Discovering and Ranking Semantic Associations over a Large RDF Metabase

Chris Halaschek, Boanerges Aleman-Meza, I. Budak Arpinar, Amit P. Sheth
LSDIS Lab, Computer Science, University of Georgia

Motivation

1. Query between “Hubwoo” (company) and “SONERI” (bank) results in 1,160 semantic associations (SWETO testbed)
2. Cannot expect users to sift through resulting associations
3. Results must be presented to users in a relevant fashion...need ranking

Background

Semantic Associations

Approach

Association Rank Function

\[
\text{Score} = k_1 \times \text{Context} + k_2 \times \text{Subsumption} + k_3 \times \text{Trust} + k_4 \times \text{Rarity} + k_5 \times \text{Popularity} + k_6 \times \text{Association Length}
\]

(for flexibility, \(k\) values are assigned by user)

Semantic criteria

Context (classes/relations of interest)
Subsumption
Trust

Statistical criteria

Rarity
Popularity
Association Length

Acknowledgements

• SemDis Team – LSDIS Lab
• NSF-ITR-DM Award #0325464, ‘SemDIS: Discovering Complex Relationships in the Semantic Web’
• NSF-ITR-DM Award #0219649, ‘Semantic Association Identification and Knowledge Discovery for National Security Applications’