

THE CONVERSATION

Academic rigour, journalistic flair



Ahmet Misirligul/Shutterstock

Science fiction could save us from bad technology

September 24, 2019 1.04pm BST

The short film *Slaughterbots* depicts a near future in which swarms of micro drones assassinate thousands of people for their political beliefs. Released in November 2017 by academics and activists warning of the dangers of advanced artificial intelligence (AI), it quickly went viral, attracting over 3m views to date. It helped spark a public debate on the future of autonomous weapons and put pressure on diplomats meeting at the United Nations Convention on Conventional Weapons.

But this kind of speculative science fiction storytelling isn't just useful for attracting attention. The people who design and build advanced technology can use stories to consider the consequences of their work and ensure it is used for good. And we think this kind of "science fiction prototyping" or "design fiction" could help prevent human biases from working their way into new technology, further entrenching society's prejudices and injustices.

Authors



Alessio Malizia

Professor of User Experience Design,
University of Hertfordshire



Silvio Carta

Head of Art and Design and Chair of the
Design Research Group, University of
Hertfordshire

Slaughterbots



A bias can lead to the arbitrary preference of some categories (of results, people, or ideas) over others. For example, some people may be biased against hiring women for executive jobs, whether they are conscious of it or not.

Technology built around data that records such bias can end up **replicating the problem**. For instance, recruitment software designed to select the best CVs for a particular job might be programmed to look for characteristics that reflect an unconscious bias towards men. In which case, the algorithm will end up favouring men's CVs. And this isn't theoretical – it actually happened to Amazon.

Designing algorithms without considering possible negative implications **has been compared to** doctors “writing about the benefits of a given treatment and completely ignoring the side effects, no matter how serious they are”.

Some tech firms and researchers are trying to tackle the issue. For example, Google drew up a set of **ethical principles** to guide its development of AI. And UK academics have launched an initiative called **Not-Equal** that aims to encourage greater fairness and justice in the design and use of technology.

The problem is that, publicly, companies tend to deliver only a positive vision of the potential consequences of near-future technologies. For example, driverless cars are often portrayed as solving all our transport issues from cost to safety, ignoring the increased **dangers of cyberattacks** or the fact they could encourage people to walk or cycle less.

The difficulty in understanding how digital technologies work, especially those that are heavily driven by obscure algorithms, also makes it harder for people to have a complex and comprehensive view of the issues. This situation produces a tension between a reassuring positive narrative and the vague suspicion that biases are embedded to some degree in the technologies around us. This is where we think storytelling through design fiction can come in.

Stories are a natural method of thinking about possibilities and complex situations, and we have been hearing them all our lives. Science fiction can help us speculate on the impact of near-future technologies on society, as *Slaughterbots* does. This can even include issues of social justice, like the way certain groups, such as refugees and migrants, can be **excluded from digital innovations**.

Revealing the (possible) future

Design fiction stories provide a novel way for designers, engineers and futurists (among others) to think about the impact of technology from a human perspective and link this to possible future needs. With a mixture of logic and imagination, design fiction can reveal aspects of how technology may be adopted and used, starting conversations about its future ramifications.

For example, the short story “Crime-sourcing” explores what might happen if AI was to use crowdsourced information and a criminal database to predict who might commit a murder. The researchers found that because the database was full of people in minority ethnic groups who, for social reasons, were statistically more likely to reoffend, the “crime-sourcing” model was more likely to wrongly suspect minorities than white people.

You don’t have to be a talented writer or make a slick film to produce design fiction. Brainstorming activities involving cards and storyboards have been used to develop design fiction and help develop the storytelling process. Making workshops that used these kinds of tools more common would enable more engineers, entrepreneurs and policymakers to use this method of assessment. And making the resulting work publicly available would help to expose potential biases in technologies before they affect society.

Encouraging designers to create and share more stories in this way would ensure the narrative that underpins new technology wouldn’t just present a positive picture, nor an extremely negative or dystopian one. Instead, people will be able to appreciate both aspects of what is happening around us.



Science fiction prototype Algorithm bias