



ELSEVIER

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

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

Overcoming the pitfalls of ontology authoring: Strategies and implications for tool design [☆]

Markel Vigo ^{*}, Samantha Bail, Caroline Jay, Robert Stevens

School of Computer Science, University of Manchester, Manchester, UK

ARTICLE INFO

Article history:

Received 11 February 2014

Received in revised form

17 June 2014

Accepted 21 July 2014

Communicated by N.F. Noy.

Available online 30 July 2014

Keywords:

Ontologies

Semantic web

Authoring tools

ABSTRACT

The process of authoring ontologies appears to be fragmented across several tools and workarounds, and there exists no well accepted framework for common authoring tasks such as exploring ontologies, comparing versions, debugging, and testing. This lack of an adequate and seamless tool chain potentially hinders the broad uptake of ontologies, especially OWL, as a knowledge representation formalism. We start to address this situation by presenting insights from an interview-based study with 15 ontology experts. We uncover the tensions that may emerge between ontology authors including antagonistic ontology building styles (definition-driven vs. manually crafted hierarchies). We identify the problems reported by the ontology authors and the strategies they employ to solve them. These data are mapped to a set of key design recommendations, which should inform and guide future efforts for improving ontology authoring tool support, thus opening up ontology authoring to a new generation of users. We discuss future research avenues in light of these results.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

1. Introduction

The use of ontologies for capturing the knowledge of a domain of interest has grown significantly, with areas such as bioinformatics and healthcare modelling and sharing information in ontologies of varying size and complexity. One of the central ideas of the Semantic Web is that of sharing, linking, and reusing data from multiple sources; as a consequence, the availability of (high quality) semantically described data sources and thus the uptake of Semantic Web technologies are important, not only to the success of the Semantic Web, but also to applications in which rich domain descriptions can play a role.

Considering the inherent complexity of a knowledge representation formalism such as OWL, convincing domain experts, and thus, potential ontology authors, of the usefulness and benefits of using ontologies is one of the major barriers to broader ontology adoption (Hepp, 2007). The following excerpt, which was extracted from the interviews that inform the study presented here, is a clear example of the current situation:

“A domain expert has to be totally convinced that ontologies are the right way of modelling knowledge in a domain, and then has to consistently work for a period of time in order to be self-sufficient. In the initial stages, when they start doing the

modelling, you need a knowledge engineer to hold their hand [...] the moment the knowledge engineer disappears they will not carry on with it because it's much easier to get in databases or Excel than to do all this.” (Participant 13)

There are several factors that may prevent the uptake of ontologies including the cognitive complexity of ontology languages such as OWL (Horridge et al., 2011), the steep learning curve of these languages (Warren et al., 2014) or the required minimum expertise to model very specific domains (Randall et al., 2011). One of the key factors, in addition to the ones mentioned, that contribute towards the acceptance of a new technology is the availability of adequate tool support that allows both novice and expert users to create high quality artefacts that meet their functional requirements. While ontology authoring tools have seen many improvements over recent years, most notably demonstrated by the popularity of Protégé as an integrated ontology development tool (Cardoso, 2007), we still lack a good understanding of the effectiveness of existing tools, or what kind of tool support ontology authors need to successfully create, explore, reuse, refactor, and debug ontologies. The quote above is indicative of the complexity of the ontology authoring process, and how typical ontology engineering tools are unsuccessful in making the process of authoring ontologies easier and more accessible; a shortfall which affects not only the uptake of ontology by new users, but also the quality of artefacts created by existing ontology engineers.

[☆]This paper has been recommended for acceptance by N.F. Noy.

^{*} Corresponding author.

E-mail address: markel.vigo@manchester.ac.uk (M. Vigo).

In order to gain insights about the problems encountered in the authoring process we carry out an interview study with 15 ontology authors. We deliberately target ontology experts for two reasons: firstly, their understanding of problems is more comprehensive and richer; and secondly, they employ sophisticated authoring strategies that may be indicative of latent problems of tools. With the help of a qualitative analysis software tool, transcribed interviews were thematically analysed in an open coding fashion. A codebook of themes was agreed among the authors and an analysis of reliability was conducted to discard any source of bias (see the study in Section 3). The identified problems and strategies may not be representative of the whole spectrum of ontology authors though: novices may struggle with additional problems to the ones we identify; also, the participants were mostly drawn from life sciences and thus, some of the problems might be domain specific. We argue that emergent design recommendations will also entail benefits for novices and across disciplines as there are commonalities—of a different magnitude—between the problems faced by different types of users. For instance, the cognitive complexity of OWL is challenging for all users even if it is a more severe barrier for non-experts.

We believe that, in order to improve the experience of ontology authors, it does not suffice to only analyse their interaction with ontology engineering tools; instead, we need a more *holistic* view of the ontology authoring process and lifecycle. In Section 4.1 we first focus on the different approaches that exist to authoring ontologies, including the different ontology building styles (what we here call ‘definition-driven’ ontology development vs. manually crafted hierarchies). We then need to gain a clear picture of the problems ontology authors encounter when developing ontologies, common tasks with which they struggle, and the tools and strategies they employ to overcome such problems (see Section 4.2). As a result, the contributions of this study are the following:

- We raise awareness about the latent tensions about ontology building styles in Section 5.1. While we do not take a stance in whether having a common understanding and strategy is right or not, having this awareness may help to remove frictions in teams.
- By mapping the problems and strategies to design insights, we suggest how the next generation of authoring tools could cater for the needs of broader audiences (see Section 5.2).
- In Section 5.3 we enumerate a set of operable design recommendations that emerge from the acquired insights.
- We evaluate state-of-the-art authoring tools against the design recommendations (see Section 5.4).

While some of our findings confirm existing assumptions about the ontology engineering process, such as the lack of tools for helpful exploration and visualisation of ontologies, others are more surprising. Amongst others, both preventing errors and debugging erroneous ontologies are still considered to be challenging tasks when constructing ontologies, while support for judging the quality of ontologies in the context of their functional requirements, in the spirit of unit tests in traditional software engineering, is not widespread. Furthermore, as reasoning is used frequently to check the correctness of ontologies, on-the-fly reasoning capabilities seem to become increasingly relevant.

2. Related work

Over the years, the usability of the tools of the trade has not been a major concern in the ontology engineering community. Pioneering work exists in the realm of knowledge representation systems, where it is indicated that explanation of inferences,

adequate reporting of errors, and performance are key to the usability and success of such tools (McGuinness and Patel-Schneider, 1998).

One other systematic approach to the hitherto usability analyses of ontology authoring tools included an evaluation framework containing 26 dimensions to compare six ontology authoring tools (Duineveld et al., 2000). A set of dimensions was used to heuristically assess usability qualities through a number of statements such as “Evaluate the clarity of the interface” or “Is there a good overview of the ontology?”, whereas another set of questions articulated the features supported by tools in terms of “Is it possible to use multiple inheritance?” or “Does the tool check new data for consistency?”. The main conclusion of this work was that authoring tools of that time (circa 2000) were unusable for domain and ontology experts due to difficulties learning to use the tools and their poor usability.¹

When it comes to ontology authoring tools, literature surveys are the most comprehensive and replicable method by which tools are classified and assessed against evaluation frameworks. Katifori et al. (2007) provide a detailed analysis discussing how authoring tools implement ontology visualisations such as zoom features or 3D navigation techniques. They suggest that visualisation should not be used as an isolated technique, but complemented with efficient search and navigation mechanisms. Also, they suggest that future ontology visualisation tools should be able to remove clutter and lead ontology authors to their goals. Usability, interoperability, and portability are some of the dimensions included in a framework to classify semantic authoring tools of textual content (Khalili and Auer, 2013). In this study, faceted browsing, faceted viewing, single point of entry interface, and inline editing were identified as the features that increase the usability of semantic content authoring tools. It is remarkable that out of the 175 papers analysed in the survey only 2 involve end-users in order to evaluate the user interface. In both cases enquiry methods were employed a posteriori in order to provide evidence about user acceptability of the evaluated tools. We do not know whether this small extent of user involvement can be generalised to general purpose ontology authoring tools. However, it can be considered an indicator of low involvement of users in the development process.

In a landscape where user tests are scarce, there are some notable exceptions of ontology authoring tool evaluation with users. In addition to evaluating tools against predefined criteria (stability, availability, and extendibility, to name a few), Lambrix et al. (2003) presented a study in which 8 non-expert users carried out a number of basic tasks including ontology loading, entity addition, modification, and removal. After completing their tasks, users were given questionnaires in order to quantify the relevance of tools to accomplish tasks, tool efficiency, user attitude towards tools, and learnability aspects. They concluded that none of the tools stood up for all tasks, but each of them had their weaknesses and strengths.²

A later user study from 2006 involved 28 participants with basic OWL knowledge carrying out three tasks with TopBraid Composer³ and Protégé (Dzbor et al., 2006). Tasks included creating classes and properties, as well as adding subsumption, equivalence, and range axioms. Afterwards, participants were given questionnaires to measure the effectiveness, efficiency, and user experience with the tools. Dzbor et al. (2006) conclude that

¹ Tools evaluated were Ontolingua, WebOnto, Protégé Win, OntoSaurus, ODE and KADS22.

² Again, this research dates back to 2003, so it is unknown how this analysis applies to the current situation. Evaluated tools were Protégé 2000, Chimaera, DAG-Edit and OilEd.

³ <http://www.topbraidcomposer.com/>

Download English Version:

<https://daneshyari.com/en/article/6861113>

Download Persian Version:

<https://daneshyari.com/article/6861113>

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