

# The Alan Turing Institute

## Research Associate, Safe and Ethical AI

### THE ALAN TURING INSTITUTE

There has never been a more significant time to work in data science and AI. There is recognition of the importance of these technologies to our economic and social future. The technical challenge of keeping our data secure and private has grown in its urgency and importance. At the same time, voices from academia, industry, and government are coming together to debate how these technologies should be governed and managed.

The Alan Turing Institute, as the UK's national institute for data science and artificial intelligence, plays an important part in driving forward advances in 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 is considered to have laid 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 with government funding from EPSRC and five founding universities, 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 500 researchers, a rapidly growing team of in house research software engineers and data scientists and a business team.

### BACKGROUND

At The Alan Turing Institute, the AI programme's goal is to advance world-class research into artificial intelligence, its applications and its implications for society, building on the wealth of expertise and knowledge across our academic network.

Our programme expertise draws on the cross-disciplinarity at the heart of the Turing; bringing together technical experts to forge ahead with breakthroughs in fundamental AI research and its applications, with social scientists, ethicists, legal experts, industry, and policy-makers who are considering the implications of AI in real-world applications.

The programme will contribute to leadership around this area, drawing on our links with industry and government to help ensure that the UK remains at the forefront of scientific innovation while building an ethical and regulatory framework for the use of AI that prevents misuse and inappropriate discrimination.

Our aim is to build a leading centre of excellence for technical work on safe and ethical AI, in line with the government's Industrial Strategy and in step with global demand for research and guidance in this domain.

### ROLE PURPOSE

We are seeking a full-time post-doctoral Research Associate (RA) to work as part of the AI Programme, on one or more of the following themes of the Turing Safe and Ethical AI activities:

- Fairness – including measuring and mitigating inappropriate bias against subgroups.
- Transparency and Interpretability - improving our understanding of how algorithmic systems operate with a view to their applications in the real world.
- 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.
- Privacy - ensure data is **secure** from attack and respecting **privacy** while enabling the benefits of appropriate data sharing.

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

# The Alan Turing Institute

We are looking for ambitious, strong researchers to contribute to these areas. A key consideration is how we can try to ensure good performance not only of the AI system, but human-machine teams. Hence, experience and expertise in HCI or related areas will be beneficial.

Applicants with strong theoretical backgrounds and/or applied are welcome. In addition, we recognise that in real-world settings, there will often be interdependencies and tensions among the above goals and performance. In order to make constructive progress, we will examine settings in context, working with practitioners and domain experts. Application domains will likely include healthcare, criminal justice and finance.

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 as part of the AI Programme, alongside Adrian Weller, Ricardo Silva, Yarin Gal 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.

This is an exciting opportunity to help develop strong technical underpinnings while engaging with real-world practitioners, broader stakeholders and policy makers.

## DUTIES AND AREAS OF RESPONSIBILITY

- To undertake research related to the 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 community as appropriate.
- To travel as necessary to present work and 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 an inclusive manner conducive to the success of the research project as a whole.
- To adhere to modern principles of reproducible and ethical data science in carrying out their responsibilities.
- To ensure compliance with secure handling of data and health and safety in all aspects of work.
- To contribute to the life of the Institute and support its community

Other duties:

- Teaching may be required as part of collaboration work

Please note that job descriptions cannot be exhaustive, and the post-holder may be required to undertake other duties, which are broadly in line with the above key responsibilities. This job description is written at a specific time and is subject to changes as the demands of the Institute and the role develop.

# The Alan Turing Institute

## PERSON SPECIFICATION

<b>Skills and Requirements</b> Post holders will be expected to demonstrate the following	<b>Essential (E)</b>  <b>Desirable (D)</b>	<b>Tested at application(a)</b>  <b>Tested at interview (i)</b>
<b>Education/Qualification</b>		
Research Associate: PhD or equivalent experience in machine learning or in a related technical discipline.	<b>E</b>	<b>(a)</b>
Assistant level: Near completion of a PhD or equivalent level of qualification in machine learning or in a related technical discipline.	<b>E</b>	<b>(a)</b>
<b>Knowledge and Experience</b>		
Experience in machine learning or in a related technical discipline.	<b>E</b>	<b>(a) and (i)</b>
Track record of publications in leading international conferences or journals, as appropriate to career stage.	<b>E</b>	<b>(a)</b>
Publication record in leading international conferences or journals, as appropriate to career stage.	<b>E</b>	<b>(a)</b>
Previous Post Doctoral experience is beneficial	<b>D</b>	<b>(a)</b>
Expertise in work in Fairness, Transparency, Privacy or Robustness.	<b>D</b>	<b>(a) and (i)</b>
Expertise in HCI or related areas.	<b>D</b>	<b>(a) and (i)</b>
Interest and/or knowledge of healthcare or criminal justice or finance.	<b>D</b>	<b>(a) and (i)</b>
Familiarity with existing AI ethics and safety methods associated with their area(s).	<b>D</b>	<b>(a) and (i)</b>
Experience of collaboration with other academic disciplines.	<b>D</b>	<b>(a)/ (i)</b>
<b>Communication</b>		
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.	<b>E</b>	<b>(a) and (i)</b>
Ability to adapt the style of communication to the audience and ensures understanding.	<b>E</b>	<b>(a) and (i)</b>
Ability to communicate and liaise between collaborators from different interdisciplinary backgrounds.	<b>E</b>	<b>(a) and (i)</b>
<b>Liaison and Networking</b>		
Ability to collaborate successfully with colleagues in a multidisciplinary environment.	<b>E</b>	<b>(i)</b>

# The Alan Turing Institute

A desire to collaborate with experts in social sciences, the law and others.	E	(a)
Participates in networks within the organisation or externally to share knowledge and information in order develop practice or help others learn.	E	(a)/ (i)
A desire to collaborate with experts in social sciences, the law and others.	D	(a)/ (i)
<b>Service Delivery</b>		
Ability to keep accurate and up to date knowledge of services available in own and related areas of work.	E	(a)/ (i)
Ability to work across the Research Team/Programme, to contribute and assist in a number of research activities.	E	(i)
<b>Decision Making</b>		
Independently makes decisions which are low risk and that mainly affect themselves or a small number of people and are guided by regulation and practice.	E	(a)/ (i)
Works with others to make collaborative decisions for their team.	E	(a)/ (i)
Recommends and advises on available options for decisions that affect operational processes, taking into account any risks.	D	(a)/ (i)
<b>Initiative and Problem Solving</b>		
Ability to use own judgement to analyse and solve problems.	E	(a)/ (i)
Considers possible solutions to identify solutions which offer wider benefits. Obtains evidence to support thinking.	E	(a)/ (i)
Ability to lead one's own work, including planning and execution, and to prioritise work to meet deadlines.	E	(a)/ (i)
<b>Analysis and Research</b>		
Ability to organise working time, take the initiative, and carry out research independently, under the guidance of the PI.	E	(i)
<b>Other Requirements</b>		
Commitment to meeting deadlines	E	(a)/ (i)
Flexible attitude towards work	E	(i)
Commitment to EDI principles and to the Organisation values	E	(i)

# The Alan Turing Institute

## Our Values

The Alan Turing Institute is committed to equality diversity and inclusion and to eliminating discrimination. All employees are expected to embrace, follow and promote our [EDI Principles](#) and [Rules of the Game](#)



**Respect** – We treat everyone with respect, dignity and kindness and acknowledge the experiences, skills and contributions of others.

**Trust** - We communicate openly and honestly to support an environment where we have trust in each other.

**Transparency** – We seek to ensure that everyone understands the how and the why of our decisions and actions. We take on board to feedback when those decisions are challenged.

**Inclusivity** – We are committed to continuously learning how to be more inclusive by listening to those facing barriers.

**Leadership** – We recognise creating an inclusive, diverse and equitable institute requires leadership from all. We stand up and speak out when change is needed.

**Integrity** – We recognise that how we work is as important as our outputs and seek to exemplify best practice in all our decisions.

## APPLICATION PROCEDURE

If you are interested in this opportunity, please click the apply button below. You will need to register on the applicant portal and complete the application form including your CV, covering letter and contact details for your referees. If you have questions about the role or would like to apply using a different format, please contact them on 0203 862 3340, or email [recruitment@turing.ac.uk](mailto:recruitment@turing.ac.uk).

**CLOSING DATE FOR APPLICATIONS: Sunday 01 August 2021.**

## TERMS AND CONDITIONS

This full-time post is offered on a 2 year fixed-term basis with the possibility of extension. The ideal start is October 2021 however an earlier start date can be arranged. The annual salary is £37,000-£42,000 for Research Associate plus excellent benefits, including flexible working and family friendly policies, <https://www.turing.ac.uk/work-turing/why-work-turing/employee-benefits>

Candidates who have not yet been officially awarded their PhD will be appointed as Research Assistant within the salary range £34,500 per annum

## 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 and maternity, religion or belief, sex and sexual orientation.

Reasonable adjustments to the interview process will be made for any candidates with a disability.

***Please note all offers of employment are subject to obtaining and retaining the right 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](mailto:HR@turing.ac.uk).***