

The Alan Turing Institute

Research Associate, Safe, Explainable and Trustworthy 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 so-called fourth industrial revolution. 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

Project BlueBird - Advancing Probabilistic Machine Learning to Deliver Safer, More Efficient, and Predictable Air Traffic Control

Project BlueBird is a partnership between NATS and The Alan Turing Institute to develop the fundamental science to deliver the world's first AI system to control a section of airspace in live trials.

The collaborative research will take a hierarchical approach to air traffic control (ATC) by developing a digital twin alongside a multiagent machine-learning control system for UK airspace. The partnership will develop technical approaches to deploy trustworthy AI systems, considering how safety, explainability and ethics are embedded within our methods, so that we can deliver new tools which work in harmony with human air traffic controllers in a safety-critical environment.

Project BlueBird is arranged into three interconnected scientific research themes (RTs) and underpinned by a cross-cutting theme (XCT). The three scientific Research Themes are:

1. Probabilistic Digital Twin of UK Airspace
2. Machine Learning Tactical and Strategic Flight Control
3. Safe, Trustworthy and Explainable ATC

This role will be part of Research theme 3: Safe, Trustworthy and Explainable Air Traffic Control. Each scientific theme involves two academics and two NATS domain experts.

Alongside NATS and Turing, the research programme includes partner universities of Exeter and Cambridge. The research programme is led by Professor Tim Dodwell (Turing AI Fellow at Turing & University of Exeter) with Dr Evelina Gabasova (Turing Principal Research Data Scientist), Prof Mark Girolami (Turing Programme Director for Data Centric Engineering programme), Prof Richard Everson (Professor of Machine Learning at University of Exeter and Turing Fellow), Dr Edmond Awad (Assistant Professor at University of Exeter), Dr Adrian Weller (Turing Programme Director for the AI programme), and our partners at NATS including Dr Richard Cannon (Research Lead) and their wider research team.

The Alan Turing Institute

ROLE PURPOSE

We are seeking a Research Associate to work as part of Project BlueBird to deliver a human-centred, data-driven approach to explainable and trustworthy AI.

To enable effective human-AI cooperation, this role will develop tools to provide useful explanations of AI decisions as well as communicate the level of uncertainty around those decisions.

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 Human-Computer interactions (HCI) or related areas will be beneficial. Applicants with strong theoretical and/or applied backgrounds are welcome. 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.

The Research Associate will lead projects as part of a research theme. The aim of the theme is to develop effective tools that promote successful human–AI cooperation in safety-critical tactical control. The successful applicant will work with Dr Edmond Award (Exeter & Turing) and Dr Adrian Weller (Cambridge & Turing).

DUTIES AND AREAS OF RESPONSIBILITY

- To undertake research related to the Research Theme of Safe, Trustworthy and Explainable Air Traffic Control.
- 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, NATS and their 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.

Please note that job descriptions cannot be exhaustive, and the postholder 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		
Skills and Requirements	Essential (E) Desirable (D)	Tested at application (a) Tested at interview (i)
Post holders will be expected to demonstrate the following:		
Education/Qualification		
PhD or equivalent qualification/experience in Computer Science, Artificial Intelligence, or Human-Computer Interaction or a related field of study.	E	A
Be a nationally recognised authority in the subject area.	D	A
Knowledge and Experience		
Possess sufficient specialist knowledge in the Artificial Intelligence Ethics (AI Ethics), and Explainable Artificial Intelligence (XAI) to develop/follow research	E	A/I
Experience of defining the research direction in collaboration with Project Investigators as appropriate to career stage.	E	A/I
The ability to initiate, develop and deliver high quality research aligned with the research strategy as indicated by the PI and any industrial stakeholders, and to publish in peer reviewed journals and conferences.	E	A/I
Experience of undergraduate/postgraduate teaching and supervision.	D	A/I
Knowledge and interest in explainable AI.	E	A/I
Communication		
Excellent communication skills with the ability to present complex information and conceptual ideas to a range of audiences.	E	A/I
The ability to write research reports and papers in styles accessible to both academic and lay audiences.	E	I
Project Delivery		
Able to identify sources of funding, generate income, obtain consultancy projects and develop external networks for future activities.	E	A/I
Demonstrated ability to collaborate actively within the Institution and externally to complete research projects and advance thinking.	E	A/I
Able to balance the pressures of research, administrative demands and competing deadlines to meet objectives.	E	I
Decision Making		
Work with others to make collaborative decisions that may be operational or strategic and impact immediate team or work area.	E	I
Ability to make independent decisions which are low risk and that mainly affect themselves or a small number of people and are guided by regulation and practice	E	I
Initiative and Problem Solving		
Uses judgement to analyse and solve problems and take action to prevent recurrence of problems.	E	I
Consider possible solutions to identify those which offer wider benefits and obtain evidence to support thinking	E	I

The Alan Turing Institute

Analysis and Research		
Ability to plan and implement rigorous analysis plans.	E	I
Identify and use a range of standard sources to gather and analyse routine data and produce reports that can be interpreted by others.	E	I
Understand when additional data is required and identifies appropriate sources. Produces reports that identify key issues and findings.	E	I
Other Requirements		
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 Our Values.

Our values

- Trust**
We create an environment where we have trust and can be trusted
- Inclusivity**
We expect our Turing community to contribute to a culture that is inclusive and free of barriers
- Respect**
We all have different roles, priorities and challenges but our shared purpose is the same
- Leadership**
Leadership is everyone's business; Turing leaders set the right tone and lead by example
- Transparency**
Everyone should understand the how and the why of our decisions and actions
- Integrity**
We are all ambassadors for the Turing's mission of changing the world for the better

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 and covering letter. If you have questions about the role or would like to apply using a different format, please contact us on 020 3970 2148 or 0203 862 3340, or email recruitment@turing.ac.uk.

CLOSING DATE FOR APPLICATIONS: Monday 16 January 2023 at 23:59

TERMS AND CONDITIONS

This full time post is offered on a fixed term basis until May 2026. The annual salary is £38,850-£46,200 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 at a salary of £36,236 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.

We are committed to building a diverse community and would like our leadership team to reflect this. We therefore welcome applications from the broadest spectrum of backgrounds.

We are committed to making sure our recruitment process is accessible and inclusive. This includes making reasonable adjustments for candidates who have a disability or long-term condition. Please contact us at adjustments@turing.ac.uk to find out how we can assist you.

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.