

The Alan Turing Institute

SAFE AND ETHICAL AI 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 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.

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

Main purpose of the role:

- 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.

We want the successful Fellow to:

- Generate and pursue original research ideas, help design and conduct a successful programme of investigation and develop innovative, world-class research.
- Publish research in high quality peer-reviewed national and international journals. Present research results at national and international meetings, conferences, seminars and workshops.
- Take part in knowledge exchange and/or translation activities as appropriate; e.g. collaborative working with the Turing's partners, or with government departments/policy-makers, and on public engagement, policy events, etc.
- Build research collaborations within the broad spectrum of activity at the Turing.
- Develop research collaborations nationally and internationally to support the research programme, facilitate and participate in crosscutting research collaborations as appropriate.
- Support and contribute to the development of research links between the Turing and its partners, and the wider AI community as appropriate.
- Maintain their own professional development, with support from the Turing.

PERSON SPECIFICATION

The successful candidate will have:

ESSENTIAL

- Evidence of commitment, passion and vision in Safe and Ethical AI and potential to become a world leader.
- Demonstrated understanding of the national and international research landscape across AI disciplines and/or domains.
- Track record of obtaining funding to support research.
- Has a track record of outstanding research and in delivering impact.
- Shows a 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.

TERMS AND CONDITIONS

The Safe and Ethical AI Research Fellows will be employed by the Alan Turing Institute and based primarily at the Turing Headquarters in the British Library, London.

Contract type: Full time, Fixed Term Contract up to 3 years

Remuneration: Competitive with excellent benefits <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
- a list of publications (max 2 pages).

If you have questions or would like to discuss the Fellowship further with a member of the Institute's HR Team, please contact them on 0203 862 3394 or <mailto:jobs@turing.ac.uk>
Contact Aida Mehonic on aiprogramme@turing.ac.uk with questions the role.

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, gender reassignment, marital and civil partnership status, pregnancy, religion or belief or sexual orientation. Reasonable adjustments to the interview process can also be made for any candidates with a disability.

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.