A 'neuroshield' could protect citizens from artificial intelligence, argues neuroscience expert

July 24 2023, by Avery Ruxer Franklin

Credit: CC0 Public Domain

There's an urgent need to support citizens with a system of digital self-defense, argues a neuroscience expert from Rice University's Baker

Institute of Public Policy.

Steps to regulate [advanced artificial intelligence](#) (AI) and AI-enhanced social media are needed to protect people from AI "hacking" our [interpersonal relationships](#) and [collective intelligence](#), says Harris Eyre, fellow in [brain health](#) at the Baker Institute.

"Although such technology brings the entire world to our devices and offers ample opportunities for individual and community fulfillment, it can also distort reality and create false illusions," he writes in the new report. "By spreading dis- and misinformation, social media and AI pose a direct challenge to the functioning of our democracies."

Currently, deep fakes are already causing concern as the country heads into an election season. Eyre argues that there's an urgent need to design neuroscience-based policies to support citizens against AI—such as a "neuroshield."

"The way we interpret the reality around us, the way we learn and react, depends on the way our brains are wired," he writes. "It has been argued that, given the rapid rise of technology, evolution has not been given enough time to develop those regions of the neocortex which are responsible for higher cognitive functions. As a consequence, we are biologically vulnerable and exposed."

The neuroshield would involve a threefold approach: developing a code of conduct with respect to information objectivity, implementing regulatory protections and creating an educational toolkit for citizens.

The report argues that cooperation between publishers, journalists, media leaders, opinion makers and brain scientists can form a "code of conduct" that supports objectivity of information. Eyre explains that interpreting facts lies within the realm of social and political freedom,

but undeniable truths need to be protected.

"As neuroscience demonstrates, ambiguity in understanding facts can create 'alternative truths' that become strongly encoded in our brains," he explains.

A toolkit developed with neuroscientists could protect cognitive freedom while also protecting people—especially young people on social media—from disinformation. The toolkit's prime objective would be to help people learn how to do their own fact-checking and fight back against the brain's susceptibility to bias and disinformation. For example, Google currently has campaigns in other countries that show short videos on social media sites and highlight the way that misleading claims can be made.

Yet, self-governance and exclusive reliance on a code of conduct can create an uneven playing field, Eyre argues.

"It is critical for both policymakers and brain scientists to advance this policy approach," he says. "The proposed European AI Act is an example of foreseeing how AI model providers can be held accountable and maintain transparency. By closely involving neuroscientists in planning and rolling out the Neuroshield, the U.S. can ensure that the best existing insights about the functioning of our cognition are taken into account."

More information: Report: Introducing the 'Neuroshield'—A Policy Approach to Protect Citizens from the Risks of AI

Provided by Rice University