address trustworthy AI challenges	• Ensure that AI systems come with trustworthy AI certificates, ensuring that AI systems have been audited to address issues such as explainability, robustness, fairness, privacy, etc. • Ensure that AI providers and big only platforms i) apply "trustworthiness by design" principles when developing AI systems, ii) provide periodic public reports on how they address trustworthy requirements and particularly explain where they come short, iii) make public any instances in which their systems demonstrably failed to comply with trustworthy AI requirements, leading to negative impacts on media companies or media users using such systems, iv) share information about their data and algorithms with independent researchers and independent authorities that could act as auditors/testers. • Issue guidance on development of trustworthy and explainable AI solutions and relevant certifications by independent authorities that can act both as advisors as well as enforcers. • Ensure that AI providers and big platforms enable independent research on their services and products to analyse potential impact
Lack of instruments for media organisations to audit, assess the ethicality and trustworthiness of AI systems	

	and risks. • Invest in practical solutions that allow media practitioners that do not have a background in AI to critically engage with and assess the ethicality and trustworthiness of AI. • Ensure continued investment in training. Organisational structures in the media industry are needed to create space for continuous learning and keeping up with state-of-the-art research.
Business needs not aligned with ethical concerns	• Incentivise organisations to prioritise ethical, societal and environmental considerations, for instance, by setting up specific KPIs.
A need for standardised data documentation	• Ensure that the media companies adopt transparent and accountable practices in their dataset creation, enabling auditing of the data they produce to fuel AI models. The AI research community, in a multi-disciplinary approach involving interpretive methodologies in semiotics and information science, human computer interaction (HCI) and software engineering, has started to conceptualise the questions around dataset creation for ML and how to design transparent and accountable dataset creation.[148,149]
Transparency of AI models and workflows documentation	• Ensure that the media companies document in a standardised way their model creation process by using existing tools[150]already partly adopted by the AI community.[151] Companies must be enjoined to follow such processes and produce the said documentation when acting as a public media stakeholder.
Lack of combination of legal and technical "templates" to simplify dealing	• Incentivise and fund the development of "templates" for dealing with common data protection issues, using legal and technical means, tailor-made for the needs of media organisations applying AI.

[148] B. Hutchinson et al., "Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure," (2021), Available: http://arxiv.org/abs/2010.13561

[149] M. Pushkarna, A. Zaldivar, and O. Kjartansson, "Data Cards: Purposeful and Transparent Dataset Documentation for Responsible AI.", (2022), Available: http://arxiv.org/abs/2204.01075

[150] M. Mitchell et al., "Model Cards for Model Reporting," in Proceedings of the Conference on Fairness, Accountability, and Transparency, (2019), pp. 220–229. doi: 10.1145/3287560.3287596.

[151] Model Cards, 'The value of a shared understanding of AI models', (last accessed on 25th July.), https://modelcards.withgoogle.com

with data protection	
A need for space for experimentation to support policy	• Invest into creating experimentation and validation environments, such as sandboxes, that would bring interdisciplinary, cross-sector actors and, importantly, provide a direct link to and inform policymakers and regulators.

6.3. Initial Policy Recommendations Addressing the Identified Challenges for Academia and Researchers

In this subsection, we present the policy recommendations addressing the challenges for academia and researchers identified in subsection 3.3.

Table 16: Initial Policy Recommendations Addressing the Identified Challenges for Academia and Researchers

Challenge	Initial Policy Recommendations
Challenges related to the lack of data and data access	
Lack of real-world data to train AI systems for the media sector	• Promote the development of national or European clusters of media companies and AI research labs that will focus on specific topics of wider societal impact. In the context of such initiatives, the clusters will pursue the development of public datasets for AI research, the development of standard and transparent mechanisms for the formation of bilateral agreements for data sharing between media industries and AI researchers, the establishment of benchmarking datasets and competitions for testing AI algorithms, etc. • Promote the European AI on demand platform and its marketplace as the go-to place for sharing AI datasets while also providing incentives for developing datasets or sharing data.
Lack of quality data	• Ensure that AI researchers and practitioners adopt best practices of data management that guarantee the highest quality of datasets, enabling reusability and accountability. • Ensure that the use of synthetic data be clearly disclaimed on applications that can be considered as high risk, such as deep fakes. There should be a clear definition of what is considered a dangerous application of synthetic data and when it is necessary to inform media stakeholders.

Lack of data for developing synthesis detection and Privacy Enhancing Technologies	• Incentivize / fund activities that check and "clear" older datasets with respect to data protection by getting consent, to make them usable for synthesis detection and privacy enhancing technologies (PET) development. • Incentivize / fund activities for creating new datasets for synthesis detection and PET development.
Lack of common understanding that (AI) systems and tools can address privacy aspects without the need to sacrifice utility or performance	• Incentivize the development and use of privacy enhancing tools, especially considering usability aspects, for R&D projects and services. • Incentivize the communication of cases in which alleged conflicts between privacy and utility could be resolved using technology.
Challenges for AI and disinformation research	
Lack of common best practices and standards for disinformation analysis	• Establish R&D calls to develop a coherent best practice document and tool set for content verification in the media sector, using and combining existing information.
A need for sustainable R&D for disinformation analysis	• Establish *recurring* R&D calls that include media forensics and media disinformation analysis development to ensure that technologies can cope with the overall technology development that is available for attackers.

Lack of sufficient API/data access to tackle disinformation	• Address the clear need for a legally binding data access framework at the EU level that provides researchers with access to a range of different types of platform data. • Further promote and expand initiatives such as Media Data Space, or the EDMO's draft Code of Conduct on how platforms can share data with independent researchers while protecting users' rights.
Challenges related to competitiveness and resources	
Tension between aiming at AI Excellence and complying with Trustworthy AI desiderata and requirements	• Ensure sufficient public funding to research projects on AI for the media. • Ensure fair and straightforward access of European researchers to large compute infrastructure and relevant datasets. • Create a "safe regulatory space" for researchers that provides clear guidance on the limits of ethical AI research without over-constraining freedom of thought and innovation.
Open platforms for AI research and evaluation	• Incentivise the use / implementation of open APIs and public ontologies by public funding to improve interoperability and evaluation of tools developed within funded projects • Prepare (regular) R&D calls for the development and maintenance of application-specific open platforms and ontologies, especially in domains with large impact for the public (e. g. disinformation analysis)

6.4. Initial Policy Recommendations for Legal and Societal Challenges Identified

In this subsection, we present the policy recommendations addressing the legal and societal challenges identified in subsection 3.4.

Table 17: Initial Policy Recommendations for Legal and Societal Challenges Identified

Challenge	Initial Policy Recommendations
Complexity of Legal Landscape	
Plethora of policy	• Ensure a practical guidance document for the compliance with relevant regulations when applying AI in the media sector.

initiatives and a complex regulatory landscape	• Consider an instrument focusing specifically on the use of AI in the media to ensure legal certainty and answer the calls from society and professionals. • Have a place where a mapping of relevant legislation for the development, the use of AI systems in media applications is easily accessible with different categories of media actors/sectors represented, so each one knows their main obligations and rights according to the different frameworks. • Ensure during the drafting process of regulatory initiatives that the different existing instruments, along with ongoing proposals, would function together, allowing to have a global and systemic approach to this complex regulatory landscape. • Address the need to ensure consistency to improve harmonisation and legal certainty. Ensuring more coordination and better communication on the interplay between the different instruments seems necessary.
Compliance with the General Data Protection Regulation	• Solve the terminologies inconsistencies when it comes to AI and GDPR, especially for interpretable, explainable and transparent. • Define a formal framework for reliability, transparency and fairness in AI in the media applications. • Develop a multidisciplinary definition of interpretable AI which can be adopted in both the social and the computer sciences. • Resolve the issue of the power asymmetries for AI development with limited private companies' monopoly over data and develop AI inclusive policies and regulations. This can be achieved by putting citizen-centred innovation with class action, activism, and whistle-blowers schemes. • Provide more guidance on how to produce, find, use and re-use clean data sets used to develop AI applications in the media sector. • Provide best practices/guidelines on AI and GDPR (in a comprehensive manner) which will address practical questions and doubts faced by media staff.
Monolithic policy regulations	• Ensure that regulations do not create barriers for organisations to innovate areas that are seen as high-risk. • Address the need for mechanisms to prevent and effectively address misuse of AI technologies in media (e.g. deepfakes) but also mechanisms to incentivise their use for social good.
Conceptual and definitional challenges	

The role of 'media'	• Reflect on the role of media and 'new media' in the digital era, e.g. whether online intermediaries should be considered media and should they benefit from the legal privileges and protections that are enjoyed by the press and broadcasting. • Consider adopting a general European definition of 'media' under the upcoming European Media Freedom Act.
The so-called 'media exemption' in content moderation	• Reflect on the treatment of traditional media in the online environment, which should be reflected in legal norms. • Consider clarifying the DSA platforms' responsibilities with regard to content moderation towards media organisations, i.e. how platforms should deal with lawful content under the editorial control and legal liability of the publisher (or broadcaster).
The place of media in the AI Act	• Clarify the scope of the AI Act vis-a-vis media sector. • Consider extending the scope of Art. 52(3) of the AI Act to contain an obligation to make the deep fake identification information undeletable in case of transfer or further modification of the material in order not to lose track of the deepfake's original information. The transparency requirements could include more precisions on what should be communicated (the type of information), when (at which stage this should be revealed) and how.
Fundamental rights and societal challenges	
AI-driven Manipulation and Propaganda	• Revise the wording of Art. 5 of the AI Act. • Ban political microtargeting, regardless of whether it utilises personal or non-personal data. "Second-hand-manipulation" conducted by non-political figures, which is not obvious on its face, should also be considered in such a prohibition. • Revise the Human Rights Law framework concerning freedom of thought to make the rights application clear and tangible in practice. This could involve drafting a secondary legislation to explain what this right entails and how it can be enforced to create preventive (ex-ante) measures against manipulation, regardless of its frontiers and technology used.
AI bias and discrimination against underrepresented or	• Ensure that AI systems should come with AI fairness audit reports certificates, ensuring that the system has been extensively tested/audited to minimise the risk of bias. • Design transparent processes to audit AI systems for bias/discrimination. This can be done by independent authorities on EU level.

vulnerable groups	• Ensure that AI providers and big platforms provide clear info/public reports on the data used for training AI systems and how they have considered/addressed AI fairness requirements. • Ensure that AI providers and big platforms enable independent research on their services and products to analyse potential impact and risks. • Consider regulating AI-based systems such as machine translation systems (for language translation) or decision systems. • Support easy and fair access of academic/research institutes to large-scale compute and data management infrastructures.
Filter-bubbles in recommender systems	• Incentivize / fund the development, use and evaluation of recommenders tailored to avoid filter-bubble problems.
Transparent communication	• Promote research on interpretable and explainable AI. • Research on transfer learning discusses the importance of using other metrics (footprint, human cost…) to assess models, beyond the task performance metrics.
Intellectual Property Challenges	
Copyright challenges of AI use in media	• Ensure that AI systems are not granted authorship or any similar status for AI-generated-works, nor should they be considered copyrightable, to be able to encourage human creativity, flourish and expand the public domain, and incentivize AI research and development. • Avoid creating a separate type of protection such as sui generis rights for AI-generated works. • Ensure that the usage of copyright protected works for AI training purposes is not deemed infringing, in order to incentivize research and innovation.

References

Legislation and legal proposals

Treaty on European Union (TEU), OJ C 326, consolidated version on 26.10.2012, p. 13–47, ELI: http://data.europa.eu/eli/treaty/teu_2012/oj

Treaty on the functioning of the European Union (TFEU), OJ C326, consolidated version on 26.10.2012, *p. 47–390, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012E%2FTXT*

Charter of Fundamental Rights of the European Union, OJ 326, 26 December 2012, p.391-407, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT

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, 4.5.2016, p. 1–88, ELI: http://data.europa.eu/eli/reg/2016/679/oj

Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union (Text with EEA relevance.), OJ L 303, 28 November 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32018R1807

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, 7 August 1985, p. 29–33, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A31985L0374

Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation, OJ L303/16, 2 December 2000, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2000:303:0016:0022:EN:PDF

Council Directive 000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin, OJ L180/22, 19 July 2000, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2000:180:0022:0026:EN:PDF

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 ('Directive on electronic commerce'), OJ L 178, 17 July 2000, https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031

Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety (Text with EEA relevance), OJ L 11, 15 January 2002, p. 4–17, https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32001L0095

Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), OJ L201, 31 July 2022, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02002L0058-20091219

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, 28 November 2018, https://eur-lex.europa.eu/eli/dir/2018/1808/oj*

Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, OJ L 130, 17.5.2019, p. 92–125

European Commission, Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC COM/2020/825 final.

European Commission, Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act) COM/2020/842 final.

European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts COM/2021/206 final.

European Commission, Proposal for a regulation of the European Parliament and of the CouncIL on European data governance (Data Governance Act) COM/2020/767 final

European Commission, Proposal for a regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act), COM/2022/68 final, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN

Case-law

Belgian DPA, Decision nr 13/2022 of 27 January 2022, DOS-2018-04433.

CCH Canadian Ltd v Law Society of Upper Canada, [2004] 1 SCR 339, 2004 SCC 13 https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/2125/index.do

Policy documents

Council of Europe, 'CAHAI Feasibility Study' (17 December 2020), available at: https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da.

Council of Europe, 'Study on the impact of digital transformation on democracy and good governance' (26 July 2021), CDDG(2021)4 Final, available at: https://rm.coe.int/study-on-the-impact-of-digital-transformation-on-democracy-and-good-go/1680a3b9f9

Council of Europe, 'Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes' (2019), available at: https://search.coe.int/cm/pages/result_details.aspx?ObjectId=090000168092dd4b.

Council of Europe, 'Conference of Ministers responsible for Media and Information Society Artificial intelligence – Intelligent politics Challenges and opportunities for media and democracy' (2021), https://www.coe.int/en/web/freedom-expression/media2021nicosia

Council of Europe, 'Final Declaration of the Conference of Ministers responsible for Media and

Information Society and resolutions on freedom of expression and digital technologies, on the safety of journalists, on the changing media and information environment, on the impacts of the COVID-19 pandemic on freedom of expression 10-11 June 2021', https://rm.coe.int/final-declaration-and-resolutions

Council of Europe, 'Guidance note on content moderation. Best practices towards effective legal and procedural frameworks for self-regulatory and co-regulatory mechanisms of content moderation', (2021) accessible at : https://rm.coe.int/content-moderation-en/1680a2cc18

Council of Europe, 'Recommendation CM/Rec(2022)13 of the Committee of Ministers to member States on the impacts of digital technologies on freedom of expression', (2022), available at : https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680a61729

Council of Europe, 'Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems', (2020), available at : https://search.coe.int/cm/pages/result_details.aspx?ObjectId=09000016809e1154

Best Practice Manual for Digital Image Authentication, ENFSI-BPM-DI-03, available at: https://enfsi.eu/wp-content/uploads/2021/10/BPM_Image-Authentication_ENFSI-BPM-DI-003-1.pdf

European Commission, 'Press release: European Media Freedom Act: Commission launches public consultation', (2022), https://ec.europa.eu/commission/presscorner/detail/en/IP_22_85

Khan. I., "Disinformation and freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression", A/HRC/47/25 13 April 2021, available at: https://undocs.org/A/HRC/47/25.

EDMO, 'Report of the European Digital Media Observatory's Working Group on Platform-to-Researcher Data Access', (2022), https://edmo.eu/wp-content/uploads/2022/02/Report-of-the-European-Digital-Media-Observatorys-Working-Group-on-Platform-to-Researcher-Data-Access-2022.pdf.

Europol, "European Union Serious and Organised Crime Threat Assessment (SOCTA) 2021. A corrupting influence: the infiltration and undermining of Europe's economy and society by organised crime", available at: https://www.europol.europa.eu/publications-events/main-reports/socta-report

Europol, "EU Policy Cycle - EMPACT 2022+ Fighting crime together" available at: https://www.europol.europa.eu/crime-areas-and-statistics/empact

OSCE Representative on Freedom of the Media, 'Policy paper on freedom of the media and artificial intelligence' (2020), https://www.osce.org/files/f/documents/4/5/472488.pdf

OSCE Representative on Freedom of the Media, ' Policy paper on AI and freedom of expression in political competition and elections', (2021), https://www.osce.org/representative-on-freedom-of-media/483638

OSCE Representative on Freedom of the Media, ' Spotlight on Artificial Intelligence and Freedom of Expression: A Policy Manual', 20 January 2022, https://www.osce.org/representative-on-freedom-of-media/510332

Academic Resources

Bruns A., “After the ‘Apicalypse’: Social Media Platforms and Their Fight against Critical Scholarly Research” [2021] Disinformation and Data Lockdown on Social Platforms 14.

Buolamwini J. and Gebru T., “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification,” [2018], Proceedings of the 1st Conference on Fairness, Accountability and Transparency, PMLR 81:77-91, 2018.

Calo R. et all, ‘Telling Stories on Culturally Responsive Artificial Intelligence‘, (2020), University of Washington Tech Policy Lab, accessible at http://techpolicylab.uw.edu/wp-content/uploads/2022/04/Telling_Stories_Pages_4-4-22.pdf

Carys C., “The AI-Copyright Challenge: Tech-Neutrality, Authorship, and the Public Interest”, (December 14, 2021). Ryan Abbott (ed.) Research Handbook on Intellectual Property and Artificial Intelligence (Edward Elgar Press, 2022 Forthcoming), Osgoode Legal Studies Research Paper, available at SSRN: https://ssrn.com/abstract=4014811.

Carys C., "Copyright, Communication and Culture: Towards a Relational Theory of Copyright Law" (2011). Books. 53, available at: https://digitalcommons.osgoode.yorku.ca/faculty_books/53.

Craig C.J., ’The AI-Copyright Challenge: Tech-Neutrality, Authorship, and the Public Interest‘ (2021) in Ryan Abbott (ed), Research Handbook on Intellectual Property and Artificial Intelligence (Edward Elgar Press, forthcoming 2022), available at : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4014811

Coe P., “Redefining ‘Media’ Using a ‘Media-as-a-Constitutional-Component’ Concept: an Evaluation of the Need for the European Court of Human Rights to Alter Its Understanding of ‘Media’ within a New Media Landscape” [2017] 37 Legal Studies 25.

Denton E. and others, “On the Genealogy of Machine Learning Datasets: A Critical History of ImageNet” (2021) 8 Big Data & Society 205395172110359.

Eskens S., “A right to reset your user profile and more: GDPR-rights for personalized news consumers”, International Data Privacy Law, Volume 9, Issue 3, August 2019, Pages 153–172, https://doi-org.kuleuven.e-bronnen.be/10.1093/idpl/ipz007.

Fabbrizzi S., Papadopoulos S., Ntoutsi E., and Kompatsiaris I., “A Survey on Bias in Visual Datasets,” [2021], available: http://arxiv.org/abs/2107.07919.

Freelon D, “Computational Research in the Post-API Age” [2018] 35 Political Communication 665.

Jobin A., M. Ienca & E. Vayena, 'The global landscape of AI ethics guidelines', [2019], Nature Machine Intelligence.

Goyal P. and others, “Fairness Indicators for Systematic Assessments of Visual Feature Extractors” [2022] 2022 ACM Conference on Fairness, Accountability, and Transparency.

Goyal N., Kivlichan I., Rosen R., and Vasserman L., “Is Your Toxicity My Toxicity? Exploring the Impact of Rater Identity on Toxicity Annotation” [2022], available at: http://arxiv.org/abs/2205.00501

Hagendorff T., ’The Ethics of AI Ethics: An Evaluation of Guidelines‘, (2020). *Minds & Machines,* **30**, 99–120

Helberger N., Eskens S., van Drunen M., Bastian M., Moeller J. for the Council of Europe, 'Implications of AI-driven tools in the media for freedom of expression', (2019), https://rm.coe.int/coe-ai-report-final/168094ce8f

Helberger N., Diakopoulos N., (2022) "The European AI Act and How It Matters for Research into AI in Media and Journalism", Digital Journalism, DOI: 10.1080/21670811.2022.2082505.

Hutchinson B. et al., "Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure," (2021), Available: http://arxiv.org/abs/2010.13561

Koltay A., "New media and freedom of expression: Rethinking the constitutional foundations of the Public Sphere", [2019] Oxford, UK: Hart Publishing.

Leerssen P., 'Platform research access in Article 31 of the Digital Services Act: Sword without a shield?', VerfBlog, (2021), https://verfassungsblog.de/power-dsa-dma-14/ , DOI: 10.17176/20210907-214355-0.

Mitchell M. et al., "Model Cards for Model Reporting," in Proceedings of the Conference on Fairness, Accountability, and Transparency, (2019), New York, NY, USA ,pp. 220–229. doi: 10.1145/3287560.3287596.

Mouzhi Ge, Delgado-Battenfeld C., and Jannach D. [2010], "Beyond accuracy: evaluating recommender systems by coverage and serendipity", in Proceedings of the fourth ACM conference on Recommender systems (RecSys '10). Association for Computing Machinery, New York, NY, USA, 257–260. https://doi-org.kuleuven.e-bronnen.be/10.1145/1864708.1864761.

Napoli Philip M. and Caplan R., "When Media Companies Insist They're Not Media Companies and Why It Matters for Communications Policy", [2016], available at SSRN: https://ssrn.com/abstract=2750148.

M. Pushkarna, A. Zaldivar, and O. Kjartansson, "Data Cards: Purposeful and Transparent Dataset Documentation for Responsible AI.", (2022), Available: http://arxiv.org/abs/2204.01075

Sag M., 'The New Legal Landscape for Text Mining and Machine Learning', (2019), in Journal of the Copyright Society of the USA, Vol. 66 p.291, accessible at : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3331606

Sambasivan N. and others, "Everyone Wants to Do the Model Work, Not the Data Work': Data Cascades in High-Stakes AI" [2021] Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.

Sap M. and others, "The Risk of Racial Bias in Hate Speech Detection" [2019] Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics.

Simon Felix M., "Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy", [2022] Digital Journalism, DOI: 10.1080/21670811.2022.2063150.

Tambini D., "Media Freedom", [2021], Polity Press 2021.

Tawfik, M., "History in the Balance: Copyright and Access to Knowledge. Radical Extremism' to 'Balanced Copyright'" [2010], in Canadian Copyright and the Digital Agenda, available at: https://scholar.uwindsor.ca/lawpub/17.

Tufekci Z., “Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls” [2014], Proceedings of the 8th International AAAI Conference on Weblogs and Social Media, available at: https://arxiv.org/abs/1403.7400.

Vermeulen, B. P., “Freedom of Thought, Conscience and Religion (Article 9)” [2006], In P. van Dijk, F. van Hoof, A. van Rijn, & L. Zwaak (Eds.), Theory and Practice of the European Convention on Human Rights (pp. 751-771). (4th edition).

Wang A, Narayanan A and Russakovsky O, “Revise: A Tool for Measuring and Mitigating Bias in Visual Datasets” [2020] Computer Vision – ECCV 2020 733.

Williams, M. L., Burnap, P., & Sloan, L. (2017). Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users’ Views, Online Context and Algorithmic Estimation. Sociology, 51(6), 1149–1168. https://doi.org/10.1177/0038038517708140.

Xianhong H., Neupane B., Echaiz L.F., Sibal P. and Rivera M. for UNESCO, ’Steering AI and advanced ICTs for knowledge societies: a Rights, Openness, Access, and Multi-stakeholder Perspective’, (2019), https://unesdoc.unesco.org/ark:/48223/pf0000372132

Other

AI4Media project: D2.1 Overview & Analysis of the AI Policy Initiatives on EU level.

AI4Media project: D2.3, led by CERTH, “AI technologies and applications in media: State of Play, Foresight, and Research Directions”, (2022)

All Top Everything, ’Top 10 Biggest Video Game Companies in the World’, (last accessed on 8th August 2022), https://www.alltopeverything.com/top-10-biggest-video-game-companies/

Alliance of Independent Press Councils of Europe (AIPCE) and Media Council in the Digital Age project, ’Survey about self-regulation bodies andchallenges of digital age‘, (2020), (last accessed on 4th August 2022), https://presscouncils.eu/Survey-of-professional-journalists-about-self-regulation-bodies-and-challenges-of-digital-age

Alliance of Independent Press Councils of Europe (AIPCE) & Media Council in the Digital Age project, ’Code of Ethics’, (last accessed 4th August 2022), https://presscouncils.eu/Code-of-Ethics

Baujard T., Tereszkiewicz R., de Swarte A. and Tuovinen T., ’Entering the new paradigm of artificial intelligence and series A Study commissioned by the Council of Europe and Eurimages’, *2019. https://rm.coe.int/eurimages-entering-the-new-paradigm-051219/1680995331*

BR, ’Our AI Ethics Guidelines’, 2020, (last accessed on 8th August 2022), https://www.br.de/extra/ai-automation-lab-english/ai-ethics100.html

Bublitz J. C., Douglas T., “Manipulative Influence via AI Systems and the EU Proposal for Regulation of Artificial Intelligence”, [2021], available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/F2665640_en.

CLAIRE project: What is regulatory sandboxing for AI?, https://claire-ai.org/brainfoods4/.

DARPA website, (last accessed 10th August 2022), https://www.darpa.mil/program/media-forensics

DARPA website, (last accessed 10th August 2022), https://www.darpa.mil/program/semantic-forensics

Donald W. Reynold Journalism Institute, (last accessed on 4th August 2022), https://rjionline.org/

Dutkiewicz L., "From the DSA to Media Data Space: the possible solutions for the access to platforms' data to tackle disinformation", 2022, European law blog, available at: https://europeanlawblog.eu/2021/10/19/from-the-dsa-to-media-data-space-the-possible-solutions-for-the-access-to-platforms-data-to-tackle-disinformation/

EBU, "Position Paper: The Digital Services Act must safeguard freedom of expression online", available at: https://www.ebu.ch/files/live/sites/ebu/files/News/Position_Papers/open/2022/220118-DSA-media-statment-final.pdf

EBU, "Protecting Media Content Online: A Decisive Moment", [2021], available at: https://www.ebu.ch/news/2021/10/protecting-media-content-online-a-decisive-moment

EDMO & the Institute for Data Democracy & Politics (The George Washington University), 'Report of the European Digital Media Observatory's Working Group on Platform-to-Researcher Data Access', (2022), available at : https://www.google.com/url?q=https://edmo.eu/wp-content/uploads/2022/02/Report-of-the-European-Digital-Media-Observatorys-Working-Group-on-Platform-to-Researcher-Data-Access-2022.pdf&sa=D&source=docs&ust=1662130030233908&usg=AOvVaw0hQHiaBjhPj6xIxiIfS8Ty

EGDF & ISFE, 'EGDF and ISFE Position paper, European Commission Consultation on Digital Principles', (2021), https://www.isfe.eu/wp-content/uploads/2021/09/EGDF-ISFE-Position-Paper-Digital-Principles-September-2021.pdf

Ethical Journalism Network, 'About the EJN', (last accessed on 4th August 2022), https://ethicaljournalismnetwork.org/who-we-are

European Centre for Press & Media Freedom (ECPMF), 'Governance', (last accessed on 4th August 2022), https://www.ecpmf.eu/about/governance/

European Federation of Journalists (EFJ), 'About EFJ', (last accessed on 4th August 2022), https://europeanjournalists.org/about-efj/

European Federation of Journalists (EFJ), 'Ethics', (last accessed on 4th August 2022), https://europeanjournalists.org/policy/ethics/

EU Disinfo lab, "Fact-checkers and experts call on MEPS to reject a media exemption in the DSA" [2021], available at: https://www.disinfo.eu/advocacy/fact-checkers-and-experts-call-on-meps-to-reject-a-media-exemption-in-the-dsa/

EUI Centre for Media Pluralism and Media Freedom, ' The Media Pluralism Monitor 2022 (MPM2022)', available at : https://cmpf.eui.eu/mpm2022-results/

European Agency for Safety and Health at work, 'Safety and Health legislation', https://osha.europa.eu/en/safety-and-health-legislation

Excavating AI website, (last accessed 7th July 2022), https://excavating.ai

Federal Trade Commission, 'Using Artificial Intelligence and Algorithms', (2020), https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-and-algorithms

Garvie C., "Garbage in, garbage out: face recognition on flawed data", Georgetown Law, Center on Privacy and Technology, 2019, available at: https://www.flawedfacedata.com/.

Haapanen L., 'Media Councils and Self-Regulation in the Emerging Era of News Automation', (2020), Published by the Council for Mass Media in Finland, https://presscouncils.eu/userfiles/files/Emerging%20era%20of%20news%20automation%20Haapanen.pdf

Hao K., "We read the paper that forced Timnit Gebru out of Google. Here's what it says." (December 4, 2020), MIT Technology Review, available at: https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/.

Impress, 'Code Review 2020-2022', (last accessed on 4th August 2022), https://www.impress.press/standards/code-review.html

International Federation of Journalists (IFJ), 'About', (last accessed on 4th August 2022), https://www.ifj.org/who/about-ifj.html

International Federation of Journalists (IFJ), 'IFJ Global Charter of Ethics for journalists', (2019) https://www.ifj.org/who/rules-and-policy/global-charter-of-ethics-for-journalists.html

International Game Developers Association (IGDA), 'About US', (last accessed on 8th August 2022), https://igda.org/about-us/

International Game Developers Association (IGDA), 'Core Values &Code of Ethics', (last accessed on 8th August 2022), https://igda.org/about-us/core-values-and-code-of-ethics/

Juntunen L. for the Council for Mass Media in Finland, 'Digital Challenges to ethical standards of journalism. Responses and needs of European Media Councils', (2022), https://presscouncils.eu/New-report-explores-digital-challenges-to-ethical-standards-of-journalism

Luscombe R., "Google Engineer Put on Leave after Saying AI Chatbot Has Become Sentient", [2022], The Guardian, available at: https://www.theguardian.com/technology/2022/jun/12/google-engineer-ai-bot-sentient-blake-lemoine&gt.

Media Wise website, (last accessed on 4th August 2022), http://www.mediawise.org.uk/

Microsoft, 'Empowering impactful responsible AI practices', (last accessed on 8th August 2022), https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6

Microsoft, 'Microsoft Responsible AI Impact Assessment Guide', June 2022, (last accessed on 8th August 2022), https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4ZzOI

Microsoft, 'Microsoft responsible AI principles', (last accessed on 8th August 2022), https://www.microsoft.com/en-us/ai/our-approach?activetab=pivot1%3aprimaryr5

Model Cards, 'The value of a shared understanding of AI models', (last accessed on 25th July.), https://modelcards.withgoogle.com

Pan European Game Information, 'The PEGI Code of Conduct', (last accessed on 8th August 2022), https://pegi.info/pegi-code-of-conduct

Raad voor Journalistiek, 'About us', (last accessed on 4th August 2022), https://www.rvdj.be/pagina/english-version

S+T+ARTS website, (last accessed 4th August 2022), https://starts.eu/

Tambini D., "Public service media should be thinking long term when it comes to AI", (12 May 2021), LSE blog, available at: https://blogs.lse.ac.uk/medialse/2021/05/12/public-service-media-should-be-thinking-long-term-when-it-comes-to-ai/.

Tampere University, 'Ethicnet, Journalism Ethics', (last accessed on 4th August 2022), https://research.tuni.fi/ethicnet/

Tencent, ' Integrity Policy', available at https://www.tencent.com/en-us/integrity-policy.html

Tencent website, section employees (last accessed 10th August 2022); https://www.tencent.com/en-us/employees.html

Thomson Reuters, 'Our AI Principles', (last accessed on 8th August 2022), https://www.thomsonreuters.com/en/artificial-intelligence/ai-principles.html

Vanleeuw R., "Processing of public Twitter data: as free as a bird?" [2022], CiTiP blog, available at: https://www.law.kuleuven.be/citip/blog/processing-of-public-twitter-data-as-free-as-a-bird/.

Ventura P., 'Algorithms in the Newsrooms. Challenges and Recommendations for Artificial Intelligence with the Ethical Values of Journalism.', (2021), Published by the Catalan Press Council, https://fcic.periodistes.cat/wp-content/uploads/2022/03/venglishDIGITAL_ALGORITMES-A-LES-REDACCIONS_ENG-1.pdf

Warc, 'Tencent proposes an ethical framework for AI', 2019, (last accessed on 10th August 2022), https://www.warc.com/newsandopinion/news/tencent-proposes-an-ethical-framework-for-ai/en-gb/42036

AI4media

ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu