The Alan Turing Institute

RESEARCH ASSOCIATE / SENIOR RESEARCH ASSOCIATE – AI EVALUATION AND ASSURANCE

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 purpose is to make great leaps in data science and AI research to change the world for the better. Its goals are to advance world-class research and apply it to national and global challenges, build skills for the future by contributing to training people across sectors and career stages, and drive an informed public conversation by providing balanced and evidence-based views on data science and AI.

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

The Alan Turing Institute's Fundamental Research mission exists to make the advances in data science and AI on which the Institute's other missions and the wider field depend: new methods, new theory, and the foundations that let AI systems be built and used with confidence. Within it, the Trustworthy and Assured AI theme asks how we can know that an AI system deserves the trust placed in it, exploring questions such as how to evaluate what a system does, how to turn evaluation results into assurance-grade evidence, and how to assemble that evidence into an argument that a designer, regulator or operator can act on. The theme works across domains, with a current emphasis on national security and resilient critical national infrastructure. It also maintains the Turing's Trustworthy and Ethical Assurance (TEA) platform - an open-source tool for building and sharing structured assurance arguments.

This post is funded initially through a research partnership with an industry partner in the defence and national security sector. The partnership concerns the evaluation and assurance of machine-learning components that must operate under tight operational constraints, such as limited time and compute, and where the evidence produced by evaluation and explanation methods has to be combined into an argument that assurance and certification practice can accept.

This role sits within the Turing's Fundamental Research mission and the Trustworthy and Assured AI theme. The post holder will lead the theoretical strand of the partnership's work, contribute to its evaluation studies and groundwork, and help shape the theme's longer-term research on AI evaluation and verification.

ROLE PURPOSE

The Research Associate will lead a scoped research project on the combination of evaluation evidence about machine-learning systems. The project addresses a gap in current assurance practice: evaluation and explanation methods each produce a result, but a result is not yet a claim. Whether two results count as two pieces of evidence or one repeated depends on whether the methods can fail independently, and no settled account exists of how such results combine, how their reliability should discount them, or what the combined verdict licenses a designer to claim. The post holder will develop that account, drawing on evidence theory, probabilistic reasoning and current work on machine reasoning and human interpretation, and illustrate it on results from the theme's evaluation platform. Alongside the theoretical work, they will contribute to the partnership's evaluation studies, to the mapping of evaluation evidence onto defence and civil-aviation assurance frameworks, and to the design of protocols for eliciting operational requirements from domain experts.

Beyond the partnership, the role is intended to grow into a longer-term programme with the theme lead on AI evaluation and verification, covering topics such as evaluation methodology for language-model and agentic systems, uncertainty quantification, and the integration of evaluation evidence into structured assurance cases.

The post holder will work with Dr Christopher Burr (Principal Research Scientist and Theme Lead for Trustworthy and Assured AI) and report directly to him. They will also work with a research software engineering team that builds and runs the theme's evaluation infrastructure. Internally they will be expected to collaborate with colleagues across Fundamental Research and other missions (e.g. Defence and National Security). Externally, the main stakeholders are, initially, the industry partner and, through it, government stakeholders in defence and national security. The role involves working in a defence and national security context and will require security vetting.

The role will make use of the Trustworthy and Ethical Assurance (TEA) Platform, an open-source web application developed by the Trustworthy and Assured AI theme, which enables project teams to build structured, graphical assurance cases demonstrating how goals like fairness, explainability, safety, and sustainability have been achieved in data-driven technologies.

HOW YOU WILL MAKE AN IMPACT

- Publish and disseminate high-quality research papers and publications detailing research outputs and project case-studies.
- Cultivate strong relationships and work collaboratively with internal stakeholders, academic experts and research partners across the Turing and wider Turing community.
- Communicate and present research outputs to diverse stakeholders, including project partners, through conferences, events, sprint reviews, workshops, meetings and press opportunities as appropriate.
- Undertake required administration tasks (e.g. organising workshops, roundtables, peer review).
- Contribute to the life of the Institute and support a diverse and inclusive community through embracing the Turing values.
- Adhere to and promote principles of reproducible and ethical data science and ensure secure handling of data and health and safety in all aspects of work.
- Undertake high-quality research, contributing to the evaluation and assurance of machine-learning systems within the Trustworthy and Assured AI theme and Fundamental Research.
- Lead the research and writing of a peer-reviewed paper setting out a formal account of how evaluation and explanation evidence about machine-learning systems can be combined, illustrated with results from the theme's evaluation platform.
- Design and analyse evaluation studies of machine-learning components, including choice of metrics, experimental protocols and statistical analysis, for explanation methods and for systems operating under latency and hardware constraints, working with research software engineers who implement and run them.
- Map evaluation evidence onto assurance and certification frameworks and onto structured assurance arguments.
- Design protocols for eliciting requirements from domain experts and operators, and complete the ethics applications and data protection assessments needed to run them.

- Draft specifications and handling plans for datasets curated with partners, and contribute to the design of the theme's evaluation platform.
- Contribute to the theme's longer-term research on AI evaluation and verification, including evaluation methodology for agentic and language-model systems.
- Contribute to project reports and grant applications for follow-on funding.
- Work within the security requirements of a defence and national security context, including the handling of partner material.

ADDITIONAL DUTIES FOR SENIOR RESEARCH ASSOCIATE LEVEL

- Provide intellectual and technical leadership for the research project, defining and refining the research direction with the Principal Investigator and wider team.
- Provide senior intellectual leadership for the integration of evaluation and explanation evidence into a formal research account, reviewing and shaping the associated peer-reviewed publications.
- Lead engagement with project partners and stakeholders on research priorities, assurance requirements and the interpretation of evaluation evidence, deputising for the Principal Investigator where appropriate.
- Lead the preparation of proposals, applications and reports for follow-on funding and partnership activity.
- Coordinate research delivery across relevant researchers and research software engineers, identifying dependencies, risks and decisions required to achieve successful outcomes.
- Provide leadership, support and mentorship to junior researchers, interns, students or other early-career colleagues where required.
- Play a leading role in shaping the Trustworthy and Assured AI theme’s longer-term research direction on AI evaluation, verification and assurance.

OTHER DUTIES

- Teaching may be required as part of collaboration work.
- Travel may be necessary to meet requirements of the role

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.

PERSON SPECIFICATION		
Skills and Requirements Post holders will be expected to demonstrate the following:	**Essential (E)** **Desirable (D)**	**Tested at application (a) / interview (i)**
A PhD (or equivalent experience and/or qualifications) in a relevant area (e.g. computer science, machine learning, statistics, formal methods, or philosophy of science with demonstrable technical skills in programming and machine learning). Equivalent experience may include substantial industry experience in the evaluation, testing or assurance of AI/ML systems.	E	A
Research experience in one or more of: evaluation of machine-learning systems, explainable AI, formal reasoning under uncertainty (probabilistic or Bayesian methods, evidence theory), or AI assurance.	E	A/I
Experience designing and running evaluations of machine-learning models, including choice of metrics, experimental design and statistical analysis.	E	A/I
Experience with quantitative data analysis and competency in Python and the scientific machine-learning stack.	E	A/I
Experience with Git and GitHub for version control and collaborative development, and experience of working on reproducible research code with software engineers.	E	A/I
Excellent written and verbal communication skills including the ability to present complex or technical information to diverse stakeholders outside the research community	E	A/I
Evidence of high-quality publication(s) in a relevant field (published or in-press) commensurate with your career stage	E	A
Ability to collaborate successfully with colleagues in a multidisciplinary environment within the organisation/externally to share knowledge and information in order develop practice or help others learn	E	A/I
Willing to work in a defence and national security context and able to undergo the security vetting the role requires. Must be eligible for SC clearance	E	I
Commitment to EDI principles and to the Organisation values (outlined below)	E	I
Ability to negotiate and influence others	E – Senior Research Associate	I
Demonstrable experience of leading research projects	E – Senior Research Associate	A/I
Ability to carry out research independently and take the lead on writing up findings as they emerge	E – Senior Research Associate	A/I
Ability to take the lead on research direction in collaboration with the PIs	E – Senior Research Associate	A/I
Previous experience overseeing, supervising or mentoring more junior colleagues' work	E – Senior Research Associate	I
Familiarity with assurance methodologies, safety cases, or structured argumentation frameworks (e.g. GSN, CAE).	D	A
Knowledge of standards and guidance for AI in safety-critical or regulated sectors (e.g. JSP 936, EASA guidance on machine learning, ISO/IEC 42001).	D	A/I
Experience of research in a defence, security or other regulated context.	D	A/I

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 3862 3565 or email recruitment@turing.ac.uk.

TERMS AND CONDITIONS

This full-time post is offered, on a permanent basis. The annual salary range is £46,415 - £52,494 for Research Associate level / £57,000 - £59,000 for Senior Research Associate level. Our offers include excellent benefits, including flexible working and family friendly policies, which can be seen here.

Candidates who have not yet been officially awarded their PhD will be appointed as Research Assistant at a salary of £43,445 (pro-rated to working hours) per annum.

Eligibility for Security Check (SC) clearance is a requirement for this role. Eligibility criteria and further information on the process can be found on the UK Government security vetting website.

The Alan Turing Institute is based at the British Library, in the heart of London's Knowledge Quarter. We expect staff to come to our office at least 4 days per month. Some roles may require more days in the office; the hiring manager will be able to confirm this during the interview.

EQUALITY, DIVERSITY AND INCLUSION

We value diversity of background, experience, and perspective, and are proud to be an inclusive employer. We warmly encourage applications from all backgrounds, particularly from groups currently under-represented in our sector. If you feel passionate about this role but don't meet every single requirement, please apply – we recognise that great candidates may bring strengths beyond the criteria listed.

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 recruitment@turing.ac.uk to advise us 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***.***

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.

