



# ROADMAP ON AI TECHNOLOGIES & APPLICATIONS FOR THE MEDIA INDUSTRY

## SECTION: “ETHICAL, SOCIETAL, ECONOMIC, ENVIRONMENTAL CONCERNS & RISKS – A LOOK INTO THE FUTURE OF AI FOR THE MEDIA”



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

[info@ai4media.eu](mailto:info@ai4media.eu)

[www.ai4media.eu](http://www.ai4media.eu)

## Authors

**Anna Schjøtt Hansen** (University of Amsterdam)  
**Tobias Blanke** (University of Amsterdam)

This report is part of the deliverable D2.3 - “*AI technologies and applications in media: State of Play, Foresight, and Research Directions*” of the AI4Media project.

You can site this report as follows:

F. Tsalakanidou et al., Deliverable 2.3 - AI technologies and applications in media: State of play, foresight, and research directions, AI4Media Project (Grant Agreement No 951911), 4 March 2022

This report was supported by European Union’s Horizon 2020 research and innovation programme under grant number 951911 - AI4Media (A European Excellence Centre for Media, Society and Democracy).

The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf.

## Copyright

© Copyright 2022 AI4Media Consortium

This document may not be copied, reproduced, or modified in whole or in part for any purpose without written permission from the AI4Media Consortium. In addition to such written permission to copy, reproduce, or modify this document in whole or part, an acknowledgement of the authors of the document and all applicable portions of the copyright notice must be clearly referenced.  
All rights reserved.



## 1 Ethical, societal, economic, environmental concerns & risks – a look into the future of AI for the Media

In AI4Media deliverable D2.2 “*Initial white paper on the social, economic, and political impact of media AI technologies*”, an overview of the possibilities and challenges that AI poses for the media sector across the media cycle were presented based on an extensive literature review of both industry reports and scholarly journal articles. In the white paper, the most prevalent societal concerns and risks regarding AI for media are identified, which include:

- Concerns of biases and discrimination;
- Concerns of media (in)dependence and commercialisation;
- Concerns of increased inequality to access to AI;
- Concerns of labour displacement, monitoring and profession transformations;
- Concerns of privacy, transparency, accountability and liability;
- Concern of manipulation and misinformation as an institutional threat;
- Concerns of environmental impacts of AI.

In the following subsections, these concerns are briefly summarised with a clear look towards what practises will be important to counteract potential negative societal impacts of AI for the media industry.

### 1.1 Concerns of biases and discrimination

The risk of **biases** and how such biases might lead to **discriminatory practices** remain a reoccurring concern across uses of AI in the media sector. AI is here often discussed as a double-edged sword by, on one side often being seen as tools to mitigate both conscious and unconscious biases in human judgement and decision-making, such as mitigating existing media biases. Thereby, offering positive societal impacts relating to, for example, more diversity in coverage, which could improve the public debate and political awareness of previously overlooked societal issues. On the other hand, AI is also built by humans who make decisions on what data to include in the training dataset (which might reflect existing societal biases) and how to design the AI system (e.g., by using standard algorithmic models or deciding on including certain metrics), which can replicate or even enhance existing biases by reinforcing certain ways of ‘knowing’ and ‘seeing’ in these systems<sup>1,2</sup>.

In the media sector, the implications of biases are perhaps less severe in their immediate effects compared to other sectors (e.g., law enforcement, economic or public sector) as they do not have direct consequences for the individual. However, AI could induce long term negative social and political impacts by, for example, maintaining certain **gender representations and racially discriminatory patterns** in the coverage, which could lead certain societal groups to feel underrepresented in the media landscape and disconnect from the public debate as well as produce a skewed portrayal of certain societal topics.

---

<sup>1</sup> Campolo, A. et al. (2017) AI Now Report 2017. AI Now Institute at New York University. Available at: [https://ainowinstitute.org/AI\\_Now\\_2017\\_Report.pdf](https://ainowinstitute.org/AI_Now_2017_Report.pdf)

<sup>2</sup> Littman, M.L. et al. (2021) Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report | One Hundred Year Study on Artificial Intelligence (AI100). Available at: <https://ai100.stanford.edu/2021-report/gathering-strength-gathering-storms-one-hundred-year-study-artificial-intelligence>



Equally, problematic feedback loops from recommender systems could produce the risk of very individualised and closed off media consumption patterns (i.e. “filter bubbles”), which could in the future negatively affect **political fragmentation and polarisation** or support the circulation of radicalising online content. Last, the proved biases and discriminatory effects in content moderation systems, for example on Facebook, as well as the limited availability of such tools in certain languages could prove to have highly negative impacts on the public debate online and contribute to highlighting already dominant voices and further **undermine minority voices** online (e.g., racial or gender). Long-term effects that are critical since media accounts, while increasingly contested, remain considered as representations of ‘the reality’ by many of its audiences<sup>3</sup>. AI induced biases and discrimination pose a serious long-term issue for media institutions that cannot be left for the future. Rather mitigative measures must be put in place. Here, we turn to the general discussion on how to minimise the negative impacts of biases (while of course never completely unavoidable, just as with humans).

The following practices are highlighted as **mitigative measures of biases and discrimination** as a result of AI in the media sector:

- **The need for more domain and social and/or cultural expertise in the development process of AI systems for media.** All AI projects in the media sector should strive for diversity in the team (e.g., in terms of backgrounds, ethnicities or gender) to ensure that the decisions made regarding datasets, classification or metrics are made on a well-founded and reflective basis. Critically, domain knowledge should be prioritised together with social and cultural knowledge in qualifying these decisions.
- **The need for more awareness of the biased ‘human nature’ of AI amongst media professionals and audiences.** To counteract effects of an overly belief in ‘algorithmic objectivity’, it will be important to create more critical awareness of the potential of biases in AI systems amongst both audiences who as a baseline find AI more credible and amongst media professionals who also continue to sustain the idea of neutrality in AI.
- **The need for new best practises on how to produce equitable AI use in the media sector.** Currently, the examples of AI projects promoting data justice are scarce. If the sector is to begin a conversation on ways to achieve this, examples of best practises will be needed. This could be in the form of industry research collaborations.
- **The need for sector specific open-source and non-commercial datasets for training AI systems.** As many AI projects today rely on (commercial) open-source datasets, another way to mitigate the potential negative impacts of biases in AI systems is through the development of open-source domain specific datasets, which have been critically examined by a diverse team.
- **The need for best practises and policies of ‘diversity by design’.** Currently, limited knowledge and best practices exist on how to evaluate whether, for example, a recommender system is successful – not only in a commercial sense. It would be beneficial to have best practices on how to make such decisions without benchmarking with, for example, purely commercial actors and how to include domain specific measures of

<sup>3</sup> Reese, S.D. and Shoemaker, P.J. (2016) ‘A Media Sociology for the Networked Public Sphere: The Hierarchy of Influences Model’, *Mass Communication and Society*, 19(4), pp. 389–410. doi:10.1080/15205436.2016.1174268



diversity in the projects (e.g., filling the gaps of user knowledge etc.). This could also be illustrated through concrete policies on diversity by design.

## 1.2 Concerns of media (in)dependence and commercialisation

Another concern that runs across the different media industry cycles is how the use of AI might induce an **increased commercialisation of media organisations at the expense of societal responsibility** as they become more deeply embedded into the platform economy<sup>4,5</sup>. This is not to state that media organisations have not always had a commercial side; private media organisations are a business and PSM's still need to provide legitimisation for their funding by, for example, illustrating their viewership. However, historically these two parts of media organisations have been separated, but over the last 50 years that separation has crumbled<sup>6,7</sup>. AI has proved to further intensify this classic conflict between editorial and commercial side of media organisations by, for example, pushing the limits of the data tracking practises pursued in media organisations<sup>8</sup> or by shifting power to commercial departments who more unquestioned than previously can affect how decisions are made through their knowledge of the infrastructures (e.g., AI or data)<sup>9</sup>.

Some of the potential impacts discussed regarding media organisations concerns how the increased valorisation of data and particularly audience data might **impoverish the overall media landscape by affecting what forms of media content is produced**, by for example valorising certain genres (e.g., more sensationalist content) and de-valorising content that is not 'clickable' but of societal importance – also due to the importance of content circulating well on, for example, social media or ranking high in Google News. Another way to view the potential negative impacts of particularly platformisation is discussed in relation to how it places **immense power in the hands of very few companies**<sup>10</sup> – in a European context these are the five 'tech giants' Google, Facebook, Microsoft, Apple and Amazon who predominately provide both data and other technical infrastructures for, among other, the media sector. The potential negative impacts of this power imbalance are discussed widely as it places enormous amounts of societal influence in the hands of a few commercial actors, who do not bear a societal responsibility beyond upholding legislation.

Beyond providing concrete infrastructure and providing important intermediary functions in content distribution, these tech giants are also becoming increasingly vital economic patrons in

<sup>4</sup> Lindschow, K. (2016) Exploring Digital News Publishing Business Models: A Production Network Approach. PhD Thesis. Copenhagen Business School. Available at: <https://research.cbs.dk/en/publications/exploring-digital-news-publishing-business-models-a-production-ne>

<sup>5</sup> Sørensen, J.K. and Van den Bulck, H. (2020) 'Public service media online, advertising and the third-party user data business: A trade versus trust dilemma?', *Convergence*, 26(2), pp. 421–447. doi:10.1177/1354856518790203.

<sup>6</sup> Willig, I. (2010) 'Constructing the audience: a study of segmentation in the Danish press', *Northern Lights: Film & Media Studies Yearbook*, 8(1), pp. 93–114. doi:10.1386/nl.8.93\_1

<sup>7</sup> Willig, I. (2021) 'From audiences to data points: The role of media agencies in the platformization of the news media industry', *Media, Culture & Society*. doi:10.1177/01634437211029861.

<sup>8</sup> Trow, J. (2016) *Media Today: Mass Communication in a Converging World*. 6th edn. New York: Routledge. doi:10.4324/9781315681726.

<sup>9</sup> Schjøtt Hansen, A. and Hartley, J.M. (2021) 'Designing What's News: An Ethnography of a Personalization Algorithm and the Data-Driven (Re)Assembling of the News', *Digital Journalism*, 0(0), pp. 1–19. doi:10.1080/21670811.2021.1988861.

<sup>10</sup> Bird, E. et al. (2020) The ethics of artificial intelligence: Issues and initiatives. EPRS | European Parliamentary Research Service. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS\\_STU\(2020\)634452\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf)

providing support for digital innovation in the media sector. Both Google and Facebook have developed funding schemes in support of innovation, namely the Facebook Journalism Initiative (FJP) (now Meta Journalism Project) and first Google's Digital News Innovation Fund (DNI), followed by Google's News Initiative (GNI). All promising to further the digital innovation at media organisations and ensure a sustainable future for the sector. These innovation programs further strengthen the economic and technological dependence between media organisations and the platforms or digital intermediaries and potentially threaten the *independence of media organisations*, putting at risk their societal accountability function.

The following practises are highlighted as **mitigative measures of increased dependence and commercialisation** as a result of AI in the media sector:

- **The need for responsible domain specific infrastructures to support the development of responsible AI.** Due to the high reliance on commercialised and platform infrastructures in the development of AI in the media sector, it will be important to develop alternative infrastructures, such as the European Media Data Space<sup>11,12</sup>, that perhaps better accommodate the European values and are specific to the media sector.
- **The need for a critical awareness of economic 'patrons' of the media sector and how they affect the development in the media sector.** Currently, limited research exists on the role of 'media patrons' and how they affect the future of the media sector. It will be important that more research is conducted, but also that researchers in fact can get access to these processes, as that is currently highly difficult.
- **Funding schemes oriented towards EU values.** In order to counteract the growing role of (US based) platforms in stimulating development, it will be important to develop similar funding schemes that better encompass EU values and the societal function of media.

### 1.3 Concerns of increased inequality in access to AI

Another related concern regarding AI in the media sector is the *inequalities in access to AI* solutions by users and AI infrastructure by media organisations. We start with the latter as it more directly relates to the discussion above. Charlie Beckett in a report from 2019<sup>13</sup> highlights how AI is unevenly distributed in the media sector, where particularly local and regional media with smaller budgets are lacking behind, which can reinforce the existing inequalities in the media sector. The inequalities relating to AI could further amplify this trend by further increasing the divide between local and regional as well as niche media organisations and large economically more secure media organisations. As discussed above, AI in fact holds promises to reinvigorate the 'local journalism' through the potential scaling of automated content to cover small events and sports. However, this requires that regional and local actors break the barrier of gaining access to such tools.

<sup>11</sup> European Media Data Space Project: <https://europemedialab.eu/media-policy-europes-media-lab/european-media-data-space-project>

<sup>12</sup> European Commission, Staff working document on data spaces (23 Feb. 2022): <https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces> (pp. 36-37)

<sup>13</sup> Beckett, C. (2019) New powers, new responsibilities: A global survey of journalism and artificial intelligence. Polis, London School of Economics and Political Science. Available at: <https://blogs.lse.ac.uk/polis/2019/11/18/new-powers-new-responsibilities>





Alexander Fanta and Ingo Dachwitz in a report from 2020<sup>14</sup> also show how the funding by the Google DNI Fund, at least in the German context, is oriented towards already large (commercial) media organisations, while smaller start-ups, niche or non-profit organisations are less funded, illustrating how they perhaps selectively stimulate the innovation, making the rich richer, rather than diversifying the access to AI. Again, illustrating how this powerful intermediary elite can stimulate the access to AI in a certain direction, amplifying existing inequalities in the sector<sup>10</sup>. This increasingly uneven access to AI could have serious social and economic impacts in society by **diminishing diversity in the available media offering** as certain media organisations unable to leverage the power of AI might be unable to remain competitive.

We now return to the first point mentioned, namely how access to AI services by users is also **highly unequal across languages**. Training data, for example, often only exist in English or tools are predominately being developed for English. The tools available in other languages also often perform significantly worse due to inherent biases in the datasets. The benefits of AI are, therefore, not shared across the globe, and particularly the divide between the global North and South is growing with the increased use of AI across all sectors and therefore also when it comes to media<sup>497</sup>. The social and political implications of this are vast regarding several AI applications for media, for example, content moderation in large diverse countries like India will suffer under this as they generally will have less efficient content moderation and it is to be expected that many minority languages will experience worse performance, potentially keeping them from partaking in the public debate. There will be a need to place increased focus on this inequality of access to ensure that media diversity is sustained and that the benefits of AI become shared, not only by those who are already in privileged positions.

The following practises are highlighted as **mitigative measures against the increased inequality in access to AI** in the media sector:

- **The need for funding schemes and initiatives focusing on media diversity.** It will be important to counteract the trend in private funding schemes where established media organisations remain the main beneficiaries of funding for innovation. To not further the increasing competitive divides in the media sector, funding should be specifically oriented towards furthering media diversity.
- **The need for an increased focus on global AI divides and their consequences.** In general, more knowledge is needed on the severity of the AI divide between the global North and South. It will be important to explore the extent of the issue and its implications further.
- **The development of AI models for diverse languages or language adaptive models.** In order to improve the overall access to AI benefits, AI models for large foreign and minority languages should be developed together with adaptive models that can be more efficiently reused for other languages.

#### 1.4 Concerns of labour displacement, monitoring and profession transformations

One of the most discussed impacts of AI has been regarding the labour market and the **prospect of mass job loss** when tasks become increasingly automated<sup>1</sup>, with sometimes very high

<sup>14</sup> Fanta, A. and Dachwitz, I. (2020) Google, the media patron. How the digital giant ensnares journalism. preprint. SocArXiv. doi:10.31235/osf.io/3qbp9



estimates of the job losses to be expected, such as 38 million people in the US being in high risk of losing their job because of automation<sup>10</sup>. While the fear of displacement has become more nuanced, both since full automation of many jobs still lies far in the future<sup>1</sup> and as studies have shown that while some jobs will disappear, others will emerge with the growing AI industry<sup>10</sup>. There is also a palpable fear of displacement in the media sector and a few examples of how AI had in fact led to layoffs of ‘human’ staff. However, in the media sector the fear of displacement discussion has also been nuanced to focus more on the changes AI might impose on the profession and how to maintain the legitimacy of the profession.

One of the impacts of AI has been an emphasised **focus on ‘technical’ or ‘data oriented’ media professionals**<sup>15,16</sup>, so an upskilling, rather than displacement, but also how the technology and data focus is increasingly legitimised through managerial shifts<sup>17</sup> and results in media professionals becoming more ‘disposable’ compared to employees with technical skills who are often much harder to recruit<sup>9</sup>. This is, therefore, producing new asymmetries in the labour market of the media sector, where certain types of jobs and skills are being devaluated. This might require new policies (economic and social) to address this potential societal gap created due to AI<sup>1</sup>.

Beyond, impacting what skills are considered important, AI has also been a contributing factor, as part of the overall **datafication of the media sector**, to the increased importance of data in performance evaluations and recruitment processes, as highlighted by Angele Christin<sup>18</sup>. In the AI Now Institute report from 2017, this is highlighted as a general negative social impact of AI, as the increased reliance on such, often non-transparent, tools have impoverished the working conditions in many places by, for example, **amplifying the power asymmetry between employer and employee** or by impacting recruitment processes<sup>1</sup>. These professional changes can negatively affect mental health through, for example, dynamic data visualisations and reminders of goals, which might place stress on the individual to perform or even overperform, due to its importance for keeping one’s job<sup>1</sup>. These systems are also generally developed with the employer and not employee in mind, placing the impacts mainly on the individual<sup>19</sup>. Currently, little research is available on the degree of this problem in the media sector, where the focus has more been more on how AI affected production and distribution patterns of media content and not on how AI-enhanced datafied work practices affected media professionals.

The following practises are highlighted as **mitigative measures against the concerns of labour displacement, monitoring and profession transformation** in the media sector:

- **The need for more research and policies addressing potential displacement patterns as a result of AI.** As the increased reliance on AI might result in certain jobs disappearing (e.g., routine tasks) in the media sector as well as across other sectors, providing a societal

<sup>15</sup> Lewis, S.C. and Usher, N. (2013) ‘Open source and journalism: toward new frameworks for imagining news innovation’, *Media, Culture & Society*, 35(5), pp. 602–619. doi:10.1177/0163443713485494.

<sup>16</sup> Lewis, S.C. and Usher, N. (2014) ‘Code, Collaboration, and the Future of Journalism’, *Digital Journalism*, 2(3), pp. 383–393.

<sup>17</sup> Young, M.L. and Hermida, A. (2015) ‘From Mr. and Mrs. Outlier to Central Tendencies’, *Digital Journalism*, 3(3), pp. 381–397. doi:10.1080/21670811.2014.976409.

<sup>18</sup> Christin, A. (2018) ‘Counting clicks: Quantification and variation in web journalism in the United States and France’, *American Journal of Sociology*, 123(5), pp. 1382–1415.

<sup>19</sup> Crawford, K. et al. (2019) *The AI Now Report 2019*. AI Now Institute at New York University. Available at: [https://ainowinstitute.org/AI\\_Now\\_2016\\_Report.pdf](https://ainowinstitute.org/AI_Now_2016_Report.pdf)



problem of unemployment. It will, therefore, be important that societal mechanisms and policies are developed to handle the citizens who will be left jobless and in need of specific upskilling.

- **The need for an increased focus on data and AI in media education.** The changes in the media professions also require action from the educational sector who must support students in developing the right skills for the labour market, including increased skills in data science and in understanding how AI systems work.
- **The need for more research on how AI is changing labour conditions and power asymmetries in the media sector.** It will be important to understand how the introduction of AI is enhancing already increasing workplace asymmetries, for example, through the use of performances measurements and with what impacts on the individual and society.

#### 1.4.1 Privacy, transparency, accountability and liability

A plethora of new concerns regarding AI relate to the **users' rights to both privacy and transparency** in 'who' they are interacting with, but also to how the introduction produces new questions of **accountability and liability** was brought forward in the above review. A discussion that is also echoed in the wider discussions of AI<sup>20,21,1,19,10</sup>. The potential social or economic impacts that tracking practices might have on individuals have been highly discussed, for example, in relation to how facial recognition technologies, can allow the identification of individuals across contexts - and even their moods and sexual orientation<sup>10,20</sup>, raising questions of individuals' right to privacy in public spaces. Or when the first cases of people being fired based on, for example, GPS data or when data from a pacemaker was used to geographically locate a citizen, leading to his conviction of arson, as a court case recently set precedence for<sup>10</sup>. All these uses of the increasing amounts of trackable personal data have raised serious questions of how to protect the data rights of individuals.

Much like the above discussion on biases, the individual impacts of the data tracking practices related to the use of AI in the media sector, might be much less severe than the examples given above. However, **untransparent and potentially excessive tracking practices** by media organisations could have fatal consequences for the trust in these organisations, impeding them from fulfilling the societal task of providing public information. It will, therefore, be important to increase transparency in data use by media organisations, where an overview of the use of data is made easily available and more important the user can react based on this data and retract their consent for certain uses.

Article 52 from the upcoming AI Act points to the right of individuals to know whether they are interacting with an AI system. This is again a wide-reaching discussion that both emphasises the right to know if it is an AI, but also, for example, in the context of healthcare, whether one can request not to have an AI involved in the process of diagnosis<sup>22</sup>. For the media sector, as

<sup>20</sup> Whittaker, M. et al. (2018) AI now report 2018. AI Now Institute at New York University New York. Available at: [https://ainowinstitute.org/AI\\_Now\\_2018\\_Report.pdf](https://ainowinstitute.org/AI_Now_2018_Report.pdf)

<sup>21</sup> Ada Lovelace Institute (2021) Algorithmic accountability for the public sector: Learning from the first wave of policy implementation. London, UK: The Ada Lovelace Institute. Available at: <https://www.adalovelaceinstitute.org/report/algorithmic-accountability-public-sector/>

<sup>22</sup> Ploug, T. and Holm, S. (2020) 'The right to refuse diagnostics and treatment planning by artificial intelligence', Medicine, Health Care, and Philosophy, 23(1), pp. 107–114. doi:10.1007/s11019-019-09912-8.



discussed, the disclosure practices are highly differentiated amongst media organisations, illustrating the need for more harmonisation of how media organisations should approach this new challenge. This includes **transparency in disclosing when an AI has been involved in the process** of producing or curating content, but also in how the system came to its decision. This will be important for users, citizens, media and researchers to be able to **hold the systems accountable** in cases of harmful decisions.

Another related question that is highly discussed is the question of **liability regarding AI systems**, because one thing is disclosing that an AI, for example, produced a piece of content, but it is another to determine liability, because many new actors are now involved in this question (e.g., external service providers, in-house developers etc.). Currently, there are still no clear policies or guidelines on this question, which could have negative impacts on media organisations or individual media professionals. Equally, as Lewis et al.<sup>23</sup> point to, there is also a risk that the current regulation will allow a loophole for AI produced content in the case of, for example, personal deformation suits (i.e., lawsuits regarding a false statement made about a person). It will, therefore, become highly important to develop more clarity for media organisations on how to act on this question, and how to translate any legal accountability obligations into organisational practices and internal divisions of responsibility between editors, journalists, data, and economic departments.

The following practises are highlighted as **mitigative measures against the concerns of privacy, transparency, accountability and liability** regarding AI in the media sector:

- **The need for more best practises of responsible data practices in the media sector.** As the extensive use of data continues to grow in the media sector, it will be vital that new best practises are developed to support responsible data strategies that protect the rights of the individual.
- **The need for best practices and policies regarding disclosure of AI systems for the media sector.** As the question of who produced or curated an article is no longer limited to, for example, journalists, editors, and producers, it will be vital that new guidelines for how to disclose the utilisation of AI in these processes are developed to protect the individual's right to transparency.
- **The need for explainable AI solutions that can help users understand how the AI system works and makes its decision.** As users increasingly are partly serviced by AI systems in their media experience, it is important that they have access to understandable explanations of what the system does and on the basis of what data, to uphold their right to, for example, object to the way the decision was made.
- **The need for clearer regulation and guidelines on the liability question regarding AI.** There is a need to help media organisations navigate the liability question that arises from the use of AI systems.

## 1.5 Manipulation and misinformation as an institutional threat

<sup>23</sup> Lewis, S.C., Sanders, A.K. and Carmody, C. (2019) 'Libel by Algorithm? Automated Journalism and the Threat of Legal Liability', *Journalism & Mass Communication Quarterly*, 96(1), pp. 60–81. doi:10.1177/1077699018755983.



There is also a growing concern amongst the media organisations regarding **manipulation of content and misinformation**. While this was not related specifically to their own work, the negative impacts of the growing amounts of misinformation were seen as highly detrimental to the **trust in the media sector**, as for example evident in the survey by Georg Rehm<sup>24</sup>. Making it an institutional threat to the existing media landscape, whose legitimacy is increasingly contested as part of this development. Equally, the fact checking genre and independent fact checking institutions as a result of this growing problem also become a new part of the media landscape. This discussion is, therefore, also slightly different as many of the AI systems that are utilised to mitigate such misinformation are developed by social media platforms or to assist fact checking organisations.

The increasing focus on removing misinformation with the assistance of AI systems also raises important discussions regarding **freedom of expression**, as new guidelines for appropriate forms of censorship must be discussed as well as the potential risks of false positive and negatives in these processes and the lack of complaint mechanisms or satisfactory explanations of why content was deleted<sup>25,26</sup>. Equally, as the practice of fact checking, and particularly AI assisted fact checking grows, these practises must also be more explored, as this remains a highly subjective practice, but which is gaining societal importance. Here both the need for more transparency in the workings of the AI systems used to identify misinformation will become important, particularly as they become intertwined with fact checking organisations through strategic partnerships, such as the ones initiated by Facebook. The topic of disinformation is discussed in the section “*AI for counteracting disinformation*” of this Roadmap.

The following practises are highlighted as **mitigative measures against the increased threat of manipulation and misinformation** for the media sector:

- **The need for mitigative and adaptive AI systems to counteract misinformation.** To protect the legitimacy of media organisations and the integrity of the online deliberative spaces, it will be important to develop AI systems to assist in content moderation and fact checking efforts. These must be highly adaptive to be effective and counteract adversarial tactics by groups who spread misinformation.
- **The need for more transparency in moderation systems and AI fact-checking systems.** Currently, the AI systems used to identify misinformation on social media platforms remain non-transparent in their workings and the people who experience consequences do not always have access to a satisfying explanation of why, for example, their profile was deleted or to a complaint mechanism. As many fact checkers are today part of strategic partnerships with Facebook, the need to be transparent will become even more important to sustain legitimacy in these institutions that now serve an important societal function.
- **The need for more knowledge on fact-checking as a social practice and its effects in the deliberate landscape.** As fact-checking becomes an important societal function in societies,

<sup>24</sup> Rehm, G. (2020) The use of artificial intelligence in the audiovisual sector: concomitant expertise for INI report : research for CULT Committee. LU: EU Publications. Available at: <https://data.europa.eu/doi/10.2861/294829> (Accessed: 24 November 2021).

<sup>25</sup> Llansó, E. et al. (2020) Artificial intelligence, content moderation, and freedom of expression. Annenberg Public Policy Center of the University of Pennsylvania.

<sup>26</sup> Gillespie, T. (2020) ‘Content moderation, AI, and the question of scale’, Big Data & Society, 7(2), p. 2053951720943234. doi:10.1177/2053951720943234

it will be important to gain more in-depth knowledge in how they construct ‘factual’ accounts as well as what the consequences of potentially countering epistemologies of the truth might mean for the deliberative space and societal polarisation.

## 1.6 Environmental impacts of AI

A concern that has not really been touched upon but is becoming increasingly important in the more general discussion of AI is the environmental impacts of AI. This was for example highly on the agenda at the 2021 Global Partnership of AI summit in Paris. Training AI systems is ‘computation heavy’ and leads to large amount of carbon emissions; the storage of data is also an energy intensive process while the hardware needed requires large amounts of natural resources<sup>10</sup>. While AI is also considered to be one of the key solutions to dealing with climate change by providing the basis for ‘smarter’ solutions across a range of ‘carbon heavy’ industries (e.g., agriculture, energy etc.), the impacts on the environment from AI also needs to be further understood so that it is possible to make sensible decisions on when the benefits of AI will surpass the impact on the environment. This remained limitedly present in the current discussion of AI for media but should increasingly be addressed by media organisations who embark on AI projects. The negative environmental impacts also highlight the need for further open-source AI models that can help more than one media organisation. The issue of the environmental impact of AI is further discussed in the section on “*Climate crisis and AI*” of this Roadmap.

The following practises are highlighted as **mitigative measures against the concerns of environmental impacts of AI** in the media sector:

- **The need for best practises in mitigating the environmental impacts of AI in the media sector.** To help media organisations consider the environment, some guidelines or best practises should be developed to help guide their development processes.
- **The need for more open-source AI models for media.** To minimise the environmental impacts of AI, more collaborative and open-source AI projects should be developed across the media sector.

## 1.7 Summing up

As shown above, there is a need to take action to prevent potential negative societal impacts from affecting society when developing and deciding upon AI applications for the media sector. This requires mitigative actions that span from developing funding schemes that can induce the development of responsible and value sensitive AI to concrete initiatives in the processes of development and implementation, which must be considered in future AI applications and the surrounding discussions for the future of the sector.





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

[info@ai4media.eu](mailto:info@ai4media.eu)

[www.ai4media.eu](http://www.ai4media.eu)