

# Abstract

## The Ethical Dangers of Public Data

Harry Thompson

University of Dublin, Trinity College, 2022

Supervisor: Prof. Declan O'Sullivan

Governments around the world are increasingly making available datasets on a variety of topics affecting their citizens. Throughout the Covid-19 pandemic governments have released case numbers and vaccination records. This paper explores Linking public datasets together, and investigating how Linked Open Data could potentially affect people's lives. An open data linking application was developed to form a concrete use case to answer & explore the research question: *"To what extent can public datasets be linked together in what could be considered an ethically questionable matter ?"*. To investigate the Potential ethical issues that could arise with public data. Three datasets were linked in the areas of: Housing prices, Social demographics, & Vaccination statistics. These datasets were linked together using freely available software to: Convert CSV data to an RDF triplestore, using an R2RML mapping based on a bespoke ontology. The RDF triples from the mapping were utilized to create a knowledge graph in GraphDB, and a set of competence questions in SPARQL were used to explore the connections between the datasets with a User Interface based in python. The links between the datasets were used to explore data ethics & public data, their uses and potential effects on wider society. It is clear that as society progresses to be increasingly data driven, we need to be aware of data ethics, and its role in our society.