AI watch: from deepfakes to a rock star humanoid

Artificial intelligence is either going to save humanity or finish it off, depending on who you speak to. Either way, every week there are new developments and breakthroughs. Here are some of the AI stories that have emerged in recent days. • The consumer champion Martin Lewis has urged the government to take action against AI-powered generative deepfakes after he found that scammers were using an artificially generated version of him to defraud consumers. Lewis posted a fake video on Thursday of him apparently backing an Elon Musk project, and warned that without action against similar videos lives would be ruined. He tweeted: "This is frightening, it's the first deep fake video scam I've seen with me in it. Government and regulators must step up to stop big tech publishing such dangerous fakes. People will lose money and it will ruin lives." The president of Microsoft, Brad Smith, said last month he expected tech firms to launch an initiative for watermarking AI-generated content, which would be one necessary step against fraudsters. • Your Twitter feed not working? Blame AI. Musk, one of the siren voices on the rapid pace of AI development, said the technology was partly the cause of his decision to limit views of posts last weekend. The Twitter owner, who has joined calls for a hiatus in building powerful AI systems, said the platform was being affected by companies "scraping" tweets from the site to train AI programs. AI tools such as chatbots rely on vast amounts of data to construct the models that underpin them, with Musk claiming the scraping was putting pressure on Twitter's servers (which store and process the data behind Twitter posts), so limits on viewing tweets were imposed. However, one former Twitter executive said blaming data scraping for the move did not "pass the sniff test". • Two authors are suing the company behind the ChatGPT chatbot in another data-scraping row. Mona Awad, whose books include Bunny and 13 Ways of Looking at a Fat Girl, and Paul Tremblay, author of The Cabin at the End of the World, are suing San Francisco-based OpenAI in the US, claiming that their works were unlawfully "ingested" and "used to train" ChatGPT. Such lawsuits will add to the pressure on AI firms to be transparent about the data used to train their models. • The historian and author Yuval Noah Harari warned that "trust will collapse" if AI-powered fake accounts proliferate unchecked on social media. Speaking in Geneva at the annual United Nations AI for Good summit this week, he said tech executives should face the threat of jail sentences if they do not take measures against bot accounts. "What happens if you have a social media platform where … millions of bots can create content that is in many ways superior to what humans can create – more convincing, more appealing," he said. "If we allow this to happen, then humans have completely lost control of the public conversation. Democracy will become completely unworkable." The ability of generative AI – the catch-all term for AI tools that can rapidly mass-produce convincing text, image and voice – to create disinformation is a common cause of alarm among experts. • As the summit's title suggested, however, it also made the case for positive uses of the technology as humanoid robots turned up in force at Geneva. Ai-da, an artist robot, offered opinions about art while Desdemona, a

rock star humanoid, performed with a human backing band. Another AI-powered robot, Nadia, was presented as an alternative to human carers for the sick and elderly people – and has been used at a home for older people in Singapore, playing bingo and talking to residents.