

Student Name: Sarah Alzahrani

Supervisor Name: Prof. Declan O'Sullivan

Dissertation title: Towards Autonomic Uplift of Information

MSc in Computer Science (Networks and Distributed Systems)

2015

Abstract

In the past few years, there has been a considerable amount of growth in semantic web adoption. A large amount of the information available on the web has been modelled using various types of data model, each having a different form of standardization, format and rules governing the data. Thus, users are increasingly choosing to uplift data into Resource Description Framework (RDF) format, which is a standard model for data interchange on the web. However, because data schemas and ontologies that are used to describe data in RDF may change frequently over time, the need for an approach to represent the mapping statements in an analyzable form has arisen.

Most current research effort in this area is focused on demonstrating generic solutions for mapping, or translating data that exists in different formats from the standard data model. However, not much attention has been given to the representation of the mapping statements in the previous/current solutions. In this study, focus is placed on achieving the mapping process in an autonomic fashion. In other words, representing query-based mapping (XQuery queries) using RDF (Resource Description Framework), which appears (after significant investigation) to be a novel idea in the area. Therefore, this representation will enable the manipulation and analysis of mapping statements.

Two techniques are proposed to perform the mapping from XQuery to RDF, which are *manual* and *automatic* mapping approaches. The two approaches are based on one solution path that begins with the extraction of XML documents using XQuery to uplift data into the standard RDF data model. Then, the XQuery that is used for the mapping process itself is transformed into RDF to allow for further reasoning and analysis of the query. Finally, an evaluation approach was formulated in order to assess the two mapping approaches.