Semantic-based data integration and exchange for a research group

Erick García1,2, Natalia Villanueva-Rosales1,2,*

1Department of Computer Science, 2Cyber-SHARE Center of Excellence, *Faculty Mentor
eugarcia87@miners.utep.edu, nvillanuevarosales@utep.edu
The University of Texas at El Paso

Abstract

Disseminating products resulting from research efforts for validation, reproducibility, and reuse is an important task of a research group. Sharing resources using natural languages, like English, are good for humans but hard for computers to process. Sharing resources using machine languages, like JSON, are great for machines but hard for humans to understand. Having a one-stop place where resources of research groups can be dynamically populated from other sources and share information for both, humans and machines, would save time and also increase reusability of their resources. In this paper, we describe our approach to disseminate the products of the iLink group through a web-based application designed for both humans and machines.

Background

• The Semantic Web is a set of standard languages and tools promoted by the World Wide Web Consortium for exposing machine understandable data to be shared and reused. [1]
• Web-Services are applications that enable the communication between two electronic devices over a network. [1]
• Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. It facilitates discovery of relevant information. [2]

Research Question

Can semantic, web-based approaches, e.g., ontologies, enable the creation of systems that automatically integrate information about a research group and share information for both humans and machines?

Objective

Design a web-based application for the iLink research group that can be:
• dynamically populated from heterogeneous sources,
• mobile-friendly for human consumption,
• use web standards for interoperability of other systems, and
• share resources using semantic annotations (meta-data) for machine processing. (Fig. 1)

Methodology

1. Design of a high-level data model (E/R diagram) that covers information retrieved from other sources and metadata
2. Evaluation for frameworks for web development that supports the creation of responsive mobile-, human-friendly website and the use of standards for interoperability
3. Development of a front- and back-end of a research group website through the use latest web-based technologies.
4. Comparison of metadata standards and vocabulary to share research group’s resources with semantic annotations.
5. Semantic Annotation of all the information shared through the website.

Results

• The iLink research group’s dynamic website was created. This website is:
  i) dynamically populated from other sources, ii) based on Model-View-Controller architecture, iii) based on standard web-languages and technologies including Bootstrap, Code Igniter, HTML5 and JSON.
• The information shared on the Link’s website is semantically annotated which means it is fully understandable for both humans and machines. (Fig. 2 & 3)

Conclusions

Evaluation of the system based on:
• Discoverability of resources by machines using metadata.
• Usability for human users.
• The ability of search engines understanding the websites data so there will be more accurate results to what the user is searching.
• The semantic annotation was done in RDFa instead of Microformats. or Microdata has the advantage of being able to create your own vocabulary if it does not already exist.

References

[4] Link Website: http://iLink.cybershare.utep.edu/