Trend Maker AI – An Overview On Bias and Other Risks to Our Information Consumption Habits

Caius Tudor Luminosu
UPT – Politehnica University Timisoara, Romania
caius.luminosu@upt.ro

Cosmin Constantin Baias
UPT – Politehnica University Timisoara, Romania
cosmin.baias@upt.ro

Anca Draghici
UPT – Politehnica University Timisoara, Romania
anca.draghici@upt.ro

Abstract

As artificial intelligence (AI) has become a major trend in research as in almost all other areas of our lives, it is necessary to start to study to what measure this trend is creating new trends that will influence us in ways we are just beginning to realise.

This paper is meant to provide an overview of how much AI is taking up the role of a trend maker of the media we consume, a curator of the information this media is created from and an outlook on the risks derived from this. The use of the internet in our information intake has more clearly than before shown how we select the media we derive this information from as well as amplified some negative habits regarding how we handle media and information, such as getting trapped in echo chambers and overlooking bias if it fits our own convictions or worldview.

AI is the next evolutionary step in handling information. Its high capabilities in this business hold the risk of amplifying these negative habits while also making us less conscious of their existence, thus increasing their negative effects on people's lives and on society itself. As we will try to show, there are some ways to positively influence the development and deployment of AI systems to obtain a more favourable outlook.

Keywords: AI, bias, media, echo chamber, principles for AI

INTRODUCTION

More and more news about the real or presumed capabilities of AI systems has made headlines since the beginning of the second decade of the 21st century, so it becomes increasingly important to have a clear picture of these digital products. Some of these are already in commercial use, while a lot of others are experimental but open to the public. DALL-E, ChatGPT, GitHub Copilot, and Sora are just some intense mediatised examples due to their spectacular and highly visible results. Their effects on how we interact, communicate, and work with one another are impossible to predict right now as the technology is relatively new. It is only certain that such revolutionary technologies will have a definite impact on our society. This raises the issue of regulation.

The European Union (EU), as one of the largest and most effective international organisations, has started to approach the topic of AI, as it regards this set of technologies as potentially beneficial for its citizens. This is reflected in the proposed so-called EU AI Act, by which the Union is trying to create legislation to regulate the production and functioning of this technology to produce the desired effects on society. The EU does not stand alone, however, as the United States of America has already enacted legislation pertaining to the topic, and other relevant states, like P. R. China and Japan, are developing their own regulatory approaches.

Defining AI

To define AI, the EU proposes a technology-neutral and broad definition to encompass future development that is unforeseeable now. Art. 3 para. (1) of the AI act states that "'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches […] and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with".

Looking at the legal definition AI is considered software that can generate results which have an impact on human communication and thus on the way society works. Therefore, it is vital that we learn how this impact looks like in the media we consume.

Figure 1: Types of AI
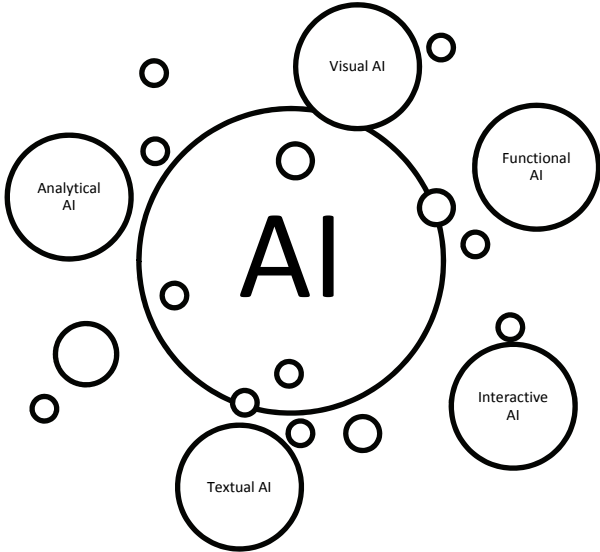
Adapted from and see classification by Sarker, 2022

There are many types of AI systems currently in use and development. The above figure provides for orientation only a general classification (Sarker, 2022). In practice, to provide applicable solutions, combinations of various AI types and AI techniques such as Machine Learning, Deep Learning, Natural Language Processing, and others are deployed. In typical interaction, users will thus encounter more than one type of AI, but it is of no concern to them, as the effects are the important aspect. Also, legal definitions are formulated to encompass all types of AI systems.

Uses of AI in (digital) media

The use of AI pertains naturally only to media available in digital format. Media that is not accessible in digital form will be referenced only by AI systems. However, due to the ever-increasing digitalization, it seems likely that in future, the overwhelming part of media consumed will be digital and, as such, susceptible to AI use.

There are several advantages of AI use in media. These are mainly of an economic nature, as the development and deployment of AI systems are rather costly. Some advantages of AI use in digital media are:

- AI Improves marketing and monetization efficiency.
- AI helps to keep track of consumer trends by analysing user behaviour to place ads and recommendations.
- Legal compliance
- AI helps remove offending information and is used for bias control.
- AI improves user experience and interaction in digital media.

AI has become extremely valuable in digital marketing, and optimising AI tools is a competitive advantage. AI helps to understand customer needs and interests and enables companies to act very quickly based on these findings as well as to understand user behaviour across different platforms. By this, they can match customer interests with appropriate content. Also, companies must be able to interpret large amounts of data generated by users to enable personalization and to enhance user experience as fast as possible (Alkhayyat & Ahmed, 2022)

AI systems can fulfil different roles regarding media. The use cases usually require custom-made solutions that help with market differentiation for the enterprise that deploys the system. Multiple tasks in different roles are normally filled by the use of AI. Due to the multitude of use scenarios for AI systems, we have selected three use cases as examples. However, many others can be found, especially in news content and video streaming platforms, such as Youtube.

Use-case 1: In social media, depending on the market objectives of the media product, different types of AI fill different roles. In Figure 2, Rangaiah from the market research company Analyticsteps.com, provides an overview of the use of AI in three major and widely used social networks, each acting in a different market: Facebook, Twitter and Linkedin and others.

Figure 2: Uses of AI in social media
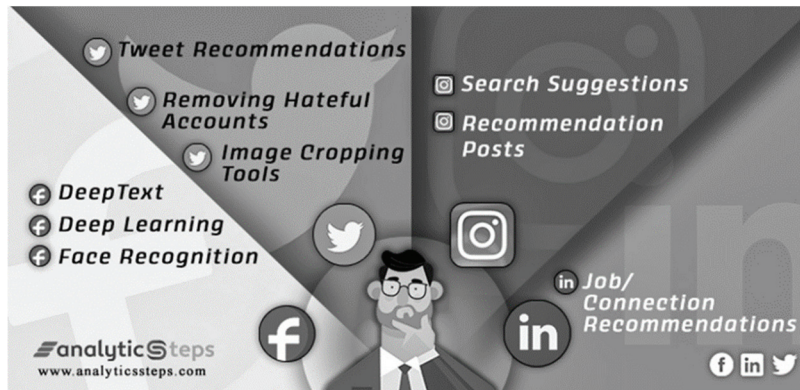
(Source: Analyitcsteps.com, Rangaiah, 2021)

Use-case 2: "Automated journalism"

More and more news articles meant to convey information or mere facts on new events are written by AI. These systems are being employed by news outlets, such as DPA, Reuters etc. and provide the advantage of speed and optimal use of human resources.

Also, there is a growing market for tools for journalists: AI recommends visual content used in articles (such as Panels by Getty Images). Tools for data visualization are employed where there is a need to have clear overviews of data: AI creates interactive data visualization, placed directly into relevant articles. When interacting with users, chatbots are used to curate forums and to sort out offending posts.

Use-case 3: media consumption, streaming of media

Large streaming platforms (Netflix, Spotify and others) employ AI systems for content personalization, generating recommendations and managing user data. AI is also used for the automatic generation of subtitles and tagging of content.

Media bias

Media bias is a term that has yet to obtain a unanimous definition. (Lichter, 2018) For the purpose of this chapter, we can define bias as a phenomenon that "relates to outcomes which are systematically less favourable to individuals within a particular group and where there is no relevant difference between groups that justifies such harms". (Turner Lee et al., 2019) This definition does refer to all types of bias, including the one found in media. When referring to media, we include all types of digital media, not only news but all forms of journalism and media that can be found in the digital realm.

Why is media bias of interest? It is digital media that is shaping public opinion and, by extension, the opinions and beliefs of individuals, thereby having an increasingly great influence on society itself. In societies that rely on the public expression of opinion, such as democracies, it is vital that the shaping of opinion is facilitated by access to unbiased information.

Media bias comes in different forms, depending on many factors and contexts, such as news aggregators, recruitment tools, search engines, and facial recognition – including uses in marketing, ads personalization, and others. (Turner Lee et al., 2019) Initial studies focused on media bias from a journalistic point of view, analysing only news media. However, the findings related to this type of media can be easily interpreted in the context of other forms of media. Many definitions of media bias

types exist (see D'Alessio & Allen, 2000), such as intentional and unintentional, ideological or spin (reporting bias), and coverage, gatekeeping and statement being widely accepted. (Hamborg et al., 2018; Cowell, 2023) Figure 3 is intended to provide a short overview of the scope of this introduction to media bias.

Figure 3: Types of media bias

Types of media bias

Placement or Selection Bias

Layout Placement

The editorial staff decides the importance of a topic by its placement in an article. A story can be "buried" by placing it in a section that is less read.

Commercial or Selection

Editors select stories that draw larger audiences of readers in order to meet sponsor demands.

Sensationalism Selection

Media focus on stories that emphasize fear, anger, and excitement.

Temporal Placement

Editors give follow-up and clarifying stories less prominent placement because news is expected to be current and timely.

Visual Selection

The selection of images can skew audience perception of a story's importance and the events reported; therefore, visuals are used to attract readers' attention.

Reporting Bias

Reporting biases occur when an article is written with a particular tone or "spin" so that readers will perceive it in a certain way without applying skepticism or comparing the piece to other news outlets with a different ideology or perspective.

Situational Bias

Complexity

A bias that occurs when an article is written using the who, what, when, where, why and how rubric. However, in real life there are complexities that may not conform to those guidelines.

Geographic

This bias is seen when an article factors in the diversity of story through cultural and social issues. Readers located outside of that locale may have different reactions to the same story.

Definition

A bias reflected by the way in which words take on different meanings depending on the context of use and the background/culture of the reader.

Stereotyping

This bias can be used to understand groups and situations which are not a regular part of our lives. However, classifying and categorizing people or events can affect that way in which a story is perceived.

Source: Cowell, 2023

Hamborg et al. present an overview (see Figure 4) of the motives and sources of media bias reported to different types of media bias stemming from the different phases of media creation. Knowledge of these motives can help the consumers of media, the users, to identify bias and interpret the information from their own perspective. Many external and internal factors of the media creator influence reporting or spin bias, as well as placement or selection bias and situational bias. As stated above, knowledge of the existence and type of potential bias encountered can be useful for the users of media to reflect on the information provided and the context to create an opinion. Especially in the context of a democratic society, it is vital that members of the public, the citizens, can correctly process the information provided and be able to create their own rational opinions.

Figure 4: Motives and sources of media bias
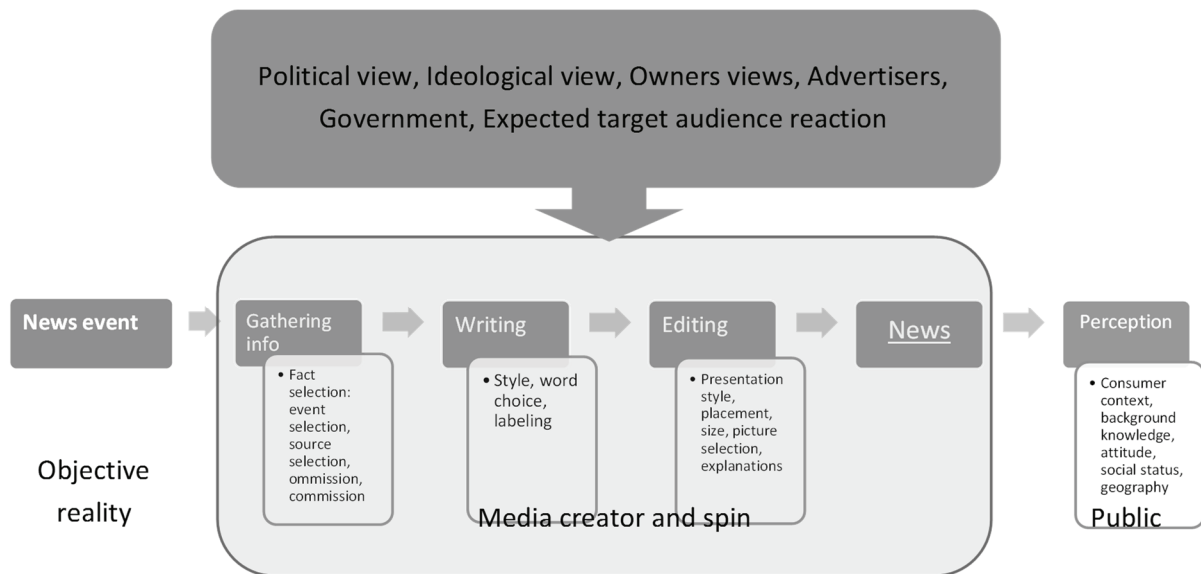
Adapted from Hamborg et al., 2018.

AI use in controlling bias

The previous sections provide an overview of the extensive and growing use of AI systems in the way we perceive information by consumption of different types of media. Some types of AI can prove useful on the side of media creation, while others can also be used in controlling bias and other unwanted phenomena regarding media. Natural language processing and other types of AI that can analyse text inputs can prove very useful for removing offensive or illegal content. Such systems also help combat plagiarism.

Hamborg et al. (2018) analyse and propose different AI-based approaches to controlling media bias, referring to different phases of media coverage, from event selection to editorial processes.

At the event selection phase, "automated approaches need to (1) find articles relevant to the question being analysed (we describe relevant techniques later in this subsection, see the paragraphs on news aggregation), (2) link articles to baseline data or other articles, and (3) compute statistics on the linked data." (Hamborg et al., 2018). This can be done by event detection, topic modelling, document clustering, and news aggregation.

At the source selection phase, human intervention is still crucial, as there are few AI-based approaches to the issue of the trustworthiness of sources. Anti-plagiarism-related software is one such solution. On social media, the users select their own sources that reflect their own opinions, thereby increasing the risk of creating echo chambers. "An echo chamber is an environment where a person only encounters information or opinions that reflect and reinforce their own. Echo chambers can create misinformation and distort a person's perspective, so they have difficulty considering opposing viewpoints and discussing complicated topics. They're fuelled in part by confirmation bias, which is the tendency to favour info that reinforces existing beliefs." (GCFGlobal 1, 2023). Regarding commission and omission, there seem to be no specific AI systems able to discern the redactional bias or "spin" of media. (Hamborg et al., 2018) It remains to be seen if large language model systems, such as GPT, could provide some support in this field. At the present time, some useful approaches to combat media bias are still being worked on in computer sciences but are still to be implemented. "Methodologies and models of media

bias in the social sciences can help computer scientists make the automated approaches more effective. Likewise, the development of automated methods to identify instances of specific forms of media bias can help make content analysis in the social sciences more efficient by automating more tasks." (Hamborg et al., 2018)

Turner Lee et al. propose a self-regulatory approach consisting of creating and implementing ethical and legal frameworks, not only by media creators but also by algorithms, respectively, AI creators. Such a framework would be enforced by audits and public bias statements, thus achieving what the authors call algorithmic hygiene (Turner Lee et al., 2019).

Risks of AI use in media

Scenarios of risks in AI are very numerous to imagine. Still, being aware of these risks is crucial for both those who employ and use such systems as well as for consumers of media. Some risks are considered almost unavoidable now, such as filter bubbles. It is worth mentioning that AI technology is relatively new, and no exhaustive research on the negative aspects of AI use has been published yet.

Algorithms keep track of what media products users access. Then, these algorithms propose to users content. By repeating this process, more similar and more likely-to-be-accessed content is displayed. "This process can lead to the creation of a filter bubble. Being in a filter bubble means these algorithms have isolated [the user] from information and perspectives [the user hasn't] already expressed an interest in, meaning [the user] may miss out on important information" and any information that is deemed irrelevant to them. (GCFGlobal 2, 2023)

Another issue is that AI bias control can easily transform into opinion censorship by commercial or state actors. The People's Republic of China Social Credit System, a system implemented by an authoritarian regime to keep track, reward and punish the behaviour of its citizens in every moment of their lives, is heavily reliant on AI systems. Censorship and perpetual monitoring of media and media consumption are made possible only by using such tools. (see Liang et al. Artificial Intelligence +, 2021)

Misuse, malicious use, or abuse of data by means of AI tools by organized crime groups and state actors, including, but not limited to, deep fakes, manipulation through fake news, and automatic sanctioning after permanent monitoring and control are also problems that are being enabled by AI. (see Liang et al.) Online organized crime groups are already exploiting AI tools in order to produce fake emails and fake information in different languages, by means of which they are able to commit frauds, identity thefts, steal data, extort ransom for hijacked systems and other cybercrimes. (see Trend Micro Research et al., 2020)

Even unintentional errors, such as errors in AI filtering of content can lead to serios consequences, not only loss of comfort (e.g. companies locking out users after accidentally tagging uploaded cloud content as offending).

On a psychological level, excessive use of AI at the user level can create detrimental effects such as opinion polarization and echo chambers, both of which affect the behaviour of the individual. Such problems can also severely affect society, especially when the functioning of the democratic state depends on the formation and expression of individual opinions. (see Einav et al., 2022)

Mitigating the risks of AI – regulation and self-regulation

Presenting the numerous advantages and undeniable benefits of AI use leads us to the discussion of the potential risks and pitfalls of AI. As with any tool created by man, the presented negative sides of these systems are not inherent but the result of intentional or unintentional misuse, unreflected or unregulated use. The main approach to tackling these issues could be, above all, regulation (Krönke, 2020). Regulation is the first step, which creates a framework in which the stakeholders will find it easier to identify and apply the proper guidelines that pertain to their use scenario.

Some states and international organisations have started to work on guidelines on which to base regulations, presumably by laws. At present, there is no international agreement or treaty on the use of AI. It is likely that when such a treaty or convention is negotiated, it will pertain to potential military use of AI systems and not to civilian use cases. This will leave important aspects of the use of AI unregulated, mainly in communication, and will tend to neglect the importance of communication and thus of potential projection of soft power – influence and manipulation.

The OECD adopted its AI principles in May 2019. Other related initiatives, like that of the G20 group of states, followed:

- Inclusive growth, sustainable development, and well-being
- Human-centred values and fairness
- Transparency and explainability
- Robustness, security, and safety
- Accountability

The OECD recommendation for its member states is that "Governments should work closely with stakeholders to prepare for the transformation of the world of work and of society. They should empower people to effectively use and interact with AI systems across the breadth of applications, including by equipping them with the necessary skills." (OECD, principle 2.4).

On a state level, the US has pioneered the field and laid the groundwork for an encompassing regulatory framework by the National Artificial Intelligence Initiative (NAII), established by the National Artificial Intelligence Initiative Act of 2020 (NAIIA), that entered into force on January 1st 2021. By this act, an institution called the National Artificial Intelligence Initiative Office (NAIIO) has been established and "tasked to:

- Provide technical and administrative support to the Select Committee on AI (the senior interagency committee that oversees the NAII) and the National AI Initiative Advisory Committee;
- Oversee interagency coordination of the NAII;
- Serve as the central point of contact for technical and programmatic information exchange on activities related to the AI Initiative across Federal departments and agencies, industry, academia, non-profit organizations, professional societies, State and tribal governments, and others;
- Conduct regular public outreach to diverse stakeholders; and
- Promote access to technologies, innovations, best practices, and expertise derived from Initiative activities to agency missions and systems across the Federal government."

As shown above, the creation of the NAIIO as a support and coordination institution is a key part of an entire state structure meant to oversee the implementation of the provisions of the national AI legislation.

A similar approach, albeit at a supra-statal level, is being taken by the EU. The Union has created at the initiative of the European Commission, the EU Ethics Guidelines for Trustworthy AI principles (EU, 2023):

- "Human agency and oversight,
- Technical robustness and safety,
- Privacy and data governance,
- Transparency,
- Diversity, non-discrimination, and fairness,
- Environmental and societal well-being,
- Accountability."

Due to its supra-statal nature, the EU is moving at a slower pace than the US, as it needs to initiate and moderate a debate between representatives of the member states regarding AI legislation. This is very useful, as the use of AI transcends borders. This is also the main argument for creating international regulations for AI use. However slim the chance for this might be now. As of March 2024, the EU AI act mentioned here has passed the EU Parliament in its first reading. This will be followed by proposals by the Commission and the Council before final adoption.

A complementary approach to state or international regulation is self-regulation by the main stakeholder that creates and operates AI systems. At the present time, most of the large companies involved in the development and use of AI have adopted their own guidelines or principles. The main difference between regulation by state authority and self-regulation is in the very nature: the former is done by legislation, while the latter is of an ethical nature and thus not binding and subject to modification at any time. Legislation can be modified and adapted but only by procedure, which in democratic states involves the expression of the will of the citizen. Setting a framework by legislation, which also means a minimum standard for self-regulation by the developers and users of AI, would be essential for the functioning of democratic societies.

As an example, we include the Microsoft responsible AI principles (Microsoft, 2023)

- "Fairness - AI systems should treat all people fairly.
- Reliability & Safety - AI systems should perform reliably and safely.
- Privacy & Security - AI systems should be secure and respect privacy.
- Inclusiveness - AI systems should empower everyone and engage people.
- Transparency - AI systems should be understandable.
- Accountability - People should be accountable for AI systems."

Another company that invests in the development and use of AI is Google. They "will assess AI applications in view of the following objectives". AI should be:

- "Be socially beneficial.
- Avoid creating or reinforcing unfair bias.
- Be built and tested for safety.

- Be accountable to people.
- Incorporate privacy design principles.
- Uphold high standards of scientific excellence.
- Be made available for uses that accord with these principles."

Also, the consulting and IT services company Accenture encourages the industry to "design and deploy responsible AI solutions that are ethical, transparent, and trustworthy". "Responsible AI enables the design, development and deployment of ethical AI systems and solutions. Ethical AI acts as intended, fosters moral values and enables human accountability and understanding. […] fundamental criteria include soundness, fairness, transparency, accountability, robustness, privacy, and sustainability." (Accenture, 2023)

Conclusion

AI is not a trend, even if it is trending in the news, but a tool that is rapidly being adopted by industries that create media that we consume. As media is information, the problem of bias and other unwanted or potentially dangerous effects on society arise to attention. AI can be used for bias control, but it can also be a tool that amplifies bias (see echo chambers). There can be many forms of digital, AI-enhanced bias in recruitment tools, search engines, and facial recognition – including uses in marketing and ad personalization. All forms of bias can be reduced or avoided by implementing ethical and legal frameworks, but also self-regulation of algorithm creators (regular audits, bias statements) and education of both the public and algorithm creators and operators, achieving algorithmic hygiene (Turner Lee et al., 2019)

It becomes evident that all stakeholders, meaning first-place states and, private business actors, the media-creating industry, are increasingly aware of the need to develop and enact regulations regarding the development and use of AI. Some states and most companies involved have already enacted or proposed frameworks of guidelines. It is to be observed that state regulations will set the limits and private actors can choose to improve upon them by means of ethical self-binding principles. International regulations of AI, while very useful, as software services cross state boundaries, are still pending and are rather unlikely to be created at this point. Until such time has arrived, the concept of algorithmic hygiene and AI ethical principles should be explored on a self-regulatory basis. Lastly, the education of media consumers is probably more important than ever, but this is to be explored by further research, as it lies outside the confines of this paper.

Acknowledgements

REFERENCES

Alkhayyat, A., Ahmed, A., (2022). The impact of artificial intelligence in digital marketing, https://www.diva-portal.org/smash/get/diva2:1663148/FULLTEXT01.pdf Accessed 15.02.2023

Artificial Intelligence + (2021). China's 'Social Credit' System, https://www.aiplusinfo.com/blog/chinas-social-credit-system/ Accessed 15.02.2023

Clarke, R. (2019). Principles for responsible AI. https://tech.humanrights.gov.au/sites/default/files/inline-files/4A%20-%20Roger%20Clarke.pdf Accessed 15.02.2023

Cowell, L., HUHS Library Media Center (2023). Media Bias Uncovered: Overview http://libguides.huhs.org/mediabias Accessed 15.02.2023

D'Alessio, D., Allen, M.: Media bias in presidential elections: a meta-analysis. J. Commun. 50, 133–156 (2000)

Einav, G., Allen, O., Gur, T., Maaravi, Y., Ravner, D. (2022). Bursting filter bubbles in a digital age: Opening minds and reducing opinion polarization through digital platforms, Technology in Society, Volume 71, 2022, 102136, ISSN 0160-791X, https://doi.org/10.1016/j.techsoc.2022.102136

European Union, Digital Single Market, Ethics Guidelines for Trustworthy AI, available for download from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai Accessed 15.02.2023

GCFGlobal 1 (2023). What is an echo chamber? https://edu.gcfglobal.org/en/digital-media-literacy/what-is-an-echo-chamber/1/ Accessed on 15.02.2023

GCFGlobal 2 (2023). How filter bubbles isolate you https://edu.gcfglobal.org/en/digital-media-literacy/how-filter-bubbles-isolate-you/1/ Accessed on 15.02.2023

Hamborg, F., Donnay, K., Gipp, B. (2019). Automated identification of media bias in news articles: An interdisciplinary literature review. International Journal on Digital Libraries. 20 (4): 391–415. http://dx.doi.org/10.1007/s00799-018-0261-y

Krönke, C. (2020). Artificial Intelligence and Social Media in Wischmeyer, T., & Rademacher, T. (Eds.). (2020). Regulating Artificial Intelligence. http://dx.doi.org/10.1007/978-3-030-32361-5

Liang, F., Das, V., Kostyuk, N., Hussain, M. (2018). Constructing a Data Driven Society: China's Social Credit System as a State Surveillance Infrastructure. Policy and Internet. 10. http://dx.doi.org/10.1002/poi3.183

Lichter, R. (2018). Theories of Media Bias. In Kenski, Kate; Jamieson, Kathleen Hall (eds.). The Oxford Handbook of Political Communication. Oxford Handbooks Online. Oxford; New York: Oxford University Press. doi:10.1093/oxfordhb/9780199793471.013.44

OECD (2019). Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Analyticsteps.com, Rangaiah, M. (2021). Role of Artificial Intelligence (AI) in Media Industry, https://www.analyticssteps.com/blogs/role-artificial-intelligence-ai-media-industry Accessed on 15.02.2023

Sarker, I.H. (2022). AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems. SN COMPUT. SCI. 3, 158. https://doi.org/10.1007/s42979-022-01043-x

Trend Micro Research, United Nations Interregional Crime and Justice Research Institute (UNICRI), Europol's European Cybercrime Centre (EC3), (2020). Malicious Uses and Abuses of Artificial Intelligence. https://www.europol.europa.eu/publications-events/publications/malicious-uses-and-abuses-of-artificial-intelligence