

# Publishing Data of Ferdowsi University of Mashhad as Linked Data

Samad Paydar<sup>1</sup>, Mohsen Kahani<sup>1</sup>, Behshid Behkamal<sup>1</sup>, Mahboobeh Dadkhah<sup>1</sup>, Elaheh Sekhavaty<sup>2</sup>

<sup>1</sup>Web Technology Lab., Dept. of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

<sup>2</sup>Payam Noor University of Tehran, Tehran, Iran

samad.paydar@stu-mail.um.ac.ir, kahani@um.ac.ir, behkamal@stu-mail.um.ac.ir, mb.dadkhah@stu-mail.um.ac.ir, sekhavaty@tehran.pnu.ac.ir

**Abstract**—In this paper, the process of publishing a subset of data of Ferdowsi University of Mashhad as Linked Data is discussed. A framework developed for this reason is presented. Different parts of the framework are discussed and the linking procedure is described. Analysis of the experimental results and a discussion of problems and challenges are also presented.

**Keywords**- *Linked Data; RDF; Semantic Web; University*

## I. INTRODUCTION

Linked Data principles, first outlined by Tim Berners-Lee in 2006 [2], provide simple guidelines for publishing, interlinking, and accessing structured data in a uniform machine-understandable format. Linked Data is based on well-known Web technologies like URI and RDF, and enables creating the Web of data, the data infrastructure required for realization of the Semantic Web vision [3]. Linking Open Data (LOD)<sup>1</sup> project is the most important project initiated by Linked Data community to bootstrap the Web of Data by publishing well-known web datasets as Linked Data and creating a huge Linked Data cloud.

One of the current themes in Linked Data projects is to publish a dataset based on Linked Data best practices [14], adding it to the LOD cloud.

In this paper, experience of publishing the data of Ferdowsi University of Mashhad (FUM) as Linked Data is discussed. Throughout this paper, we use “FUM-LD” term to refer to the Linked Data version of FUM data. The paper is structured as follows: In section 2, an overview of some similar works is presented. Section 3 discusses some motivations of the current work. Section 4 is dedicated to describing the publishing approach used in this project. Section 5 presents results of the project. Some challenges and problems are discussed in section 6. Finally, section 7 is conclusion of the paper.

## II. RELATED WORK

There are some works in the literature, dedicated to discussing experiences of publishing Linked Data datasets. Here, we shortly discuss some of these works.

In [6] a work on publishing the Eurostat data, containing statistical data about the European Union, is discussed. For this purpose the *riese*<sup>2</sup> system is developed which comprises a module for converting the Eurostat data into RDF and creating the interlinks, and a Web server including a scripting environment that fills predefined templates with the values from the (static) RDF/XML representation in order to generate an RDFa representation of the themes and the data tables.

In [8], the applicability and potential benefits of using Linked Data to connect drug and clinical trials data sources and interlinking them with existing Linked Data sources is discussed.

LinkedMDB [7] deals with publishing data of major movie datasets (e.g. IMDb, OMDb) as linked data. LinkedMDB includes links to several external datasets, like FreeBase, DBpedia, RottenTomatoes, YAGO, Geonames and lingvoj.

Another work [9] describes the conversion of an economics thesaurus to RDF/SKOS.

To the best of our knowledge, there is no work in the literature focused on publishing university data as linked data. Also there is too much little data represented in Farsi language on the LOD cloud (e.g. there is little Farsi data on DBpedia). There is no work regarding publication of Farsi datasets as Linked Data.

## III. MOTIVATIONS

Data of universities and their activities is important to many web users like students, researchers and teachers. Such data, if published as Linked Data and linked to appropriate datasets (e.g. general datasets like DBpedia, or special datasets like DBLP or ACM), provides valuable benefits through enabling different scenarios in answering users' information need. For

<sup>1</sup>

<http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>

<sup>2</sup> <http://riese.joanneum.at>