A big data framework to facilitate modelling efforts against COVID19

Dr Minh Kieu, University of Auckland, minh.kieu@auckland.ac.nz
Prof Alison Heppenstall, University of Leeds. a.j.heppenstall@leeds.ac.uk

Project Abstract:

Individual-based modelling, for instance contact-tracing through a mobile application is needed for an effective modelling in efforts against COVID19, and to recover from the negative impacts of the virus outbreak. However, governments over the globe have a rising concern over privacy and exchange of personal data. Many researchers, while working with individual data, also do not fully comprehend the implications of regulations such as the EU's GDPR and the US' Privacy Shield Framework in protecting individuals' rights to privacy and enhance data protection.

This project will develop a transparent, modularised and secured data sharing platform that aims to work with multiple legal requirements regarding individual data, and adaptively develop a code of conduct that will connect data owners and researchers who want to use individual data. The idea is to develop a data platform and a set of data synthesis modelling algorithms in order to de-identify the individual data submitted from data owners into synthetic, or simulated data that retain the distributions in the population, but each individual data points do not represent a real person. In other words, we preserve the ability to run queries on the encrypted data. Analysts can ask questions that link together personal data, and access data that are very similar to the real data but they only ever see anonymised or aggregated results. All the processing procedures are recorded and auditable by a third regulators or courts. The algorithms are also modularised and adaptive, e.g. the level of anonymisation can be dependent on the risks of re-identification. We will also develop a code of conduct to guide data owners and users to ethically and effectively use individual data, while nurturing the connection between them.

This project has to start now, because the access to individual data is the main hurdle to many COVID19-related research in epidemiology, public health, social simulation and engineering. This data platform is not a solution to all existing privacy questions (like consent, but it will be an infrastructure that allows researchers to access individual data and policy makers to access cutting-edge models for their data, while minimising a number of critical risks related to personally identifiable information and maintaining appropriate regulations across multiple legal contexts.

About the PIs

Minh Kieu is a Lecturer in Intelligent Transport System at the Department of Civil and Environmental Engineering, University of Auckland. Minh is an expert in big data analytics and computer simulation, especially individual-based simulation such as agent-based modelling.He has been contributing to the development and delivery of large research projects, such as an ERC Horizon 2020 in the UK (https://dust.leeds.ac.uk/), and the Premier Innovation Initiative in Australia (http://adait.io/inno-pii.html).

Alison Heppenstall is a Professor in Geocomputation at the School of Geography and Leeds Institute of Data Analytics (LIDA) at the University of Leeds, and an ESRC-Alan

Turing Fellow at the Alan Turing Institute, UK. Alison has extensive experience with large research projects that require a consortium between multiple universities, such as Systems Science in Public Health and Health Economics Research (SIPHER); Behavioural, ecological and socio-economic tools for modelling agricultural policy (BESTMAP); Consumer Data Research Centre (CDRC); and especially Quantifying Utility and Preserving Privacy in Synthetic Data (QUIPP) – a joint project with the Alan Turing Institute that is aims to generate synthetic versions of sensitive data sets that contain all the relationships and preserve individual privacy.

People who have showed their interests:
Prof Mark Gahegan, University of Auckland, New Zealand
Prof David O'Sullivan, Victoria University of Wellington, New Zealand
Prof Nicolas Malleson, University of Leeds, UK
Dr Koen van Dam, Imperial College London, UK
Dr Jason Thomson, University of Melbourne, Australia

Interested researchers should contact Minh & Alison directly, cc'ing Debs as the WUN Coordinator.

Minh Kieu, minh.kieu@auckland.ac.nz
Alison Heppenstall, a.j.heppenstall@leeds.ac.uk

cc: Deborah McAllister (WUN Coordinator, Auckland) d.mcallister@auckland.ac.nz