Bernie Sanders, Elon Musk and White House seeking my help, says 'godfather of AI'
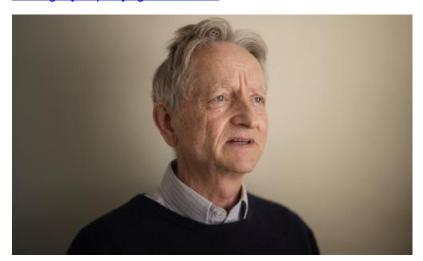
The man often touted as the godfather of artificial intelligence will be responding to requests for help from Bernie Sanders, Elon Musk and the White House, he says, just days after quitting Google to warn the world about the risk of digital intelligence. Dr Geoffrey Hinton, 75, won computer science's highest honour, the Turing award, in 2018 for his work on "deep learning", along with Meta's Yann Lecun and the University of Montreal's Yoshua Bengio. The technology, which now underpins the AI revolution, came about as a result of Hinton's efforts to understand the human brain – efforts which convinced him that digital brains might be about to supersede biological ones. But the London-born psychologist and computer scientist might not offer the advice the powerful want to hear. "The US government inevitably has a lot of concerns around national security. And I tend to disagree with them," he told the Guardian. "For example, I'm sure that the defense department considers that the only safe hands for this stuff is the US defense department – the only group of people to actually use nuclear weapons. "I'm a socialist," Hinton added. "I think that private ownership of the media, and of the 'means of computation', is not good. "If you view what Google is doing in the context of a capitalist system, it's behaving as responsibly as you could expect it to do. But that doesn't mean it's trying to maximise utility for all people: it's legally obliged to maximise utility for its shareholders, and that's a very different thing." Hinton has been fielding a new request to talk every two minutes since he spoke out on Monday about his fears that AI progress could lead to the end of civilisation within 20 years. But when it comes to offering concrete advice, he is lost for words. "I'm not a policy guy," he says. "I'm just someone who's suddenly become aware that there's a danger of something really bad happening. I wish I had a nice solution, like: 'Just stop burning carbon, and you'll be OK.' But I can't see a simple solution like that." In the past year, the rapid progress in AI models convinced Hinton to take seriously the threat that "digital intelligence" could one day supersede humanity's. "For the last 50 years, I've been trying to make computer models that can learn stuff a bit like the way the brain learns it, in order to understand better how the brain is learning things. But very recently, I decided that maybe these big models are actually much better than the brain. "We need to think hard about it now, and if there's anything we can do. The reason I'm not that optimistic is that I don't know any examples of more intelligent things being controlled by less intelligent things. "You need to imagine something that is more intelligent than us by the same degree that we are more intelligent than a frog. It's all very well to say: 'Well, don't connect them to the internet,' but as long as they're talking to us, they can make us do things." Even outside of the existential risk, Hinton has other concerns about the rapid growth in power of AI models, citing the influence of Cambridge Analytica backer Robert Mercer on political campaigns on both sides of the Atlantic. "Bob Mercer and Peter Brown, when they were working at

IBM on translation, understood the power of having a lot of data. Without Bob Mercer, Trump might well have not yet got elected. "And Bob must have understood the power of manipulation that big data could give you, and so I think already had terrible consequences there." Authoritarian governments, he says, are the biggest red flag that suggests that humanity will not be able to get hold of the risks of AI before it's too late. "This stuff helps authoritarian governments in destroying truth, or manipulating electorates. And having to deal with these threats, in a situation where Americans can't even agree to not give assault rifles to teenage boys, that's not a hard thing to think about. "In Uvalde [the 2022 massacre of 21 people at an elementary school in Texas], there were 200 policemen who didn't dare go through a door because the guy on the other side had an assault rifle and was shooting children. "And yet, they can't decide not to ban assault weapons. So a totally dysfunctional political system like that is just not the right system to have to deal with these threats."