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.

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 diverse 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. OntoGame[5] 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;

Related Content

Optimal User Association of LTE/Wi-Fi/Wi-Gig Bands in 5G Cellular Networks
Aziza Ibrahim Hussainand Ahmed Zakaria Sayed (2021). *International Journal on Semantic Web and Information Systems (pp. 22-40).*
www.irma-international.org/article/optimal-user-association-of-ltewi-fiwi-gig-bands-in-5g-cellular-networks/277080

Managing Large Amounts of Data Generated by a Smart City Internet of Things Deployment
Jorge Lanza, Pablo Sotres, Luis Sánchez, Jose Antonio Galache, Juan Ramón Santana, Verónica Gutiérrezand Luis Muñoz (2016). *International Journal on Semantic Web and Information Systems (pp. 22-42).*
www.irma-international.org/article/managing-large-amounts-of-data-generated-by-a-smart-city-internet-of-things-deployment/164483

Semantic Search on Unstructured Data: Explicit Knowledge through Data Recycling
Alex Kohn, François Bryand Alexander Manta (2012). *Semantic-Enabled Advancements on the Web: Applications Across Industries (pp. 194-213).*
www.irma-international.org/chapter/semantic-search-unstructured-data/64023

An Architecture for Restful Web Service Discovery Using Semantic Interfaces
José Renato Villela Dantasand Pedro Porfirio Muniz Farias (2020). *International Journal on Semantic Web and Information Systems (pp. 1-24).*
www.irma-international.org/article/an-architecture-for-restful-web-service-discovery-using-semantic-interfaces/244185

Coronavirus Pandemic (COVID-19): Emotional Toll Analysis on Twitter
Jalal S. Alowibdi, Abdulrahman A. Alshdadi, Ali Daud, Mohamed M. Dessoukyand Essa Ali Alhazmi (2021). *International Journal on Semantic Web and Information Systems (pp. 1-21).*
www.irma-international.org/article/coronavirus-pandemic-covid-19/277079