

Hot Clicks: We Should Let AI Argue with Itself

Rounding up IT and advanced tech-related news impacting government and industry.

[Amanda Ziadeh](#)

Mon, 05/07/2018 - 15:03

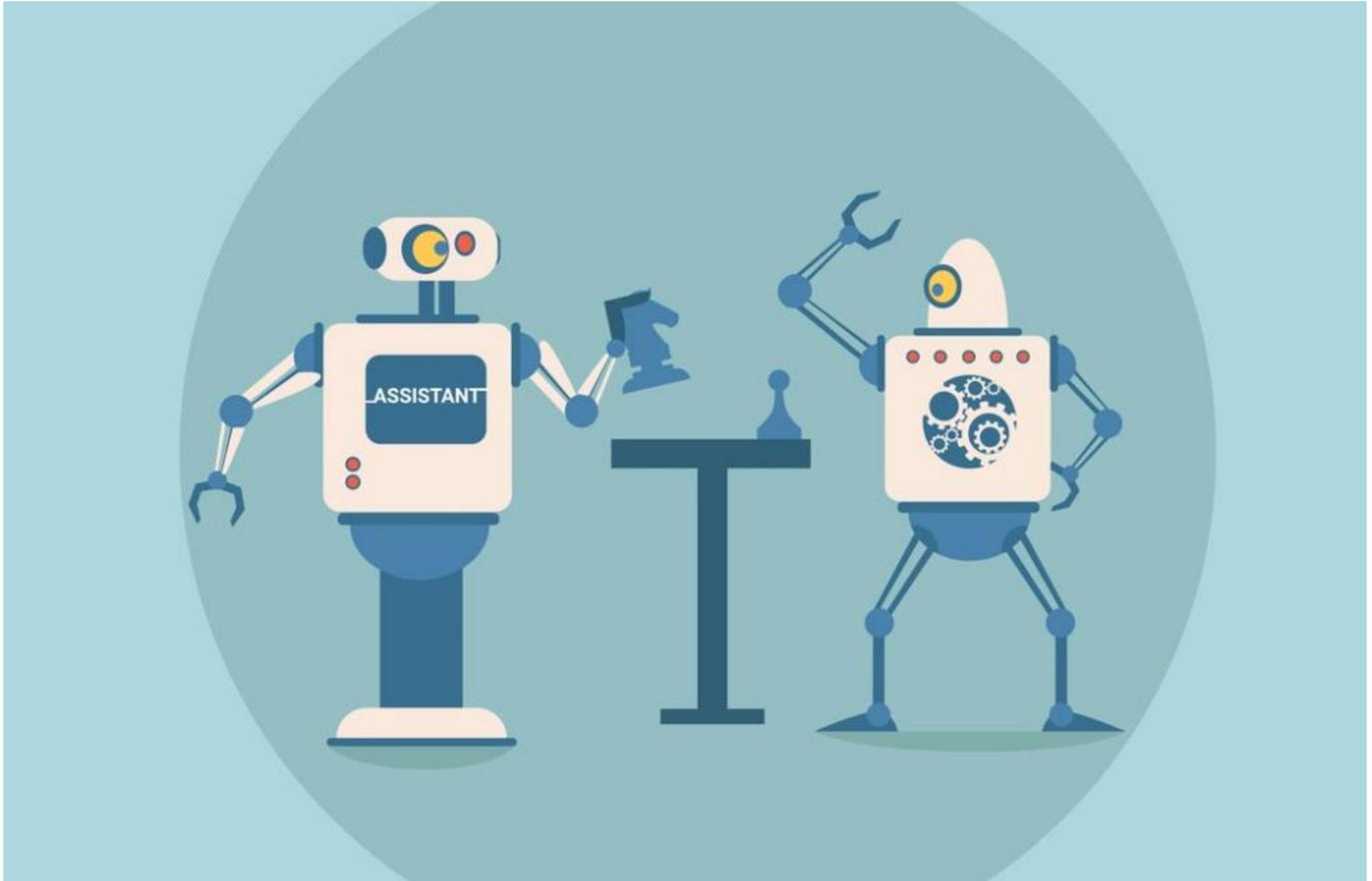


Illustration: gmast3r/iStock

The researchers at nonprofit OpenAI (think SpaceX CEO Elon Musk) suggest the best way to make sure artificial intelligence behaves is to have AI systems outwit one another, so a human observer can judge the system's intentions. When AI systems train themselves, it can result in unexpected habits (like when a computer game AI agent found out how to cheat its way to a higher score). Sometimes, a person can watch and see how it happened, but if the program is doing something really complex, it may take having two systems explain their actions and challenge one another instead.

So, this approach allows AI systems to perform more advanced cognitive tasks while sticking to human preferences. For example, if an AI system is designed to defend against hackers, as a way to prevent it from doing anything unethical, it can be challenged to explain and debate its logic for a complex action to another AI system using natural language, while a person observes.

But this will take more advanced technology than we have available, so OpenAI is exploring the concept with a few simple examples. It even created a website where two people can try play the roles of the debating AI systems, and a third plays the judge. [MIT Technology Review](#)

Is Facebook's Dating Service Legit?

The social media giant's CEO Mark Zuckerberg announced last week Facebook is launching a dating feature intended for long-term relationships, not just hookups. Whether Facebook is seeking a comeback from Cambridge Analytica, privacy experts say this new dating service can actually end up in fake profiles and catfishing, even more data gathering and dating fraud.

The service will allow people 18 and older to create a dating profile separate from their main profile and invisible to their friends. It'll show matches based on common interests, dating preferences, location and mutual friends, and people can choose to be "interested" or to "pass" on the match.

But privacy and advertising experts say the service can actually expose users to the worst of the internet, considering Facebook already has problems with fake accounts, scams and strangers. And this will add a lot of sensitive information to what Facebook already collects. Could it just create a larger target and increase the social media site's abuse? Will it really just appease to advertisers? And how will Zuckerberg police the collection and use of that information to ensure safety? Only time will tell. [The Washington Post](#)

What's Wrong with Augmented Reality

AR is getting bigger and better as technology advances. It's already being used in education, entertainment, health care, consumer electronics, gaming, industrial machines and more — allowing humans to explore and immerse in a digital realm.

In fact, the AR and virtual reality combined market value could reach \$108 billion by 2021, but because the tech is still new, it comes with risks that should be dealt with before we go much further. There are security and privacy concerns, dangers to physical safety, and ongoing cyberthreats from criminals and nation-state actors that could seriously slow down the success of AR if not managed. And without this awareness and self-regulation, AR could also fall under strong government oversight, which can also slow innovation.

So, how can we address some of the AR challenges before they cause trouble? Through awareness, by establishing industry standards and a comprehensive feedback system amongst the AR/VR communities, and by developing and investing responsibly. [TechCrunch](#)

NASA's Getting Closer to a Space Nuclear Reactor

The agency has been trying to take a nuclear reactor to space for 50 years, and announced last week it successfully completed its first test of the Kilopower reactor, a design in development since 2012. It's an experimental portable nuclear power plant, and this is important — because NASA plans to one day power permanent settlements for humans on Mars and the moon, and to launch high-energy robotic missions to outer space.

Each Kilopower reactor is the size of a refrigerator and can deliver up to 10 kilowatts of energy once deployed to space (enough energy to power 10 American homes). It's estimated four Kilopower reactors could sustain a "decent-sized" outpost on the moon or Mars for 10 years. '

And this has been years in the making; testing on the single kilowatt prototype started in November, and the testing program studied the device in all kinds of critical failure scenarios before completing its full-power test, which kept the reactor online for 28 hours. [Motherboard](#)

NIH's DNA Data Mission

After more than three years of planning, the National Institutes of Health started its million-person genetic survey, asking for volunteers to share decades of health records and dozens of test tubes of blood and urine, all in the name of science. This gigantic precision health initiative started under the Obama administration in 2015, and it's called All of Us. The goal? To collect detailed health data from 1 million Americans, so scientists can better understand disease and personalized treatments.

The program launched May 6, and anyone 18 or older living in the U.S. can donate their data (and so far, 45,000 people have already started the process, thanks to the May 2017 beta phase). The recruitment sites sync a bunch of data, from electronic health records and behavioral surveys to genetic reports and fitness device data.

Launching the infrastructure was a process, too. The All of Us full project budget for its 10-year run is \$1.455 billion, which includes the 298 enrollment sites NIH plans to launch by the end of 2018 (it has 120 up so far) and funding for a national biobank for sample storage. [Wired](#)

[View printer friendly version](#)

[AI](#)

[NIH](#)

[NASA](#)

[Elon Musk](#)

[Facebook](#)

[augmented reality](#)

[virtual reality](#)

[data](#)