



School of Computing, Engineering, and the Built Environment Edinburgh Napier University

PHD STUDENT PROJECT

Application instructions:

Detailed instructions are available at :
<https://www.napier.ac.uk/research-and-innovation/doctoral-college/how-to-apply>

Prospective candidates are encouraged to contact the Director of Studies (see details below) to discuss the project and their suitability for it.

Project details

Supervisory Team:

- DIRECTOR OF STUDY: Dr Simon Wells (Email: s.wells@napier.ac.uk)
- 2ND SUPERVISOR: tbc

Subject Group: Computer Science

Research Areas: Artificial Intelligence

Project Title: Advancing Explainable Artificial Intelligence

Project description:

There have been significant recent advances in the effectiveness of machine intelligence systems based upon neural networks, deep learning, and machine learning techniques. However questions have been raised about the nature of the reasoning occurring in these systems, how opaque that reasoning is, and how that impacts people's ability to trust the resulting systems. However patterns of human reasoning has been well studied over the last two millenia and a particularly useful approach has been that of argumentation schemes which capture stereotypical patterns of presumptive reasoning as used by people in the form of defeasible arguments.

This project will involve a detailed study of both existing machine intelligence systems and of argumentation schemes with the goal of understanding the stereotypical patterns of human & machine reasoning and advancing the state of the art of explainable AI. The novel contribution will be an executable model of argumentation schemes and the development of a correspondence theory relating

argumentation schemes to machine reasoning. This theory will be used to demonstrate how human and machine reasoning patterns are related, and how the symbolic and numeric approaches inherent to each can be augmented to facilitate the construction of explainable, justifiable, and composable AI systems.

Candidate characteristics

Education:

A second class honour degree or equivalent qualification in Software Engineering, Computer Science, Machine Learning, Artificial Intelligence

Subject knowledge:

- Artificial Intelligence

Essential attributes:

- Knowledge of Artificial Intelligence
- Good written and oral communication skills
- Strong motivation, with evidence of independent research skills relevant to the project
- Good time management