

PhD Studentship in Big Data for MIDAS (Meaningful Integration of Data, Analytics and Services)

Applications are invited for the Department for the Economy (DfE) funded PhD studentship tenable in the Faculty of Computing & Engineering at the Magee Campus.

Project Summary:

Ulster University has announced that it will lead a major multi-million pound research project designed to harness the power of big data to better inform public policy and improve health and wellbeing outcomes across Europe. The Meaningful Integration of Data, Analytics and Services (MIDAS) project was awarded over 4.5 million euros in funding from the European Union's Horizon 2020 programme (<http://www.midasproject.eu/>).

The Ulster University led project will develop a pioneering digital platform for healthcare policy makers. This will allow these decision-makers to tap into unstructured and unconnected healthcare data to better inform policy, reduce costs and improve health and wellbeing of the population. The MIDAS platform will investigate connecting patient data from European health authorities with individual data collected from apps, sensors and social media. Complying with the highest standards of data protection and ethics, the data will be analysed on the MIDAS platform, which provides a tool for policy makers to benchmark, simulate and forecast outcomes of healthcare policy decisions.

Globally, populations are ageing; by 2050 it is estimated that more than 2 billion people will be aged over 60 years¹. This demographic shift has been accompanied by an increase in cognitive dysfunction which ranges from mild cognitive impairment to dementia². Certain dietary patterns or nutritional components are recognised to have a beneficial role, thus offering potential strategies to prevent or delay the onset of cognitive dysfunction in ageing^{4,5}. It has been estimated that a 5-year delay in the onset of cognitive dysfunction would reduce the population prevalence projections for dementia by 50%, which would have significant implications in terms health resources and for society generally, this PhD will begin by looking closely at this problem and the data. Small but effective dietary modifications could have major impacts on the quality of life of older people and their families. (Poly)phenols are one such category^{6,7}, but linking these to cognition and other health outcomes is problematic because of their transient appearance in plasma⁸. Procyanidins, abundant in foods such as tea, cocoa, grapes, nuts and berries, are of specific interest among (poly)phenols, and are almost exclusively metabolised by colonic bacteria to phenolic acids and valerolactones. In the VALID project (see, we will use plasma gamma-valerolactones⁹ which are stable rather than transient plasma biomarkers of procyanidin-rich foods and demonstrate their association with cognitive function and markers of metabolic health, in an ageing European population. The VALID project will validate gamma-valerolactones as novel biomarkers that are modulated by dietary (poly)phenols and associated with favourable cognitive and health outcomes in the ageing using Data Analytical methodologies. We will utilise data and perform new analysis on samples from the TUDA study, a unique resource designed to assess nutrition and ageing in 5,186 adults aged 60-102 years recruited from the island of Ireland. All participants have had their cognitive function assessed using 3 separate tests of cognition (Mini-Mental State Exam, Frontal Assessment Battery and the Repeatable Battery for the Assessment of Neuropsychological Status) and biomarkers of metabolic health status. Apart from the existing bio-banked TUDA (Trinity, Ulster, Department of Agriculture Study) study, we will access data

and samples from 'TUDA 5+', a follow-up study of 500 participants from the original cohort 5 years after their initial investigation, to determine the role of procyanidin-rich foods in preventing cognitive decline over a 5-year follow-up period. The TUDA and subsequent VALID datasets are potential large heterogeneous nutritional time series data sets, if connected could offer great opportunities for piloting and development of a Data Mining platform, virtual data connection layer, using data mining CRISP methodology^{12, 13, 14} to identify and extract correlations and patterns between the (poly)phenols and cognitive function and markers of metabolic health.

Consequently, this PhD involves novel solutions to aggregating heterogeneous time series datasets into a comprehensive platform, using data from the existing TUDA cohorts and data to be generated as part of the VALID project. The platform will provide effective time series interrogation, analysis of patterns and drift along with visualisation of results allowing the discovery of patterns and trends that are of pertinent to public health analysis. It will involve using APIs, new approaches to heterogeneous data management and integration, new approaches to data mining and using effective data visualisation.

Entrance Requirements:

Candidates should have ordinary UK residence to be eligible for both fees and maintenance. Non UK residents who hold ordinary EU residence may also apply but if successful will receive fees only. All applicants should hold a first or upper second class honours degree in Computer Science or a cognate area. Applications will be considered on a competitive basis with regard to the candidate's qualifications, skills experience and interests. Successful candidates will enrol as of April 2017, on a full-time programme of research studies leading to the award of the degree of Doctor of Philosophy.

The studentship will comprise fees together with an annual stipend of £14,296 and will be awarded for a period of up to three years subject to satisfactory progress.

If you wish to discuss your proposal or receive advice on this project, please contact:

Dr Michaela Black (mm.black@ulster.ac.uk)

Procedure

For more information on applying go to ulster.ac.uk/research
Apply online ulster.ac.uk/applyonline

The closing date for receipt of completed applications is 24th February 2017

Interviews will be held in March 2017

Additional Info:

References: 1) United Nations. World Population Ageing. (2013). 2) World Health Organisation, Mental and behavioural disorders (2015). 3) World Health Organisation. Dementia: A public health priority. (2012). 4) McGarel et al., Proc Nutr Soc. 74:46-55 (2015). 5) Lamport et al., Nutr Rev. 72:774-89 (2014). 6) Canevelli et al., Nutrients. 4:8 doi:10.3390 (2016). 7) Kennedy et al., Adv Nutr. 5:515-33 (2014). 8) Del Rio et al., Antioxid.Redox.Signal., 18:1818-1892 (2013). 9) Curti et al., Adv.Synth.Catal.,357:4082 (2015). 10) Goodrich et al., J. Chromatogr.B, 958:63-74 (2014).11) Barrett, M.A., et al., Big data and disease prevention: From quantified self to quantified communities. Big Data, 2013. 1(3): p. 168-175.12) <http://spark.apache.org/docs/latest/ml-guide.html>.13) <http://thunder-project.org/14>) <http://www.sv-europe.com/crisp-dm-methodology/>