

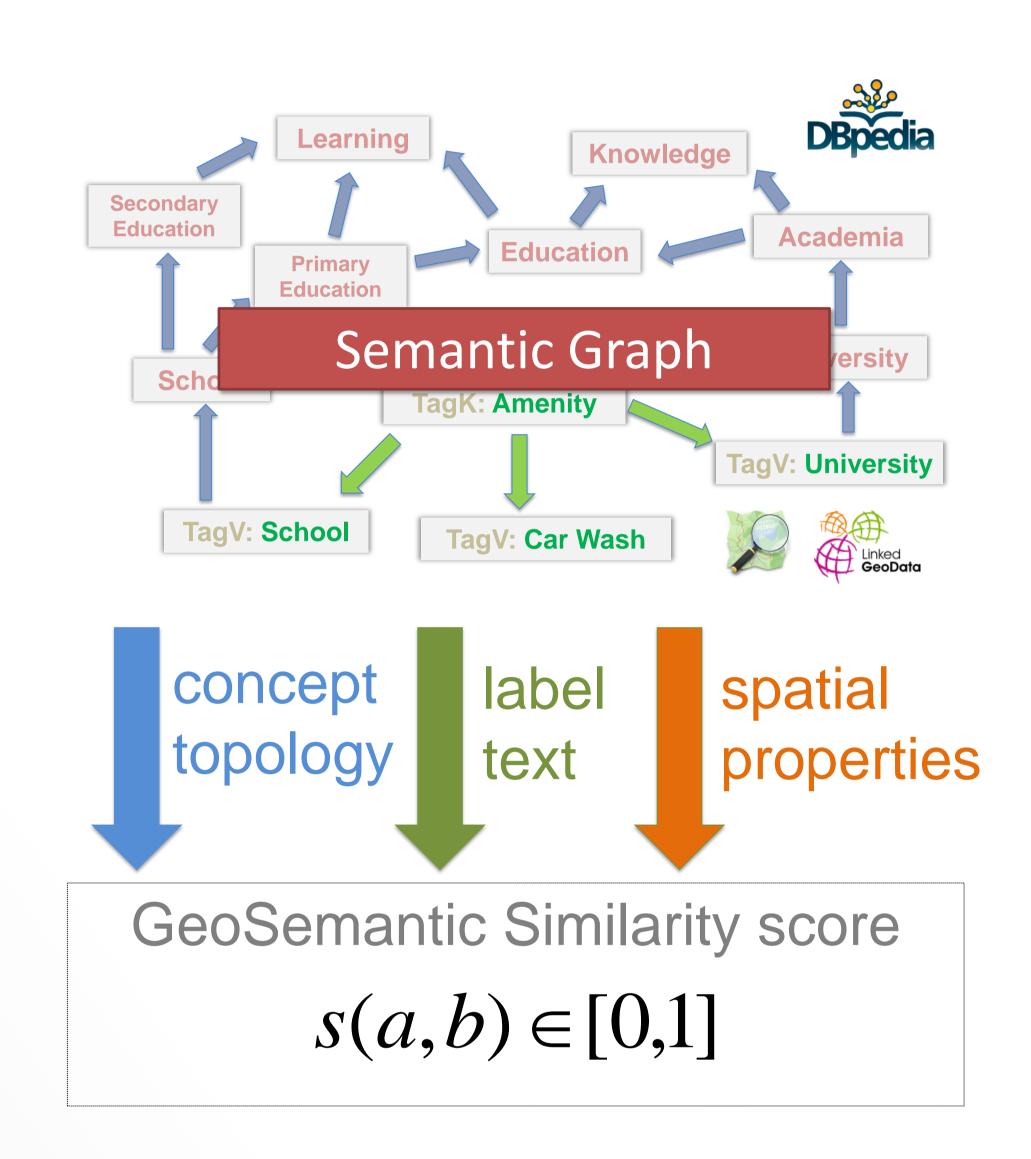


Mining Distributed Ontologies for a Semantic Similarity Measure in the Geo Web

Andrea Ballatore, PhD Student, andrea.ballatore@ucd.ie
Michela Bertolotto, PI, michela.bertolotto@ucd.ie
School of Computer Science and Informatics, UCD

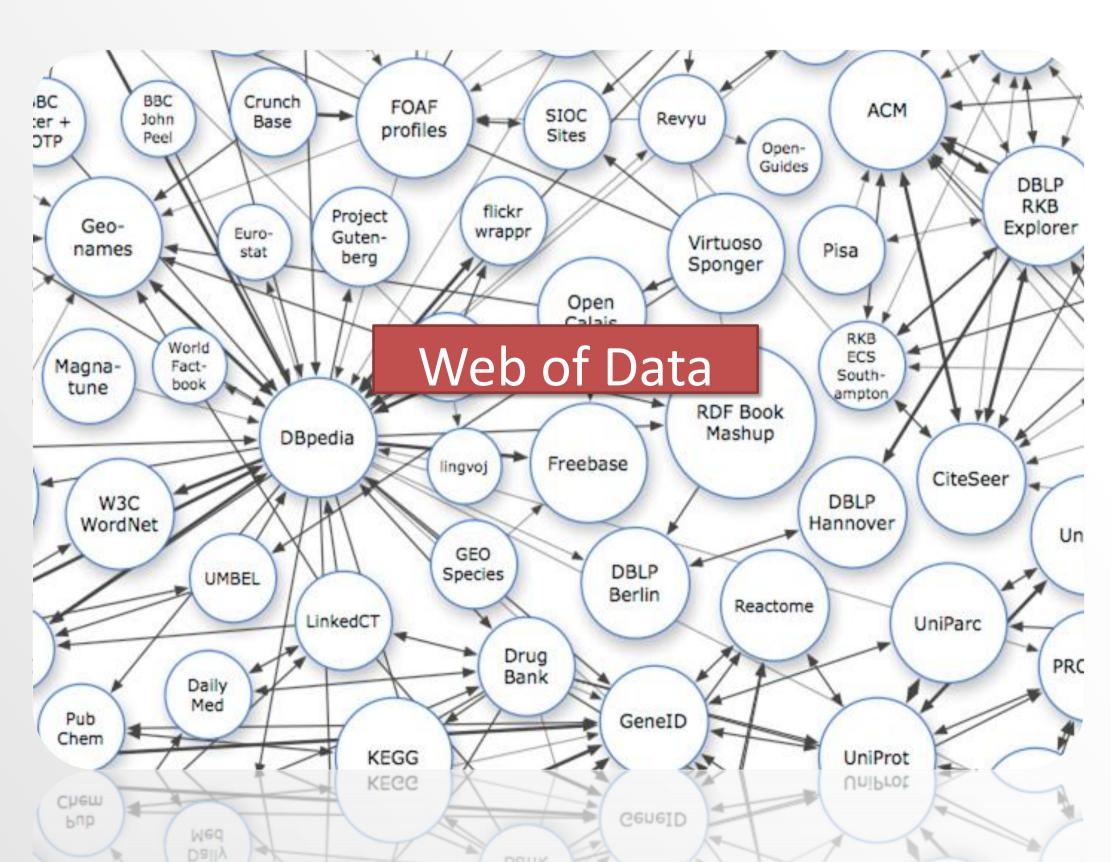
Motivation

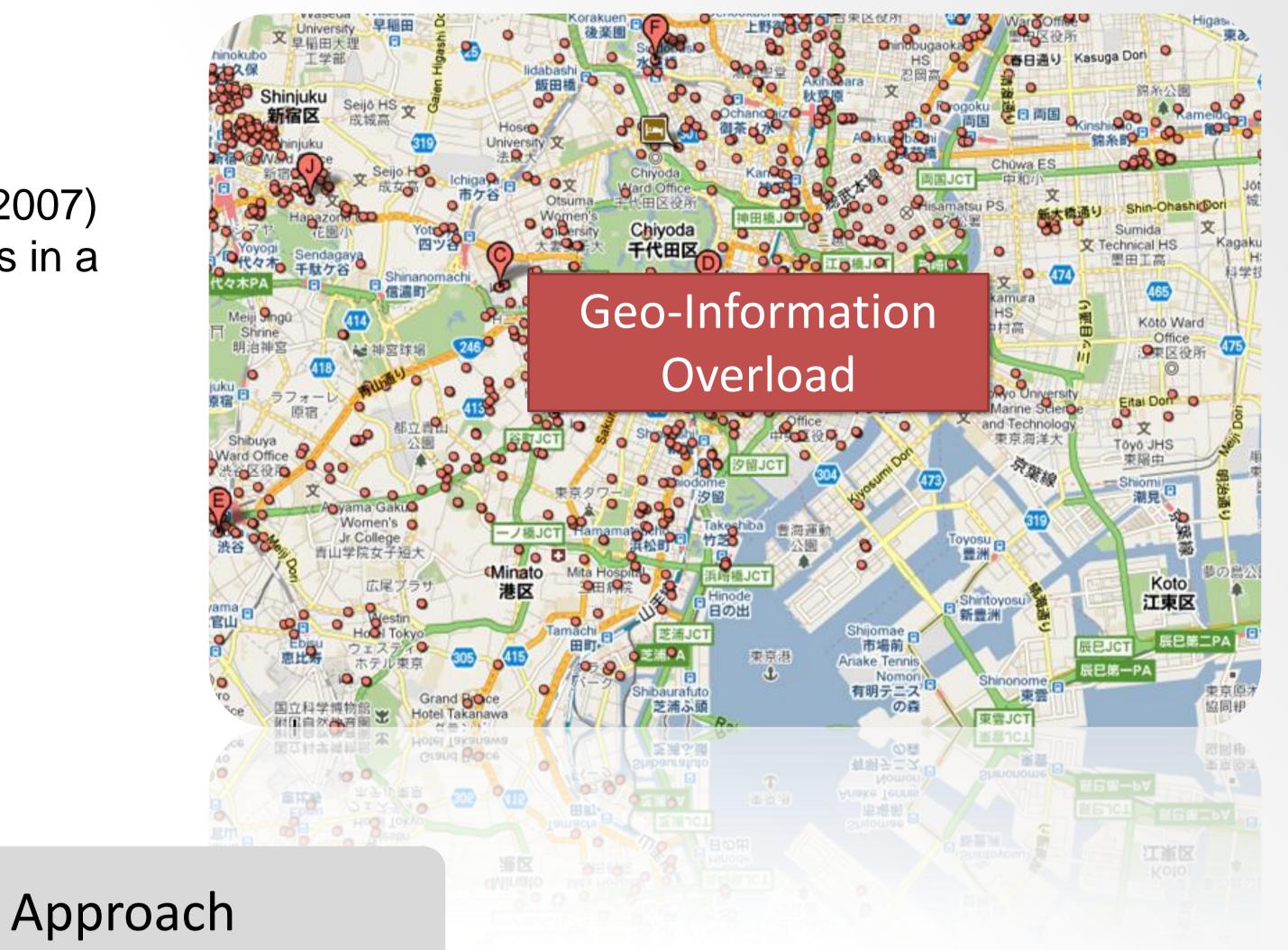
- Semantic Geo Web (Egenhofer, 2002)
- Volunteered Geographic Information (Goodchild, 2007)
- Emergent semantics: concepts are defined by users in a bottom-up process (Mika, 2005)
- Reduce Geo-Information Overload



Applications

- GeoData mining and clustering
- Semantic Geographic Information Retrieval (GIR)
- Semantic Map Personalisation
- Data Integration and Ontology alignment
- Geo-Data Discovery

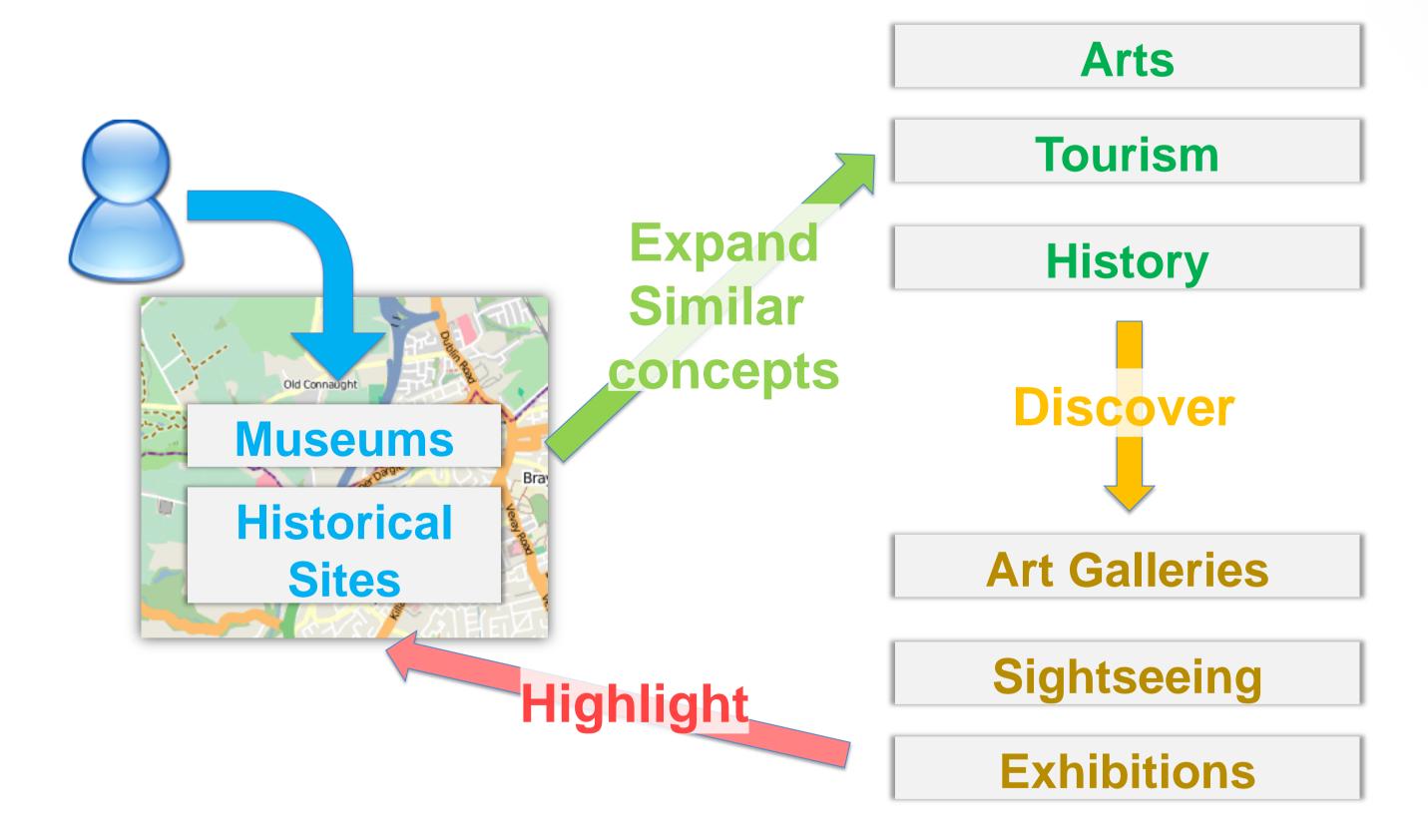




A weighted combination of:

- 1. Network Similarity: Co-Citation algorithms on edge structure.

 Topological position of the geographic concept in the semantic graph
 - Co-Citation
 - SimRank (Jeh and Widom, 2002)
- 2. Textual Similarity: bag-of-words similarity (Corley, 2005)
 - Knowledge-based: Wordnet-based word similarity
 - Corpus-based: TF-IDF
- 3. Spatial Similarity: physical aspects
 - Object area
 - Location (proximity)



Ongoing & Future Work

- Evaluation on existing similarity datasets
- Case study on OpenStreetMap and on the Web of Data (DBPedia, LinkedGeoData)
- Creation of a Geo-similarity golden standard

Research presented in this poster was funded by a Strategic Research Cluster Grant (07/SRC/I1168) by Science Foundation Ireland under the National Development Plan. The authors gratefully acknowledge this support.













