

Module 3: The use of Artificial Intelligence to detect harmful practices during elections

[00:00:18] **Speaker 1** Hello! My name is Osama Alijaber, and I'm very happy to be here with you today. I am the digital democracy specialist at UNDP, Regional Hub for the Arab States. So, in the first module of the MOOC, you heard about issues brought up by the use of Artificial Intelligence for social media. I'm going to briefly summarize its impact online in today's session before talking about the use of A.I. to detect harmful practices such as misinformation, hate speech, and online propaganda during elections, but also how national authorities can fight back against these harmful practices.

[00:01:01] In 2020, we created every day 2.5 quintillion data bytes, and in the last 11 years, the volume of data generated, copied, and consumed worldwide grew by almost 5100%. The big data environment is one of the ways to look at the issue of information overload we are experiencing, resulting in a chaotic news environment and increasing the chances of viewing misinformation, hate speech, etc.. Nevertheless, it also gives big tech companies access to large amounts of data to power their own AI module, including but not limited to their own content moderation algorithms. Algorithms that can predict the content you will engage with the most and target you back with targeted advertising. While social media platforms rely on different AI-powered algorithms for content moderation, they are limited in many ways, mainly due to the accuracy and reliability of these algorithms and sometimes their [00:02:11] bias. [0.0s] AI development has promoted many discussions regarding how national authorities utilize machine learning to tackle the challenges they face, especially during elections. As misinformation during the electoral campaign, the weaponization of hate speech, and the targeting of voters to fake accounts and bots represent critical challenges to electoral integrity, they also contribute to undermining and distorting public debate and informed decision-making during elections. This can be very much dangerous, especially for young people, who can be manipulated during elections influencing their will and choices on Election Day. Therefore national stakeholders, including Electoral Management Bodies, civil society organizations, community-based organizations, and media institutions, can assess whether they should and can take advantage of AI in their scope of work and mandate to help combat information disorder online, including misinformation, hate speech and online violence against the disadvantaged group. However, one of the main questions that always comes up when talking about using AI is whether we can really trust the judgment and results of an AI system, especially with acknowledging that AI technology may inherit human biases due to the biases in the training dataset. And there are many, many examples of that. But there are other questions related to what AI systems can do and what they cannot do: Who should use AI to address harmful content? Can it support the core mandate of an institution or group? Who has the human and financial capacity to use AI without preventing other core activities? And finally, how to really navigate ethical discussions surrounding AI. Electoral Management Bodies, in general, have extensive tasks and managing the electoral process and may not be in the best place to take up such an extensive and sensitive task. Or they may not even have the legal mandate to monitor the electoral campaign. However, the Electoral Management Bodies that have the supervisory power to regulate the environment for the elections might benefit from these enhanced AI capacities provided, as long as they strictly abide by international and national human rights laws. It's also important for national stakeholders to examine all of these questions before deciding on whether to build their own model and potentially their own AI strategy. Also, it's important, too, to ensure safe, trusted, and inclusive AI systems by enhancing collaboration between government, industry, academy, and civil society organizations.

[00:05:29] AI systems development should ensure three main key principles, which are: fairness and accountability. and transparency. Fairness of an AI system means that national stakeholders need to examine clearly what AI can bring to the table, whether it's really fair to apply machine learning in the context of the first place, and if the system could violate the privacy and rights of a specific individual race, gender, and group. Accountability, on the other hand, is another essential principle, especially in ensuring an accountable system with an AI system going in the wrong direction, but also when an AI outcome contradicts citizen opinion and interests. Finally, AI transparency. Citizens need to know really that an AI system is being used, but also, there is a need to have transparency about the algorithms themselves. It's always important to remember that the quality of AI system output depends very much on its input data. Hence, cleaning the training data sets from conscious and unconscious bias assumptions allowed us to make better data-driven decisions. In this sub-module, the Supervisory Commission for Elections in Lebanon will introduce you to [00:06:55] **e-monitor blood, which** [0.5s] is an AI-powered algorithm that the UNDP build to monitor digital media as part of wider electoral support and assist the electoral process. As part of its core mandate, the judges of the Supervisory Commission for Elections in Lebanon, supported by a team of technicians and monitors, detected electoral violations, political polarization, misinformation, disinformation, hate speech, online violence against women, and campaignings finance online, in addition to their observation offline.