

Semantic Keyword Search in Linked Data

Andrea Calì, Leonardo Coaccioli, Mirko Michele Dimartino,
Riccardo Frosini and Federico Pastori

University of London, Birkbeck College
Oxford-Man Institute of Quantitative Finance, University of Oxford
Università Roma Tre

International Keystone Conference 2015

Coimbra, Portugal, 9th September 2015

- keyword search sometimes is not enough
- need for considering the **semantics** of terms

- keyword search sometimes is not enough
- need for considering the **semantics** of terms
- **query expansion** according to semantic criteria
 - extension of keyword search
 - [Guha, McCool, Miller WWW 2003]
 - [Rocha, Schwabe, de Aragão, WWW 2004]

- RDF graph traversal [Catarci et al. ECAI 2004]
- keyword/concept mapping
- RDF graph patterns
- ...

Food Market inefficiency

- Intermediaries to reduce friction in the market [Shi & Siou 2010; Gehrig 1993]
- In the food market intermediaries buy from producers
 - large warehouses, retail stores
 - large distances, long chain, high consumer prices, high waste

RealFoodTrade (RFT)

- marketplace for food
- sellers are the **producers**
- **no middleman** — wholesalers do not take part

The buyer

- geo-located sales
- flash stand and flash market
- Semantic Search to match demand and supply with:
 - ★ domain ontology (human-made) by FAO
 - ★ Linked Data

Our approach

- We combine **DBpedia** with the domain ontology together
- Representation as **3-dimensional vector**
- **Vector Space Model** to compute similarity
 - ★ cosine between vectors as similarity degree
 - *⟨species, genus, family⟩* as vector properties.

Results for tench

- 1 tench
- 2 bighead carp
- 3 blacknose dace
- 4 California roach
- 5 catla
- 6 chiselmouth
- 7 common carp
- 8 common dace
- 9 desert dace
- 10 fathead minnow

- RFT: the Web to improve workers' life

- **RFT**: the Web to improve workers' life
- Linked Data sets proved useful
 - ★ graph navigation
 - ★ integration of domain ontologies with Linked Data

- **RFT**: the Web to improve workers' life
- Linked Data sets proved useful
 - ★ graph navigation
 - ★ integration of domain ontologies with Linked Data
- Potential for significant **socio-economic impact**
 - ★ lower end prices
 - ★ higher profit for fishermen

- **RFT**: the Web to improve workers' life
- Linked Data sets proved useful
 - ★ graph navigation
 - ★ integration of domain ontologies with Linked Data
- Potential for significant **socio-economic impact**
 - ★ lower end prices
 - ★ higher profit for fishermen

Application

- Tracking of seafood chain (provenance)
 - ★ gather and interpolate geographic data
- Entend to other markets (agriculture etc.)
- Multilingual and colloquial names

Recommendation

- Incorporate **learning** into the system
 - ★ feedback from user behaviour
 - ★ personalized recommendations
 - ★ ...

Acknowledgments

- Camilo Rodríguez Beltrán, Univ. del Desarrollo
- Patricio Durán, fisherman
- Karen Croxson, McKinsey & Co.
- Thomas W. Lynch, Reasoning Technologies Ltd