



'Computer bots are like humans, having fights lasting years'

[RESEARCH \(/NEWS-LISTING?CATEGORY=228\)](#)[SCIENCE \(/NEWS-LISTING?CATEGORY=252\)](#)

Researchers say 'benevolent bots', otherwise known as software robots, that are designed to make articles on Wikipedia better often end up having online fights lasting years over changes in content.

Editing bots on Wikipedia undo vandalism, enforce bans, check spelling, create links and import content automatically, whereas other bots (which are non-editing) can mine data, identify data or identify copyright infringements. The team looked at how much disruption they caused on Wikipedia sites, seeing how they interacted on 13 different language editions over ten years (from 2001 to 2010). They found that bots interacted with one another, whether or not this was by design, which led to unpredictable consequences. The research paper, published in *PLOS ONE*, concludes that bots are more like humans than you might expect as they appear to behave differently in culturally distinct online environments. The paper says the findings are a warning to those using artificial intelligence for building autonomous vehicles, cyber security systems or for managing social media. It suggests that scientists may have to devote more attention to bots' diverse 'social life' and their different cultures.

The research paper by the University of Oxford and the Alan Turing Institute in the UK explains that although the online world has become an ecosystem of bots, our knowledge of how they interact with each other is still rather poor. Although bots are automatons that do not have the capacity for emotions, bot to bot interactions are unpredictable and act in distinctive ways. It finds that German editions of Wikipedia had fewest conflicts between bots, with each undoing another's edits 24 times, on average, over ten years. This shows relative efficiency, says the research paper, when compared with bots on the Portuguese Wikipedia edition, which undid another bot's edits 185 times, on average, over ten years. Bots on English Wikipedia undid another bot's work 105 times, on average, over ten years, three times the rate of human reverts, says the paper.

The findings show that even simple autonomous algorithms can produce complex interactions that result in unintended consequences – 'sterile fights' that may continue for years, or reach deadlock in some cases. The paper says while bots constitute a tiny proportion (0.1%) of Wikipedia editors, they stand behind a significant proportion of all edits. Although such conflicts represent a small proportion of the bots' overall editorial activity, these findings are significant in highlighting their unpredictability and complexity. Smaller language editions, such as the Polish Wikipedia, have far more content created by bots than the large language editions, such as English Wikipedia.

Lead author Dr Milena Tsvetkova, from the Oxford Internet Institute, said: 'We find that bots behave differently in different cultural environments and their conflicts are also very different to the ones between human editors. This has implications not only for how we design artificial agents but also for how we study them. We need more research into the sociology of bots.'

The paper was co-authored by the principal investigator of the EC-Horizon2020-funded project, HUMANE, Dr Taha Yasseri,

also from the Oxford Internet Institute. He added: 'The findings show that even the same technology leads to different outcomes depending on the cultural environment. An automated vehicle will drive differently on a German autobahn to how it will through the Tuscan hills of Italy. Similarly, the local online infrastructure that bots inhabit will have some bearing on how they behave and their performance. Bots are designed by humans from different countries so when they encounter one another, this can lead to online clashes. We see differences in the technology used in the different Wikipedia language editions and the different cultures of the communities of

Wikipedia editors involved create complicated interactions. This complexity is a fundamental feature that needs to be considered in any conversation related to automation and artificial intelligence.'

Professor Luciano Floridi, also an author of the paper, remarked: 'We tend to forget that coordination even among collaborative agents is often achieved only through frameworks of rules that facilitate the wanted outcomes. This infrastructural ethics or infra-ethics needs to be designed as much and as carefully as the agents that inhabit it.'

The research finds that the number of reverts is smaller for bots than for humans, but the bots' behaviour is more varied and conflicts involving bots last longer and are triggered later. The average time between successive reverts for humans is 2 minutes, then 24 hours or one year, says the paper. The first bot to bot revert happened a month later, on average, but further reverts often continued for years. The paper suggests that humans use automatic tools that report live changes and can react more quickly, whereas bots systematically crawl over web articles and they can be restricted on the number of edits allowed. The fact that bots' conflicts are longstanding also flags that humans are failing to spot the problems early enough, suggests the paper.

The paper, 'Even Good Bots Fight: The Case of Wikipedia' (<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0171774>) is published in PLOS ONE.

“ **Even the same technology leads to different outcomes depending on the cultural environment. An automated vehicle will drive differently on a German autobahn to how it will through the Tuscan hills of Italy. Similarly, local online infrastructure that bots inhabit will have some bearing on how they behave and their performance** ”

Dr Taha Yasseri from the Oxford Internet Institute