Sources	Nature		Date	May 2021

Potential scale of impact	Certainty of outcome	Impact horizon		
★★★★	★★☆☆	H1	H2	**H3**

AI systems and algorithms have been developed as individual agents. Work involving multiple agents has not yet tackled the hard problems of cooperation. Getting this right is the key to ensuring that AI systems integrate with and assist the human labour force - rather than displace it.

AI needs social understanding and cooperative intelligence to integrate well into society. The next wave of development will demand AI systems that interact in rapid and complex ways with each other and with humans: on pavements and roads, in consumer and financial markets, in e-mail communication and social media, in cybersecurity and physical security. Autonomous vehicles or smart cities that do not engage well with humans will fail to deliver their benefits, and might even disrupt stable human relationships.

The key to this next wave of development will be moving from autonomous AI systems - what we have now - to the cooperative AI systems we need in the future.

Two key and immediate areas of R&D are:

- **AI-AI cooperation** which involves developing AI systems that can bargain and negotiate with each other. Current work is built on cooperative games where AI systems can learn what 'team mates' are thinking and planning - and adapt their approach to deliver a mutually beneficial outcome

- **Human-AI cooperation** which builds on existing practice in manufacturing. AI alignment - agents that act in accordance with human intentions, preferences and values - is crucial for cooperation. Progress here will require advances in understanding human language, gestures and activities, and ad hoc teamwork. Increasingly research will need to move into real environments in order to develop effective and high performing human–machine teams.

CO-OPERATIVE AI

Human AI cooperation is key to ensuring AI enhances people's jobs rather than displacing them

Home Page
www.10xhorizonscan.online