

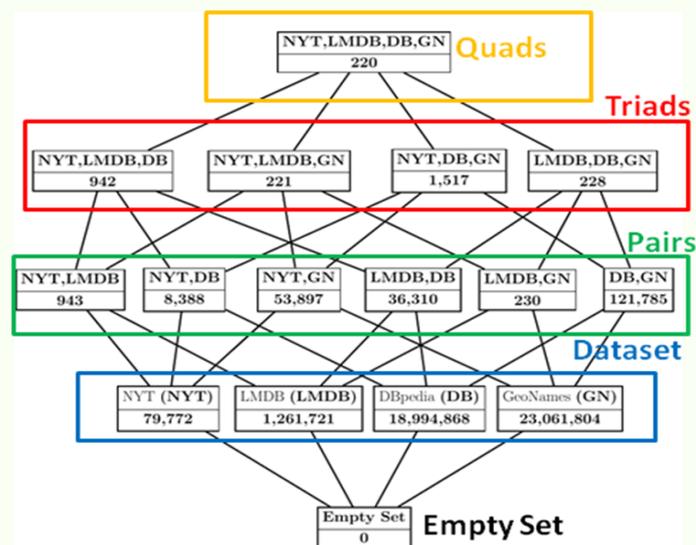


LODsyndesis: Connectivity of LOD Datasets

Overview

A big number of datasets has been published according to the principles of Linked Data and this number keeps increasing. Although the ultimate objective is linking and integration, it is not currently evident how connected the current LOD cloud is, and there are no services that take into account all datasets of LOD. For instance, it is not currently easy to find all URIs that correspond to one entity (say Aristotle), it is not currently easy to find all datasets that are connected to your own dataset, etc. Such tasks are not trivial since the owl:sameAs relationships model an equivalence relation and therefore its transitive closure has to be computed and this presupposes knowledge of all datasets.

To tackle such problems we have designed and developed LODsyndesis (<http://www.ics.forth.gr/isl/LODsyndesis/>). LODsyndesis provides query services and measurements that are useful for several important tasks like (a) object co-reference, (b) dataset discovery, (c) visualization, (d) connectivity assessment and monitoring. In addition, this page provides measurements that concern the commonalities of Linked Datasets.



Measurements between any set of datasets

Target Domains

LODsyndesis provides query services and measurements (by using special novel indexes & algorithms) that are useful for several important tasks like (a) object co-reference, (b) dataset discovery, (c) visualization, (d) connectivity assessment and monitoring.

Description

LODSynthesis provides measurements about the number of common real world objects between any set of datasets (of the current LOD cloud), where the number of common real world objects is the number of classes of equivalence of URIs after taking the symmetric and transitive closure of the set of owl:sameAs relationships from all datasets.

The web page of LODSynthesis provides:

- a list of answerable queries for dataset discovery and global entity lookup
- an active SPARQL Endpoint where one can run the aforementioned queries
- a link to all measurements which have been published in datahub
- a link to an interactive 3D visualization that exploits some of these measurements
- a service for global entity lookup
- a service for dataset discovery

Such services could greatly aid the continuous semantic integration of datasets across domains which in turn could boost data-intensive scientific discovery. The indexes and algorithms behind LODSynthesis are described in the paper entitled “On Measuring the Lattice of Commonalities Among Several Linked Dataset” (pdf available at <http://www.vldb.org/pvldb/vol9/p1101-mountantonakis.pdf>) which presented in the prestigious VLDB 2016 (42th International Conference on Very Large Data Bases) conference in September 2016.

The diagram illustrates the services provided by LODSynthesis, organized into four main categories:

- Object Coreference:** "Give me all the URIs of Aristotle" (accompanied by a portrait of Aristotle).
- Visualizations:** Shows a 3D visualization of a network and a hierarchical tree structure.
- Connectivity Assessment and Monitoring:** Includes a line graph and a bar chart showing connectivity metrics.
- Dataset Discovery:** "Give me the K most connected Datasets to my Dataset" (accompanied by a network diagram connecting FishBase, Freebase, Wikidata, DBpedia, and GeoNames).

Below the diagram is the text "Services provided by LODSynthesis". To the right is the "LODSynthesis Logo", which features a colorful network of nodes and edges above a red horizontal bar and a black basketball hoop.

Additional Information

More information about LODSynthesis and its services (object coreference, dataset discovery, visualization and connectivity assessment) can be found at <http://www.ics.forth.gr/isl/LODSynthesis>



LODSynthesis website

Contact details: **Yannis Tzitzikas and Michalis Mountantonakis**
tzitzik@ics.forth.gr, mountant@ics.forth.gr
www.ics.forth.gr/isl