



Semantic-based data integration and exchange for a research group

Erick Garcia^{1,2}, Supervised by: Dr. Natalia Villanueva Rosales^{1,2}

¹Department of Computer Science, ²Cyber-ShARE Center of Excellence

egarcia87@miners.utep.edu, nvillanuevarosales@utep.edu

The University of Texas at El Paso

Motivation

- Maintaining information about research group efforts is a tedious manual task.
- 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 (Fig. 1).
- 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.

Methodology

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

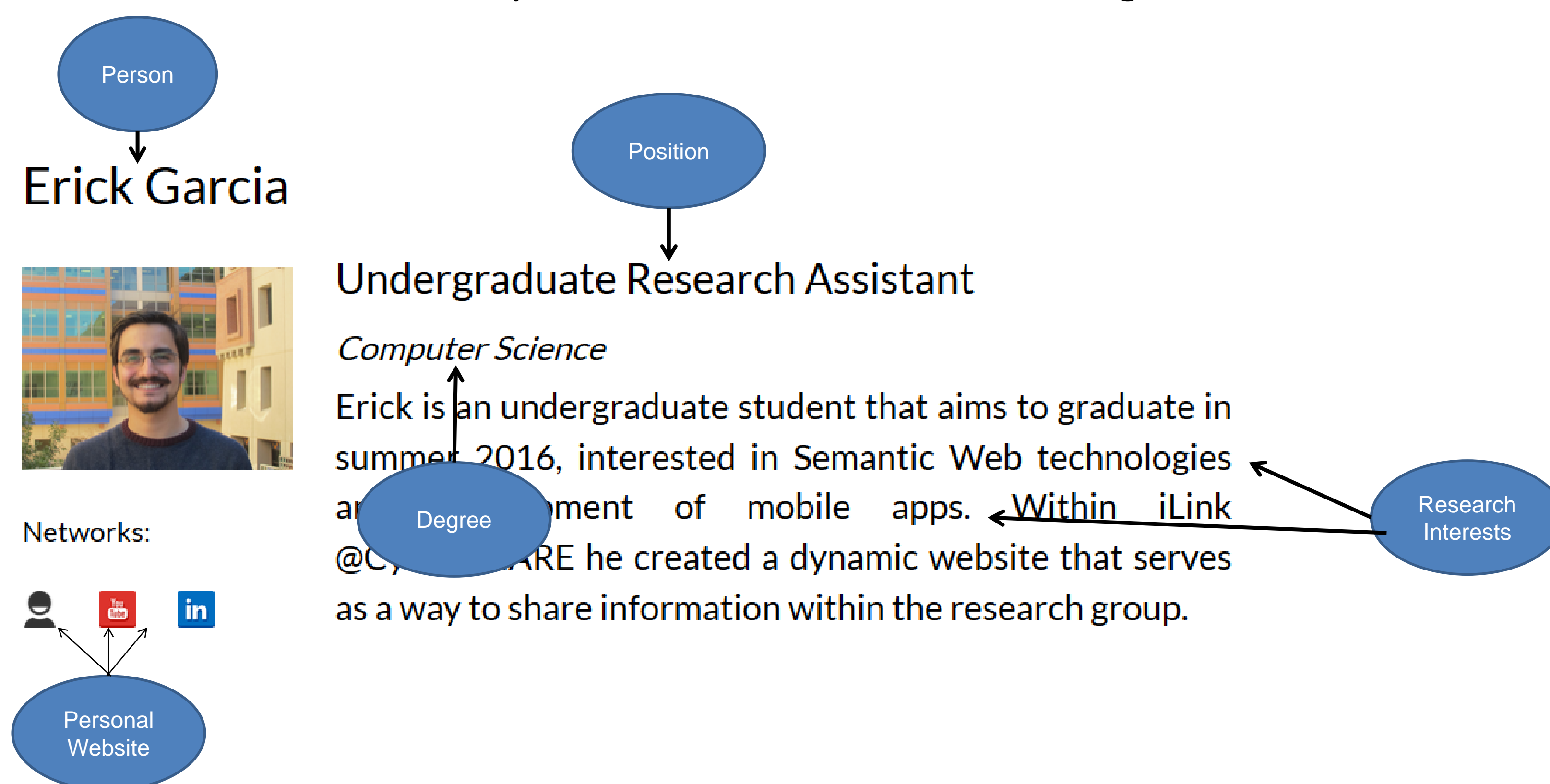


Figure 1. Differences between data that can be understood by humans vs. data that can be understood by machines.

6. Evaluate the web-based system.

Future Work

Evaluation of the system based on:

- Discoverability of resources by machines using metadata.
- Usability for human users.

References

- [1] Understanding Metadata (pp. 4-10). (2004). Bethesda, NISO Press National Information Standards Organization
- [2] Semantic Web, Retrieved April 15, 2015 from <http://www.w3.org/standards/semanticweb/data>
- [3] FOAF, Retrieved April 15, 2015 from <http://www.foaf-project.org/>.
- [4] iLink Website: <http://iLink.cybershare.utep.edu/>

Background

- The **Semantic Web** is a standard promoted by the World Wide Web Consortium which aims at converting the World Wide Web from the current model largely unstructured documents and data into a common machine understandable framework that allows data to be shared and reused. [2]
- **Web-Services** are client and server applications that enable the communication between two electronic devices over a network. [2]
- **Metadata** is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data. It is really important because it facilitates discovery of relevant information. [1]

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 :

- i) dynamically populated from heterogeneous sources,
- ii) mobile-friendly for human consumption,
- iii) use web standards for interoperability of other systems, and
- iv) share resources using semantic annotations (meta-data) for machine processing.

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 iLink's website is prepared for semantic annotations to be fully understandable for both humans and machines. We are currently identifying controlled vocabularies to describe the information on the website, including Friend of a Friend[3] and schema.org

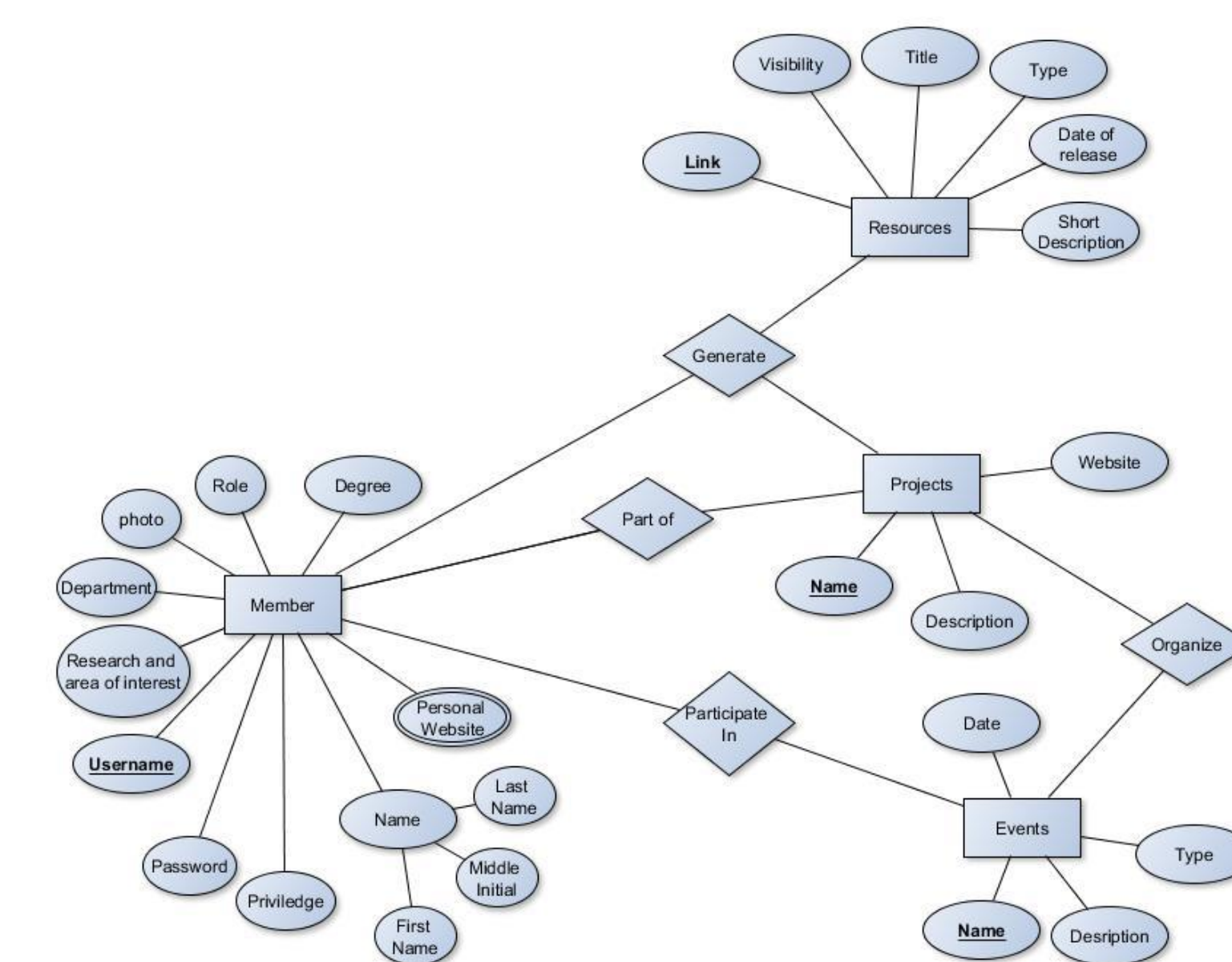


Figure 2. Entity Relationship diagram representing iLink's information