AI4media

ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

ROADMAP ON AI TECHNOLOGIES & APPLICATIONS FOR THE MEDIA INDUSTRY

SECTION: "SUMMARY OF AI APPLICATIONS, TECHNOLOGIES AND CHALLENGES 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 www.ai4media.eu

Author	**Filareti Tsalakanidou** (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).

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.

Summary of AI applications, technologies and challenges for the media

In the previous sections (*Overview-AI applications for the media sector* & *Overview - AI research & technology trends*), we tried to offer an in-depth overview of the different AI applications and AI research trends that can have an impact on the media sector, by presenting the findings and insights from a selected number of roadmaps, surveys, white papers, articles, etc. This presentation revealed a landscape where the opportunities for the use of AI are enormous while the variety of tasks across the media supply chain that AI can improve, assist, automate, expand or create is limitless. AI can have a truly transformative influence on the media sector, reinventing the business model of media organisations, establishing new ways of work and increasing the productivity and creativity of the workforce, and finally transforming and enhancing the user experience across platforms.

In the following, we will attempt to summarise the most prominent AI-enabled applications for the media sector. Figure 1 presents the relevant applications per industry, focusing on news, film/TV/streaming, music, games, social media, advertising, and publishing. As can been seen, each sector has its own unique needs (e.g. fact-checking for newsrooms or AI-based casting for films) but most of the applications featured in Figure 1 aim to satisfy similar needs, e.g. the need for content personalisation or automated content creation (whether this is music, films, ads, books or games), better recommender systems, enhanced understanding of users, etc.

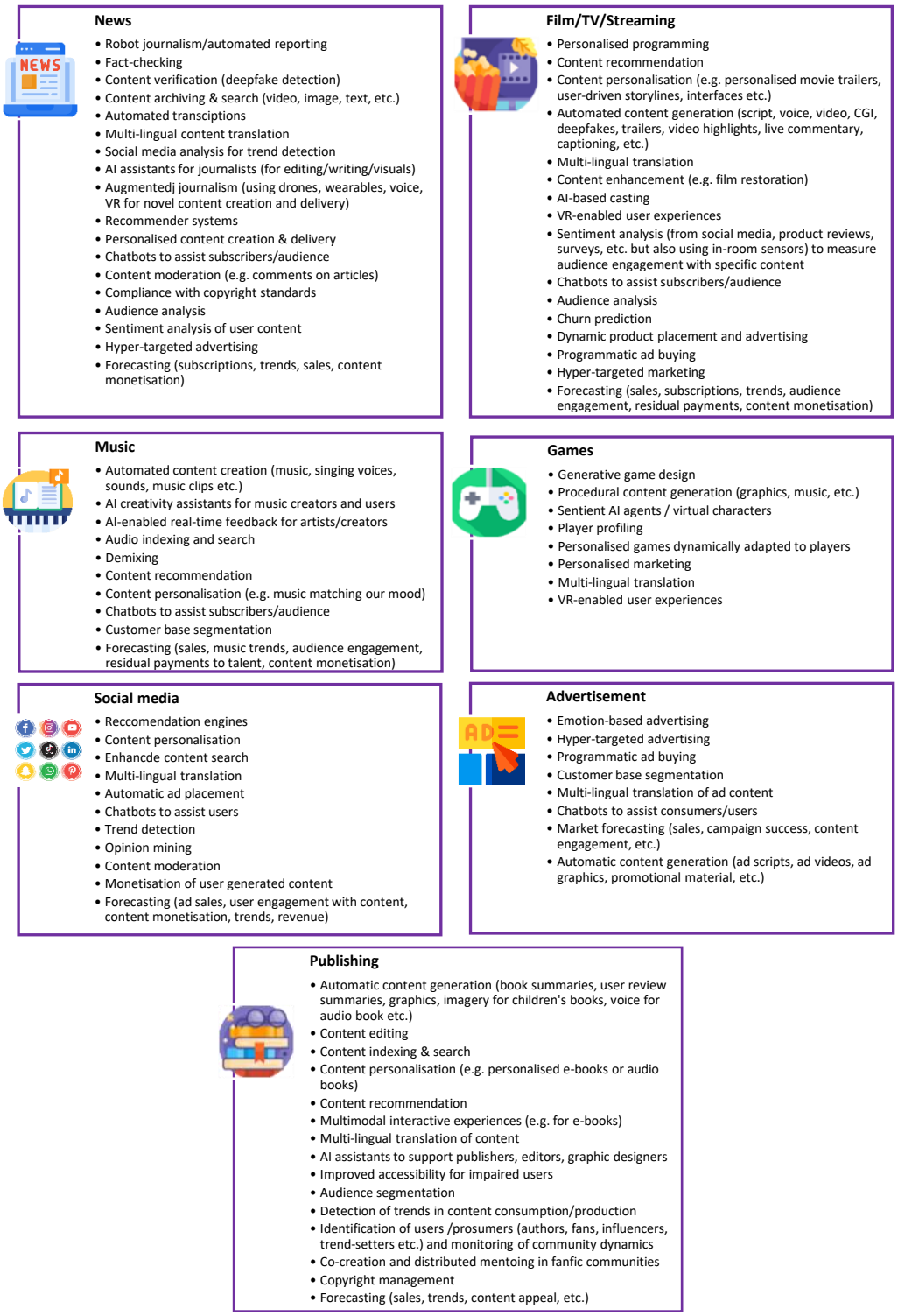


Figure 1: A summary of AI applications for the media and entertainment industry.

AI applications for the media sector

The list below summarises applications of AI that are already having or can have a significant impact in most media industry sectors, addressing common needs and shared aspirations about efficient workplace automation, enhanced content, enhanced user experiences, better understanding of user needs, better marketing[1]:

Smart recommender systems that will analyse vast amounts of user and content data (also combining them with trend analysis or info about current state of affairs in the neighborhood and the world) to recommend content that matches user preferences or current mood or even needs that they did not know they had.

Content personalisation that will allow media companies to offer content not to audiences of billions but to billions of individual users with their own unique preferences, experiences, needs and moods. The trend is clearly more and more personalisation, going beyond simple user preferences derived from past actions and behaviours and adopting more elaborated approaches that also take in mind what happens in the moment to the user herself and in her environment. Personalisation involves content presentation (e.g. writing style of news digests), interaction with content (e.g. user-driven film storylines), and personalisation of content itself (e.g. music that matches our mood, game visuals that match our aesthetics, personalised movie trailers that will excite us.

Automated content creation (text, image, video, audio, VR), aiming to improve productivity and enhance creativity of media professionals but also to offer new ways of creative expression to the users. The potential applications of this technology are unlimited: robot journalism from automatic headlines to article writing, procedural content creation for games, creation of new music, deepfakes for the film industry, automated TV commentary for sports events, graphics for ads, book/film/podcast summaries but also creation of new art (image, video, music) by talent and simple users.

Enhanced content search for media professionals and users to allow efficient real-time retrieval of relevant content and thus improve exploitation and monetisation prospects. This will require automatic metadata tagging and extraction based on advanced video, audio and text analysis of content (detection and recognition of faces, voices, objects, places, dates etc.) that will make content discovery easy and accurate while also enabling other functionalities like automatic content moderation or content recommendations. Search will not be limited to text queries but will support voiced questions, audio, images, sketches etc. Again, the applications are numerous: users find the online content they search for with less effort, journalists can search large audiovisual archives to find historical information or social media and other outlets' websites, music composers can find audio excerpts to inspire them, book publishers can search the vast volume of titles produced every year, etc. Efficient content indexing and search means that the content can be more easily monetised.

[1] The icons used in this and the next page are from flaticon.com and vecteezy.com

Audience analysis aiming to capture user behaviour and interactions with content and increase insights on audience needs. By understanding their audience and what makes them happy, media companies can more effectively monetise their content through personalisation of services/content as well as high-accuracy ad targeting. Audience analysis is based on large amounts of data, including user behaviour but also user generated content like posts, reviews or likes. A new promising direction in audience analysis is the use of sensors that can accurately and in real-time record the emotional reactions of users to content, allowing in-depth analysis of user interest and engagement (e.g., Which parts of the film were the most entertaining or emotionally moving? What storyline plots created bigger engagement?).

Social media analysis to detect trends about what is interesting right now or what people think about a specific issue. Trend detection allows media companies and professionals to react in real-time to what is happening to the world or their users and adapt their content and services accordingly.

AI assistants that can help media professionals do their job more efficiently but also assist the audience when using media services or trying to access content. In the first case, AI assistants support professionals in everyday tasks (e.g. an AI assistant that helps journalists by suggesting story topics based on trend analysis of social media, by searching the archives to find statements of a politician on a specific subject, by suggesting visuals to accompany stories, by validating content veracity, by answering questions etc.) or even replace them in doing tedious and boring tasks that limit creativity (e.g. an assistant that produces daily financial news summaries allowing the journalist to focus their efforts into meaningful commentary on current financial situation). In the second case, AI assistants in the form of chatbots or talking digital characters are increasingly used to help the users to find the content they want, to answer user questions, or help users play a game. Advances in NLP promise to make this kind of communication much more natural and effective.

Multi-lingual translation to help the media distribute and promote their content and services more efficiently, reaching audiences around the globe. This will allow content to be offered in the user's native language, users to communicate with each other without language barriers and professionals to search, analyze and use content from different languages. This development will also be a decisive step towards democratising access to content and media services as well as participation in the digital online world.

Forecasting to facilitate accurate predictions with regard to a variety of subjects such as content engagement, emerging trends, user behaviour, sales, subscriptions, content monetisation, churn, ad revenue, etc. Predictive analytics are essential for everyday tasks (e.g. organisation of marketing campaign and TV programming) but also for designing long-term strategies.

Hyper-targeted advertisement and programmatic ad buying. Advertisement is essential for the survival of the media industry both for promoting their own content and services aiming to create revenue from their content but also as the promoters of third-party goods, which results in revenue from brands that wish to advertise their

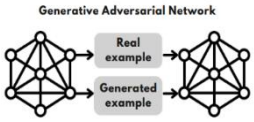
products through the media company's channels. By analyzing vast amounts of user, content, operational and other (e.g. societal, financial, environmental) data, media companies but also those who advertise in them can target users in real-time with highly effective personalised ads that will make them loyal subscribers or customers. In addition, real-time analysis of audience dynamics and of available ad space across multiple channels can enable automated buying and selling of ad inventory and real-time marketing campaign optimisations.

Compliance with copyright standards. Copyright in media and creative industries is a huge issue that can cause a lot of headaches to media professionals and have serious legal implications. AI can transform this area by automatically analysing relevant legal documents or online copyright information thus saving thousands of hours for the legal department but also allowing media professionals to quickly identify IPR issues of online content.

AI technology trends for the media sector

To transform the media industry through the aforementioned applications well-known AI and ML technologies such NLP, RL, GANs etc. are already employed to harness the data and deliver the envisaged functionalities. In the following, we summarise a list of AI technologies that hold the greatest potential to realise the media's vision for AI. The list is obviously not exhaustive (and may be biased towards AI4Media's research activities and use cases) but reflects efficiently the most important current and future trends identified by AI experts and industry stakeholders in the surveys, reports and articles presented in the previous *Overview* sections[2].

Reinforcement learning: RL techniques are currently used to train AI agents that can play increasingly complex games, beating human champions. Their ability to learn goal-oriented behaviours through reward and punishment strategies will be increasingly used to develop AI agents for different applications, including AI agents that assist and learn from humans and their environment (for office automation tasks but also as companions in the digital world), smart recommenders that learn from user interaction with content and recommend content that will maximise user engagement and satisfaction or chatbots that self-improve.



Generative AI: Generative AI is used to automatically create new content by utilising existing content. Technologies like GANs, transformers, VAEs or diffusion models will be increasingly used by the media industry to create new high quality text, image, video and audio. The range of potential applications is without limits: deepfakes for the film industry, music composition, game assets, script creation, film preservation etc.

Transformers: After revolutionising natural language processing thanks to their computational efficiency, transformers are set to also dominate the computer vision field by allowing training at large-scale with vast image databases to enhance the performance of image classification algorithms. This trend will accelerate progress in visual

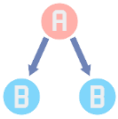
[2] The icons used in this and the next page are from flaticon.com and vecteezy.com

search algorithms that are increasingly important for efficient content indexing, search and recommendation.

Few shot or no shots learning: Big breakthroughs in AI have been enabled mainly by the development of models trained with massive amounts of data. However, it is not always plausible to have such amounts of data. Few shot learning techniques rely on limited available examples to learn models, which could come handy e.g. in the case of building NLP models for rare languages where limited data is available or voice cloning from few audio samples.

Emotion AI: Emotion AI includes technologies that learn and recognise human emotions. This knowledge can be used to help machines recognise a human's emotional state and adapt their behaviour to human emotion in addition to human action but also to teach machines to exhibit human emotion. The aim is to make human-machine communication natural, imitating the way humans would communicate with each other, but also to use knowledge about human emotion to offer better personalised services. The applications are numerous: chatbots, sentient virtual characters, targeted advertisement, understanding the audience, content recommendation or delivery based on user mood, etc.

NLP and multi-lingual NLP: NLP has witnessed a true revolution during the last few years with large language models like GPT-3. NLP is expected to become increasingly mainstream in the media business through applications such as conversational agents and virtual characters, creative writing, robot journalism, interactive storytelling, voice search for image/video/audio, sentiment analysis in social media, voice dubbing, or multi-lingual translation. Multi-lingual translation in particular will be a real breakthrough, breaking language barriers and allowing, on one hand, content creators to reach new audiences worldwide but also to exploit creatively the wealth of content available online (which is currently out of reach because it is in other languages) and, on the other hand, helping audiences and users to communicate freely and benefit equally (and more democratically) by the content created all over the world.



Causal AI: The *20-Year Community Roadmap for AI Research* makes several mentions to this technology which helps to move the needle beyond correlations to the identification of causal relationships. Causality is still an unfulfilled aspiration of AI which however receives increasing interest. Causal AI can have a significant impact on the media, since it will allow understanding the why and examining the what if: why users like some content more than other, why the recommendation algorithm provided this content, what would happen if a different film casting decision would be made, etc. The applications are many: explainability of AI decisions, better content personalisation, prediction of trends, prediction of content engagement, better content recommendations, more natural interaction with virtual characters (that understand why we act in a specific way and respond accordingly) etc.

AI at the edge: This is a highly emerging area of research in AI that aims to facilitate processing of data closer to the end-user device (e.g. in the user's mobile phone). This will significantly enhance data privacy and processing speed. Interesting applications include learning models (e.g. for content personalisation or development of AI agents) that learn

continuously from the user by collecting sensor data from the home environment as well as real-time user feedback without allowing the collected data to leave the user's device.

Quantum computing: This emerging technology still has a long way to go. Its potential however would be game-changing for the media industry. These unforeseen capabilities of processing power will facilitate and accelerate machine learning using huge volumes of data, allowing for example social media analysis at scale, development of large multilingual language models and video game characters that behave hyper-realistically.

AI that learns like humans: Deep neural networks are inspired by biological neural networks of animals. Despite revolutionising AI and bringing major advances in machine cognition, DNNs are still a long way from enabling human-like intelligence especially with regard to humans' ability to easily generalise and learn new tasks with minimum training from infancy. An emerging trend in AI focuses on trying to understand the human cognitive system and to imitate human brain capabilities. Bio-inspired learning is expected to transform the way we interact with machines, allowing the development of AI agents with human-like intelligence and capabilities. Besides obvious applications in gaming and the Metaverse, such human like capabilities will have a profound impact in many different applications from search of content to AI-enabled creativity.

Other interesting trends include the **fusion of AI and cloud** that will allow collection of data from the cloud, self-training of models with this data, and availability of learnt knowledge in the cloud for others to exploit. Also, **convergence of AI with blockchain technology** has the potential to improve training data annotation and model training.

Challenges of AI adoption in the media sector

The previous analysis revealed the vast potential of AI to bring positive change to the media industry sectors. However, the works examined in the previous *Overview* sections highlighted that with high potential also come significant challenges and risks. We briefly summarise the main challenges identified below[3]:

AI explainability: Currently, AI systems are mostly black boxes without being able to explain why they recommended a product or predicted the success of a film or made a moderation decision. In order to fully adopt and trust such AI systems, media professionals but also users need to understand how such systems work. Explainable AI aims to do just that, increasing transparency and increasing trust and adoption of AI-enabled applications.

AI robustness: performance of AI algorithms may be hindered by many reasons, including malicious adversarial attacks but also poor performance when dealing with data different from those they were trained with. To ensure robustness, tools that help fortify AI models against attacks, predict new types of attacks, and ensure that the models perform as well in the real-world as they do in a sandbox are increasingly necessary.

[3] The icons used in this and the next page are from flaticon.com and vecteezy.com

AI bias and AI fairness: AI systems often exhibit bias against specific groups of people, including racial bias, gender bias, etc. due, for example, to prejudiced hypotheses made when designing the models or due to problems of diversity and representation in training data. AI bias can lead to bad business decisions or discriminate against groups of users. A prominent example for the media sector is bias that may be embedded in large language models. Such models are trained with swaths of Internet data, which are by definition produced in the biggest or richest countries, in languages with higher linguistic footprint, and by communities with large representation, or mainly by men[4], thus resulting in models that fail to capture changing social norms or the culture of minorities and underrepresented groups and which will eventually discriminate against such groups or produce language that is not attuned to changing social norms[5,6]. The gigantic volume of data also makes it hard to audit such models for embedded bias. To address this problem, new techniques and new frameworks have been proposed aiming to enhance AI-fairness and minimise bias.

Privacy concerns: AI applications like recommender systems or content personalisation are based on the collection of vast amounts of data about user's preferences, behaviours, actions, as well as user generated content. Obviously, this creates a lot of concerns about privacy and how this data may be used. To address such concerns, the EU has proposed regulations like the GDPR while companies are starting to explore solutions that will enhance the privacy of the users and their data.

Data for AI training: Many of the examined reports highlighted the need for large volumes of real and high-quality data for training AI models for the media industry.

AI skills: One of the reasons hindering the adoption of AI in the media industry is the lack of relevant skills by media professionals and challenges in recruiting AI experts. To overcome this obstacle AI training and education are necessary for media professionals as well as raising awareness about AI and its potential across an organisation. Collaboration of the media industry with academia/research but also with other media organisations or industries on AI topics of common interest would also be beneficial.

AI strategy*:* Another issue concerning the experts is that still many media companies do not have a clear AI strategy that will allow them to efficiently adopt AI in the workplace, recruit or train staff, make investments in specific technologies, pursuit useful collaborations and fully exploit AI's potential for the media.

---

[4] Such models usually get trained with data scrapped by sources like Wikipedia or Reddit where women are significantly under-represented. According to this Guardian article, women are less than 20% of the contributors of Wikipedia; according to Statista, women represent only 37% of Reddit users worldwide.

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

[6] E. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In. Conf. on Fairness, Accountability, and Transparency (FAccT '21), March 2021.

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