



Contents lists available at ScienceDirect

Web Semantics: Science, Services and Agents on the World Wide Web

journal homepage: www.elsevier.com/locate/websem

Visual query interfaces for semantic datasets: An evaluation study[☆]



Guillermo Vega-Gorgojo^a, Laura Slaughter^b, Martin Giese^{a,*}, Simen Heggestøy^a, Ahmet Soylu^c, Arild Waaler^a

^a Department of Informatics, University of Oslo, Norway

^b Oslo University Hospital, Oslo, Norway

^c Faculty of Computer Science and Media Technology, Norwegian University of Science and Technology, Gjøvik, Norway

ARTICLE INFO

Article history:

Received 24 September 2014

Received in revised form

1 October 2015

Accepted 11 January 2016

Available online 4 March 2016

Keywords:

Semantic search

Visual query interfaces

User studies

Usability

ABSTRACT

The rapid growth of the Linked Open Data cloud, as well as the increasing ability to lift relational enterprise datasets to a semantic, ontology-based level means that vast amounts of information are now available in a representation that closely matches the conceptualizations of the potential users of this information. This makes it interesting to create ontology based, user-oriented tools for searching and exploring this data. Although initial efforts were intended for tech users with knowledge of SPARQL/RDF, there are ongoing proposals designed for lay users. One of the most promising approaches is to use visual query interfaces, but more user studies are needed to assess their effectiveness. In this paper, we compare the effect on usability of two important paradigms for ontology-based query interfaces: form-based and graph-based interfaces. In order to reduce the number of variables affecting the comparison, we performed a user study with two state-of-the-art query tools developed by ourselves, sharing a large part of the code base: the graph-based tool OptiqueVQS^{*}, and the form-based tool PepeSearch. We evaluated these tools in a formal comparison study with 15 participants searching a Linked Open Data version of the Norwegian Company Registry. Participants had to respond to 6 non-trivial search tasks using alternately OptiqueVQS^{*} and PepeSearch. Even without previous training, retrieval performance and user confidence were very high, thus suggesting that both interface designs are effective for searching RDF datasets. Expert searchers had a clear preference for the graph-based interface, and mainstream searchers obtained better performance and confidence with the form-based interface. While a number of participants spontaneously praised the capability of the graph interface for composing complex queries, our results evidence that graph interfaces are difficult to grasp. In contrast, form interfaces are more learnable and relieve problems with disorientation for mainstream users. We have also observed positive results introducing faceted search and dynamic term suggestion in semantic search interfaces.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The increasing availability of Linked Data is changing the ways that developers design for interaction with Web content. Heath [1] outlines this shift in metaphor, away from the current

document-centric Web, to one in which users are interacting with things (data and objects) and the connections between them. This is especially true for interfaces for retrieving information from the Web. Existing Linked Data can be used in everyday tasks, such as making decisions about a car purchase or researching the potential success of opening a new organic-food shop. However, it is unclear how average Web users can find and digest Linked Data in order to fulfill their information needs, i.e. without requiring specific knowledge of RDF and SPARQL. The ubiquitous document retrieval style interfaces for organizing and locating Web pages (i.e. documents) are not meant for finding Linked Data. One of the challenges of designing for the Semantic Web includes finding new ways to allow people to use this content [2]. Unfortunately, most tools available are not easily used by searchers having few-to-no technical skills [3]. Most are SPARQL query interfaces that require

[☆] This research has been partially funded by the Norwegian Research Council through the Semicolon II project (<http://www.semicolon.no/>), and the European Commission through the Optique (FP7 GA 318338) and BYTE (FP7 GA 619551) projects. The authors also want to thank the participants in the user study.

* Corresponding author. Tel.: +47 22852737.

E-mail addresses: guiveg@ifi.uio.no (G. Vega-Gorgojo), laura.slaughter@gmail.com (L. Slaughter), martingi@ifi.uio.no (M. Giese), simenheg@ifi.uio.no (S. Heggestøy), ahmet.soylu@ntnu.no (A. Soylu), arild@ifi.uio.no (A. Waaler).

a user to learn how to write SPARQL syntax to access triplestore endpoints.

The lack of intuitive, non-technical, and novice-friendly search interfaces essentially blocks many ordinary users from gaining access to published Linked Data. To overcome this problem, there have been attempts at developing Linked Data search tools that hide SPARQL syntax from the user. Of these tools, the present types of search statement input styles include: natural language, keyword-based, form or template-based, and other visual approaches such as graphs. Some user interfaces (UIs) employ only one style of interaction at input and others use multiple (mixed) styles. For example, K-search [4] presents an ontology tree view combined with a form-based query entry interface. Results of queries, the retrieved RDF content, are often presented as tables and lists, or as graphical data objects.

Although various types of search interfaces have been developed and reported in the literature, there is a lack of empirical evidence of the effectiveness of these approaches [5]. Innovative search interfaces for querying RDF triples are often described technically and with little discussion of design for human interaction. It is not often that they have conducted usability tests. Elbedweihy et al. [6] published one of the few papers on evaluating search query approaches for the Semantic Web. They state the importance of working towards a comprehensive evaluation framework that provides guidance to developers, including design criteria for task type (e.g. search for facts) and user type (e.g. domain experts lacking technical skills), stating “there are very few studies that have focused on assessing the usability of semantic search systems”.

As we have established, user studies are scarce and this motivates our work, which pursues the proposal of a portable search tool for mainstream users, the non-savvy searcher having little to no knowledge of Semantic Web technical standards, looking for facts and having well-defined searches (as opposed to ill-defined or vaguely defined information needs). For example, users such as journalists who need to know how many people in Norway voted in the last election within a specific geographical region.

The test case for the project was based on a freely accessible government data site, a Linked Open Data version of the Norwegian Company Registry.¹ Based on the prior work discussed above, we identified form-based interfaces as best for the mainstream user [6], and the PepeSearch interface was therefore constructed as a multi-class form, taking inspiration from facet-based interfaces. This design is compared to the graph-based interface of OptiqueVQS [7] that displays the underlying classes and relationships in a visual query-building environment intended for domain experts.

In this paper we present the results of a comparison study between OptiqueVQS* and PepeSearch. We collect user feedback on both interfaces, asking participants to complete specific search tasks and to report their satisfaction with the tools. Our contribution is to determine whether graph-based and form-based interfaces are effective for non-tech users in terms of retrieval performance and usability. We hypothesize that for mainstream users, form interfaces will outperform graph interfaces, and that satisfaction scores will be higher. We look at interface strengths and weaknesses, user feelings of being in control, and disorientation. In addition, we investigate user confidence in the result sets. This study fits into a larger framework that will bring further knowledge to the community of developers working on Semantic Web challenges [2,8], addressing some of the current shortcomings of understanding the average user’s perspective and needs for Semantic Web tools.

We organize the paper by first presenting relevant background on the definition of mainstream users and semantic search, as well as previously developed interfaces and evaluation studies that are related to our design. We then describe the two interfaces in detail before moving on to report the methods. Results are presented in detail followed by a discussion section that includes both limitations and plans for future work.

2. Background

2.1. Semantic search for mainstream users

The Semantic Web is moving out of the stage where only programmers and those experts with the technical skills necessary for working with raw RDF are using it. Efforts to expose data openly to citizens and customers for analysis and reuse have been made. Examples of early adopters of Linked Data for the mainstream have included open government data, such as Data.gov in the USA [9]. News media have also worked toward opening up information sources, for instance the Guardian’s Open Platform.²

Although there are now many open data sites for mainstream consumption, there exists a significant gap in support for non-technical lay users. Dadzie [10] states that “in the Web of Data, people should not be required to learn SPARQL or to have extraordinary technical skills to access data”. The first step in creating tools for the non-technical is to understand these potential users and what tasks they need to accomplish. There are surprisingly few papers that discuss the characteristics of the intended audience for Linked Data.

Battle [11] proposed a framework of user types and their intended tasks specific in the context of the Semantic Web. Users are grouped in three high-level categories: (1) end-users, (2) content curators, and (3) ontologists. It is the first category of end-users that we are interested in. These users are defined as “ordinary people who are either seeking information or trying to accomplish something in the course of their everyday life or work”. They could be news seekers, patients, graduate students, medical researchers, or car buyers. Their range of domain knowledge of the subject matter could be anywhere along the continuum from low to high, and the main tasks are information seeking tasks, information-synthesis tasks, and action-oriented tasks. The usability studies described above, carried out by [12,13], cite this definition as the basis for their *casual user* category.

Dadzie et al. [10] distinguishes three types of users in the following categories: (1) lay users or mainstream, (2) domain experts, and (3) tech users. Similar to Battle’s end-user definition, lay users are computer literate, have some searching skills to find resources, and are looking to do everyday tasks such as making comparisons while shopping. However, Dadzie’s work defines lay users as not having much domain knowledge, and so this grouping would not include, for example, medical researchers or graduate students. Instead, there are two other categories for the non-mainstream: tech users who understand Semantic Web technologies, and another for domain experts who make use of “sophisticated domain-specific analysis tools”. For the purposes of this study we adhere to this classification, focusing our interest on mainstream users.

While different definitions of *semantic search* can be found in the literature [14, ch. 3], we refer here to finding results to information needs by using Semantic Web data. This is probably the most common interpretation and, hence, it does not prescribe a particular query specification mechanism—existing approaches

¹ <http://www.brreg.no/>.

² <http://www.theguardian.com/open-platform>.

Download English Version:

<https://daneshyari.com/en/article/561750>

Download Persian Version:

<https://daneshyari.com/article/561750>

[Daneshyari.com](https://daneshyari.com)