

Ministers not doing enough to control AI, says UK professor

Publication Date: 2023-05-13

Author: Harry Taylor

Section: Technology

Tags: Artificial intelligence (AI), Consciousness, news

Article URL: <https://www.theguardian.com/technology/2023/may/13/ministers-not-doing-enough-to-control-ai-says-uk-professor>



One of the professors at the forefront of artificial intelligence has said ministers are not doing enough to protect against the dangers of super-intelligent machines in the future. In the latest contribution to the debate about the safety of the ever-quicken development of AI, Prof Stuart Russell told the Times that the government was reluctant to regulate the industry despite the concerns that the technology could get out of control and threaten the future of humanity. Russell, a lecturer at the University of California in Berkeley and former adviser to the US and UK governments, told the Times he was concerned that ChatGPT, which was released in November, could become part of a super-intelligent machine that could not be constrained. “How do you maintain power over entities more powerful than you – for ever?” he asked. “If you don’t have an answer, then stop doing the research. It’s as simple as that. “The stakes couldn’t be higher: if we don’t control our own civilisation, we have no say in whether we continue to exist.” After the release of ChatGPT to the public last year, which has been used to write prose and has already worried lecturers and teachers over its use in universities and schools, the debate has intensified over its safety in the long-term. Elon Musk, the Tesla founder and Twitter owner, and the Apple co-founder Steve Wozniak, along with 1,000 AI experts, wrote a letter to warn that there was an “out-of-control race” going on at AI labs and called for a pause on the creation of giant-scale AI. The letter warned the labs were developing “ever more powerful digital minds that no one, not even their creators, can understand, predict or reliably control”. There is also concern about its wider application. A House of Lords committee this week heard evidence from Sir Lawrence Freedman, a war studies professor, who spoke about the concerns on how AI might be used in future wars. Google’s rival, Bard, is due to be released in the EU later this year. Russell himself previously worked for the UN on how to monitor the nuclear test-ban treaty, and was asked to work with Whitehall earlier this year. He said: “The Foreign Office ... talked to a lot of people and they concluded that loss of control was a plausible and extremely high-significance outcome.” “And then the government came out with a regulatory approach that says: ‘Nothing to see here ... we’ll welcome the AI industry as if we were talking about making cars or something like that’. “I think we got something wrong right at the beginning, where we were so enthralled by the notion of understanding and creating intelligence, we didn’t think about what that intelligence was going to be for,” he said. “Unless its only purpose is to be a benefit to humans, you are actually creating a competitor – and that would be obviously a stupid thing to do. “We don’t want systems that imitate human behaviour ... you’re basically training it to have human-like goals and to pursue those goals. “You can only imagine how disastrous it would be to have really capable systems that were pursuing those kinds of goals.”