Using ontologies

Understanding the user experience

Paul Warren[1], Paul Mulholland[2], Trevor Collins[3], Enrico Motta[4]

Knowledge Media Institute, The Open University
Milton Keynes, Buckinghamshire, MK7 6AA, U.K.
(1) paul.warren@cantab.net, (2) paul.mulholland@open.ac.uk,
(3) trevor.collins@open.ac.uk, (4) enrico.motta@open.ac.uk

Abstract. Drawing on 118 responses to a survey of ontology use, this paper describes the experiences of those who create and use ontologies. Responses to questions about language and tool use illustrate the dominant position of OWL and provide information about the OWL profiles and particular Description Logic features used. The survey revealed a considerable range of ontology sizes and analysis suggests a classification into two broad groups; one where ontologies have very few individuals, the other in which the number of individuals is more commensurate with the number of classes. The survey also reports on the use of ontology visualization software, finding that the importance of visualization to ontology users varies considerably. Pattern use is also examined in detail, drawing on further input from a follow-up study devoted exclusively to this topic. Evidence suggests that pattern creation and use are frequently informal processes and there is a need for improved tools. An analysis of the purposes for which ontologies are used suggests a classification into four categories of users: *conceptualizers*, *integrators*, *searchers* and *multipurpose users*. It is proposed that the categorisation of users and user behaviour should be taken into account when designing ontology tools and methodologies. This should enable rigorous, user-specific use cases.

Keywords: ontology use; ontology size; ontology visualization; ontology patterns; Description Logics; OWL profiles.

1 Introduction

According to Guarino et al. (2009), the notion of ontology in computer science was originally defined by Gruber (1993), as “an explicit specification of a conceptualization”. Over the two decades since this original definition, the use of ontologies has become widespread across a range of application areas. The Handbook on Ontologies (Staab & Studer, 2010) has chapters describing applications in bioinformatics, cultural heritage and recommender systems, besides knowledge management generally. As will be illustrated later, the use of Description Logics (DLs), specifically the variants of OWL, has become the dominant paradigm. With this has come an understanding of the computational properties of the various DLs and the development of efficient reasoners, e.g. see Möller and Haarslev (2009) and Motik (2009). This work has fed into tool development, e.g. the development of the Protégé ontology editor[1] and a variety of ontology visualization tools, e.g. see Katifori et al. (2007). During this time practitioners in particular domains have reported on their experience, e.g. Stevens et al. (2007). Others have undertaken relatively small scale studies of ontology users. For example, Vigo et al. (2014a) report on an interview study of 15 ontology authors and make a number of recommendations for improvement to ontology tool design. In another paper Vigo et al. (2014b) describe a Protégé plug-in to harvest information about the authoring process and thereby enable this process to be studied in detail. Our survey complements these kinds of studies by using a questionnaire to reach a large number of ontology users. The survey also complements the work of researchers who have analyzed actual ontologies. For example, Power and Third (2010) report on the usage of OWL language features based on the analysis of about 200 ontologies. In addition, Khan and Blomqvist (2010) analyzed 682 online ontologies to determine the frequency of occurrence of content patterns from the ODP portal[2]. In contrast, our approach is concerned particularly with the subjective experience of ontology users. Our aim is to understand that experience as a step towards improving it and making it more effective.

The remainder of this paper is organized as follows. Section 2 describes the survey, how it was conducted, and gives some information about the respondents. In section 3 we look at the various reasons for using ontologies. Section 4 discusses ontology languages and ontology tools. Section 5 discusses ontology size and relates size to aspects of ontology use. Section 6 discusses the respondents' experiences with visualization tools and reports a range of attitudes to visualization. A key aid in the creation of ontologies is the use of ontology patterns. Section 7 discusses this and also reports on a subsequent survey specifically into pattern use. Section 8 then records some final comments from respondents, and section 9 draws some conclusions.

2 The survey and the respondents

The survey was conducted during the first three months of 2013 using the *Survey Expression*[3] tool. Responses were obtained using a number of contacts and relevant mailing lists. The latter included: the ontolog-forum[4], the U.K. Ontology Network[5], the Semantic Web for Life Sciences group and the Description Logic group on LinkedIn, lists maintained by the Open Knowledge Foundation[6], and the internal mailing list within the authors' university department. In all there were 118 respondents. In general, respondents only answered a subset of the questions. However, most questions resulted in several tens of responses. The survey aimed to improve the

[1] protégé.stanford.edu
[2] http://ontologydesignpatterns.org/
[3] www.surveyexpression.com
[4] ontolog-forum@ontolog.cim3.net
[5] ontology-uk@googlegroups.com
[6] okfn-{en|scotland|nl}@lists.okfn.org

understanding of how ontology languages and tools are being used. The goal was to use this understanding to identify themes for future research into making ontology use more effective. In line with these aims and goals, the survey requested information about the respondents, their ontologies, their ontology tools, the ontology languages and language features they use, and their use of ontology patterns.

Respondents were asked to categorize themselves by sector. 116 respondents provided this information, with the following distribution: academic (45%); from research institutes (25%); industrial (17%); and other (13%). They were also asked to give their primary application area. All 118 respondents provided this information and the distribution was: biomedical (31%), business (9%), engineering (19%), physical sciences (7%), social sciences (5%); and other (30%). The 'other' category included computer science and information technology and, to a lesser extent, humanities. 115 people responded to a question about the length of time they had worked with ontologies. 62% had over five years' experience and only 6% had less than one year. More detailed information about the respondents, plus other additional material complementing this paper, are provided in Warren (2013).

3 Purposes

Respondents were asked for which purposes they used ontologies. There were eight options plus 'other'. Multiple responses were permitted and there were in all 341 responses from 73 respondents, which reduced to 332 responses from 72 respondents when 'other' was ignored. Thus, ignoring 'other', the average number of responses per respondent was 4.6. Table 1 shows the eight options with the text as used in the survey and the percentage of responses from the 72 respondents. The table also provides a two letter code which is used in the subsequent discussion.

Table 1. Purposes for using ontologies; percentages of 72 respondents

Code	Text in survey	percentage respondents
CM	Conceptual modelling, e.g. formally defining a domain	72%
DI	Data integration, i.e. merging a number of databases	72%
SC	Defining knowledgebase schemas, e.g. as a means of storing and retrieving information	65%
LD	Linked data integration, e.g. linking data from different public knowledgebases	64%
KS	Knowledge sharing, e.g. between individuals in an organisation	56%
HD	Providing common access to heterogeneous data, i.e. providing a common schema for data access	56%
OS	Ontology-based search, i.e. using ontologies to refine search	50%
NL	Supporting natural language processing	26%

A correlation analysis revealed significant[7] positive correlations between all pairs from DI, KS, LD and SC, with two-sided p values ranging from 0.0006 for DI and SC to 0.0296 for DI and KS. These four responses all related broadly to knowledge sharing. The only other significant correlations were a positive correlation between DI and OS ($p = 0.0356$) and a negative correlation between HD and NL ($p = 0.0496$).

Maximal predictive clustering was also performed to categorize the 72 respondents to this question. This is a clustering technique suitable for binary data in which each cluster has a 'predictor' vector with components zero or one. The criterion to be maximized is the total number of agreements between each group member and the group predictor. The criterion value was 432 for two clusters and stabilized at close to 500 for eight, nine and ten clusters[8]. A four group classification (criterion = 465) is discussed here because moving to five clusters only increased the criterion by four and led to a cluster with only five members.

The four categories of users were:

- 16 respondents with a predictor comprising only CM and with an average of 2.2 responses. These are termed *conceptualizers*. Conceptualizers may be interested in using ontologies for modelling, rather than manipulating quantities of data. They might, for example, be using reasoning to identify inconsistencies in a model.
- 12 respondents with a predictor comprising DI, HD, LD, SC and with an average of 4.3 responses. These are termed integrators. Integrators may be more interested in integrating data from various sources, e.g. a variety of databases.
- 11 respondents with a predictor comprising CM, LD, OS, SC and with an average of 3.9 responses. These are termed searchers. Like the integrators, their predictor includes LD and SC. However, whereas the integrators' predictor includes DI and HD, that for the searchers includes CM and OS. Searchers are more likely to be interested in ontological search, e.g. over the linked data cloud. None of the searchers expressed an interest in heterogeneous data. Perhaps they saw the linked data cloud as being rendered homogeneous by the mechanism of RDF.
- 33 respondents with a predictor comprising all the response options except NL. The average number of responses for this group was 6.2. These will be referred to as *multipurpose users*.

4 Languages and tools

Of the 65 respondents to a question about which languages they used, 58 indicated OWL, 56 RDF and 45 RDFS. The other two predefined options, OIL and DAML+OIL received no responses. 11 indicated 'other', which included the Open

[7] Throughout this paper 'significant' is taken to mean at the 95% level.

[8] When each point is in a cluster of one, then the maximum of 576 is achieved, i.e. eight responses x 72 respondents.

Biological and Biomedical Ontology format[9], query languages, plus other more specialist languages.

The dominance of OWL was also indicated by the response to a question about which ontology editors are being used. Respondents were given a choice of 12 editors, plus ‘other’. Multiple responses were permitted. 63 respondents replied to the question and figure 1 shows the tools for which there was more than one response, indicating the split between OWL and non-OWL[10]. OBO-Edit and Neurolex were amongst the ‘other’ category; all the others shown were listed in the questionnaire.

When asked which OWL profiles were used, there were 133 responses from 56 respondents, indicating considerable multiple use of profiles. The range of responses is shown in figure 2. Respondents who used DLs were asked to indicate which DL features they used. The choice of features and the responses are shown in table 2. It is noteworthy that a number of people were using the more specialist features, e.g. the four object property characteristics at the bottom of the list (inverse functional, reflexive, asymmetric and irreflexive). A study into the usability of some of the most commonly used DL constructs by Warren et al. (2014) has used the information in this table as one of the sources to identify which constructs to investigate.

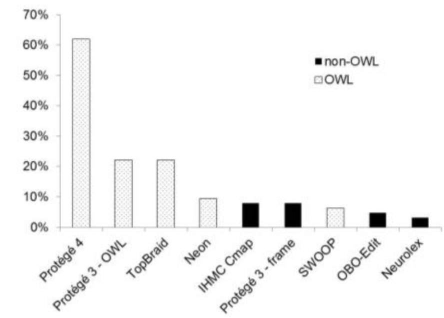


Fig. 1. Usage of ontology editors; percentage of 63 respondents

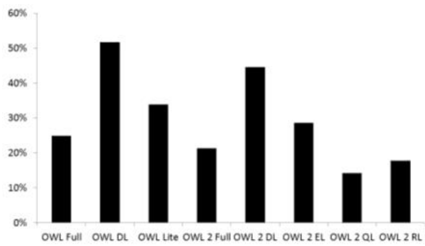


Fig. 2. Usage of OWL profiles; percentage of 56 respondents

When invited to make general comments about ontology languages, there were a number of specific suggestions for language extensions. There were also general comments, e.g. the difficulty of design decisions such as classes versus individuals and classes versus properties; and the difficulty of grasping the implications of open world reasoning – one respondent commented that it would be “great if OWL semantics would (partially) support closed world reasoning”. Another comment seemed to relate to the difficulty of rigorous modelling, stating that it was not always possible to characterize “strongly semantically”.

Amongst the comments on the state of ontology editors were a number referring to the need for the kind of functionality normally found in other system development tools, e.g. auto-completion, version control, and distributed development features. One respondent noted the need for different tools for domain experts and for those

[9] http://www.obofoundry.org/

[10] Note that, because of the possibility of multiple responses, the percentages total to more than 100%. This is true of a number of other figures and tables in this paper.

“formalizing the ontology”. Another commented that “Protégé is not suitable for working together with domain experts”.

Table 2. Usage of DL features; percentage of 47 respondents

DL feature	
object property domain	79%
object property range	77%
disjoint classes	74%
datatype properties	72%
intersection of classes	70%
transitive object properties	68%
object subproperties	68%
union of classes	66%
existential restrictions	66%
inverse object properties	64%
functional object properties	64%
universal restrictions	60%

DL feature	
hasValue restrictions	60%
cardinality restrictions	51%
symmetric object property	51%
functional datatype property	51%
datatype subproperties	49%
complement of a class	47%
qualified cardinality restriction	43%
inverse functional object prop	36%
reflexive object property	30%
asymmetric object property	26%
irreflexive object property	17%

5 Ontologies and ontology sizes

Respondents were asked to list their most commonly used ontologies, to a maximum of five. 69 people responded and the most common responses were: Dublin Core (49% of respondents), FOAF (29%), Dbpedia (19%), the Gene Ontology (17%) and SKOS (16%). After this came a category comprising the respondents’ own ontologies (14%).

For their most commonly used ontologies, they were then asked about the size of those parts of the ontologies with which they actually worked. Specifically, they were asked for the number of classes, individuals, properties, top-level classes and the depth of the hierarchy. This led to 125 responses from 40 respondents. Figure 3 illustrates the enormous range in the size of the ontologies with which respondents were working. Note that the distribution of the number of individuals has a particularly long tail, as to a lesser extent does that for the number of classes.

Many of the ontologies with a very large number of classes were in the biomedical domain. For example, the response in the range ‘1,000,001 to 3 million’ represented the set of ontologies known as the Open Biomedical Ontologies (OBO)[11], which include the Gene Ontology. SNOMED-CT[12] was one of the responses in the range 300,001 to 1 million. Many of the ontologies of depth greater than ten were also in

[11] http://www.obofoundry.org
[12] see http://www.ihtsdo.org/

the biomedical domain. They include, for example, the OBO ontologies. Outside the biomedical domain, CYC[13] was an example of an ontology with depth more than ten.

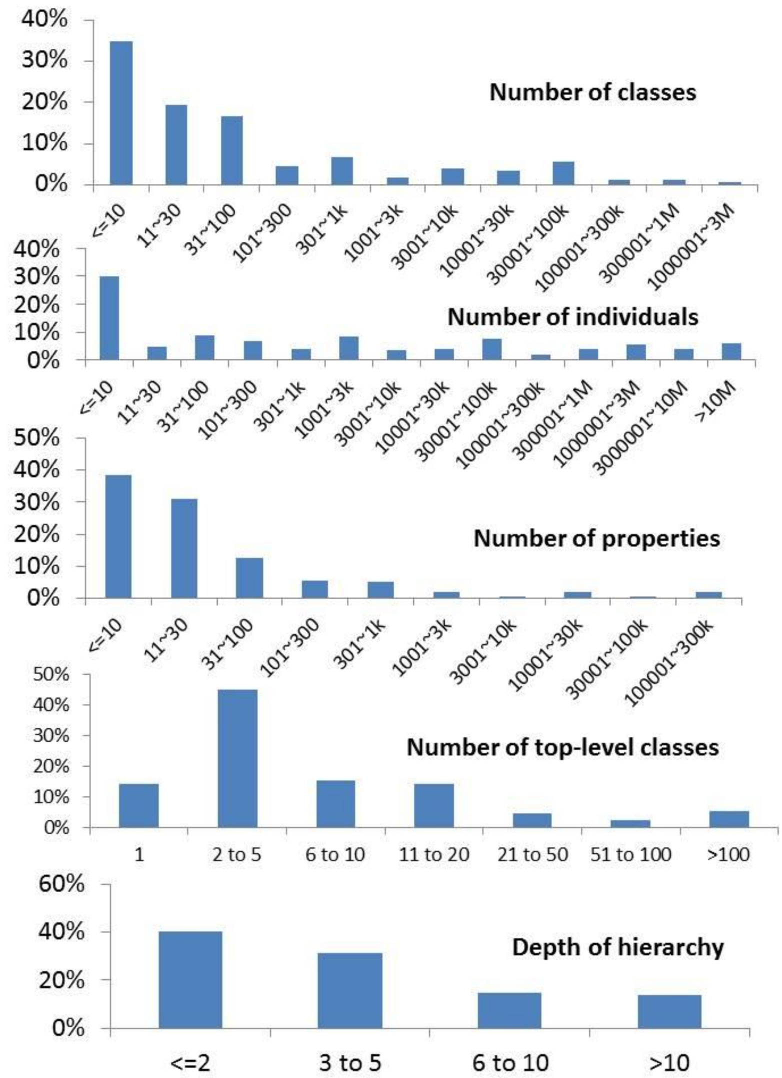


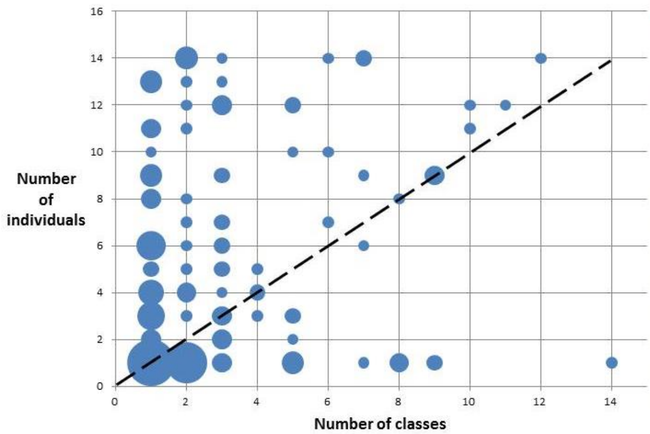
Fig. 3. Ontology size; showing percentage of responses in each size category.

It was hypothesized that the size of ontologies might be related to the category of user, as described in section 3. As already noted, respondents were able to provide the number of classes, individuals, properties, and top-level classes, and the depth for up to five of their commonly used ontologies. The maximum of each of these dimen-

[13] http://www.cyc.com/

sions was taken for each respondent and used as the variate in a series of Kruskal-Wallis one-way analysis of variance tests, the factor being the respondent's category membership. No significant results were obtained, the p-values being: for classes (0.077), individuals (0.332), properties (0.164), top-level classes (0.796) and depth (0.056). It was noticeable that for each measure the conceptualizers' ontologies were on average smaller, except with regard to the number of top-level classes where the conceptualizers were on average the largest. To investigate further, a Kruskal-Wallis test was performed comparing the conceptualizers with the other three categories combined. Significance was achieved for depth ($p = 0.020$), number of classes ($p = 0.040$), but not for number of properties ($p = 0.083$), number of individuals ($p = 0.134$) or number of top-level classes ($p = 0.730$). In summary, the conceptualizers use significantly fewer classes and significantly shallower ontologies than the non-conceptualizers.

A Spearman's rank correlation applied to each pair from the dimensions: number of classes, properties, top-level classes and depth (i.e. excluding number of individuals) showed a high degree of positive correlation; the highest p-value was 0.007 (depth versus number of top classes). The number of individuals, on the other hand, was only significantly correlated with the number of properties ($p < 0.001$). This suggests that the data can be represented by two dimensions; the number of individuals and some representative of the other four dimensions. Figure 4 shows a plot of the number of individuals versus the number of classes. The most striking feature of the plot is the empty area at the bottom right. The ontologies can be regarded as comprising two groups. In one group the number of individuals is greater than ten and in the majority of cases greater than the number of classes. In the other group the number of individuals is in the smallest category of zero to ten, whilst the number of classes occupies a wide range.



scale	range
1	0–10
2	11–30
3	31-100
4	101-300
5	301–1,000
6	1,001-3,000
7	3,001-10,000
8	10,001-30k
9	30,001-100k
10	100,001-300k
11	300,001-1M
12	1,000,001-30M
13	3,000,001-10M
14	> 10M

Fig. 4. Number of individuals versus number of classes; size of each bubble represents number of points; scale is shown at right; dashed line represents equal number of classes and individuals.

Where the respondents providing the data points in figure 4 could be assigned to categories, this was done. The resultant percentage breakdown of the two data groups is shown in table 3 for each of the user categories and overall. These two dimensions are not independent (p = 0.018 on a Pearson's χ^2 test). Note that the vast majority of searchers' ontologies have more than ten individuals, conforming with the intuition that searchers work with knowledgebases of significant size.

Table 3. Split of ontology groups between user categories

	Multipurpose users	Integrators	Conceptualizers	Searchers	Overall
0-10 individuals	28%	48%	37%	9%	30%
>10 individuals	72%	52%	63%	91%	70%

6 Visualization

Respondents were asked which ontology visualization tools they used. Figure 5 shows all those tools for which there was more than one response and indicates which of these tools are incorporated into Protégé. Figure 6 gives the percentage breakdown between the various alternative answers to the question 'how useful do you find visualization?', demonstrating a wide range of views. No significant relationship could be discerned between the perceived usefulness of visualization and the size of ontologies being used. However, the ability of each respondent to describe up to five ontologies makes this analysis difficult. Ideally one would wish to know the attitude towards visualization of a particular ontology. A Kruskal-Wallis test revealed no significant difference in the attitudes of the four categories of users (p = 0.818).

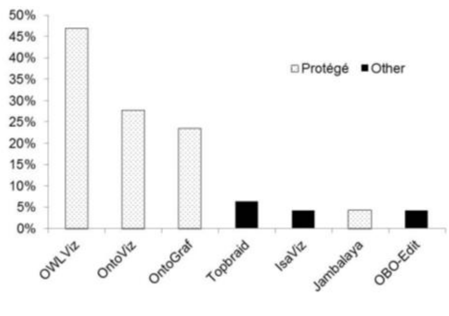


Fig. 5. Usage of ontology visualization tools; percentage of 47 respondents

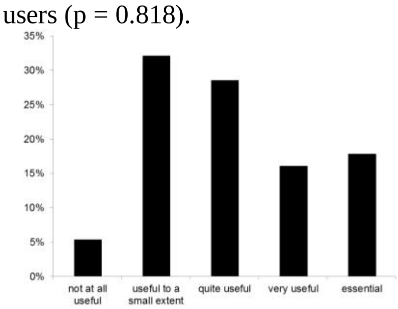


Fig. 6. Perceived usefulness of visualization; percentage of 56 respondents

One respondent wanted to be able to visualize "schema with huge amounts of instance data in order to analyze the effects of changes in real time". Another respondent noted that "visualization is, especially for the end-user, really hard and not task-specific". This echoes comments in other fields, e.g. see Maletic et (2002) al. discussing visualization in software engineering.

7 Patterns

7.1 Original survey

One section of the survey was concerned with ontology patterns. No definition of 'ontology pattern' was provided, leaving respondents free to interpret the concept in any way they chose. The OntologyDesignPatterns.org[14] pattern library provides a classification into structural, correspondence, content, present and lexico-syntactic. Falbo et al. (2013) describe this classification and also an alternative classification approach related to the ontology development phase at which the pattern is used.

There were 35 responses to a question asking from where ontology patterns were obtained. Table 4 shows the percentage of respondents indicating each of the categories. The category 'other' included the OBO library, see section 5, and the W3C. One point of note is the bias in the biomedical community not to use the generic libraries cited or the Protégé wizard. Of the 14 responses from nine people in the biomedical community, none were for the Protégé wizard and one each, from the same respondent, were for the two libraries. A Pearson χ^2 test revealed that this was significant ($p = 0.039$). One biomedical researcher noted "usually I do generate patterns myself".

Respondents were also asked why they used patterns. Table 5 shows the percentage of respondents indicating each of the categories. Two of the three responses in the 'other' category were really restatements of categories provided. The remaining response was to enforce design restrictions, specifically to "enable complex restrictions of context and time of, for example, roles and statuses".

Table 4. Sources of ontology patterns; 35 respondents; 61 responses; %age respondents

Own mental models	46%
Own or colleagues' collections	37%
ODP Portal (OntologyDesignPatterns.org)	34%
ODP public catalogue (http://www.gong.manchester.ac.uk/odp/html)	17%
Protege wizard	14%
Other	26%

Table 5. Reasons for using patterns; 33 respondents; 113 responses; %age respondents

Enforce consistency	61%
Make ontology development easier	61%
Encourage best practice	55%
Make more comprehensible	55%
Reduce modelling mistakes	52%
Speed up ontology development	42%
Restrict language features	9%
Other	9%

[14] http://ontologydesignpatterns.org

Of the 32 respondents to a question about the use of tools for creating, editing and using patterns, 20 used no tools "other than standard ontology editing tools". Five of the respondents used the patterns plug-in to Protégé 4 and the remainder used a variety of tools, some specifically developed for ontologies, others generic. Respondents were also asked how they used patterns, specifically whether they imported patterns or recreated them, possibly modified. The great majority (20) indicated only the latter of these two options, five indicated only the former and four indicated both. There were five responses in the 'other' category, including "fully integrated into the tool" and the use of templates. Taken together, the answers to these two questions indicate the informal way in which patterns are used, with patterns often being recreated with ordinary ontology tools.

Respondents were asked for general comments on their experience with using patterns. One respondent commented that the "best patterns are rather simple, not very complex, basic". A researcher in the biomedical domain expressed the view that there are "seldom some available patterns out there for us to use". This may be because the required patterns are frequently domain-specific rather than generic. Another respondent called for better tool support, stating that "tools should suggest suitable patterns". One comment was about the difficulty of understanding patterns: "initially hard to learn, but provide required functionalities"; this suggests the need for better ways of representing patterns in human-readable form.

7.2 Follow-on patterns survey

A smaller survey was subsequently undertaken specifically to investigate the use of ontology patterns. The survey was broadcast on the same mailing lists as the original survey, and also to some of the respondents to the original survey who expressed an interest in ontology patterns. 13 respondents provided information. A detailed report on the survey results is provided by Warren (2014).

Respondents were again asked how they used patterns. This time a more detailed set of options was provided. Table 6 shows the options and the percentage of respondents indicating each option, out of a total of 13 respondents. Respondents who used pattern creation tools were asked to specify the tools. Responses were: Extreme Design (XD) Tools, see Presutti et al. (2009) and Blomqvist et al. (2010); Tawny OWL, see Lord (2013); Excel and XSLT; and an own unpublished tool.

Table 6. Pattern usage; 13 respondents; 19 responses; %age respondents

Use patterns as examples and recreate modified	54%
Import patterns, e.g. as OWL files	46%
Use patterns as examples and recreate unmodified	31%
Generate with a tool specifically designed for pattern creation	31%

Respondents were asked how they identified the need for a pattern, with the options "by noticing repeated use of identical or similar structures" and "by systematically analyzing the ontologies I work with, e.g. using a tool". Of the 13 responses there were ten in the first category, two in the second and one who indicated both. Thus over three-quarters of the respondents were identifying patterns informally.

Table 7 shows the response to the question "how do you create and store patterns?". The most striking feature is that over half make use of diagrams. It is also interesting that the two respondents from the biomedical domain both used formal languages, but no other technique. Conversely, none of the other eleven respondents used formal languages.

Table 7. Creating and storing patterns; 13 respondents; 21 responses

Technique	**percentage respondents**
Diagrammatically	54%
Using an ontology editor and storing as, e.g. OWL files	31%
Not written down, from memory	23%
Using a formal language, e.g. MOS, in a text editor	15%
Using an informal language, e.g. English	15%
Other ("own UI to own DBMS", "XSLT source", "to application instance data")	23%

Respondents were also asked whether they identified "erroneous or ineffective patterns to avoid (sometimes called antipatterns)". One respondent had a list which was used in teaching "including AND and OR abuse, part as membership, cause as entailment, broader as subsumption, lack of connection between n-ary projections". Other responses included: "properties that have the same name as a class", "modelling roles as subclasses of people", and "use the class hierarchy as a topic hierarchy". One respondent used OOPS[15] which scans an ontology for common pitfalls, see Poveda-Villalón et al. (2012).

Respondents were asked to provide information about their typical patterns. Up to five possibilities were permitted. The specific information requested was: number of classes, number of individuals, number of properties, and domain specificity. For the first three of these the available categories were: 0; 1-2; 3-5; 6-10; 11-30; 31-100; greater than 100. The categories of domain specificity are shown in table 8, which also shows the range of pattern sizes for each level of specificity, based on 17 responses from 11 respondents. This suggests a correlation between domain specificity and pattern size. A Spearman's rank correlation confirmed that there was a significant correlation between the number of classes and domain specificity ($p = 0.015$). There was no significant correlation between specificity and number of individuals ($p = 0.055$) and specificity and the number of properties ($p = 0.136$), nor between any of the three dimensions of size.

[15] http://oeg-lia3.dia.fi.upm.es/oops

Table 8. Relationship between domain specificity and pattern size

	no. of classes	no. of individuals	no. of properties
generic	1 to 10	0	3 to 10
suited to a broad discipline, e.g. anatomy	3 to 10	0 to > 100	1 to 10
suited to a specific discipline, e.g. human anatomy	6 to 100	0 to > 100	3 to 100
specific to my work	6 to > 100	0 to 30	3 to 100

Respondents were asked about the difficulties they experienced using patterns. Two commented on the need for documentation and examples. Other comments included the difficulties of finding the right pattern, and on pattern generation and integration with existing ontologies. One respondent noted that when ontologies are imported, information about patterns is not available. Taken with another comment about the complexity of visualization when several patterns are used simultaneously, this suggests that it would be useful to have editor facilities for viewing patterns embedded in an ontology, be they architectural, content or any other type of pattern.

8 Final comments

At the end of the ontology user survey respondents were asked for any final comments on their experiences with using ontologies. A number of the resultant comments related to the difficulty of modeling with ontologies. One referred to the difficulty of designing classes, which had taken the respondent "many years of learning". Another, working in the biomedical domain, called for a more mature discipline of ontology design: "Ontologies should be built towards use cases and answering biological questions and this is not always the case. Engineering practices in the domain are rarely applied and immature."

Related to the point already made about the need for different tools for ontology specialists and domain experts, one respondent noted "…tool support for non-experts working with ontologies / knowledgebases is generally poor". In a follow-up conversation, one of the respondents noted that visualization needs to be domain-specific. In the same follow-up conversation, the respondent identified two significant problems. For the ontologist, there is the difficulty of understanding the reason for incorrect entailments, e.g. see the work of Horridge et al. (2011) and Nguyen et al. (2012). For the domain specialist the major problem is searching and navigating the ontology.

On the positive side, one respondent commented on the experience of using ontologies: "I couldn't build what I build without them".

9 Conclusions

This survey has revealed a wide variation in the types of ontologies being used, in the purposes for which they are used, and in the user experience. It has identified that users have difficulty with design decisions, e.g. related to the open word assumption. This suggests the need for methodologies and tool features aimed at guiding designers. The use of patterns is often very informal. Again, this calls for methodologies and tool features to assist the user. There is a wide variation in attitudes to visualization. There is scope here for improvements to tools. This includes tools for visualizing patterns, given the importance given to pattern visualization by some of the respondents to the pattern survey.

In designing tools and methodologies, there needs to be more awareness of the specific targeted end-users and of their goals in using ontologies. The importance of distinguishing between ontology specialists and domain experts has already been made. From the data in our survey we have also suggested a split into four categories of user. We do not propose this as the last word in user categorization. Indeed, it should be viewed as part of a tradition of user and application categorization, starting with the generic categorization of Uschold and Jasper (1999) and continuing with the categorizations specific to biology and medicine made by Shah and Musen (2009) and by Stevens and Lord (2009). Our point is that future developments need to be built on a better understanding of the specific requirements of different user groups. As part of this, precisely defined use cases need to be created prior to the development of new tools and methodologies. This will support development and lead to more precise criteria by which to evaluate tools and methodologies.

The survey has shown that DLs are the dominant paradigm for designing ontologies. More needs to be done to understand, and hence mitigate, user difficulties with DLs. The authors have reported an initial study, see Warren et al. (2014). This work is continuing.

References

1. Blomqvist, E., Presutti, V., Daga, E., & Gangemi, A. (2010). Experimenting with eXtreme design. *Knowledge Engineering and Management by the Masses*, 120–134.
2. Falbo, R., Guizzardi, G., Gangemi, A., & Presutti, V. (2013). Ontology patterns: clarifying conceptsa and terminology. Presented at the Workshop on Ontology and Semantic Web Patterns - WOP2013, Sydney, Australia. Retrieved from http://ontologydesignpatterns.org/wiki/WOP:2013
3. Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199 – 220. doi:http://dx.doi.org/10.1006/knac.1993.1008
4. Guarino, N., Oberle, D., & Staab, S. (2009). What is an Ontology? In *Handbook on ontologies* (pp. 1–17). Springer.
5. Horridge, M., Bail, S., Parsia, B., & Sattler, U. (2011). The cognitive complexity of OWL justifications. *The Semantic Web–ISWC 2011*, 241–256.
6. Katifori, A., Halatsis, C., Lepouras, G., Vassilakis, C., & Giannopoulou, E. (2007). Ontology visualization methods—a survey. *ACM Computing Surveys (CSUR)*, *39*(4), 10.
7. Khan, M. T., & Blomqvist, E. (2010). Ontology design pattern detection-initial method and usage scenarios. In *SEMAPRO 2010, The Fourth International Conference on Advances in Semantic Processing* (pp. 19–24). Retrieved from http://www.thinkmind.org/index.php?view=article&articleid=semapro_2010_1_40_50071

8. Lord, P. (2013). The Semantic Web takes Wing: Programming Ontologies with Tawny-OWL. *arXiv Preprint arXiv:1303.0213*. Retrieved from http://arxiv.org/abs/1303.0213
9. Maletic, J. I., Marcus, A., & Collard, M. L. (2002). A task oriented view of software visualization. In *Visualizing Software for Understanding and Analysis, 2002. Proceedings. First International Workshop on* (pp. 32–40). IEEE.
10. Möller, R., & Haarslev, V. (2009). Tableau-based reasoning. In *Handbook on Ontologies* (pp. 509–528). Springer.
11. Motik, B. (2009). Resolution-Based Reasoning for Ontologies. In *Handbook on Ontologies* (pp. 529–550). Springer.
12. Nguyen, Power, Piwek, & Williams. (2012). Measuring the understandability of deduction rules for OWL. Presented at the First international workshop on debugging ontologies and ontology mappings, Galway, Ireland. Retrieved from http://oro.open.ac.uk/34591/
13. Poveda-Villalón, M., Suárez-Figueroa, M. C., & Gómez-Pérez, A. (2012). Validating ontologies with oops! In *Knowledge Engineering and Knowledge Management* (pp. 267–281). Springer.
14. Power, R., & Third, A. (2010). Expressing OWL axioms by English sentences: dubious in theory, feasible in practice. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters* (pp. 1006–1013).
15. Presutti, V., Daga, E., Gangemi, A., & Blomqvist, E. (2009). eXtreme design with content ontology design patterns. In *Proc. Workshop on Ontology Patterns, Washington, DC, USA*. Retrieved from http://ceur-ws.org/Vol-516/WOP2009_Proceedings.pdf#page=90
16. Shah, N., & Musen, M. (2009). Ontologies for formal representation of biological systems. In *Handbook on Ontologies* (pp. 445–461). Springer.
17. Staab, S., & Studer, R. (2010). *Handbook on ontologies*. Springer.
18. Stevens, R., Egaña Aranguren, M., Wolstencroft, K., Sattler, U., Drummond, N., Horridge, M., & Rector, A. (2007). Using OWL to model biological knowledge. *International Journal of Human-Computer Studies*, *65*(7), 583–594.
19. Stevens, R., & Lord, P. (2009). Application of ontologies in bioinformatics. In *Handbook on Ontologies* (pp. 735–756). Springer.
20. Uschold, M., & Jasper, R. (1999). A framework for understanding and classifying ontology applications. In *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5), Stockholm, Sweden*.
21. Vigo, M., Jay, C., & Stevens, R. (2014). Design insights for the next wave ontology authoring tools. Presented at the ACM SIGCHI Conference on Human Factors in Computing Systems, CHI 2014, Toronto, Canada: ACM Press.
22. Vigo, M., Jay, C., & Stevens, R. (2014). Protege4US: harvesting ontology authoring data with Protege. HSWI2014 - Human Semantic Web Interaction Workshop, Crete.
23. Warren, P. (2013). *Ontology Users' Survey - Summary of Results* (KMi Tech Report No. kmi-13-01). Retrieved from http://kmi.open.ac.uk/publications/pdf/kmi-13-01.pdf
24. Warren, P. (2014). *Ontology patterns: a survey into their use* (KMi Tech Report No. kmi-14-02). Retrieved from http://kmi.open.ac.uk/publications/pdf/kmi-14-02.pdf
25. Warren, P., Mulholland, P, Collins, T, & Motta, T. (2014). The usability of Description Logics: understanding the cognitive difficulties presented by Description Logics (pp. 550–564). Presented at the ESWC 2014, Crete: Springer.