How can we understand AI?

Published November 2024

Dr Milena Tsvetkova,
Associate Professor of
Computational Social Science,
Department of Methodology, LSE

AI is transforming our world. But as the relationship between humans and machines evolves, a new approach is needed to study its impacts, argues **Milena Tsvetkova**.

In 2017, Dr Milena Tsvetkova, then a postdoctoral researcher at the Oxford Internet Institute, co-authored a paper on surprising anomalies found in interactions between groups of small, artificially intelligent (AI) bots.

"These AI bots were designed by different Wikipedia editors to carry out simple maintenance tasks such as fixing specific grammatical errors," she explains. "We found that while they were not designed to interact, they would undo each other's work – one bot would make their change, and another bot would come along and amend it. In some cases, we found backwards and forwards between the same bots had continued over years."

Dr Tsvetkova, now Associate Professor in the Department of Methodology at LSE, believes that it was this project that sparked her interest in AI behaviour. **In new collaborative research**, she further explores the consequences of both expected and unintended interactions between AI machines, and now also between humans and AI machines.

"We have a multiplicity of AI machines being integrated into our social world, forming networks of complex relationships between machines and humans," Dr Tsvetkova says. "As our paper highlights, robotic trains and cars drive us around, auction bots outbid us for purchases, ChatGPT answers our questions, and social media bots feed us information and news. We've reached a point in AI and human interactions where machines are now social actors."

For Dr Tsvetkova and her colleagues, the acknowledgement of machines as social actors opens doors for the extension of an existing academic discipline – a new "sociology of humans and machines".

A new sociology of humans and machines

The field of sociology traditionally studies human society and social patterns as well as phenomena that emerge from human behaviour and interactions. For almost two centuries, sociology has focused on humans as the only entities with agency, social cognition and creative communication.

Dr Tsvetkova and her collaborators would like to disrupt the old way of thinking: "A new sociology of humans and machines will help us to address many new social challenges such as online misinformation, cybersecurity, market flash crashes, road safety, to name a few. If we want to tackle pressing issues in society, it is no longer sufficient to understand humans. We need to understand machines, and how the collective behaviour of humans and machines can be predicted," she explains.

The group's latest paper assembles evidence of situations where algorithms, bots and robots perform as autonomous entities, interacting with each other and with humans to compete, cooperate, coordinate, communicate and make decisions.

In financial markets, for example, we see high frequency trading algorithms compete with each other as they trade on the same market. "The important point here is to guarantee the diversity of these algorithms," elaborates Dr Tsvetkova. "There is a lot of competitive pressure to develop the best super intelligence to optimise trading. But because these machines are often trained on the same data and the same patterns, they converge to very similar behaviour and respond to stimulus in the same way. Herding behaviour can cause flash crashes.

66

It is only through studying cases like this that we find out more about the importance of diversity – this is important for AI developers when they build the machines, and for policymakers who must regulate them. **99**

The need for informed regulation by policymakers is highlighted when it comes to the social influence of machines on humans, in particular reference to malicious bots on social media which are designed to influence public opinion or provoke specific reactions. "This is supported by the psychological theory of the human tendency to rely on information from others to handle uncertainty and to conform," says Dr Tsvetkova. "Malicious bots on social media exert collective power through many small, individual acts of influence, causing a cascade effect on human opinion and reactions."

Developments in AI are not all bad news, however, and research can also bring to the foreground cases of innovation and exciting developments in AI. Machines, for example, have been observed taking part in collective decision-making with humans with positive results.

"The emerging field of hybrid intelligence, which investigates how and when to combine human and algorithmic decision-making, has combined clinician and

algorithmic judgments in a medical setting to improve cancer diagnoses," Dr Tsvetkova and her co-authors highlight. "Zooniverse, a prominent citizen science platform which combines crowd-based and machine classifications, found supernovae candidates in a study on stars more effectively than humans or machines alone."

Understanding artificial intelligence for better societies

The consequences of rapid new developments in AI are still becoming apparent, and with every new development there is the need to understand its implications.

"We want social scientists to become more interested in applying their theories and methodologies to human-machine social systems in order to identify what is beneficial or detrimental for societies.

"In doing so we can help AI designers understand how to engineer for effective interactions between machines, and between machines and people. We can also help policymakers to ensure they are striving for ecological diversity.

"The future of AI doesn't need to be scary," affirms Dr Tsvetkova. "Simple forms of AI are currently everywhere, and if we treat them as a type of actor, although not yet human but a type of actor nonetheless, as social scientists we have a template for how to analyse them. We have theories and methods to do so and can provide valuable and novel information for designers, engineers and policymakers. Creating a social science framework for understanding AI – a new sociology of humans and machines – will help to shape AI's development for better societies." ■

Dr Milena Tsvetkova was speaking to Molly Rhead, Media Relations Officer at LSE.

A new sociology of humans and machines by Milena Tsvetkova, Taha Yasseri, Niccolo Pescetelli and Tobias Werner was published in Nature.

Even Good Bots Fight: The case of Wikipedia (2017) is by Milena Tsvetkova, Ruth García-Gavilanes, Luciano Floridi and Taha Yasseri .

Subscribe to receive articles from LSE's online social science magazine

lse.ac.uk/rftw