Neapolis University

HEPHAESTUS Repository **http://hephaestus.nup.ac.cy**

School of Information Sciences Book chapters

2015-11

Combining Syntactic and Semantic Evidence for Improving Matching over Linked Data Sources

Christodoulou, Klitos

Springer Link

http://hdl.handle.net/11728/6277

Downloaded from HEPHAESTUS Repository, Neapolis University institutional repository

HEPHAESTUS
Research Repository

Title:	**COMBINING SYNTACTIC AND SEMANTIC EVIDENCE FOR IMPROVING MATCHING OVER LINKED DATA SOURCES**
Year:	2015-11
Author:	Christodoulou, Klitos ; Fernandes, Alvaro A.A. ; Paton, Norman W.
Abstract:	In the context of Linked Data (LD) sources, the ability to traverse links and retrieve further information can be exploited to harvest semantic annotations. Such annotations can, in turn, underpin the inference of semantic correspondences between sources. This paper shows that using semantic annotations as additional evidence of equivalence between schematic representations of LD sources can improve upon the prevalent, purely syntactic approaches. The paper both describes the construction of probabilistic models that yield degrees of belief on the equivalence of the real-world concepts represented by the data and shows how these models are crucial in underpinning a Bayesian approach to assimilating both syntactic evidence (in the form of similarity scores derived by string-based matchers) and semantic evidence (in the form of semantic annotations stemming from LD vocabularies) of equivalence. The paper presents an empirical evaluation of the techniques described. The main finding is confirmation that, with respect to equivalence judgements made by human experts, the use of the contributed techniques incurs significantly fewer discrepancies than purely syntactic approaches.

Neapolis University
HEPHAESTUS Repository https://hephaestus.nup.ac.cy