



Supporting the government's drive towards trustworthy AI

As AI systems become ever more sophisticated, they are making and supporting a growing number of decisions in areas ranging from finance and healthcare to employment, law and transport. These systems offer huge potential benefit to society and the economy, but they come with risks. Is the system free from bias? Does it comply with regulations? Are there privacy and security concerns? To help counter these risks, the government's Centre for Data Ethics and Innovation (CDEI) is leading the development of an 'AI assurance' industry that will provide tools and services to enable businesses, consumers and regulators to verify that AI systems are effective, trustworthy and legal.

The Turing has become involved in this area through **Lukasz Szpruch**, Programme Director for Finance and Economics, who was an advisor to a pilot project between University College London and the CDEI. The result of this project was a 2021 paper that summarises the **key concepts of AI assurance** and identifies some of the difficulties in this area, such as how to create standards for algorithm safety when their uses and risks vary so much between different sectors.

This paper has in turn informed the '**Roadmap to an effective AI assurance ecosystem**', a first-of-its-kind report published by the CDEI in December 2021, which sets out the steps required to build an AI assurance ecosystem in the UK.

Lukasz continues to work in this area as part of Project FAIR – a collaboration between the Turing and partners including HSBC and Accenture (also see page 25), a goal of which is to develop clear guidelines for AI assurance within the finance sector.

“For AI to be used responsibly, ethical principles need to be translated into real-world practice. This work with the CDEI will help guide the UK towards a thriving industry dedicated to that goal.”



Lukasz Szpruch
Programme Director for Finance and Economics, The Alan Turing Institute