Navigating Disinformation Landscapes within Synthetic Realities. AI a friend or a foe?

Session Chairs: Symeon (Akis) Papadopoulos, Nikos Sarris, ITI-CERTH

The goal of this panel is to discuss potential future scenarios and realities, in which AI is pervasive and offers much more advanced capabilities in terms of generating realistic multimodal content to the point that complete synthetic realities will be easy to generate. In such a future setting, it will be interesting to discuss what kind of disinformation risks emerge and whether AI could be an ally or a foe. The panel features experts from different disciplines including computer vision, natural language processing, media and journalism, and human computer interaction.

Chairs

Symeon (Akis) Papadopoulos, ITI-CERTH

Symeon Papadopoulos is a Senior Researcher (Grade C') with the Information Technologies Institute / Centre for Research and Technology Hellas, Thessaloniki, Greece. His research interests lie at the intersection of multimedia understanding, social network analysis, information retrieval, big data management and artificial intelligence. Dr. Papadopoulos has co-authored more than 30 papers in refereed journals, 10 book chapters and 100 papers in international conferences, 3 patents, and has edited two books. He has co-organised numerous workshops in the areas of multimedia and social media. He has participated in a number of relevant EC FP7 and H2020 projects in the areas of media convergence, social media and artificial intelligence, including WeVerify, HELIOS, InVID, REVEAL, SocialSensor and WeKnowIt. He currently coordinates the H2020 MediaVerse project and is a member of the coordinating team of the AI4Media Network of Excellence. During the last years, he has been leading the Media Verification team, focusing on the development of tools for the detection and analysis of multimedia- and social media-based disinformation. Since 2014, he is one of the cofounders of Infalia PC, a spin-off company of CERTH.

This event is supported by the European Union's Horizon 2020 Research and Innovation programme under grant number 951911 – AI4Media

Nikos Sarris, ITI-CERTH

Nikos is a Senior Researcher at MKLab of ITI-CERTH and a member of MeVer. After receiving his PhD from the Aristotle University of Thessaloniki Nikos worked at ATC for 15 years, where he headed the Innovation lab for 7 years and is now an Advisor on media technologies. For the last 20 years he has been working in R&D projects as a researcher, project manager and coordinator of large multinational consortia. He has been involved in many projects focusing on the semantic 'understanding' of news content and the assessment of its trustworthiness, coordinating the development of two related products: TruthNest and Truly Media, the latter in cooperation with Deutsche Welle. Nikos has been the coordinator of the EC-funded SOMA project tasked to empower the European Observatory against Disinformation and is a member in the Executive Board of the EDMO project (European Digital Media Observatory).

Panelists

Richard Fletcher, Reuters Institute for the Study of Journalism

Bio: Richard Fletcher is a Senior Research Fellow and leads the research team. He is primarily interested in global trends in digital news consumption, comparative media research, the use of social media by journalists and news organizations, and more broadly, the relationship between technology and journalism. Richard primarily works on the Digital News Project and is lead researcher and co-author of the main Digital News Report – the world's largest annual survey of global news consumption. Richard also uses this survey data to underpin comparative cross-national research into patterns of news consumption, audience fragmentation and polarisation, the effects of search engines and social media on news use, trust in the news, and paying for digital news.

Statement: Concerning journalism and the news media, artificial Intelligence clearly has the potential to be both a friend and a foe when it comes to navigating disinformation landscapes. In terms of news 'production', some news organisations are already benefitting from the use of artificial intelligence to automate the production of news articles and videos for certain topics. However, similar technologies can of course be applied to the production of disinformation, with 'deepfake' videos—although not widely-used at present—providing a particularly striking example of the potential power of synthetic media in this regard. When it comes to the 'dissemination' of news, artificial intelligence is already used by search engines, social media, and aggregators to recommend news content to users. The fear is that these technologies will become over responsive to individual users, undermining the idea of a shared news agenda, or causing people to misunderstand the world around them. But at the same time, there now exists a considerable body of research documenting the benefits to users, with increased social media use linked to higher levels of news use and more diverse news diets. Finally, although

communication researchers often think primarily about production and dissemination, a third dimension concerning the appropriate level of regulation and oversight online is emerging, raising questions about the possible risks and benefits of relying on artificial intelligence to identifying synthetic media in newsworthy content.

Maria Bielikova, Kempelen Institute of Intelligent Technologies (KInIT)

Bio: Maria is a director general at KInIT. She also conducts research focusing on human-computer interaction analysis, user modelling and personalization. Recently, she has been working in data analysis and modelling of antisocial behavior on the Web. She is a former member of the High Level Expert Group on AI established by the European Commission. Currently she is a chair of the Permanent Committee for Ethics and Regulation of AI established by the Ministry of Investments of the Slovak Republic. She was also a member of the European Commission Joint Research Center Board of Governors. Before her work at KInIT, Maria was employed at the Slovak University of Technology in Bratislava, she was a lead of the PeWe research group (Personalized Web) and director of the User eXperience and Interaction research center.

Statement. The power of AI to create and spread disinformation is indisputable. Even nowadays, we can see the inception of AI-enabled disinformation campaigns - state-of-the-art extensive language models (e.g., GPT-3) can generate a piece of text that is almost indistinguishable from genuine human-made news; the same applies for manipulated images and video deep fakes. Consequently, bots, which control dozens of social media fake accounts, can spread such disinformative content among a large number of users literally in minutes. Moreover, disinformation can be unintentionally spread by AI solutions that the social media are built on - in our recent study, we showed that the recommender system at YouTube can enclose a user into misinformation filter bubble and bursting of such bubble is possible, albeit it may be difficult for some topics. It is clear that it will not be possible to cope with the amount of online content without the help of AI. Automatic detection, interactive support of fact-checkers, early warning systems and many other AI-based applications will become absolutely critical. AI-based solutions, however, currently pose a number of challenges - lack of appropriate labelled datasets, concept drifts or adversarial responses. And last but not least, since AI-generated news, published by credible sources (e.g., reports from sport events) will become quite widespread as well, it will be even more difficult to distinguish automatically generated disinformative content from the credible one.

This event is supported by the European Union's Horizon 2020 Research and Innovation programme under grant number 951911 – AI4Media

Asbjørn Følstad, SINTEF

Bio: Asbjørn Følstad is a Senior Researcher at SINTEF. His background is in psychology and human-computer interaction, and in his research, he seeks to understand how technology can be designed and developed to fit user needs and how new technology impacts our lives. Over the recent years, Asbjørn has researched chatbot user experience and interaction design, as well as its societal and ethical implications, as part of Norwegian and European research projects. He is also involved in research aiming for automatic identification of fake news based on language characteristics. He has co-organized the Chatbots for Social Good SIG at CHI 2018 and the CONVERSATIONS workshops series on chatbot research since 2017 (with the upcoming 2021 edition being the fifth one).

Statement: The landscape of online disinformation is changing rapidly. Actors spreading disinformation and fake news take advantage of new technologies and services for this purpose and the distribution of disinformation change in consequence of mitigation actions. At the same time, technological advances bring about new opportunities for tackling on the problem of disinformation. Advances in AI likely will increase the pace of change in the disinformation landscape, for better and worse. In my research on chatbots, I note how these potentially may be used for coaching and support – suggesting an opportunity for these also to combat disinformation. Also, research initiatives on the use of AI to combat fake news may be a source of hope. At SINTEF we are part of research on how AI can be used to identify fake news from its linguistic criteria. I am hopeful that online disinformation in the age of AI may be countered through a joint effort in society, and also through clever use of technology.

This event is supported by the European Union's Horizon 2020 Research and Innovation programme under grant number 951911 – AI4Media