

# LODsyndesis: Global Scale Knowledge Services

**Track: Full Length Paper**

Michalis Mountantonakis and Yannis Tzitzikas  
Institute of Computer Science, FORTH-ICS, Greece, and  
Computer Science Department, University of Crete, Greece  
{mountant|tzitzik}@ics.forth.gr

We shall present LODsyndesis, a suite of services over the datasets of the entire Linked Open Data Cloud, that offers fast content-based dataset discovery and object co-reference. Emphasis is given on supporting scalable cross-dataset reasoning for finding all information about any entity and its provenance. Other tasks that can be benefited from these services are those related to the quality and veracity of data since the collection of all information about an entity, and the cross-dataset inference that is feasible, allows spotting the contradictions that exist and on the same time provides information for data cleaning or for estimating and suggesting which data are probably correct or more accurate. In addition we will show how these services can assist the enrichment of existing datasets with more features for obtaining better predictions in machine learning tasks. Finally, we will report measurements that reveal the sparsity of the current datasets, as regards their connectivity, which in turn justifies the need for advancing the current methods for data integration. Measurements focusing on the cultural domain will also be included, specifically measurements over datasets using CIDOC CRM, and connectivity measurements of British museum data. The services of LODsyndesis are based on special indexes and algorithms and allow the indexing of 2 billion triples in around 20 minutes using a cluster of 64 computers.