



PhD position: socio-technical challenges of delivering safe and ethical LLM-based applications in health and law

Please note: To be classed as a home student for fee purposes, candidates typically need to have unrestricted access on how long they can remain in the UK (i.e., British National, have settled, or pre-settled status, have indefinite leave to remain, etc.) and have been living in the UK for the 3 years immediately prior to studentship starting. Fee status is determined by the Admissions department at point of application.

We invite applications for a PhD award linked to the AdSoLve project (Addressing socio-technical limitations of Large Language Models (LLMs) in medical and social contexts). AdSoLve is a large multi-disciplinary project funded by UKRI and RAI UK (for details, see www.rai.ac.uk). AdSoLve is led by Prof Maria Liakata at QMUL and involves four University partners (QMUL, Nottingham, Sheffield and Warwick) and 21 external partners, including commercial companies, NHS trusts, NHS England and UK AI hubs and CDTs.

AdSoLve aims to: 1) create evaluation benchmarks (including novel criteria, metrics and tasks) for assessing ethics and safety risks of LLMs in real world applications; 2) draw on expertise in law, healthcare, AI ethics and safety to devise new machine learning methodologies to mitigate these risks. The project may also involve developing 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.

Overview

The successful applicant will be funded for 3.5 years for a PhD degree in the Department of Computer Science at Warwick University. The AI and Human Centred Computing (AI-HCC) theme brings together research at the forefront of AI – multi-agent systems, machine learning, safe and ethical AI, human-AI interaction, explainable AI, participatory design, social shaping of AI.

The Computer Science department has 50 academic staff and was ranked 4th equal in REF2021. It collaborates with Warwick Medical School, the Interdisciplinary Ethics Research Group and the Centre for Interdisciplinary Methodologies.

The student will take a socio-technical approach to exploring challenges of implementing Responsible Research and Innovation (RRI) principles in the development of LLM-based applications in healthcare and law. A key goal will be to devise recommendations on how these may be addressed. This will involve: mapping AI ethics and safety regulations, RRI practices; surveys and interviews with LLM researchers, LLM-based application development teams, end users and stakeholders, focusing on practices, tools, factors influencing RRI adoption; ethnographic studies of LLM-based applications development.

Background

Essential Qualifications

- Academic degrees:
 - Masters degree at merit level or above, BA 1st Class in Sociology of Science and Technology, Science and Technology Studies, Social Anthropology or related field. Exceptional candidates with a BSc/BEng in Computer Science, AI, Machine Learning may also be considered.
- Research skills:
 - Experience in qualitative research methodologies.

Desirable Skills

- Research experience:
 - Familiarity with ethnographic fieldwork methods
 - Ability to work in an interdisciplinary research team
- Interest in AI ethics and safety
- Interest in NLP, Machine Learning, Deep Learning, and their applications
- Interest in learning about new AI technologies

Application Information, Studentship and Eligibility

The studentship covers:

- Full time PhD tuition fees at the UK home rate.
- An annual (tax free) stipend set at the UKRI rate (£19,237 for 2024-25) per year, for 3.5 years (pro-rata for part-time students).

To be classed as a home student for fee purposes, candidates typically need to have unrestricted access on how long they can remain in the UK (i.e., British National, have settled, or pre-settled status, have indefinite leave to remain, etc.) and have been living in the UK for the 3 years immediately prior to studentship starting. Fee status is determined by the Admissions department at point of application.

How to Apply

- Application Process: Applicants are invited to submit their applications through [pgradmin.dcs.@\[warwick.ac.uk\]\(http://warwick.ac.uk\)](mailto:pgradmin.dcs@warwick.ac.uk). All applicants must submit the following:
 - BA/PG Degree transcripts
 - CV (max 2 pages)
 - Cover letter (max 4,500 characters) outlining your interest, and suitability, including how your current skills and experience are relevant to this PhD
 - Research proposal (max 500 words)
 - 2 References
 - Other (academic or not) Certificates

Queries regarding the topic and project should be directed to Prof Rob Procter (rob.procter@warwick.ac.uk). If there are any queries about the application process, refer to pgradmin.dcs@warwick.ac.uk.

Deadline for Applications: July 31st (23:59GMT)