Visualizing Ontologies – A Literature Survey

Arash Saghafi

Overview

- **21 papers were gathered – a mix of design study and method papers**
  - Graph-based Approaches
  - Multi-Method Visualization Techniques

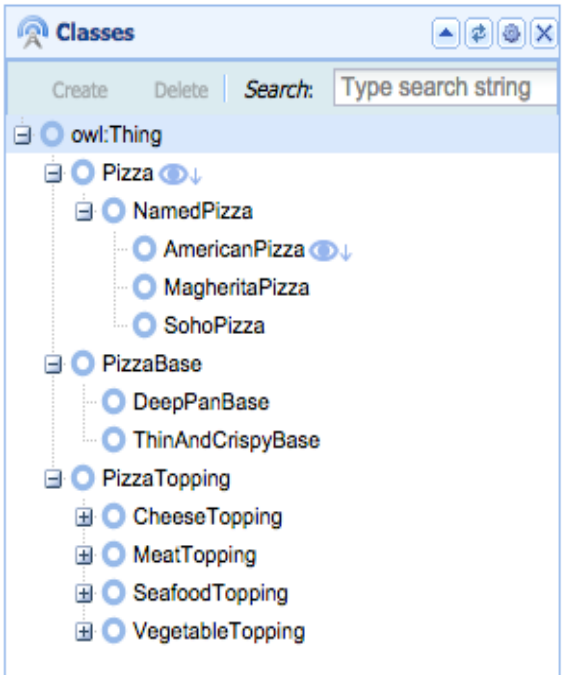
- **Four empirical studies were similar enough to justify a meta-analysis**
  - Accuracy, completion time, and user satisfaction were dependent variables in the analysis.

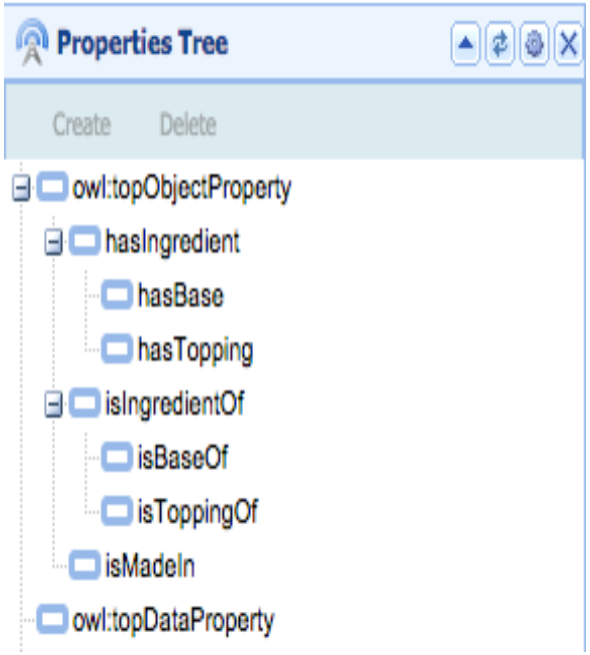
Background

- **Ontology is a "branch of philosophy which deals with the order and structure of reality in the broadest sense possible" (Angeles 1981).**
  - Domain ontology: a set of concepts, their relationship, what can happen, what can exist (Wand and Weber 2002).
- **Ontologies are used in diverse fields such as biomedical information, systems engineering, and semantic web to represent domain knowledge (or the semantic meta-data).**
  - Visualization of ontologies can improve tasks such as understanding of implicit knowledge and information alignment.
- **Katifori et al. (2007), Lanzenberger et al. (2010), and Granitzer et al. (2010) had also surveyed the literature for ontology visualization methods.**
- **The current work gathers 21 papers, from 2003 to 2014**
  - Mostly published after 2011
  - Specifically looked at papers that were not included in previous surveys.

Ontology Language and The Baseline Representation

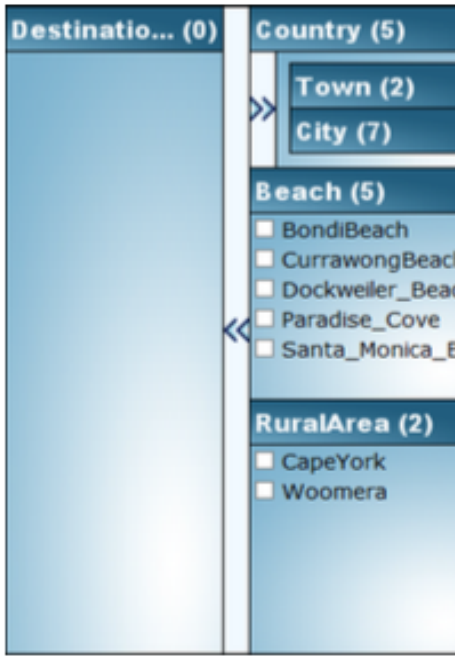
- Standard language in ontological engineering is OWL (Web Ontology Language – supported by W3C).
- Protégé is the most widely used tool for editing ontologies.
- The visualization method used in Protégé is indented lists.

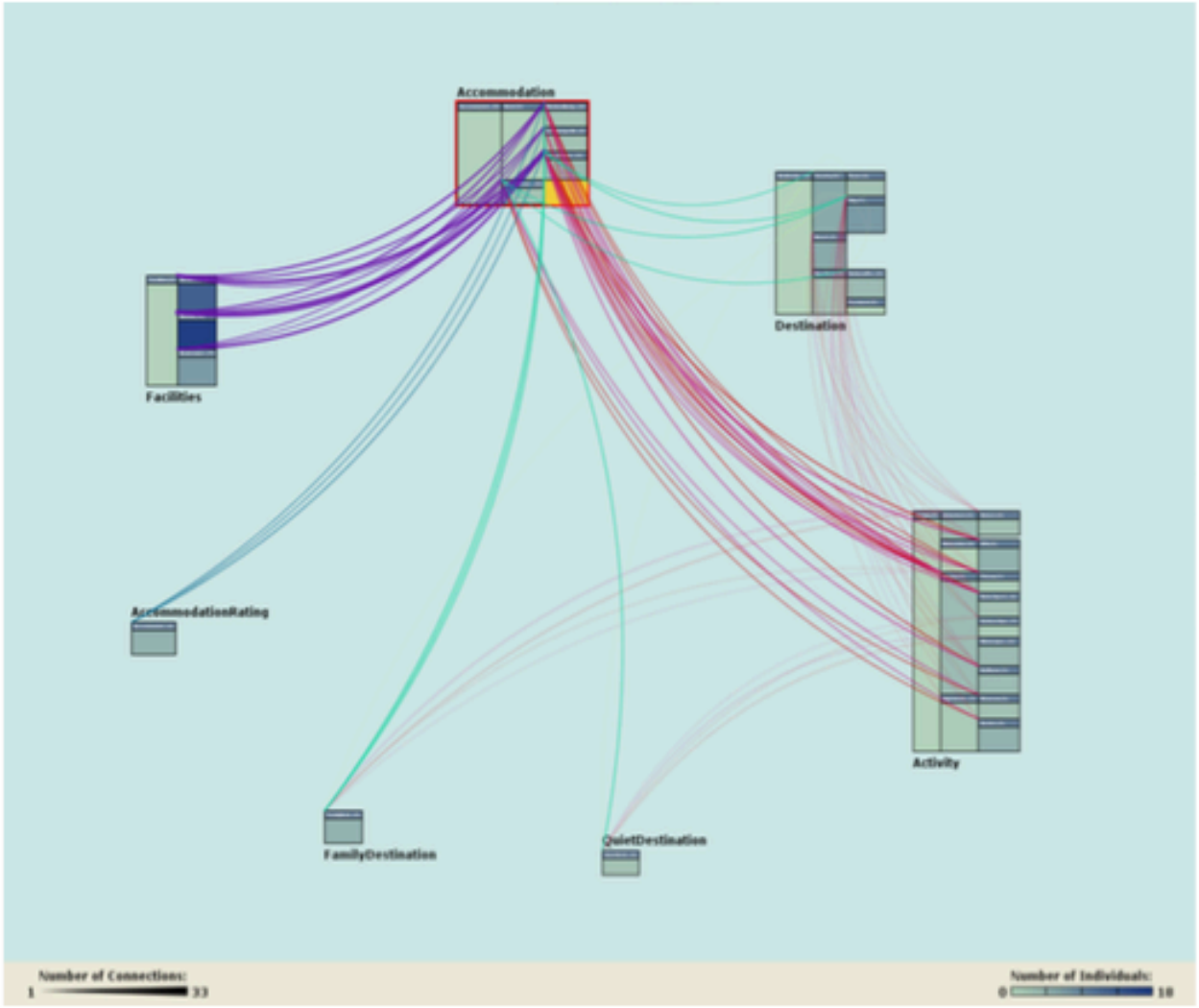




Graph-based Visaulizations

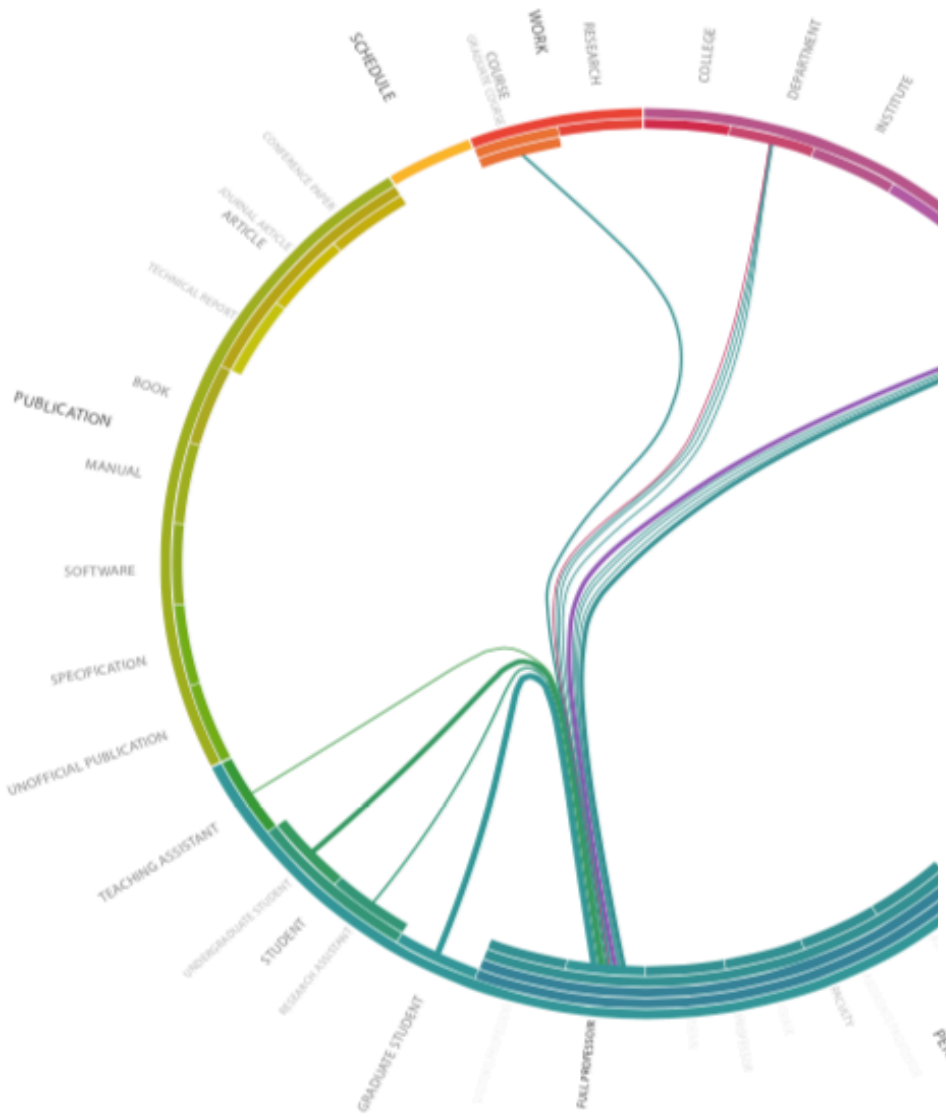
Idiom	Knoocks
What: Data	Ontolc
How: Encode	Class Links r
Why: Task	Under
Scale	Limite

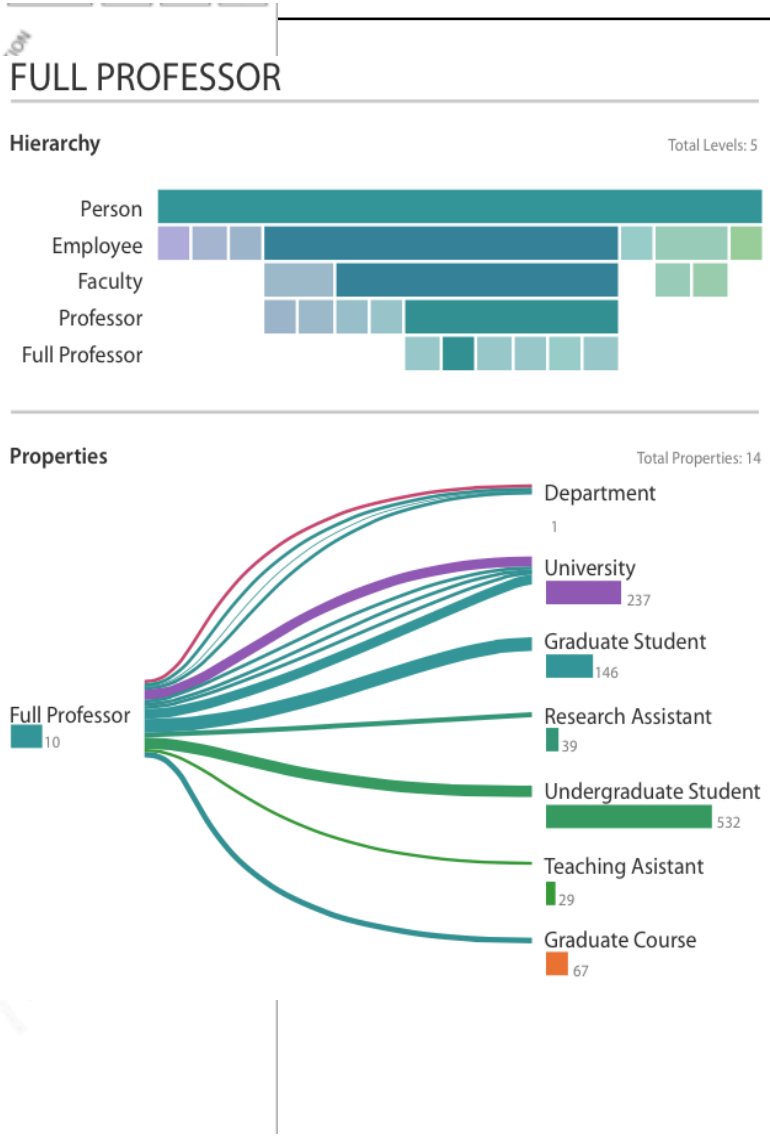




Multi-Method Visualizations

Idiom | Multi-view Ontology Visualization





Meta-analysis

- **Four studies were similar enough to justify a random effects meta-analysis**
  - Variation is attributed to sampling error, as well as Random Effects Variable (e.g. different DVs, or the experimental methods).
- **Converted reported statistics to Cohen's d**

Reference	Independent Variable	Dependent Variable	Sample Size	Effect Size (Unbiased Cohen's d)
Katifori et al. 2006	TGViz (Node-link) vs. Protégé	Accuracy	23	-0.49
Motta et al. 2011	KCViz (Node-link) vs. Protégé	Completion Time	21	1.01
		Usability Score	21	0.26
Garcia-Penalvo et al. 2012	OWL-VisMod (tree + radial layout) vs. Protégé	Accuracy	21	-0.14
		Usability Score	21	0.79
Fu et al. 2014	Node-link vs. Protégé	Completion Time	36	-1.32
		Accuracy	36	0

Discussion of Results

- **The average unbiased Cohen's d of this analysis is 0.014, with the 95% confidence interval of -1.21 to 1.22.**
  - Very weak (non-existent) effect.

- **Grouping based on Dependent Variables**

Dependent Variable	Average Unbiased Cohen's d	No. of Reported Effect Sizes	95% Credibility Interval
Accuracy	-0.21	3	-0.45 to 0.12
Completion Time	-0.16	2	-2.90 to 2.56
Usability Score	0.53	2	0.20 to 0.84

Summary and Conclusion

- **New visualizations are appealing to users because of their novelty.**

- **Indented lists are still advantageous**
  - Either due to the inherent hierarchical nature of ontological data, or prior familiarity of users.

- **Call for new visualization methods for ontologies that also improves objective performance measures such as accuracy of answers and task completion time.**

QUESTIONS?