



## **Kazakhstan's Environment-Health system, a Big Data challenge**

Claudia Vitolo (1,2,3), Bella Bella Gazdiyeva (3,4), Allan Tucker (2), Andrew Russell (1), Maged Ali (3,5), and Abraham Althonayan (3)

(1) Institute of Environment Health and Societies, Brunel University London, Uxbridge, United Kingdom, (2) Department of Computer Science, Brunel University London, Uxbridge, United Kingdom, (3) Business School, Brunel University London, Uxbridge, United Kingdom, (4) Shokan Ualikhanov Kokshetau State University, Astana, Kazakhstan, (5) Essex Business School, University of Essex, Colchester, United Kingdom

Kazakhstan has witnessed a remarkable economic development in the past 15 years, becoming an upper-middle-income country. However it is still widely regarded as a developing nation, partially because of its population's low life expectancy which is 5 years below the average in similar economies. The environment is in a rather fragile state, affected by soil, water, air pollution, radioactive contamination and climate change. However, Kazakhstan's government is moving towards clean energy and environmental protection and calling on scientists to help prioritise investments.

The British Council-funded "Kazakhstan's Environment-Health Risk Analysis (KEHRA)" project is one of the recently launched initiatives to support Kazakhstan healthier future. The underlying hypothesis of this research is that the above mentioned factors (air/water/soil pollution, etc.) affecting public health almost certainly do not act independently but rather trigger and exacerbate each other.

Exploring the environment-health links in a multi-dimensional framework is a typical Big Data problem, in which the volume and variety of the data needed poses technical as well as scientific challenges. In Kazakhstan, the complexities related to managing and analysing Big Data are worsened by a number of obstacles at the data acquisition step: most of the data is not in digital form, spatial and temporal attributes are often ambiguous and the re-use and re-purpose of the information is subject to restrictive licenses and other mechanisms of control.

In this work, we document the first steps taken towards building an understanding of the complex environment-health system in Kazakhstan, using interactive visualisation tools to identify and compare hot-spots of pollution and poor health outcomes, Big Data and web technologies to collect, manage and explore available information. In the future, the knowledge acquired will be modelled to develop evidence-based recommendation systems for decision makers in Kazakhstan.