Who is Helen Toner the Australian woman ousted from the board of OpenAI?

After a tumultuous few days at OpenAI, Sam Altman has returned to the helm. But who is the young Australian board member who was reportedly in dispute with the chief executive in the lead up to his firing? Helen Toner, along with two of the other three board members responsible for firing Altman less than a week ago, is now off the board of OpenAI. The new-look board, announced on Wednesday, is chaired by Bret Taylor, the former co-chief executive of software firm Salesforce and includes Larry Summers, the former US treasury secretary, and Adam D'Angelo, the tech entrepreneur and existing board member who also played a role in Altman's firing. The move to reinstate Altman and reform the board appears to have staved off an exodus of nearly all of OpenAI's 750-plus workforce, who had threatened to quit if Altman was not reinstated. Toner, a 30-something University of Melbourne graduate, was on the OpenAI board for two years until Wednesday and is the director of strategy at Georgetown's Center for Security and Emerging Technology (CSET). She joined the company after a career studying AI and the relationship between the United States and China. In appointing her to the board, OpenAI said that Toner was a trusted voice on the national security implications of AI and machine learning and advocated for transparency on AI issues to reduce risk. Altman said in the statement at the time that Toner "brings an understanding of the global AI landscape with an emphasis on safety, which is critical for our efforts and mission." However, it appears Toner's research work on AI safety at CSET might have been the catalyst that led to the past week of turmoil at Open AI. Sign up for Guardian Australia's free morning and afternoon email newsletters for your daily news roundup The New York Times reported this week that in the weeks leading up to Altman's firing, he and Toner had discussed an October paper she had co-authored for CSET. In the paper, OpenAI is criticised for releasing ChatGPT at the end of last year, sparking "a sense of urgency inside major tech companies", like Google, to ensure they did not fall behind and prompting competitors to "accelerate or circumvent internal safety and ethics review processes". The paper then goes on to detail how OpenAI's competitor, Anthropic, had previously decided to delay the release of its own chatbot, Claude, in order to avoid "advanc[ing] the rate of AI capabilities progress". "Anthropic had deliberately decided not to productize its technology in order to avoid stoking the flames of AI hype," the paper states. "Once a similar product (ChatGPT) was released by another company, this reason not to release Claude was obviated, so Anthropic began offering beta access to test users before officially releasing Claude as a product in March." The NYTimes report said Altman complained in an email about the paper's criticism of OpenAI and its praise of Anthropic's approach. The report said Toner defended the academic paper as an analysis of the challenges in developing AI. But Altman reportedly said it amounted to criticism from a board member and indicated it was damaging to OpenAI. At the

time of Altman's firing, the former board said he was sacked over a failure to be "candid in his communications", but it declined to give details over what communications it was referring to. Since living in the US, Toner's Australian accent appears to have faded, but as recently as 2019 Toner was not a US citizen, noting at the time that not being a US citizen prevented her from accessing some security clearances for her work. "I'm very fortunate to be Australian, which I think is possibly the least suspicious country from a US perspective," she said on the 80,000 hours podcast. "There are some clearances that are open to people from the Five Eyes countries. That's an intelligence alliance between the US, UK, Canada, New Zealand, Australia. But the types of clearances that would relevant to me, I would need to be a US citizen," she said. Toner has described herself in the past as an effective altruist, a philosophical movement that supports the maximisation of earnings in order to maximise charitable donations. Toner worked at GiveWell, an effective altruism charity, in 2015 to 2016. In a speech she gave during the time, she said her view on good management at GiveWell was encapsulated in a meeting with a co-founder who asked her what her least favourite thing about working at the organisation was. She replied it was a lack of SodaStreams in the office. The founders then bought one, she said. "It demonstrates the importance of giving people lots of opportunities to communicate," she said, and also allowed employees to give negative feedback. Toner said 95% of the charity's internal communication was done via emoji, and highlighted the thumbs up and the "TMI" emoji as two favourites to indicate approval, or something a fellow employee didn't need to know. "I think it's really, really useful to have these two options … I think they're an example of making it simple to check in and to drive a project forward by yourself but make it simple to go to others with your manager." Crikey reported this week Toner's former classmates and fellow youth United Nations participants described her as high achieving, nice and very smart. Guardian Australia put questions to Toner via LinkedIn but did not received a response by deadline. Comment was also sought from OpenAI. In her first tweet since OpenAI revealed the new board and the return of Altman, Toner said "And now, we all get some sleep."