LODsyndesis: Services for Connecting and Integrating Big Number of Datasets

Michalis Mountantonakis and Yannis Tzitzikas
{mountant|tzitzik}@ics.forth.gr
Institute of Computer Science, Foundation for Research and Technology (FORTH-ICS), GREECE, and Computer Science Department, University of Crete, GREECE

1. What is LODsyndesis?
The Information Systems Laboratory of the Institute of Computer Science of FORTH designs and develops innovative indexes, algorithms and tools to assist the process of semantic integration of data at a large scale. The suite of services and tools that have been developed are referred to as LODsyndesis.

2. The Difficulties of Semantic Data Integration at Large Scale
Different Names
Aristotle, Aristoteles, Αριστοτέλης
Different Schemas
Michalis Mountantonakis, Trikalwn 18, 71303
Different Formats
Michalis Mountantonakis
Data Quality
Which was the birth place of Aristotle?
Stagira
Athens
Data Volume
Everything produces data...

3. The Process of Global Indexing
- LODsyndesis provides indexes & measurements for 400 datasets, 412 million entities and 2 billion triples!
- LODsyndesis computes the transitive and symmetric closure of 45 million equivalence relationships!
- LODsyndesis is the largest knowledge graph of Linked Data that includes all inferred equivalence relationships!
- LODsyndesis offers novel measurements among any combination of datasets!

4. The Services of LODsyndesis
Object Coreference
Give me all the URIs of Aristotle
Dataset Discovery
Give me the K most connected Datasets to my Dataset
Dataset Enrichment with Features for ML
Create Automatically Features for any Entity

5. TRY LODsyndesis
TRY LODsyndesis: www.ics.forth.gr/isl/LODsyndesis/ for accessing all its novel services and measurements!
Follow us on twitter: https://twitter.com/lodsnysedesis