

Authors shocked to find AI ripoffs of their books being sold on Amazon

Publication Date: 2023-09-30

Author: James Tapper

Section: Technology

Tags: Artificial intelligence (AI), The Observer, Amazon, E-commerce, Intellectual property, Society of Authors, Computing, Autobiography and memoir, features

Article URL: <https://www.theguardian.com/technology/2023/sep/30/authors-shocked-to-find-ai-ripoffs-of-their-books-being-sold-on-amazon>



Publishing a book is a big occasion for any writer, and Rory Cellan-Jones is no exception. “Like any author, I obsessively check Amazon,” he said. “And this thing popped up.” The former BBC technology correspondent wrote a memoir untangling the truth about his family history. What had popped up on the Amazon website was a biography of Cellan-Jones, with a naively designed cover by someone he had never heard of. “I thought: ‘This is strange – who’s writing a biography of me?’” Cellan-Jones told the Observer. “I don’t kid myself. It’s difficult enough for me to sell books about myself, [let alone] for other people to sell books about me.” But glancing at a few passages revealed that Cellan-Jones had fallen victim to someone attempting to piggyback on his memoir by releasing a title with text apparently generated by artificial intelligence – one of an influx of AI titles since the emergence of ChatGPT enabled people to generate pages of text rather than bothering to write it. Cellan-Jones’s book, *Ruskin Park: Sylvia, Me and the BBC*, describes how he discovered a shoebox of letters from his mother detailing her love affair with his father, a BBC TV director he only met at the age of 23, and how she came to spend most of her life in a one-bedroom south London flat. It is, he said, “about growing up with a single mother and a father who wasn’t there”. The book rivaling his family detective story was “complete fantasy”, Cellan-Jones said. “There are passages about the Cellan-Joneses, an academic family sat around the table ... His father, a kindly academic; his mother, a teacher. Just complete baloney. “Then Amazon sent me an email saying: ‘You might like this.’ Their algorithm had decided this was a bloody book I would want rather than recommending my book that I’ve slaved long and hard over ... They’re effectively allowing book spam and recommending it to the very person who is most annoyed by it.” The ersatz biography and other titles by the pseudonymous author were removed by Amazon, but plenty more get through the filters intended to weed out low-quality books. It has been easy for bookspammers to release dozens of titles in a day using Amazon’s Kindle Direct Publishing (KDP) system, which enables authors to self-publish ebooks and printed books. Someone styled as “Steven Walryn” published more than 30 books, mostly nonsensical and repetitive guides on how to use camera brands, as well as a couple of fantasy romances, with 15 published on the same day in May. They were removed by Amazon last week. Amazon could not say how many books it prevents from being published or how many were taken down. In August, Jane Friedman, who writes about publishing, forced it to remove five bogus titles in her name that appeared to be AI-generated. “Amazon is clearly facing significant challenges with the influx of AI-generated products in its stores, and it appears to be playing catchup,” said Nicola Solomon, the chief executive of the Society of Authors (SoA). A few weeks

ago, the firm said publishers of new KDP publications would need to declare if they included AI-generated content and would be limited to publishing three books a day, moves welcomed by Solomon. "But these small fixes seem more designed to benefit Amazon's processes than readers and human authors," she added. The SoA wants Amazon to clearly label products as AI-generated and allow readers to filter out AI titles. The problem is similar to that encountered by musicians, who face competition for a slice of Spotify's royalty pot from people uploading white noise to the streaming service. Authors such as Margaret Atwood, Viet Thanh Nguyen and Philip Pullman are also concerned their work is being used in large language models to train AI without their consent, or any compensation or credit. Rashik Parmar, group chief executive of BCS, the chartered institute for IT, said ministers needed to introduce legislation to ensure that AI-generated material included a digital watermark so that it could be easily tracked. "The UK is uniquely placed to become the home of responsible computing," Parmar said, if legislation can help British IT professionals demonstrate a higher ethical standard. "Could 'coded in Britain' become a moniker for something special?" An Amazon spokesperson said that all publishers had to adhere to its content guidelines. "We invest significant time and resources to ensure our guidelines are followed and remove books that do not adhere to them ... While we allow AI-generated content, we don't allow AI-generated content that violates our content guidelines, including content that creates a disappointing customer experience."