PhD opportunities in addressing socio-technical limitations of LLMs

Up to 4 PhD fully funded PhD studentships are available in the *Keystone* project on addressing socio-technical limitations of Large Language Models (LLMs) in medical and social contexts. The PhD project will be part of AdSoLve (https://adsolve.github.io/), a large multi-disciplinary consortium funded by UKRI and RAi UK as part of strategic investment by the UK Government to create an international ecosystem for responsible AI research and innovation. It is led by Prof Maria Liakata at QMUL, with four University partners (QMUL, Nottingham, Sheffield and Warwick) and 21 external partners, including large, diversified companies, NHS trusts, NHS England and AI hubs and UKRI and EPSRC Centres for Doctoral Training (CDTs).

The successful applicants will be registered either at QMUL with a co-supervisor at Imperial College London or the UKRI AI Centre for Doctoral Training in Digital Healthcare (AI4Health) at Imperial College London with a co-supervisor from QMUL, depending on project/topic suitability. The studentships will cover UK tuition fees and stipends at UKRI rate (£21,237 for 2024-25) for a minimum of 3 years subject to successfully completing progression milestones.

Context

The vision of the AdSoLve consortium is to address the socio-technical limitations of LLMs that challenge their responsible and trustworthy use, particularly in the context of medical and legal use cases.

AdSoLve has two primary goals:

1) To create an extensive evaluation benchmark (including suitable novel criteria, metrics and tasks) for assessing the limitations of LLMs in real world settings, enabling our standards and policy partners to implement responsible regulations, and industry and third sector partners to robustly assess their systems. To achieve this synergy we will be running co-creation and evaluation workshops throughout the project to create a co-production feedback loop with our stakeholders.

2) Secondly, is to devise novel mitigating solutions based on new machine learning methodology, informed by expertise in law, ethics and healthcare, via co-creation with domain experts, that can be incorporated in products and services. Such methodology includes development of modules for temporal reasoning and situational awareness in long-form text, dialogue and multi-modal data, as well as alignment with human-preferences, bias reduction and privacy preservation.

The PhD research area

This PhD project will focus on the intersection of natural language processing (NLP), machine learning (ML), multimodal AI, responsible innovation, and healthcare. The successful applicant will be expected to work on one or two of the following topics during their PhD:

- State-of-the art methods for efficient and robust alignment of LLMs with human preferences for reliable generation and reduction in hallucinations
- Robust methods, tasks and metrics for fine-grained evaluation of generated content
- Augmentation of LLMs with temporal reasoning and situational awareness
- Fine-grained alignment of cross-modal LLMs
- Multi-modal representation learning for longitudinal settings
- Methods for longitudinal multi-modal monitoring for healthcare
- Methods for longitudinal synthetic language generation for data augmentation and privacy preservation.
- Methods for situation-aware, privacy-preserving, trustworthy dialogue systems for longitudinal health monitoring and self-management
- Privacy preservation and bias reduction while applying LLMs to sensitive data
- Explainable prediction and generation with LLMs
- Methods for temporally aware summarisation of real-world long-form documents including therapy sessions, social media threads, court cases.

Desired outputs include (but are not limited to) publications in top-tier NLP and ML venues and are expected to have high impact in the fields of NLP, machine learning, as well as domains such as healthcare and law, both in terms of methodological innovation and their application to real-world settings. There will also be the opportunity to work closely with researchers at the Alan Turing Institute.

The AdSoLve Keystone project and UKRI AI Centre for Doctoral Training in Digital Healthcare (AI4Health Centre)

The AdSoLve Keystone project is a large multi-disciplinary consortium project funded by UKRI and RAi UK. Details can be found at https://adsolve.github.io/).

The UKRI AI Centre for Doctoral Training in Digital Healthcare (AI4Health CDT) at Imperial College London delivers cohort-based training that integrates the development of technical skills with an appreciation for approaches to human-in-the-loop AI design that are socially and ethically acceptable. Details can be found at https://ai4health.io.

Eligibility

Essential Qualifications

- Academic degrees:
    - [MSc *or* 4-year UG degrees] with Distinction and/or first-class honours in [Natural Language Processing, Computer Science, Machine Learning, Artificial Intelligence, Engineering *or* related field]. Exceptional candidates

with a background in Medicine, Psychology, Psychiatry or Law may also be considered.

- Experience in at least one of the areas below:
  - Natural Language Processing Machine Learning/Deep Learning
- Programming Skills:
  - Strong coding skills, preferably in Python.

Desirable skills

- Programming skills:
  - Experience with machine learning frameworks based on Python such as Pytorch, TensorFlow and libraries from Huggingface
  - Familiarity with python scientific packages e.g. numpy, pandas, sklearn, scipy, matplotlib
  - Experience with version control systems
- Research Experience:
  - Experience with working in research projects
  - Publication(s) in NLP/ML venues
- Interest in/ Knowledge of Health applications & associated challenges
- Interest in learning about new AI technologies

We particularly encourage students from groups that are currently underrepresented in postgraduate science research, including black and minority ethnic students and those from a socio-economically disadvantaged background.

Application Information, Studentship and Eligibility

The studentship covers:

- Full-time PhD tuition fees at the UK home rate. For successful international candidates it may be possible to cover international fees.
- An annual (usually tax free) stipend set at the UKRI rate (£21,237 for 2024-25) per year, for a minimum of 3 years, subject to satisfying progression milestones.

To be classed as Home for tuition fee purposes, students typically need to have unrestricted access on how long they can remain in the UK (i.e. are a British National, have settled, or pre-settled status, have indefinite leave to remain etc.). The tuition fee status is determined by the university's Registry at the point of application.

How to apply

- Application Process: Applicants are invited to submit their applications through the Queen Mary University of London Application system Subjects - Queen Mary University of London (qmul.ac.uk). Scroll down and select PhD Full Time Computer Science - Semester 1 (September Start).
- Successfully shortlisted applicants considered for a project with the supervisor from Imperial would be advised to apply through the Imperial College application system,

and satisfy the admissions requirements of both Imperial College and the PhD programme of the UKRI AI Centre for Doctoral training in Digital Healthcare (AI4Health).

- All applicants must submit the following:
- 4-year UG degree/Diploma/PG Degree transcripts (translated in English, if needed)
- CV **(max 2 pages)**
- A one-side A4 statement of purpose
- Research proposal **(500 words)**
- 2 References
- Certificate of English Language (for non-UK students)
- Other (academic or not) Certificates

Queries regarding the PhD topic and project should be directed to Prof Maria Liakata (m.liakata@qmul.ac.uk)

If there are any queries about the UKRI AI Centre for Doctoral Training in Digital Healthcare, please contact Britta Ross b.ross@imperial.ac.uk

Deadline for Applications: **28 July 2024** (23:59GMT)