

# AI could cause ‘catastrophic’ financial crisis, says Yuval Noah Harari

Publication Date: 2023-11-09

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Yuval Noah Harari, Computing, Financial crisis, Economics, Banking, news

Article URL: <https://www.theguardian.com/technology/2023/nov/09/yuval-noah-harari-artificial-intelligence-ai-cause-financial-crisis>



Artificial intelligence could cause a financial crisis with “catastrophic” consequences, according to the historian and author Yuval Noah Harari, who says the technology’s sophistication makes forecasting its dangers difficult. Harari told the Guardian a concern about safety testing AI models was foreseeing all the problems that a powerful system could cause. Unlike with nuclear weapons, there was not one “big, dangerous scenario” that everyone understood, he said. “With AI, what you’re talking about is a very large number of dangerous scenarios, each of them having a relatively small probability that taken together ... constitutes an existential threat to the survival of human civilisation.” The Sapiens author, who has been a prominent voice of concern over AI development, said last week’s multilateral declaration at the global AI safety summit in Bletchley Park was a “very important step forward” because leading governments had come together to express concern about the technology and to do something about it. “Maybe the most encouraging or hopeful thing about it is that they could get not just the European Union, the UK, the United States, but also the government of China to sign the declaration,” he said. “I think that was a very positive sign. Without global cooperation, it will be extremely difficult, if not impossible, to rein in the most dangerous potential of AI.” The summit concluded with an agreement between 10 governments – including the UK and US but not China – the EU and major AI companies, including the ChatGPT developer OpenAI and Google, to cooperate on testing advanced AI models before and after they are released. Harari said one issue with safety testing of models was foreseeing all problems that a powerful system could cause. “AI is different from every previous technology in human history because it’s the first technology that can make decisions by itself, that can create new ideas by itself and that can learn and develop by itself. Almost by definition, it’s extremely difficult for humans, even the humans who created the technology, to foresee all the potential dangers and problems.” Governments cited the threat of AI systems helping to create bioweapons when trying to explain its dangers to the public, said Harari, but there were other scenarios that could be considered. The author pointed to finance as a sector ideally suited to artificial intelligence systems – “this is the ideal field for AI because it’s only data” – and a potential source of a serious AI-made crisis. “What happens if AI is not only given greater control over the financial system of the world, but it starts creating new financial devices that only AI can understand, that no human being can understand?” said Harari, adding that the 2007-08 financial crisis was caused by debt instruments such as collateralised debt obligations (CDOs) that few people understood and were thus inadequately regulated. “AI has the potential to create financial devices which are orders of magnitude more complex than CDOs. And just imagine the situation where we have a financial system that no human being is able to understand and therefore also not able to

regulate,” he said. “And then there is a financial crisis and nobody understands what is happening.” Last month, the UK government raised concerns about an advanced AI model potentially creating an existential threat by controlling and manipulating financial systems. But Harari said an AI-created financial crisis would not destroy human civilisation – “at least not directly”. He added: “It could, indirectly, if it triggers certain kinds of wars or conflicts. It’s a catastrophic risk – economic, social, political – but by itself I wouldn’t describe it as existential.” The Israeli author, who has backed calls for a six-month pause in advanced AI development and supports making artificial intelligence companies liable for damage their products cause, said the focus should not be on specific regulations and laws but on regulatory institutions with knowledge of the technology that can react quickly as new breakthroughs emerge. “We need to create as fast as possible, powerful regulatory institutions that are able to identify and react to the dangers as they arise based on the understanding that we cannot predict all the dangers and problems in advance and legislate against them in advance.” He added: “This should be the main effort, not the effort to now write some very long and complicated regulation that by the time it passes parliament or congress, might be outdated.” As part of that setup, AI safety institutes should hire experts who understood AI’s potential impact on the world of finance, said Harari. Last month Rishi Sunak announced the establishment of a UK AI safety institute, followed days later by the White House announcing plans for a similar body, with both expected to play key roles in testing advanced AI models. Speaking at the summit, Sunak said the UK needed to understand the capabilities of advanced models first before introducing legislation to deal with them. A spokesperson for the Department for Science, Innovation and Technology said a recent white paper on AI had flagged the UK’s financial regulators, the Financial Conduct Authority and Prudential Regulation Authority, as the appropriate watchdogs for AI and finance. “They understand the risks in their sectors and are best placed to take a proportionate approach to regulating AI,” said the spokesperson.