

Oxford shuts down institute run by Elon Musk-backed philosopher

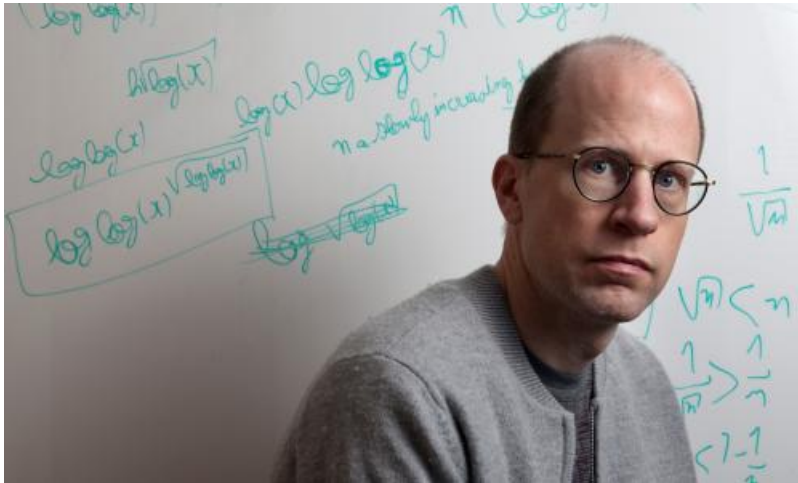
Publication Date: 2024-04-19

Author: Nick Robins-Early

Section: Technology

Tags: [Elon Musk](#), [University of Oxford](#), [Philosophy](#), [news](#)

Article URL: <https://www.theguardian.com/technology/2024/apr/19/oxford-future-of-humanity-institute-closes>



Oxford University this week shut down an academic institute run by one of Elon Musk's favorite philosophers. The Future of Humanity Institute, dedicated to the long-termism movement and other Silicon Valley-endorsed ideas such as effective altruism, closed this week after 19 years of operation. Musk had donated £1m to the FHI in 2015 through a sister organization to research the threat of artificial intelligence. He had also boosted the ideas of its leader for nearly a decade on X, formerly Twitter. The center was run by Nick Bostrom, a Swedish-born philosopher whose writings about the long-term threat of AI replacing humanity turned him into a celebrity figure among the tech elite and routinely landed him on lists of top global thinkers. Sam Altman of OpenAI, Bill Gates of Microsoft and Musk all wrote blurbs for his 2014 bestselling book *Superintelligence*. "Worth reading *Superintelligence* by Bostrom. We need to be super careful with AI. Potentially more dangerous than nukes," Musk tweeted in 2014. Bostrom resigned from Oxford following the institute's closure, he said. The closure of Bostrom's center is a further blow to the effective altruism and long-termism movements that the philosopher had spent decades championing, and which in recent years have become mired in scandals related to racism, sexual harassment and financial fraud. Bostrom himself issued an apology last year after a decades-old email surfaced in which he claimed "Blacks are more stupid than whites" and used the N-word. Bostrom – who popularized the theory that humanity may be living in a simulation, one that Musk often repeats – spoke about the closure of the institute in a lengthy final report published on its website this week. He praised the work of the center, while also saying that it faced "administrative headwinds" from Oxford and its philosophy department. "The closure is the culmination of process that's been playing out over several years," Bostrom said via email. "We were funded initially for three years, back in 2005, and then that got extended a number of times. "Eventually a pressure to conform began bearing down (we were administratively housed within the faculty of philosophy, even though the majority of our research team by this time were non-philosophers), and there was a death by bureaucracy." Bostrom added that he was touched by the number of people speaking out in support of the institute's work, and that it had been a privilege to work with his colleagues. "FHI was a special place with a unique and highly fruitful intellectual culture," Bostrom said. "I think we had a good run!" A statement on the Future of Humanity's website claimed that Oxford had frozen fundraising and hiring in 2020, and that in late 2023 the faculty of philosophy decided to not renew the contracts of remaining staff at the institute. An Oxford University spokesman said: "We regularly consider the best structures for conducting our academic research, as part of the university's governance processes. After such consideration, the decision was made to close the Future of Humanity Institute. The university recognises the Institute's important contribution to this emerging field, which researchers elsewhere across the university are likely to continue." Effective altruism, the utilitarian belief that people should focus

their lives and resources on maximizing the amount of global good they can do, has become a heavily promoted philosophy in recent years. The philosophers at the center of it, such as Oxford professor William MacAskill, also became the subject of immense amounts of news coverage and glossy magazine profiles. One of the movement's biggest backers was Sam Bankman-Fried, the now-disgraced former billionaire who founded the FTX cryptocurrency exchange. Bostrom is a proponent of the related long-termism movement, which held that humanity should concern itself mostly with long-term existential threats to its existence such as AI and space travel. Critics of long-termism tend to argue that the movement applies an extreme calculus to the world that disregards tangible current problems, such as climate change and poverty, and veers into authoritarian ideas. In one paper, Bostrom proposed the concept of a universally worn "freedom tag" that would constantly surveil individuals using AI and relate any suspicious activity to a police force that could arrest them for threatening humanity. Bostrom and long-termism gained numerous powerful supporters over the years, including Musk and other tech billionaires. Bostrom's Institute received £13.3m in 2018 from the Open Philanthropy Project, a non-profit financially backed by Facebook co-founder Dustin Moskovitz. The past few years have been tumultuous for effective altruism, however, as Bankman-Fried's multibillion-dollar fraud marred the movement and spurred accusations that its leaders ignored warnings about his conduct. Concerns over effective altruism being used to whitewash the reputation of Bankman-Fried, and questions over what good effective altruist organizations are actually doing, proliferated in the years since his downfall. Meanwhile, Bostrom's email from the 1990s resurfaced last year and resulted in him issuing a statement repudiating his racist remarks and clarifying his views on subjects such as eugenics. Some of his answers – "Do I support eugenics? No, not as the term is commonly understood" – led to further criticism from fellow academics that he was being evasive. The university launched an investigation into Bostrom's conduct following the discovery of his racist email, while other major effective altruism groups distanced themselves from him. "We unequivocally condemn Nick Bostrom's recklessly flawed and reprehensible words," the Centre for Effective Altruism, which was founded by fellow Oxford philosophers and financially backed by Bankman-Fried, said in a statement at the time. • This article was amended on 21 April 2024 to add a comment from Oxford University.