

Chapter 3

SpotTheLink: A Game-Based Approach to the Alignment of Ontologies

Stefan Thaler

STI-Innsbruck, University of Innsbruck Austria

Elena Simperl

AIFB, Karlsruhe Institute of Technology, Germany

Katharina Siorpaes

STI-Innsbruck, University of Innsbruck, Austria

Stephan Wölger

STI-Innsbruck, University of Innsbruck, Austria

ABSTRACT

A multitude of approaches to match, merge, and integrate ontologies, and more recently, to interlink RDF data sets, have been proposed over the past years, making ontology alignment one of the most active and at the same time mature area of research and development in semantic technologies. While advances in the area cannot be contested, it is equally true that full automation of the ontology-alignment process is far from being feasible; human input is often indispensable for the bootstrapping of the underlying methods, and for the validation of the results. The question of acquiring and leveraging such human input remains largely unaddressed, in particular when it comes to the incentives and motivators that are likely to make users invest their valuable time and effort in alignment tasks such as entity interlinking and schema matching, which can be domain-knowledge-intensive, technical, or both. In this chapter, the authors present SpotTheLink, a game whose purpose addresses this challenge, demonstrating how knowledge-intensive tasks in the area of the Semantic Web can be collaboratively solved by a community of non-experts in an entertaining fashion.

DOI: 10.4018/978-1-4666-0894-8.ch003

1. INTRODUCTION

A large share of tasks in knowledge engineering crucially rely on human input (Siorpaes and Simperl 2010). This applies in particular to those tasks which are acknowledged to be hardly approachable in a systematic, engineering-driven fashion; and also, though to a lesser extent, to the wide array of (semi-) automatic methods and techniques that have been proposed as an attempt to reduce the costs of knowledge-engineering projects by minimizing the need for human involvement in these projects. In this second case, despite constant progress in improving the performance of the corresponding algorithms and the quality of their results, experiences show that human input is nevertheless required, even if it just for the configuration of the algorithms, the creation of knowledge corpora the algorithms can be trained on, or the validation of (intermediary) algorithm outputs. Examples of such knowledge engineering tasks are numerous, including machine translation, information extraction, text summarization, data integration, multimedia analysis, and conceptual modeling.

One novel approach which proved successful to resolve technical tasks via human computation is based on ‘games with a purpose’ (Van Ahn 2006). The idea behind games with a purpose is simple, but effective; tasks which remain difficult to handle by computers, but which humans seem to tackle easily are hidden behind entertaining, collaborative games targeting not experts, but casual Internet users. By playing a game with a purpose, users are indirectly generating data that can be capitalized to build knowledge corpora required for the training of algorithms, thus providing a powerful example of how human and computational intelligence can be combined to address important, challenging problems. Since the original proposal by Van Ahn in 2006 games with a purpose have been applied to tasks as di-

verse as image and video annotation,¹ genetics,² natural language processing,³ and knowledge engineering.⁴

In our work we have applied the idea of games with a purpose in the area of semantic technologies. Semantic technologies exploit machine-understandable representations of data, processes and computational resources in order to create IT systems and applications that are able to better interpret and process the information needs of their users, and to interact with other systems in an interoperable way. A fundamental building block of this approach are ontologies, which are used to capture and structure knowledge in a given domain in terms of classes, instances, relationships and axioms, as a baseline for the implementation of the functionality just mentioned. OntoGame⁵ is the framework for the implementation casual games which we designed for this purpose. OntoGame capitalizes on fun and competition as two key motivators for people to willingly invest their valuable time and effort in executing specific knowledge-engineering related tasks, whose technicalities are hidden behind the game experience (Siorpaes and Hepp 2008). Compared to other similar games targeting the same or related tasks, OntoGame’s distinctive feature lies in the fact that the input of the players is translated into Semantic Web content, for instance ontologies and metadata encoded in RDFS and OWL,⁶ which are W3C standards for the representation of data on the Semantic Web. In this chapter we present one of the latest releases of the OntoGame series, called SpotTheLink, which is dedicated to the task of ontology alignment. The design principles, enabling technology and evaluation methodology we relied upon to develop SpotTheLink can be applied in a variety of human computation scenarios targeting different other knowledge-engineering tasks. We have done so ourselves with games such as Onto-Pronto, which classifies entities from the English Wikipedia according to an upper-level ontology;

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/spotthelink-game-based-approach-alignment/65686

Related Content

Blockchain-Based Lightweight Authentication Mechanisms for Industrial Internet of Things and Information Systems

Mingrui Zhao, Chunjing Shi and Yixiao Yuan (2024). *International Journal on Semantic Web and Information Systems* (pp. 1-30).

www.irma-international.org/article/blockchain-based-lightweight-authentication-mechanisms-for-industrial-internet-of-things-and-information-systems/334704

Semantic Integration of Structured and Unstructured Data in Data Warehousing and Knowledge Management Systems

Liane Haak (2012). *Semantic Technologies for Business and Information Systems Engineering: Concepts and Applications* (pp. 81-105).

www.irma-international.org/chapter/semantic-integration-structured-unstructured-data/60057

Remote Sensing Image Semantic Segmentation Method Based on a Deep Convolutional Neural Network and Multiscale Feature Fusion

Guangzhen Zhang and Wangyang Jiang (2023). *International Journal on Semantic Web and Information Systems* (pp. 1-16).

www.irma-international.org/article/remote-sensing-image-semantic-segmentation-method-based-on-a-deep-convolutional-neural-network-and-multiscale-feature-fusion/333712

Trust Metrics on Controversial Users: Balancing Between Tyranny of the Majority

Paolo Massa and Paolo Avesani (2007). *International Journal on Semantic Web and Information Systems* (pp. 39-64).

www.irma-international.org/article/trust-metrics-controversial-users/2830

Shape-Based Image Retrieval By Alignment

Enver Sangineto (2007). *Semantic-Based Visual Information Retrieval* (pp. 46-67).

www.irma-international.org/chapter/shape-based-image-retrieval-alignment/28921