Post Title: PhD Studentship in Ethics and Multilingual Communication Using Generative AI
Location: Dublin City University (DCU)
Anticipated Start Date: September/October 2023
Closing Date: 4th August 2023
Apply: https://forms.gle/Fu9rHRVp7uWi7T1a6

Why ADAPT?

- **Contribute** to the ADAPT research agenda that pioneers and combines research in AI driven technologies: Natural Language Processing, Video/Text/Image/Speech processing, digital engagement & HCI, semantic modeling, personalisation, privacy & data governance.
- **Work** with our interdisciplinary team of leading experts from the complementary fields of, Social Sciences, Communications, Commerce/Fintech, Ethics, Law, Health, Environment and Sustainability.
- **Leverage our success.** ADAPT's researchers have signed 43 collaborative research projects, 52 licence agreements and oversee 16 active commercialisation funds and 52 commercialisation awards. ADAPT has won 40 competitive EU research projects and obtained €18.5 million in non-exchequer non-commercial funding. Additionally, six spinout companies have been formed. ADAPT's researchers have produced over 1,500 journal and conference publications and nearly 100 PhD students have been trained.

As an ADAPT/DCU co-funded PhD researcher you will have access to a network of 85 global experts and over 250 staff as well as a wide multi-disciplinary ecosystem across 8 leading Irish universities. We can influence and inform your work, share our networks and collaborate with you to increase your impact, and accelerate your career opportunities. Specifically we offer:

1. Opportunity to build your profile at international conferences and global events.
2. A solid career pathway through formalised training & development, expert one-on-one supervision and exposure to top specialists.
3. A 4 year PhD postgraduate studentship with research student fees covered in full and including a non-taxable stipend (currently €19,000 per annum), along with equipment and travel funding
4. Open to EU and non-EU applicants

Context

Ethics and Multilingual Communication Using Generative AI

The term artificial intelligence (AI) has recently become synonymous with deep learning using neural networks, as is applied for machine translation (MT), speech recognition, and the use of large language models (LLMs) for text generation. These are examples of subsymbolic AI based on training data that is converted to numbers and then processed via "a stack of equations – a thicket of often hard-to-interpret operations on numbers" (Mitchell 2020, 24). As a result, systems are often referred to as an impenetrable 'black box'. Bender et al. (2021, 8) identified environmental and financial costs

behind LLMs, and further commented that the apparent "fluency and coherence" of their output raises risks because people interpret text as "meaningful and corresponding to the communicative intent of some individual or group of individuals who have accountability". However, output has been found to encode hegemonic societal views from training data, particularly from polarised content scraped from the web, producing output that is at times problematic, biased, and offensive. In order to minimise this effect, some developers use reinforcement learning using human feedback, for example OpenAI employ staff (remotely and on low pay) to make ChatGPT output less toxic (Perrigo 2023).

The proposed interdisciplinary PhD project will investigate the ethics of multilingual communication using LLMs and will be supervised by Dr Joss Moorkens (School of Applied Language and Intercultural Studies/SALIS) and Prof Bert Gordijn (Institute of Ethics). There has so far been relatively little research produced on this topic. For example, Bender et al. (2021) considered potential dangers and ethical issues related to LLMs and Mitchell and Krakauer (2022) considered the debate about whether LLMs can be said to understand content. However, these did not focus on multilingual communication and were written before ChatGPT launched on November 30th, 2022, reaching one million users in five days. The intervening months have seen some articles on the translation and translation evaluation capabilities of LLMs, along with one short article by Moorkens (2023), but now would appear to be the perfect time to systematically assess the ethical aspects of the development and use of LLMs for translation and multilingual communication in order to accommodate responsible research and innovation.

On completion of this project, the graduate should have published work that adds to the field. Experience in this topic should be highly beneficial, leading to opportunities for further research at postdoctoral level in industry or academia. There is an expectation that the student would contribute to teaching in SALIS, which should also increase employability on graduation. The project is co-funded by the ADAPT Centre and Dublin City University.

Informal enquiries may be directed to Joss Moorkens at joss.moorkens@adaptcentre.ie

References

Bender E et al. 2021 On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? FAccT '21, Canada.
Mitchell M 2020 Artificial Intelligence: A guide for thinking humans. London, Penguin.
Mitchell M, Krakauer DC 2022 The Debate Over Understanding in AI's Large Language Models. arxiv.org/pdf/2210.13966.pdf
Moorkens J 2023 Generative tools using large language models and translation. Marie Curie Alumni Association Newsletter 35, 16-18. https://www.mariecuriealumni.eu/sites/default/files/2023-06/MCAA-Newsletter-June-2023.pdf
Perrigo B 2023 OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic. Time.

Minimum qualifications:

Applicants should:

1. Hold a relevant undergraduate degree at 2.1 or first-class honours level (or equivalent) in Translation Studies, Translation Technology, Philosophy of Technology, Ethics, or an aligned

field. Ideally, they should also hold a relevant master's qualification. Candidates who are currently completing a master's qualification are welcome to apply.
2. Meet, where relevant, the English language requirements for non-native speakers of English. Link [here](#).

Preferred qualifications:

- *Demonstrable research competence from previous projects at undergraduate or postgraduate level*
- *Some understanding or experience in machine translation, machine learning, and big data*

Application Process

Applications should be submitted as a single pdf file. Each application should only consist of

1. Detailed curriculum vitae, including – if applicable – relevant publications;
2. Transcripts and results of final years of degrees,
3. The name and email contacts of two academic referees,
4. A cover letter/letter of introduction (max 1000 words). In the letter, applicants should include the following details:
   a. An explanation of your interest in the research to be conducted and why you believe they are suitable for the position.
   b. Details of your final year undergraduate project (if applicable)
   c. Details of your MA/MSc project (if applicable)
   d. Details of any relevant modules previously taken, at undergraduate and/or master's level.
   e. Details of any relevant work experience (if applicable).

Diversity

ADAPT is committed to achieving better diversity and gender representation at all levels of the organisation, across leadership, academic, operations, research staff and studentship levels. ADAPT is committed to the continued development of employment policies, procedures and practices that promote gender equality. On that basis we encourage and welcome talented people from all backgrounds to join ADAPT.

About the ADAPT Centre

ADAPT is the world-leading SFI research centre for AI Driven Digital Content Technology hosted by Trinity College Dublin. ADAPT's partner institutions include DCU, University College Dublin, Technological University Dublin, Maynooth University, Munster Technological University, Athlone Institute of Technology, and the National University of Ireland Galway. ADAPT's research vision is to pioneer new forms of proactive, scalable, and integrated AI-driven Digital Content Technology that empower individuals and society to engage in digital experiences with control, inclusion, and accountability with the long term goal of a balanced digital society by 2030. ADAPT is pioneering new Human Centric AI techniques and technologies including personalisation, natural language processing, data analytics, intelligent machine translation human-computer interaction, as well as setting the standards for data governance, privacy and ethics for digital content.

Our Research Vision

Governments and civil society are starting to recognise the need for urgent and concerted action to address the societal impact of the accelerating pace of digital content technologies and the AI techniques that underpin them. ADAPT provides an ambitious, ground-breaking, integrated research programme that assembles three interlocking Strands that together are capable of addressing this challenge. Each of these complementary and reinforcing research Strands takes one of the different perspectives on the provision of personalised, immersive, multimodal digital engagement, i.e. the individual's experience and control of the engagement, the algorithms underlying digital content processing, and the balanced governance by enterprise and societal stakeholders.

Digitally Enhanced Engagement Strand

From the individual perspective, research within this Strand will deliver proactive agency techniques that sense, understand and proactively serve the needs of individual users to deliver relevant, contextualised and immersive multimodal experiences which also offer them meaningful control over the machine agency delivering those experiences.

Digital Content Transformation Strand

From the algorithmic perspective, new machine learning techniques will both enable more users to engage meaningfully with the increasing volumes of content globally in a more measurably effective manner, while ensuring the widest linguistic and cultural inclusion. It will enhance effective, robust integrated machine learning algorithms needed to provide multimodal content experiences with new levels of accuracy, multilingualism and explainability.

Transparent Digital Governance Strand

From the enterprise and societal perspective, new structured knowledge frameworks and associated practices for AI data governance will be required to balance the needs and values of individuals, organisations and society when it comes to rich digital experiences. This requires the advancement of research in the areas of data ethics, data quality, data protection, data value, data integration, and multi-stakeholder governance models.