Project Details	
Project Code	NMH27Ex Dong
Title	AI scientists for understanding dementia modifiable risk factors
Research Theme	Neuroscience & Mental Health
Project Type	Dry lab
Summary	Early dementia prediction is an urgent problem to improve the clinical pathway for individuals who are at risk of dementia due to well-established risks such as cognitive decline, lifestyle, medical or genetic factors. This project aims to design an AI system that integrates scientific knowledge for early dementia modelling, based on large language models and causal inference. The project has two stages: first, to construct a causal knowledge graph from scientific papers and cognitive test questionnaire data, and second, to integrate the graph with transformer-based large language models and causal learning. This offers an explainable understanding of personalised risk factors.
Description	Early dementia prediction is an urgent problem to improve the clinical pathway for individuals who are at risk of dementia due to well-established risks such as cognitive decline, lifestyle, medical or genetic factors. Artificial Intelligence has the potential to offer a route to personalisation of risk profiling to detect early, mild cognitive impairment cases for their better care. There are two key challenges in early dementia modelling, which form the aims of this project: (1) the generative modelling of longitudinal, tabular, cognitive test data with external knowledge bases, and (2) the meaningful, scientific explanation of the prediction with dynamic causal inference and knowledge-based learning. The project will build on recent advances in transformer-based and generative AI modelling methods [1] for tabular data understanding, and extend them with linking to domain-specific knowledge bases [2] (e.g., constructed from the literature on cognitive tests and dementia risks). This multimodal approach is expected to enhance both prediction performance and interpretability, providing more robust insights into dementia progression and associated risk factors. The project will mainly work on the data from the PROTECT platform [3], containing a cohort of 30,000 older adults in community and primary care settings with ten-year longitudinal data. The explainability will be realised through the semantic, relational and causal connections within the data and to the prediction. A causal graph will be created through linking to the PROTECT dataset. This novel knowledge-based explanation, jointly modelled with causal inference [4] and large language models, will enable deep understanding of the longitudinal prediction and reveal biases in the model. The learned explainable knowledge may be validated by the knowledge base or enrich the knowledge base. Human-understandable explanations will finally be formed and evaluated [5]. A team led by medical experts in the supervisory team will assess the explanations for better model understanding and improvement. [1] Shmatko A, Jung AW, Gaurav K, Brunak S, Mortensen LH, Birney E, et al. Learning the natural history of human disease with generative transformers. Nature. 2025 Sept 17;1–9. doi: 10.1038/s41586-025-09529-3

	[2] Chen J, Dong H, Hastings J, Jiménez-Ruiz E, López V, Monnin P, Pesquita C, Škoda P, Tamma V. Knowledge Graphs for the Life Sciences: Recent Developments, Challenges and Opportunities. Transactions on Graph Data and Knowledge. 2023 Dec 19;1(1):5-1. doi: 10.4230/TGDK.1.1.5 [3] University of Exeter. PROTECT Study: Research to support healthy ageing and reduce dementia risk. https://www.exeter.ac.uk/research/dementia-research/research/protect/#a0. Accessed 28 Sep 2025. (Led by Prof Anne Corbett and the team) [4] Zhu Y, Zhang L, Sainsbury C, Dong F, MacLay J, Lowe DJ, Ye X. Causal counterfactual simulation for treatment decisions in multimodal lung disease data. Computers in Biology and Medicine. 2026 Aug 15;213:111807. [5] Longo L, Brcic M, Cabitza F, Choi J, Confalonieri R, Ser JD, et al. Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. Information Fusion. 2024 June 1;106:102301. doi: 10.1016/j.inffus.2024.102301
Can be completed part time?	Yes
Supervisory Team	
Lead Supervisor	
Name	Dr Hang Dong
Affiliation	Exeter
College/Faculty	Faculty of Environment, Science and Economy
Department/School	Department of Computer Science
Email Address	h.dong2@exeter.ac.uk
Co-Supervisor 1	
Name	Professor Xujiong Ye
Affiliation	Exeter
College/Faculty	Faculty of Environment, Science and Economy
Department/School	Department of Computer Science
Co-Supervisor 2	
Name	Professor Jonathan Huntley
Affiliation	Exeter
College/Faculty	Faculty of Health and Life Sciences
Department/School	University of Exeter Medical School
Co-Supervisor 3	
Name	Professor Anne Corbett
Affiliation	Exeter
College/Faculty	Faculty of Health and Life Sciences
Department/School	University of Exeter Medical School