

The Alan Turing Institute

Safe and Ethical AI Research Associate / Senior Research Associate

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

We are seeking a full-time post-doctoral Research Associate (RA) or a Senior Research Associate (SRA) to work on one or more themes of the Turing 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.

DUTIES AND RESPONSIBILITIES

- To undertake research related to Safe and Ethical AI themes.
- To produce high-quality research publications documenting the results of the research, to publish these papers in relevant peer-reviewed scientific journals of international standing, and to present these results at conferences and workshops.
- To hold regular meetings with designated members of staff and with other collaborators.
- To collaborate with and support colleagues in the development of research links between the Turing and its partners, and the wider AI community as appropriate.
- To travel as necessary to meet with external collaborators.
- To take initiative and make original contributions to the research programme wherever possible, and to contribute freely to the team research environment in a manner conducive to the success of the research project as a whole.

PERSON SPECIFICATION

ESSENTIAL

- PhD (or close to completion) or equivalent experience in machine learning or in a related technical discipline.
- Publication record in leading international conferences or journals, as appropriate to career stage.
- Familiarity with the existing AI ethics and safety methods associated with their area(s).
- Excellent written and verbal communication skills including the ability to present complex or technical information, and to communicate effectively with analysts and other stakeholders outside the research community.
- Ability to collaborate successfully with colleagues in a multidisciplinary environment.
- A desire to collaborate with experts in social sciences, the law and others.
- Ability to lead one's own work, including planning and execution, and to prioritise work to meet deadlines.
- Ability to organise working time, take the initiative, and carry out research independently, under the guidance of the PI.

DESIRABLE

- Experience of collaboration with other academic disciplines.

APPLICATION PROCEDURE

If you are interested in this opportunity, please click the apply button and register on the candidate portal to submit your CV, with contact details for your referees and a covering letter.

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 02038623394 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.

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

Salary: £35,000 – £41,000 for Research Associate; £42,000-£49,000 for Senior Research Associate (dependent on experience) and a competitive benefits package (<https://www.turing.ac.uk/work-turing/why-work-turing/employee-benefits>).

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.