

# The Alan Turing Institute

## RESEARCH ASSOCIATE, APPLIED MACHINE LEARNING (Synthetic Data focus) (x2)

### 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 several 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 Finance and Economics programme brings together leading experts in data science, machine learning, finance and the social sciences, from both academia and industry to tackle the most challenging questions by producing world-leading research with significant impact. We inform public policy and enable trusted, research-led thought leadership. The programme works closely with government and the industry to exploit the potential of new technologies in the financial sector and economic research, and to position the UK as the leader in these areas.

We have recently launched a five-year Strategic Partnership with Accenture to advanced data science and Artificial Intelligence (AI) research with a focus on delivering substantial business and societal value via: a) delivering value from AI and data; b) enabling safe and robust application of AI and; c) lowering barriers to AI adoption.

### ROLE PURPOSE

The Turing has created a new Applied Machine Learning (AML) team to support and enable the delivery of the research sponsorship pillar of the partnership. The AML team will be a small, agile team of expert data scientist, machine learners and data science software engineers working on business inspired research challenges. The team will also act as a catalyst for collaboration with academic experts from across Turing network and broader research partners, and enable a high level of interactivity with Accenture Data Science Teams and relevant stakeholders.

The team will collectively have a broad range of expertise, and in Year One will initially be deployed to work on two key themes: Synthetic Data and Privacy Enhancing Technologies.

This role will be focussed on the challenge of Synthetic Data and will be responsible for work that enables the development of tools and examples to allow sharing of private datasets with a wider range of stakeholders, while preserving privacy.

You will be expected to perform high-quality research under the supervision of the principal investigator. Specifically, you will produce breakthrough research in this nascent field of research and contribute to publishing these results in top-rated journals and at national and international conferences, as appropriate. Dedicated (at 0.8FTE) to the Accenture Turing Strategic Partnership, the AML team will also play a part in tackling the broader, and related, research challenges of the Finance and Economics programme, providing a positive feedback loop into the Accenture-Turing partnership. This role reports to the Programme Director for Finance and Economics.

# The Alan Turing Institute

## Synthetic Data

Datasets are often stored in silos spread across organisations and are not easy to share with outside entities (e.g. the academic community) or with different departments within organisations. Roadblocks to sharing are principally privacy constraints and regulatory requirements. This creates a challenge for developing, testing and monitoring complex data-driven decision-making processes.

This project will explore the use of synthetic data generators (SDGs) to produce high-quality data that preserves the statistical features of the original data set. Among other benefits, they enable users to share and link data, to work with data in safe environments, to fix structural deficiencies in data, to increase the size data, and to validate machine learning systems by generating adversarial scenarios.

### DUTIES AND AREAS OF RESPONSIBILITY

The core responsibilities of the Research Associates are as follows:

- Apply and develop new state-of-the-art data science and machine learning techniques to the business-inspired research challenges of the Accenture-Turing partnership:
  - Develop state of the art data generators for both structured and unstructured data sets.
  - Develop metrics for evaluating the utility, similarity, and privacy of synthetic data sets across multiple use cases.
  - Provide methodologies for assessing the utility and privacy trade-offs
  - Enable data sharing between organisations and different departments within an organisation.
  - Establish systems to train and validate machine learning models under adversarial scenarios.
- To establish a sound research base within the Alan Turing Institute in order to pursue individual and collaborative research of outstanding quality, consistent with making a full active research contribution in line with the research strategy outlined by the PI; -
- To write or contribute to publications or disseminate research findings using other appropriate media;
- To attend and present research findings and papers at academic and professional conferences, and to contribute to the external visibility of the Institute;
- To participate in and to develop internal and external partnerships, for example to identify sources of funding, generate income, obtain projects/datasets, or build relationships for future activities;
- To ensure compliance with secure handling of data and health and safety in all aspects of work.

#### Other duties:

The AML team may be expected to:

- Scope, pilot and deliver high quality research activity in partnership with Accenture stakeholders, and under the Direction of the Programme Director.
- Work with Accenture Data Science team to develop and refine additional business-inspired research challenges.
- Drive collaboration with academic experts and broader research partners from across the Turing and the wider Turing community.
- Contribute to the broader research aims and challenges of the Finance and Economics programme, and ensure positive feedback to the Accenture-Turing partnership.
- Present, disseminate and explain our work at internal and external events hosted by Turing and/or Accenture.
- Engage with the Skills and Training pillar of the partnership. For example, presenting at an Executive Education workshop.
- Contribute to the life of the Institute and support its community.

# 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</b>		
A PhD (or equivalent experience and/or qualifications) in a relevant area, which will include Statistics, Mathematics, Engineering, Computer Science, or related discipline.	E	A
<b>Knowledge and Experience</b>		
Strong background in one or more of the following areas: Bayesian inference, probabilistic machine learning (e.g. GANs, VAE or Energy-based models), information theory, uncertainty quantification and data assimilation methods for online calibration of time-dependent models, time-series modelling, signature and rough path theory, federated learning, privacy-enhancing technologies.	E	AI
Experience managing, structuring, and analysing research data.	E	AI
Fluency in one or more modern programming languages used in research in data science and artificial intelligence(e.g Python)	E	AI
An understanding of the importance of good practices for producing reliable software and reproducible analyses (e.g. version control, issue tracking, automated testing, package management, literate analysis tools such as Jupyter and Rmarkdown)	E	AI
Ability to rapidly assimilate new computational and mathematical ideas and techniques on the job, at a more than superficial level, and apply them successfully.	E	AI
Ability to work with software like PyTorch and TensorFlow to develop deep generative models using GANs, RNNs, and Autoencoders. Develop pipelines to evaluate the performance of the generators.	E	AI
Experience in making or evaluating the case for new projects (e.g. authoring or evaluating research proposals or business cases)	D	AI
Experience of managing, prioritising and resourcing a project portfolio, leading it to a successful conclusion	D	AI
Experience contributing to, maintaining and/or leading open source research software projects.	D	AI
Developing for high-performance computing hardware (CUDA, MPI, OpenMP)	D	AI
Experience with public cloud platforms and writing technical documentation	D	AI
<b>Communication</b>		
Excellent written and verbal communication skills, including experience in the visual representation of quantitative data, documentation of software packages or data resources, the authoring of research papers or technical reports, and giving presentations or classes on technical subjects.	E	AI
Adapts the style of communication to the audience and ensures understanding.	E	AI
Ability to communicate complex, specialist or conceptual information clearly and persuasively, presenting compelling arguments to influence and/or negotiate satisfactory outcomes.	E	AI

# The Alan Turing Institute

<b>Liaison and Networking</b>		
Participates in networks within the organisation or externally to share knowledge and information in order develop practice or help others learn.	E	AI
Networks with others with shared interests, collaborating on projects and strengthening future relations.	E	AI
<b>Service Delivery</b>		
Proactive approach to managing stakeholders and identifying opportunities for collaboration.	E	AI
Adapts services and systems to meet stakeholders' needs and identifies ways of improving standards. Learns from complaints and takes action to resolve them.	E	AI
User interface design and development with web technologies, especially for data visualisation and knowledge representation.	D	AI
Ability to work with databases and APIs for the acquisition of parameter information for models.	D	AI
<b>Decision-Making Processes and Outcomes</b>		
Ability to lead own work Independently, and 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	AI
Work with others to make collaborative decisions that may be operational or strategic and impact immediate team or work area.	E	AI
Recommend and advise on available options for decisions that affect operational processes, taking into account any risks.	E	AI
<b>Initiative and Problem Solving</b>		
Uses judgement to analyse and solve problems, and take action to prevent recurrence of problems.	E	AI
Consider possible solutions to identify those which offer wider benefits, and obtain evidence to support thinking.	E	AI
<b>Analysis and Research</b>		
Identify and use a range of standard sources to gather and analyse routine data and produce reports that can be interpreted by others.	E	AI
Design and use data gathering and analytical methods appropriate for each investigation. Recognises and accurately interprets patterns and trends. Understands when additional data is required and identifies appropriate sources. Produces reports that identify key issues and findings.	E	AI
<b>Other Requirements</b>		
Working with confidential and sensitive data for research.	E	I

# The Alan Turing Institute

Commitment to EDI principles and to the Organisation values

E

I

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

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 020 3862 3575, or email [recruitment@turing.ac.uk](mailto:recruitment@turing.ac.uk).

**CLOSING DATE FOR APPLICATIONS: 31 January 2021**

## TERMS AND CONDITIONS

This full-time post is offered on a fixed-term basis for two years. The annual salary is £35,000-£41,000 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 £32,000-£34,000 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).***