

The Alan Turing Institute

SAFE AND ETHICAL AI RESEARCH FELLOW/FELLOW

THE ALAN TURING INSTITUTE

AI technologies are increasingly deployed in ways that affect many human lives, providing remarkable opportunities to benefit society and the economy while also raising important and urgent concerns. We aim to help address these concerns by developing rigorous methods to enable AI systems to be reliably trustworthy. To improve impact, we are engaged with industry and government, as well as regulators and policy makers to discuss appropriate governance.

The Alan Turing Institute, as the UK's national institute for data science and artificial intelligence, plays an important role in helping to advance these technologies in order to change the world for the better. The Institute is named in honour of Alan Turing, whose pioneering work in theoretical and applied mathematics, engineering and computing helped to lay the foundations for modern-day data science and artificial intelligence. The Institute's goals are to undertake world-class research, apply its research to real-world problems, driving economic impact and societal good, lead the training of a new generation of scientists, and shape the public conversation around data and algorithms.

After launching in 2015, the Institute has grown an extensive network of university partners from across the UK and launched a number of major partnerships with industry, public and third sector. Today it is home to more than 400 researchers and a talented business team.

THE ROLE

The Alan Turing Institute is looking to hire a full-time established mid-career or senior academic as a Research Fellow (RF) or a Fellow to work on one or more themes of Safe and Ethical AI:

- Fairness - measuring and mitigating inappropriate bias against subgroups.
- Transparency/Interpretability - improving our understanding of how algorithmic systems operate with a view to their applications in the real world.
- Privacy - protecting privacy while enabling the benefits of appropriate data sharing (e.g. differential privacy, secure multi-party computation).
- Security – the use of AI within traditional cybersecurity, as well as protecting against newer concerns for machine learning such as 'data poisoning' attacks.
- Robustness - aiming to ensure reliably good performance (e.g. within quantifiable uncertainty bounds), even when encountering different settings from those in which a system was trained.
- Resilience - maintaining a system's performance even if parts of it go down.
- Control – as increasingly autonomous systems are developed, we must maintain appropriate control and avoid unintended behaviours.

These areas require deep theoretical research, looking for rigorous, quantifiable and verifiable guarantees. Work might involve approaches such as modelling uncertainty or causal reasoning.

Another area of focus, linking more closely to HCI and social science, is measuring the extent to which AI systems influence individuals and society, aiming to enable appropriate methods of control for users.

It will be vital to interface well with other disciplines, policy makers, industry and the public – but we expect the primary contributions of those hired to be technical. Successful candidates will work alongside Adrian Weller, Chris Russell, Ricardo Silva, Anjali Mazumder, Matt Kusner and other Turing researchers.

The Institute is interdisciplinary in its approach and includes lawyers, ethicists, social scientists and humanities researchers in addition to researchers in technical disciplines and collaboration is encouraged. A team of full-time data engineers is available to support prototype development.

Our ideal candidate will have a demonstrated ability to work in a highly collaborative manner and be enthusiastic about engaging widely across disciplinary boundaries and with industry, government and the third sector.

DUTIES AND RESPONSIBILITIES

- To develop a research programme at the Turing and conduct outstanding, creative and innovative research in Safe and Ethical AI, to develop internationally-significant outcomes through high-impact publications.
- To collaborate with others across the Turing, university partners, industry partners and the AI community beyond, towards outputs and outcomes that yield significant academic, societal or economic impact.
- To play a role in advancing the Turing's AI programme.

PERSON SPECIFICATION

ESSENTIAL

- A PhD that is relevant to a technical aspect of Safe and Ethical AI or an equivalent, demonstrated sustained intellectual leadership, in a technical area of relevance, in non-academic environments.
- Evidence of commitment, passion and vision in a technical aspect of Safe and Ethical AI
- Demonstrated understanding of the national and international research landscape across AI disciplines and/or domains.
- Track record of obtaining funding to support research.
- Track record of outstanding research and in delivering impact.
- Strong awareness of the international context of the research.
- Ability to create and promote a collegiate and collaborative approach to interdisciplinary research activities.
- Evidence of outstanding approach to identifying future research directions in AI.
- Demonstrated ability to develop, lead and support a successful research team and maximise the potential in others.
- Excellent written and verbal communication skills, including experience in publishing research papers or technical reports, and giving presentations or classes on technical subjects.

DESIRABLE

- Experience in translating research output into prototype tools.

We welcome and develop applicants from diverse backgrounds, experiences and ideas and encourage applications from individuals who may have had non-traditional career paths, who may have taken a career break or who have achieved excellence in careers outside of academia.

TERMS AND CONDITIONS

This full-time post is offered on a fixed-term contract for a period of up to 3 years

Start date: 1 October 2020, or earlier preferred, if available.

Excellent benefits package (<https://www.turing.ac.uk/work-turing/why-work-turing/employee-benefits>).

APPLICATION PROCEDURE

If you are interested in this opportunity, please click the apply button below and register on the candidate portal, including:

- your CV
- contact details for two referees
- a cover letter that outlines how you meet the person specification
- a two-page proposal outlining the research agenda you plan to pursue if successful
- if applicable, a list of publications (max 2 pages).

If you have questions or would like to discuss the role further with a member of the Institute's HR Team, please contact them on 0203 862 3394 or 020 3862 3357, or email recruitment@turing.ac.uk. Applicants who would like to submit their application in a different format, please email recruitment@turing.ac.uk.

Equality Diversity and Inclusion

The Alan Turing Institute is committed to creating an environment where diversity is valued and everyone is treated fairly. In accordance with the Equality Act, we welcome applications from anyone who meets the specific criteria of the post regardless of age, disability, ethnicity, gender reassignment, marital or civil partnership status, pregnancy, religion or belief or sexual orientation. Reasonable adjustments are available to support candidates through the application and interview process. We Talk Flexible Working.

Please note all offers of employment are subject to continuous eligibility to work in the UK and satisfactory pre-employment security screening which includes a DBS Check.

Full details on the pre-employment screening process can be requested from HR@turing.ac.uk.