

Data Interlinking with Formal Concept Analysis and Link Keys

Jérôme Euzenat

INRIA & Univ. Grenoble Alpes, Grenoble, France

Abstract. Data interlinking, the problem of linking pairs of nodes in RDF graphs corresponding to the same resource, is an important task for linked open data. We introduced the notion of link keys as a way to identify such node pairs [1]. Link keys generalise in several ways keys in relational algebra. Thus, we consider how they could be extracted from data with Formal Concept Analysis. We show that an appropriate encoding makes the notion of candidate link keys correspond to formal concepts [2]. However candidate link keys are not yet link keys as they need to be selected through appropriate measures. We discuss how the measurement and concept extraction processes may be interleaved. If time permits we will also discuss extensions of this model to residual link keys and mutually dependent link keys¹.

Keywords: Linked Data, Formal Concept Analysis

References

1. Atencia, M., David, J., Euzenat, J.: Data interlinking through robust linkkey extraction. In: Proc. 21st European Conference on Artificial Intelligence (ECAI), Praha (CZ). (2014) 15–20
2. Atencia, M., David, J., Euzenat, J.: What can FCA do for database linkkey extraction? In: Proc. 3rd ECAI workshop on What can FCA do for Artificial Intelligence? (FCA4AI), Praha (CZ). (2014) 85–92

¹ This work has been developed in collaboration with Manuel Atencia, Jérôme David and Amedeo Napoli.