The Alan Turing Institute

Research Associates (x2), Computational statistics and deep learning with applications in biomedicine (Turing-Roche Partnership)

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

The Health and Medical Sciences Programme at Turing delivers research into the theory and methods of AI, statistics, and data analytics underpinning medical and health applications that will enable scientists to do better science, without compromising respect for privacy and patient trust.

Hoffman La-Roche (Roche) as a company have been committed to improving lives since it was founded in 1896 in Basel, Switzerland. Today, Roche creates innovative medicines and diagnostic tests that help millions of patients globally and was one of the first companies to bring targeted treatments to patients.

The Alan Turing Institute and Roche have recently launched a new programme which will establish a world-leading industry and academic partnership in advanced analytics focused on enabling the transformative benefits of personalized healthcare to become a reality for patients around the world.

This strategic partnership will cover multiple activities, with the "North Star" of developing new data science methods to investigate large, complex, clinical and healthcare datasets to better understand how and why patients respond differently to treatment, and how treatment can be improved. Understanding such "treatment heterogeneity" is a problem at the forefront of modern medicine and is an essential first step toward the ambitious goal of developing a personalised healthcare. Through the partnership, multiple avenues of research on heterogeneity in patient health data will be explored, the first of which focuses on structured missing data.

ROLE PURPOSE

We are looking to recruit into two positions. These are exciting opportunities for researchers with a strong background in statistical methodology and applied data analysis, who would like to grow that skill set further as well as collaborate as part of a diverse team comprising individuals across both academia and industry. The postholders will work at the Alan Turing Institute as part of a sizeable team of post doctoral researchers at the Turing led by Prof. Ben MacArthur and Dr Owen Rackham, working in close collaboration with Roche and Genentech scientists at Welwyn and San Francisco, alongside academics from affiliated institutes, including Dr Robin Mitra at UCL who leads a theme of work on "structured missingness" (see below).

The appointed candidates will focus on the analysis of large complex biomedical datasets, which is increasingly essential for scientific research. More specifically, one researcher will focus on developing robust methods to tackle missing data; while the second will focus on developing computationally efficient methods for learning from large and complex health datasets at scale.

Missing data is a common complication which can impede the effective use of data. Very often, missing values are not randomly distributed, but rather arise with an associated structure. Such "structured missingess" (see our paper here) can occur for a variety of reasons including, but not limited to, merging of different data sources across different individuals and variables, complex and varied data collection practices, and underlying latent dependencies characterising missingness processes. The successful candidate working on this theme will develop new tools for imputation and modelling to handle structured missingness in the context of large, complex, biomedical datasets.

Large and complex health datasets are playing an increasingly important role in the development of personalised healthcare. Analysing such data typically requires adaptations of existing methodologies to allow them to be deployed robustly and efficiently in healthcare contexts. Moreover, there is now an increasing need for new tools, methods and procedures, that are attuned not only to the computational challenges associated with learning at scale from complex clinico-genomic data, but also the constraints of clinical data capture processes, and the ethical implications of deployment in the clinic. The successful candidate working on this theme will develop new tools for learning at scale from complex biomedical data. This may include areas such as optimising the properties of neural networks through a theoretical understanding of their geometric intepretation, data fusion, network modelling, or scaling up of computational methods. We are particularly interested in developing tools that combine experimental data modalities – such as genetics and transcriptiomics – with clinical information, and so knowledge or experience of these data types will be an advantage. Expertise in genetics and/or genomic modelling is particularly welcome.

Many of the motivating challenges associated with both projects arise from the Flatiron Health-Foundation Medicine clinico-genomic database that comprise highly curated data, linking clinical data derived from electronic health records to comprehensive genomic data, across a large group of individuals in the US. We intend to develop methodology that includes, but is not limited to, facilitating effective analysis of this database to advance important research goals in understanding patient and outcome heterogeneity.

The successful candidates will have hands-on experience in analysing complex large data sources and will be provided with the opportunity to develop their methodological research portfolio and expertise, and see their tools deployed to address real-world analysis challenges in both academic and industrial contexts. The postholders will work with a substantial team of researchers within the Turing-Roche partnership as well as researchers at Roche and Genentech in the UK, Basel and San Francisco and a number of senior academics associated with the Turing-Roche partnership.

This is a stand-out opportunity to closely interact with both **The Alan Turing Institute** as a prestigious, national research institute and **Roche** as one of the world's largest pharmaceutical companies, which will give you visibility of many facets of the pharmaceutical industry and support to develop a network of contacts in both academia and industry.

DUTIES AND AREAS OF RESPONSIBILITY

- Undertake work including statistical analyses and relevant methodological research.
- Collaborate on projects with researchers at Roche/Genentech.
- Develop new methods and/or adapt existing methods to address complex missing data and computational statistics problems, including (but not exclusively) motivated from the Flatiron Health-Foundation Medicine clinico-genomic database.
- Publish the results of the methodological and applied research in high-quality journals and participate in national and international conferences.
- Disseminate research outputs to the Turing-Roche partnership team as well as research in the Roche Advanced Analytics group.
- Create new, or adapt existing, R / Python packages that implement the research methods developed.
- Maintain up to date skills and knowledge of statistical methods that are of relevance to the area of investigation.

- Participate in Unit, Departmental and Faculty meetings and seminars as appropriate aimed at sharing research outcomes and building interdisciplinary collaborations.
- Maintain effective communication and accurate records.
- Maintain own continuing professional development
- Actively follow and promote Turing policies, including those relating to Equality, Diversity and Inclusivity.
- 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

OTHER DUTIES

- Teaching may be required as part of collaboration work

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 Post holders will be expected to demonstrate the following:	**Essential (E)** **Desirable (D)**	**Tested at application (a)** **Tested at interview (i)**
Education/Qualification		
A PhD (or equivalent experience and/or qualifications) in a relevant area such as Mathematics, Statistics, Computer Science, or related discipline	E	A
Research Assistant level: must be near completion of PhD in a relevant area which include Mathematics, Statistics, Computer Science, or a related discipline	E	A
Knowledge and Experience		
Have knowledge of a wide range of statistical and/or deep learning methods used to model complex data	E	A,I
Knowledge of the area of missing data and methodology to address missing data problems or knowledge and expertise in computation statistics or experience of applying deep learning with knowledge of the underlying theory	E	A,I
Excellent computational skills with the statistical software R or Python	E	A,I
Interest in adapting existing, or developing new, R or Python packages, in particular building or training deep learning models	D	A,I
Collaborating experience with individuals from industry or other disciplines	D	A,I
Publishing research in peer reviewed journals	D	A,I
Communication		
Excellent written and/or 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	E	I
Ability to communicate and collaborate within the organisation or externally to share knowledge and information in order to develop best practice or help others learn	E	A,I
Project Management & Project Delivery		
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 partnership to contribute and assist in diverse research activities	E	I
Ability to represent the partnership at practioner events and high-level meetings	E	I
Ability to establish academic collaborations nationally or internationally	E	I
Decision Making		
Ability to work with others to make collaborative decisions	E	A/I

Initiative and Problem Solving		
Ability to use own judgement to analyse and solve problems	E	A,I
Ability to consider possible solutions and identify with evidence those which offer widest benefits.	E	A,I
Ability to lead one's own work, including planning and execution, and to prioritise work to meet deadlines	E	A,I
Analysis and Research		
Ability to organise working time, take the initiative, and carry out research independently	E	A
Ability to develop and lead collaborative research and innovation	E	I
Other Requirements		
Commitment to EDI principles and to the Organisation values	E	I

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 email recruitment@turing.ac.uk.

CLOSING DATE FOR APPLICATIONS: Sunday 12 May at 23:59 (London, UK BST)

We reserve the right to close this vacancy early or to interview suitable candidates before the closing date if enough applications are received.

TERMS AND CONDITIONS

These full time posts are offered on a fixed term basis for 24 months or until 30 June 2026 (whichever is earlier) and starting as soon as possible. The annual salary range for this position is £42,893 - £48,510 depending on experience, 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 £40,148 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 teal to reflect this. We therefore welcome applications from the broadest spectrum of backgrounds.

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***.***