Workshop template

Title of the workshop:	Artificial Intelligence and Disinformation
Date:	5 March 2019 10.00-12.00 ROOM 2
Name of the organizer/presenter:	Communication and Information Sector, UNESCO Moderator: Mr. Guy Berger Presenters: Ms. Wafa Ben-Hassine - Prof. Marc-Antoine Dilhac – Dr. Divina Frau-Meigs – Mr. Cordel Green – Mr. Pierre-Albert Ruquier – Ms. Elodie Vialle
Time allotted:	2 hours
About workshop:	Developments in artificial intelligence (AI) impact on the way information is produced, disseminated, and consumed; thus (re)shaping the communications landscape. In this regard, one of the most debated issues is the harnessing of AI to create and/or amplify disinformation, meaning fraudulent information (often in news formats) that is intended for political objectives and often to create confusion or to discredit verifiable journalistic reports. Disinformation in this sense includes incarnations such as counterfeit media brands, fake accounts and "deepfake" content. Beyond information and disinformation, there is also the issue of the abuse of AI in the use of other kinds of content formats - such as entertainment and memes - in order to create and spread hate-speech/incendiary messages.

Some of the most common concerns include: the potential of AI-powered solutions to amplify the dissemination of disinformation, interlacing with AI-powered social media algorithms; the capacity of AI to micro-target and establish psychometric profiles to subtly influence and manipulate individuals and groups; the impact of non-human generated information/journalism, especially when the automatic/bot nature is not explicit; automated trolling which has consequences for safety of journalists; as well as clickbait headlines, and the misleading use of captions, visuals or statistics. The above examples can have a chilling effect on privacy, as well as freedom of expression, including the sustainable development of free, independent, professional, and pluralist media.

However, AI has at the same time the potential to enhance citizens' self-determination and the exercise of human rights. AI has also been used for identifying disinformation and hate speech, including providing automated fact-checking solutions. Leading ICT and social media companies have been slowly committing to self-regulatory solutions, in dialogue with relevant international and national stakeholders. Regulators are increasingly taking up the use of big data in AI which can lead to privacy violations and related negative implications for freedom of expression. Advocates of media and information literacy are exploring how to empower Internet users with greater understanding of attention economics and big data and the way these become interlinked with AI. Academics are investigating how AI can help understand patterns in attacks on journalists.

Objectives of the workshop:	1. Explore key ways in which AI impacts on the production and dissemination of disinformation and hate speech. 2. Present responses by IGOs, NGOs, ICT stakeholders, and media coalitions, including by using AI itself to combat disinformation and hate speech. 3. Provide recommendations for multi-stakeholder and human rights-based responses and actions
Outputs:	Recommendations on tackling AI-fostered disinformation and hate speech, feeding into UNESCO's work in the field of freedom of expression, including for the 2019 World Press Freedom Day celebrated under the theme "Media for Democracy Journalism and Elections in Times of Disinformation."
Activity layout:	This session will examine the impact of AI on the information and communication environment through the issues above. Speakers will also present some of the ongoing initiatives in this fields proposed by media coalitions, IGOs, ICT platforms, and specialized NGOs.