Related programmes and teams
Research Engineering
Tools, practices and systems
AI for science and government (ASG)

"Data safe havens are all about increasing trust, so that data providers can feel confident in supplying sensitive data to researchers. Now, we want to build a community around our system that extends far beyond the Turing."

James Robinson
Senior Research Software Engineer
The Alan Turing Institute

Data safe havens spread their wings

Researchers frequently require access to sensitive data, such as health and financial records, or information on protected characteristics such as disability and sexual orientation. To prevent unauthorised access to this data, they have to be hosted in secure computing environments. Since 2018, the Turing's Research Engineering team has been pioneering the development of cloud-based '**data safe havens**' – remotely-hosted environments that allow researchers to safely analyse their data, while taking advantage of the scale and flexibility of cloud computing.

Alongside the technical aspects of creating the digital infrastructure for these safe havens, the team has developed a policy framework so that data providers and users can easily specify the security requirements of their data. When a safe haven is created for a new project, the researchers select one of five pre-defined security levels, and the software automatically creates a safe haven which is tailor-made for that level. In this way, the system is scalable for a broad range of projects.

The safe havens are routinely used in the Turing's Data Study Groups, and also support several Turing projects involving sensitive data. The team is now talking to around 10 organisations and universities who are evaluating the system for their own potential use. In May 2020, The Health Foundation – a charity committed to bringing about better health and healthcare for people in the UK – made modifications to the system to develop its thinking around the security of cloud-based platforms for processing health data, and how such platforms could be securely accessed. The Turing team has since incorporated many of these improvements into its core system, and is ultimately aiming to make its code completely open so that organisations can create their own data safe havens from scratch.