AI4media

ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

ROADMAP ON AI TECHNOLOGIES & APPLICATIONS FOR THE MEDIA INDUSTRY

SECTION: "AI FOR NEXT-GEN SOCIAL 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 www.ai4media.eu

Authors	**Symeon Papadopoulos** (Centre for Research and Technology Hellas – Information Technologies Institute) **Nikos Giatsoglou** (Centre for Research and Technology Hellas – Information Technologies Institute)

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).

Copyright

AI for next-gen social media

Current status

Social media, such as Facebook, YouTube, Twitter, and LinkedIn, have been around for at least 15 years, and have left an indelible mark on our online communication and socialisation patterns. As an example, chatting, voice messages, posts, and other means of communication are more popular today than phone calls, and users of all ages have been persuaded to let go of inhibitions and air their opinions confidently online. Owing to their popularity, social media have also fulfilled other roles besides communication. These include the role of information or education sources, entertainment in the form of games and funny videos, online marketplaces and bartering, job search and recruiting, citizen activism, connection with companies and institutions, and many more (Figure 1). Crucially, social media have had a huge impact on politics as well, as politicians routinely use them to shape their public image and exercise rhetoric online.

Despite their impressive applications, from a business perspective, the importance of social media lies in the availability of ***user data***, which offer rich and low-cost data on human behaviour. Such data can be mined with AI and ML techniques and monetised by social platforms to sell targeted ads to advertisers. If we had to pick a single application of ML to social networks, it would be ***recommendation algorithms*** to serve targeted ads, content, or friend recommendations to users. Although recommendation is not strictly necessary for communication, it is integral to social media as a practical way to explore the sea of user generated content in a personalised way. After all, it would be hard to imagine Facebook without its algorithmically curated "wall". Recommendation algorithms are typically configured to increase the engagement and attention span of the users, encouraging them to spend more time in the social media, view ads, and make profit for the company. Unfortunately, this practice has also cast shadows over social media, as it can promote sensational content and leads to the creation of filter bubbles[1] and rabbit holes[2].

Figure 1: a) Social media use around the world in Oct. 2021 and b) main reasons for using social media[3].

[1] Pariser, Eli. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin, 2011.

[2] O'Callaghan, Derek, et al. "Down the (white) rabbit hole: The extreme right and online recommender systems." *Social Science Computer Review* 33.4 (2015): 459-478.

[3] Image taken from: DataReportal (2021), "Digital 2021 Global Digital Overview," retrieved from https://datareportal.com/global-digital-overview

From a technical perspective, recommendation algorithms[4] have made use of and fuelled interest in information retrieval, collaborative filtering, and graph processing. Information retrieval concerns the retrieval of documents from a document collection that is most relevant to a user query. In the case of recommendation, the query is the online profile and interests of a user, and the documents are the candidate contents. The objective can include other metrics beyond relevance such as novelty, diversity, and fair representation of content producers. In addition, collaborative filtering derives recommendations from users with similar interests, while graph processing factors the preferences of a user's friends. Link prediction is a specific problem where graph processing has had great success[5]. Beyond recommendation, other ML algorithms that had a big impact on social media draw from NLP, audio processing, and computer vision, and include:

- face recognition and automatic detection of people in images and videos;
- automatic cropping and captioning of images;
- speech-to-text conversion for video transcribing;
- automatic translation to conveniently access post in foreign languages;
- NLP and affective computing techniques for conversational agents, e.g., customer service bots.

Besides recommendation, social data has been an important and cheap source of social analytics and it was first used by businesses to gain feedback on brand perception and guide their marketing strategies. Monitoring the social media is called *social media monitoring* or *listening*, and allows the users of related software to monitor keywords and topics across social platforms, time periods, and demographics[6]. These techniques have also been applied in politics, to gauge the public opinion and guide political campaigns, but with highly controversial results. Specifically, the Cambridge Analytica scandal has shown how user data can be exploited to influence major political events, such as the 2016 US elections and Brexit[7], raising serious ethical concerns on user data collection and causing backlash against social media. Despite widespread condemnation over the scandal, technology is now firmly intertwined with politics, with political parties routinely using bots for political campaigns and a shady industry developing around disinformation.

[4] Kulkarni, Saurabh, and Sunil F. Rodd. "Context Aware Recommendation Systems: A review of the state of the art techniques." *Computer Science Review* 37 (2020): 100255.
[5] Wang, Peng, et al. "Link prediction in social networks: the state-of-the-art." *Science China Information Sciences* 58.1 (2015): 1-38.
[6] Stavrakantonakis, Ioannis, et al. "An approach for evaluation of social media monitoring tools." *Common Value Management* 52.1 (2012): 52-64.
[7] Bömelburg, Raphael, and Oliver Gassmann. "Cambridge Analytica: Magical Rise, Disastrous Fall." *Connected Business*. Springer, Cham, 2021. 387-396.

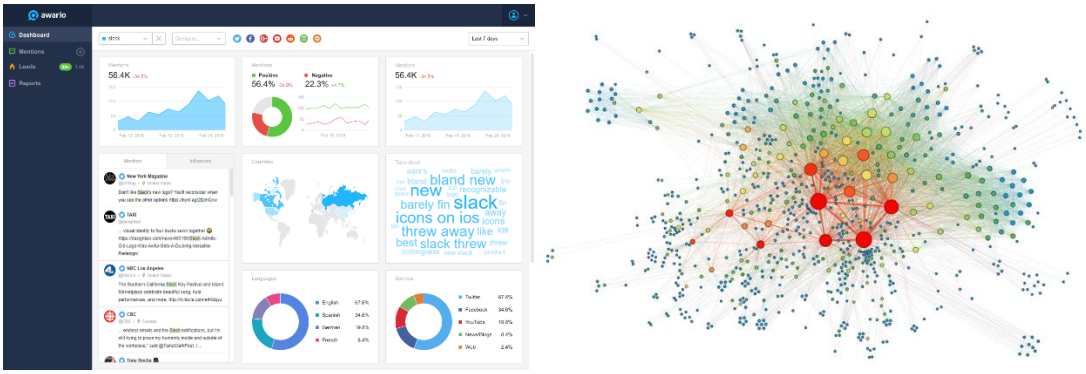

Figure 2: a) A social media analysis tool[8] (left) and social network graph visualisation[9] (right).

Despite the above misgivings, many applications of social media analysis of clear societal value have been proposed in the literature. These include:

- **opinion mining**: for capturing the public sentiment, emotion, or mood on contemporary topics in place of expensive user polling;[10]
- **smart city monitoring:** for detecting traffic jams, events, news stories, and trends;
- **environmental monitoring**: for monitoring the interaction of humans and nature, preserving biodiversity and tracking illegal trafficking;[11]
- **cultural heritage monitoring:** promoting preservation and sustainable urban development;[12]
- **mental health monitoring:** for the detection of anxiety, depression, and suicidal tendencies[13], monitoring for adverse drug reactions;[14]
- **disaster management:** for propagating information during / after a catastrophe, coordinating recovery procedures.[15]

These applications use advanced NLP techniques and computer vision to extract information from unstructured data found in social media. The information from multiple domains is then fused and processed with appropriate ML algorithms, such as neural networks, in order to extract the desired output. As was the case with the business and political applications, collecting data without compromising the user's privacy is a key challenge.

---

[8] Image source: Social Media Today - https://www.socialmediatoday.com/news/12-of-the-best-social-media-analytics-tools-and-how-they-can-help-your-bus/546568/

[9] Image source: Wikimedia - https://commons.wikimedia.org/wiki/File:SocialNetworkAnalysis.png

[10] Cortis, Keith, and Brian Davis. "Over a decade of social opinion mining: a systematic review." *Artificial intelligence review* 54.7 (2021): 4873-4965.

[11] Toivonen, Tuuli, et al. "Social media data for conservation science: A methodological overview." *Biological Conservation* 233 (2019): 298-315.

[12] Liang, Xiaoxu, Yanjun Lu, and John Martin. "A review of the role of social media for the cultural heritage sustainability." *Sustainability* 13.3 (2021): 1055

[13] Chancellor, Stevie, and Munmun De Choudhury. "Methods in predictive techniques for mental health status on social media: a critical review." *NPJ digital medicine* 3.1 (2020): 1-11

[14] Sarker, Abeed, et al. "Utilizing social media data for pharmacovigilance: a review." *Journal of biomedical informatics* 54 (2015): 202-212

[15] Saroj, Anita, and Sukomal Pal. "Use of social media in crisis management: A survey." *International Journal of Disaster Risk Reduction* 48 (2020): 101584

In addition to data mining, two important trends for social networks have been emerging which may play a role in future media analysis: i) decentralised online social networks (DOSNs)[16] and ii) the *metaverse*[17].

Regarding DOSNs, although they were initially conceived in the heyday of p2p networks in the early 2000s, they have seen a surge of interest due to the concerns raised by centralised social media and the popularity of blockchain. DOSNs intend to operate without central servers and give back control to the users over their own data. This implies letting users explicitly grant consent to applications over their data (which can be revoked at any time) and/or monetising them, often through cryptocurrencies. Key actions and proposals towards the decentralisation of the web include the 2018 decentralised web summit[18], Tim Berner Lee's Solid project[19], and Elastos[20]. A key consequence of DOSNs is that they will place strong emphasis on decentralised and federated learning techniques, which are also gathering research interest lately.

Regarding the metaverse, it refers to a virtual world running parallel to the physical one, which we will access through our avatars or digital twins. At this moment, the concept is not precise but it has been fuelled by the popularity of VR and AR technologies and the recent commitment of Facebook to implement it, after the company was renamed to Meta. The metaverse promises to be an evolution of social media and a much more immersive experience, where people will interact through their avatars, generate VR content, attend events, and engage in online purchases with bitcoins. If this vision succeeds, it will offer unprecedented opportunities for data analysis, require even more invasive data collection as well as new AI algorithms for the operation of digital twins and the management of Non Playable Characters (NPC). On the other hand, critics point out the similarities with older efforts such as SecondLife[21] (Figure 3) and the timing of the Facebook announcement when the company is facing antitrust allegations[22] and younger populations are migrating to other platforms[23]. Whatever the case, the importance of VR and holographic content in social media will undoubtedly increase.

[16] Masinde, Newton, and Kalman Graffi. "Peer-to-Peer-Based Social Networks: A Comprehensive Survey." *SN Computer Science* 1.5 (2020): 1-51

[17] Lee, Lik-Hang, et al. "All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda." *arXiv preprint arXiv:2110.05352* (2021)

[18] "Decentralized Web Summit 2018," https://decentralizedweb.net/, 2018

[19] Sambra, Andrei Vlad, et al. "Solid: a platform for decentralized social applications based on linked data." *MIT CSAIL & Qatar Computing Research Institute, Tech. Rep.* (2016)

[20] Elastos: https://www.elastos.org/

[21] Second Life: https://secondlife.com/

[22] K.A. Smith, What's Going on with the Facebook Antitrust Lawsuit? (2021): https://www.forbes.com/advisor/investing/update-facebook-antitrust-lawsuit/

[23] B. Auxier, and M. Anderson. "Social media use in 2021." *Pew Research Center* (2021).

Figure 3: A social gathering on the metaverse Second Life [24].

Research challenges

While social media analysis presents many opportunities, it also presents important challenges. These include the following:

Walled-garden data: Perhaps the most important challenge of social media analysis is the limited access to social data. Currently, social media companies keep the data collected from users at their private storage, characterised as private data silos (Figure 4). Third party access to this data is provided for non-commercial purposes through APIs, which enforce rate limiting and/or subsampling of data, and for commercial purposes by directly selling it at a premium price. This creates a competitive advantage for the social media company, which is hard to relinquish, and allows it to enforce its own rules on privacy. It is a challenge for research and society as it stifles independent research and innovation, limits applications of societal value which do not fit with the companies' business model, and hinders reaching a commonly accepted standard of user privacy. Decentralised social media and increased control of users over their data could be key to overcoming this challenge.

[24] "Live radio hour in Second Life with Draxtor Despres and Jo Yardley" by HyacintheLuynes: https://commons.wikimedia.org/wiki/File:Second_Life_11th_Birthday_Live_Drax_Files_Radio_Hour.jpg

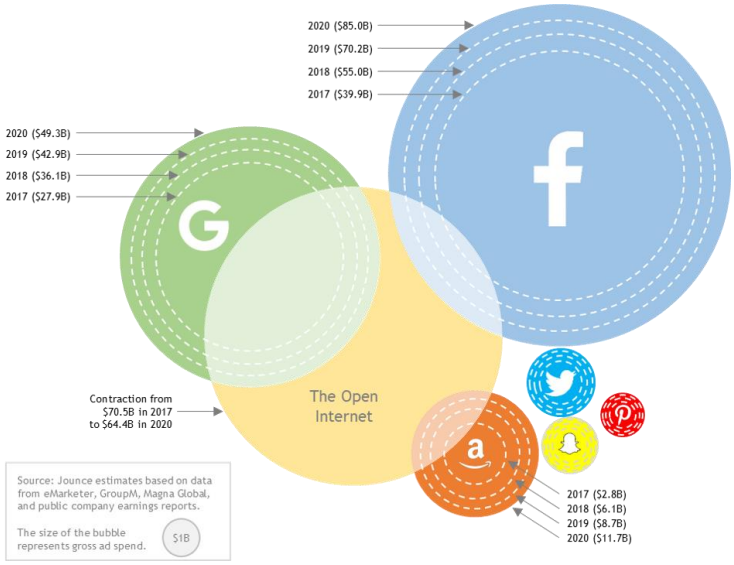

Figure 4: Walled-gardens. The image depicts the total money spent on ads in closed advertising platforms, e.g., by Google, Facebook, and Amazon, compared with open advertising networks ("The Open Internet"). The closed platforms harness first-party data from their users to offer targeted advertising services to their customers[25].

Idiosyncratic content: Social media data, especially text, is highly unstructured and noisy. This is due to the short and casual nature of posts, frequently containing spelling errors, slang, hashtags, emoticons, acronyms and abbreviations in place of detailed descriptions. Crucially, social media posts tend to depend heavily on implied context, potentially related to a niche Internet culture, and frequently contain irony and sarcasm, which is still challenging to detect with NLP algorithms. The above distinguish social from formal text, such as news and encyclopaedia articles, and necessitate extensive cleaning and advanced NLP techniques. These limitations also apply to images and videos, which are frequently processed, e.g., with digital filters that distort the original content.

Need for credibility assessment: Social media contain a lot of rumours, misinformation, and generally content of questionable nature and provenance. A further complication is that social media is often populated by bots that produce both legitimate and illegitimate content, e.g., customer service versus bulk advertising and political propaganda. Automatic credibility assessment at both the content and source level, e.g., through reputation scoring and cross-checking, will be required to use social media data confidently. Spatiotemporal analysis of information propagation on social media graphs is also important to detect false information early and even help mitigate it.

Multilinguality and multimodality: Most works on social data analysis focus on a few data modalities, with a strong bias towards English text. The reason is that text is convenient in terms of storage and availability, while English is the dominant language globally. The performance however of AI algorithms can be enhanced with the integration of other data modalities, such as audio, image, and video, as well as multilingual support, in order to understand the large non-English user-base of social media. Currently, multimodal analysis is at an early stage and suffers

[25] Image from Jounce Media 2020 Market Outlook Report (2020): https://jouncemedia.com/blog/2020/1/22/the-state-of-the-open-internet

from a lack of standard datasets. Multilingual support also faces challenges due to the linguistic diversity, breadth of dialects, and lack of training data for minority languages. From a technical perspective, non-English languages cannot be understood by models trained on English and cannot be handled accurately by naive automatic translation. The latter is due to the particularities of syntax, expression, and slang that do not map completely from one language to another. Research on multilingual models, i.e., models that understand many languages simultaneously, has only started emerging. Unfortunately, minority languages have also been overlooked by big social media companies due to the limited opportunities for profit[26,27]. For example, the leaked *Facebook Files*[28], made public in September 2021, have revealed that the company had difficulties in moderating harmful content in non-English developing countries, thus allowing misinformation and hate speech to spread. More effort is thus needed for a linguistically inclusive AI.

Lack of labelled datasets: The success of the DL paradigm was largely fuelled by supervised learning, i.e., algorithms that learn from large-scale annotated datasets. While ML algorithms have achieved marvellous feats, the quality of their output is as good as the data on which they were trained. In the case of social media analysis, data is ample but labelled data is scarce and hard to produce. For example, to create an algorithm that detects depression from user posts, users with actual depression or at least depression symptoms must be identified for ground truth. Acquiring this data requires coordination with domain experts and significant labour, which is time-consuming and expensive. In addition, labels for social issues and social media culture evolve rapidly so that supervised algorithms require frequent updates. This is a key problem of ML algorithms that could be alleviated with stronger transfer learning, unsupervised or semi-supervised algorithms, or even reinforcement learning algorithms that can adapt their learning.

Biased content: Insights from social media are inherently biased for multiple reasons. Firstly, social media users and, specifically, those who post regularly tend to belong to specific demographics, which depend strongly on the examined network. This limits generalisations to the general public. Secondly, social media are prone to amplifying harmful content and fringe views because the employed recommendation algorithms are designed to maximise user engagement and attention, achieved by promoting sensational content. This has led to the creation of echo chambers, i.e., closed communities that reinforce fringe views that end up over-representing them. Finally, social media often suffer from harmful bias and prejudice against minority groups[29,30]. These factors must be considered by AI algorithms to ensure diversity and inclusion instead of bias in their output.

[26] N. Nguyen and C. Scurato, Facebook and YouTube's refusal to moderate misinformation in global languages harms communities of color (2021): https://prismreports.org/2021/11/02/facebook-and-youtubes-refusal-to-moderate-misinformation-in-global-languages-harms-communities-of-color/

[27] D. Marinescu, Facebook's Content Moderation Language Barrier (2021): https://www.newamerica.org/the-thread/facebooks-content-moderation-language-barrier/

[28] Wall street Journal, The Facebook Files: https://www.wsj.com/articles/the-facebook-files-11631713039

[29] E. Dwoskin, N. Tiku and C. Timberg, Facebook's race-blind practices around hate speech came at the expense of Black users, new documents show (2021): https://www.washingtonpost.com/technology/2021/11/21/facebook-algorithm-biased-race/

[30] V. Vara, Why Do We Care If Facebook Is Biased? (2016): https://www.newyorker.com/business/currency/why-do-we-care-if-facebook-is-biased

Societal and media industry drivers

Vignette: Multilingual analysis of social media trends, public perception, and sentiment

Danae is a journalist working in a small local news outlet. Danae's daily work includes hunting down important trends and news from social media which may provide the basis and inspiration for Danae's articles. Traditionally, this was done manually, but Danae's company has recently acquired an AI tool that automates this process. Specifically, the AI tool aggregates data from multiple social media and detects trending topics. This has become possible since the recent EU regulation has democratised access to social media data while enforcing common standards of privacy, so that social media analysis can be conducted without exorbitant fees for third-party software. Once the AI tool detects a trending topic, it summarises and categorises it under thematic categories, such as politics, lifestyle, and sports, and presents it to Danae in near real-time or on demand. Danae can choose a topic that piques her curiosity, to see more details, such as the first post from which the topic originated and its propagation in different platforms, presented with an effective visualisation. She can also filter across different locations and time, as well as change the sensitivity of the algorithm so that it can detect more topics more quickly or save cognitive load. The algorithm can detect a variety of topics: yesterday, there were rumours on a fire at Danae's city, while today, there is gossip about the wedding of a local celebrity. Danae loves this tool since it makes her work much easier.

Recently, there have been diplomatic tensions between Danae's country and a foreign one. The relation between the two countries has been historically tricky and the current situation was sparked by a reckless exchange of the two countries' diplomats, which was unusually heated for the diplomatic level. While most news outlets focused on the content of the exchange, fuelling the situation, Danae had the idea to focus on the public perception of the event. With the help of the AI tool, she was able to estimate the public sentiment in the foreign country and found out that the public was overwhelmingly critical of the actions of their diplomat. Key to this discovery was that the AI tool could automatically translate content from foreign social media which Danae would not have understood or would have required manual translation. In addition, the tool helped Danae locate influential and extroverted social media users which were open to interviewing. It then allowed Danae to conduct an immersive interview based on VR technology, which also performed live translation. Finally, Danae gathered data from her own country and wrote a highly influential article that received widespread recognition. Danae's article played a small role in extinguishing the diplomatic tensions between the two countries and had a definite impact on Danae's career.

Future trends for the media sector

As the web becomes more and more social, the importance of social network analysis and automatic extraction of intelligence from social data will grow. The media sector will need to ***extract complex information from social media that go beyond the current practices***, for example, estimating health indexes of the environment, cultural heritage, and mental health and tracking down misinformation networks automatically. In the latter case, journalism will increasingly rely on tools that automatically detect bots and distinguish fake from legitimate content in social media, by assessing the veracity and source provenance of information. The

developed technology will be responsible for gathering information from diverse sources and presenting them in a balanced unbiased way to the journalists, thus helping them bypass their personal biases. Ultimately, this will increase the quality of information offered to the end users and citizens (more information on the topic of disinformation can be found in the section on "*AI for counteracting disinformation*").

In brief, key future trends include:

Decentralisation and democratisation of social content: Current efforts to decentralise the Web are expected to carry over to social media, empowering users with control over their personal data while simultaneously opening them to a larger audience. Online social interactions will also generalise across online applications and may eventually become public. This data will present great opportunities for analysis and will be valuable to the media sector.

Privacy and security: User privacy and security are existing trends that will grow in importance with modern applications that rely heavily on user data. Processing this data while simultaneously minimising exposure to unwarranted parties is a key challenge that future applications will need to address. An additional challenge, considering the trend to adopt unique and authentic identities online is identity theft and impersonation. Candidate solutions include advanced encryption, anonymisation, obfuscation, and differential privacy techniques all enabled by AI.

Multimodal, multilingual, and multidomain analysis: Current AI applications on social data rely heavily on English text from a single social media, typically Twitter, but this is rapidly changing. Aggregating information from multiple languages, modalities, and social domains is a natural evolution with clear benefits to the accuracy and diversity of social media analysis.

Credibility assessment algorithms: As misinformation and low content quality is typical of social media, the importance of automatic filtering, cross-referencing, source provenance detection and authoritativeness estimation will grow. These tasks will rely heavily on advanced processing on dynamic graphs and information fusion from multiple sources.

Addressing bias: As the harmful impact of bias is widely recognised and humans become increasingly aware of their susceptibility to it, there is a pressing need for automatic detection and mitigation. This need has sparked fruitful research and resources on the operationalisation of bias, which is expected to grow. The trend is important for the media industry, as the latter needs to abide by strict measures of objectivity, balanced news reporting, and journalistic ethics.

Increased immersion: The recent success of VR and AR technologies has affected social media and the announcement of the Metaverse indicates that it is an important trend. Online social interactions are thus expected to become more immersive, leading to new types of social content and applications, which can be mined by ML algorithms for novel insights. At the same time, immersion carries increased exposure and new privacy concerns that must be addressed to ensure the widespread adoption of the technology. Key opportunities for AI include the generation of avatars, landscapes, and content to populate this world, the development of sentient AI agents and affective technologies to make the interactions between humans and AI agents more human-like, and the removal of language barriers between metaverse users.

Goals for next 10 or 20 years

Based on the aforementioned trends, we can extrapolate in the near future and predict the following goals:

***Open, inclusive, and safe social web*:** Tomorrow's web will be distinctly more social. People will be authenticated by their unique online ID and will operate similarly as they would in the physical world. Proper authentication will result in safer and socially responsible conducts, while citizens will be protected from digital identity theft. Social interactions will abound and public data will be readily available for analysis, subject to user defined privacy constraints. In this world, both media and citizens will have access and benefit from insights resulting from the analysis of open social data with advanced AI techniques.

Accurate social media analysis: Based on the availability of social media data, advanced processing, and information fusion from multiple sources, social analytics will be even more accurate and useful to stakeholders, including the media industry. Journalists will have timely access to social data and ask complex questions that are impossible at present. The answers will be insulated from misinformation and bias thanks to the advanced AI algorithms but also thanks to human education and positive change in how society thinks.

Metaverse: Following the trend of increasing immersion, next generation social media will attain the vision of metaverse. Online profiles will be substituted by 3D-rendered avatars that will function as their digital twins in the virtual world. People will be able to customise their avatars as well as their surroundings with imaginative content created with the help of AI technologies. People will also participate in immersive experiences such as convening in digital cafes, attending digital concerts, and playing holographic games, and have meaningful interactions with other users as well as AI agents, which will successfully mimic human behaviour. Communication barriers will be removed with the help of advanced language processing. The virtual world will bring a serious paradigm shift and will change the way we think about the world and the media.

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