

Discrimination is a bigger AI risk than human extinction – EU commissioner

Publication Date: 2023-06-14

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Computing, European Union, Europe, news

Article URL: <https://www.theguardian.com/technology/2023/jun/14/ai-discrimination-is-a-bigger-risk-than-human-extinction-eu-chief>



Discrimination is a bigger threat posed by artificial intelligence than possible extinction of the human race, according to the EU's competition commissioner. Margrethe Vestager said although the existential risk from advances in AI may be a concern, it was unlikely, whereas discrimination from the technology was a real problem. She told the BBC "guardrails" were needed for AI, including for situations where it was being used for decisions that could affect livelihoods, such as mortgage applications or access to social services. "Probably [the risk of extinction] may exist, but I think the likelihood is quite small. I think the AI risks are more that people will be discriminated [against], they will not be seen as who they are," she said. "If it's a bank using it to decide whether I can get a mortgage or not, or if it's social services on your municipality, then you want to make sure that you're not being discriminated [against] because of your gender or your colour or your postal code." In the UK, the Information Commissioner's Office is investigating whether AI systems are showing racial bias when dealing with job applications. Regulators are concerned that AI tools will could produce outcomes that disadvantage certain groups if they are not represented accurately or fairly in the datasets that they are trained and tested on. Vestager's concerns echo some tech experts who argue that fears over existential-level risk related to AI are overshadowing more immediate risks such as AI-powered disinformation. The competition chief said calls for a moratorium on AI development, supported by Elon Musk and other senior figures, was unenforceable. AI regulation needed to be "global affair", Vestager said, but she warned that a UN-style approach would be difficult to implement. Rishi Sunak, the UK prime minister, has convened a global AI safety summit for "like-minded countries" this autumn and tech executives such as the Google chief executive, Sundar Pichai, and Elon Musk have called for global frameworks to regulate the technology. "Let's start working on a UN approach. But we shouldn't hold our breath," Vestager said. "We should do what we can here and now." The EU is working on legislation to oversee development and implementation of AI systems, which groups AI technology into four risk groups: unacceptable risk; high risk; limited risk; and minimal risk. AI systems overseeing credit scores and essential public services come into the high-risk category, meaning "clear requirements" will be set for those systems. Vestager's interjection came as the Irish Data Protection Commission blocked Google from launching its Bard chatbot in the EU over privacy concerns. The DPC, which is the chief European regulator for the California company, said it had not received sufficient information about how the tools would comply with the EU's General Data Protection Regulation. Google had intended to launch Bard in Europe this week, months after the chatbot's global release. Now, the regulator says, that will not happen. The DPC "had not had any detailed briefing nor sight of a data protection impact assessment or any supporting documentation at this

point”, the deputy commissioner Graham Doyle told Politico in an interview. A similar conflict happened in April, when the Italian regulator ordered the ChatGPT developer, OpenAI, to pause operations in the country over data protection concerns. The Italian data protection authority said that there appeared to be “no legal basis underpinning the massive collection and processing of personal data in order to ‘train’ the algorithms on which the platform relies”. OpenAI was eventually able to convince the regulator that it was in compliance and relaunched services with limited changes.