PhD Scholarships in Explainable Artificial Intelligence (XAI) & Human Thinking

Funding for PhD scholarships (stipend plus fees) is available for a collaborative project between *University College Dublin* (*Prof. Mark Keane, School of Computer Science*) and *Trinity College Dublin* (*Prof Ruth Byrne, School of Psychology and Institute of Neuroscience*) to be hosted by the *School of Computer Science* at UCD (https://www.cs.ucd.ie/) and the *Insight Centre for Data Analytics* (www.insight-centre.org).

The research focus will be on the topic of Explainable Artificial Intelligence (XAI), Explanation and Counterfactual Thinking. Artificial Intelligence (AI) is becoming more pervasive in everyday life, yet complex AI systems trained on vast amounts of data can produce decisions that appear unintelligible to humans. There is an urgent need to enable AI systems to provide explanations of their decisions to human users, to ensure trust and fairness. The PhD studentships will explore fundamental research on cognitive processes in human explanation and counterfactual thinking of relevance to the development of XAI. The successful PhD scholarship candidates will carry out cognitive science research at the interface of cognitive psychology and artificial intelligence.

Applicants should have :

- At least 2.1 grade (or equivalent) in an undergraduate or postgraduate degree in psychology, cognitive science, computer science, maths, engineering or another cognate discipline (e.g., philosophy, linguistics).
- Programming expertise is desired but not a requirement.
- Expertise in experimental design is desired but not a requirement.
- Non-native English speakers require at least IELTS 6.5 (with at least 6 in all components) or equivalent (further details on equivalence of tests are available here).

Each student will receive a scholarship valued at approximately €100,000, to include a tax-free stipend of €18,000 per year for four years, full coverage of tuition fees, funds for conference travel, and an equipment allowance. Students will also have the opportunity to gain relevant experience and earn extra income within UCD through teaching activities. The anticipated start-date is September 2019.

All enquires to either:

mark.keane@ucd.ie (see https://people.ucd.ie/mark.keane)
rmbyne@tcd.ie (see https://psychology.tcd.ie/people/rmbyrne/)

Interested candidates should read:

Byrne, R.M.J. (2019). Counterfactuals in Explainable Artificial Intelligence (XAI): Evidence from Human Reasoning. *Proceedings of the 28th International Joint Conference on Artificial Intelligence, IJCAI19*.

Kenny, E.M. & Keane, M.T. (2019). Twin-systems to explain Artificial Neural Networks using Case-Based Reasoning: Comparative tests of feature-weighting methods in ANN-CBR twins for XAI. *Proceedings of the 28th International Joint Conference on Artificial Intelligence, IJCAI19*.

Available at: https://www.dropbox.com/sh/3uuzg0mf4lcpepe/AACrBOReumQzekSGblX2UByca?dl=0