Introduction

Project Motivation
Growth of the Web of Data...
Current Estimate:
• 20 billion triples.
• 150 - 220 million interlinks.

Quality of Interlinks
Interlinks can be:
• set manually;
• generated by use of interlinking algorithms.
Most interlinks are generated automatically as data sets may contain millions of triples.
Automatic interlinks are considered ‘best effort’ matches.

User Contributed Interlinking (UCI)
An approach to improving the quality of interlinks through the inclusion of a human in the interlinking process.

Project Objective
To ascertain if the quality of Linked Data interlinks can be improved by cognitive tool support for user contributed interlinking processes.

Evaluation and Metrics

User Trial
Eight user trial participants were asked to perform interlinking tasks in relation to three given URIs. For each URI, participants were asked to:
• dereference the URI;
• edit or remove any existing inaccurate interlinks;
• compare the URI to 9 other URIs;
• create new interlinks if they thought the URIs could be linked using the predicates owl#sameAs, owl#differentFrom, dbpedia#disambiguates, rdfs:seeAlso or alternatively by adding another predicate.

Accuracy of Interlinks
The interlinks contributed by the participants were compared against a gold standard, which is a data set verified to have high quality interlinks by a linked data expert. Statistical classifications Precision and Recall were used to measure fidelity and completeness respectively.

Usability of the Demonstrator
Participants also completed a questionnaire in relation to usability of the demonstrator using the System Usability scale (SUS).

Participant Linked Data Experience
Participants were also asked how familiar they were with Linked Data
• On a scale of 1-6, the majority of participants rated themselves 4 or above with regard to Linked Data familiarity.
• Participants claimed to have levels of Linked Data experience ranging from none up to 4 years.

Results

‘Improv-A-Link’ Demonstrator Architecture

The aim of Cognitive Support for UCI processes is to provide software support that simplifies performing an interlinking task for the user.
The ‘Improv-A-Link’ demonstrator is designed to offer cognitive support to users through the inclusion of the following features:
• A clear user interface
• Decision support (URI hyperlinks launch browser)
• Access to metadata
• Filtering to display interlinks

M.Sc. in Computer Science
(Networks and Distributed Systems)

Future work
• Publish User Contributed Interlinks back to the Linked Open Data cloud.
• Add functionality to keep data cached from Sindice up to date whilst maintaining improvements to interlink quality achieved through UCI Processes.