Implementation and Evaluation of a Semantics-based User Interface for Web Gazetteers

Krzysztof Janowicz
Institute for Geoinformatics
University of Münster, Germany
janowicz@uni-muenster.de

Mirco Schwarz
Institute for Geoinformatics
University of Münster, Germany
mirco.schwarz@uni-muenster.de

Marc Wilkes
Institute for Geoinformatics
University of Münster, Germany
marc.wilkes@uni-muenster.de

ABSTRACT

As the Web moves towards structured data, new communication paradigms and user interfaces become necessary to support users beyond simple keyword search. In contrast to applications for expert users, these interfaces should hide the complexity of ontology-based systems but still offer their reasoning capabilities. Using Web gazetteers as application area, we discuss how subsumption and similarity reasoning can be integrated into modern Web interfaces and explain the made design decisions. To evaluate our approach, the results from a first human participants test are presented.

Author Keywords

Information Retrieval, Similarity, Geographic Feature Types

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces—*Graphical user interface*; H.3.3 Information Storage and Retrieval: Information Search and Retrieval

INTRODUCTION AND MOTIVATION

The major benefit of semantics-based information retrieval is the possibility to reach beyond simple keyword search, and hence to support the user in answering complex queries. While there are several promising approaches to semantics-based retrieval such as subsumption and similarity reasoning, there is surprisingly little work on how to hide the additional complexity (e.g., the underlying ontology) from the user. In addition, the results have to be presented in a way that they are also comprehensible for non-experts. The extended retrieval capabilities of the semantic Web should not be slowed down by more complex user interfaces.

Based on our previous work [14, 13], we discuss the implementation and evaluation of a semantics-based user interface for Web gazetteers. It makes use of subsumption and similarity reasoning to support the users in navigating through geographic feature types. Our idea is to combine techniques developed in conjunction with the social Web 2.0, such as search-while-you-type forms (also known as autocompletion) and tag clouds, with the reasoning capabilities of the semantic Web.

GAZETTEERS AS APPLICATION AREA

Gazetteers are place name directories. Each gazetteer record consists at least of a triple (N, F, T) where N corresponds to one or more place names, F represents one or more geographic footprints (i.e., locations), and T is the type (class) of the described feature (i.e., the representation of a real world geographic entity). Hence, gazetteers support at least two types of queries. First, they map between place names and respective footprints: $N \rightarrow F$; and second, they map between names and geographic feature types: $N \rightarrow T$. Note that a named geographic place is an abstract individual defined to refer to a physical region in space and categorized according to commonly agreed characteristics. The place name is a handle to support communication. Hence, *Place* is a social construct of interest for a particular community during a specific time span. A place may change its name, location, and type (e.g., from *city* to *ruin*) over time [10, 12].

Gazetteers are key components of all georeferenced information systems, including GIScience applications in many diverse fields, Web-based mapping services such as Google Maps^{TM}, and an increasing number of Web 2.0 applications such as plazes.com. Gazetteers are crucial for geoparsing where references to geographic locations (by their names) are recognized in texts and converted to coordinate references [10]. Most gazetteers are either accessed via a Web page or through an application programming interface (API). The geographic feature type is selected from a typing scheme, in most cases from a feature type thesaurus. The type of the described geographic feature is of special importance. For instance, a search for places named Berlin might return more than hundred places with a lot of different feature types varying from *capital* (located in Germany), over *administrative area* (a state in Germany), to *stream* (located in Colombia).

One of the major Web gazetteers is the Alexandria Digital Library (ADL) Gazetteer[1]. The ADL Web interface depicted in figure 1 includes a map to narrow down the search to a particular spatial extent, a form to enter the place name (or parts of it), as well as a scrolldown list with about 200 (or 1200

[1] http://www.alexandria.ucsb.edu/gazetteer/

Workshop on Visual Interfaces to the Social and the Semantic Web (VISSW2009), IUI2009, Feb 8 2009, Sanibel Island, Florida, USA.

if one includes the non-preferred) geographic feature types. To learn about the definitions of these types, the user has to switch to the ADL Feature Type Thesaurus (FTT) and read the (often ambiguous and brief) plain text descriptions. The thesaurus also contains information about super types and some manually selected related types (with many types remaining unrelated to others). Some gazetteers try to take up the participantion idea of the Web 2.0 and allow for user defined places and types. In both cases, the navigation through and understanding of the types is a major shortcoming [12].

The following example illustrates the navigation between feature types in ADL. A user selects *lakes* from the feature type scrolldown list using the ADL gazetteer Web interface. *Lakes* is a narrower term [2] of *hydrographic features* which is a top term in the FTT hierarchy. As the user cannot find a particular geographic feature, she decides to change the type to search for reservoirs. Unfortunately, *reservoirs* is not a narrower term of *hydrographic features*, and therefore she looks up the textual definition of *lakes* in the thesaurus. Here, *reservoirs* is specified as related term [2] and she can click on the link and read the definition of *reservoirs*. It is defined as narrower term [2] of *hydrographic structure*. This in turn is a narrower term of *manmade features* which is another top term in the ADL FTT.

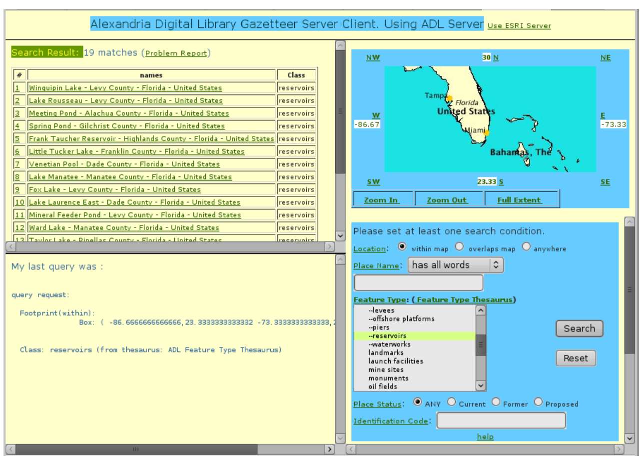

Figure 1. The Alexandria Digital Library Gazetteer Web interface showing a search for reservoirs in Florida.

As discussed by Janowicz and Keßler [12], semantics-based user interfaces in conjunction with geographic feature type ontologies would support the user in browsing through and searching in gazetteers. The alternative interface proposed in this paper aims at hiding the feature type scrolldown list (see figure 1) from the user. Instead, we propose a search-while-you-type form as well as a combination of horizontal and vertical search to improve the navigation between (related) feature types. These types will be determined based on their formal specifications. Hence, users do not have to look up potentially relevant types in the FTT by hand.

SEMANTICS-BASED INFORMATION RETRIEVAL

In contrast to mere syntactic search, semantics-based information retrieval takes the underlying conceptualization (or attribution) into account to improve searching and browsing through structured data. Three approaches can be distinguished: 1) those based on classical subsumption reasoning, 2) those based on so-called non-standard inference techniques such as computing the least common subsumer (*lcs*) or most specific concept (*mcs*) [18, 17], and finally 3) those approaches based on computing semantic similarity[2] [5, 6, 3, 14]. As depicted in figure 2, subsumption reasoning can be applied to vertical search, while similarity works best for horizontal search and should not be applied to compare sub and super types.

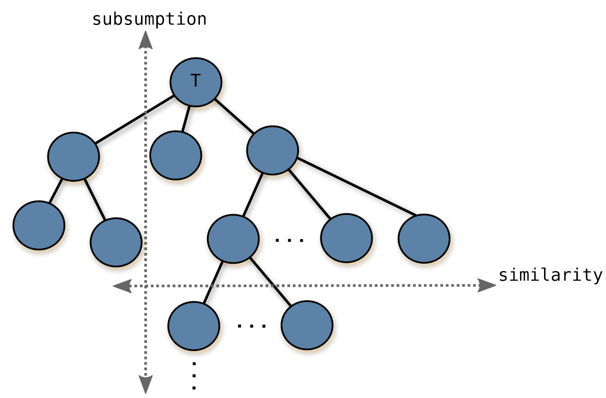

Figure 2. Vertical and horizontal retrieval within an ontology.

Formally, the result set for a user's query using subsumption reasoning is defined as $\mathbb{RS} = \{C \mid C \sqsubseteq C_s\}$; where C_s is the search (or query) concept specified by the user. As each concept returned to the user is a subsumee of the search concept, it meets the user's search criteria. This makes selecting an appropriate search concept the major challenge for end users. In fact, the search concept is an artificial construct and not necessarily the searched concept [15]. If it is too generic, i.e., at the top of the hierarchy, the user will get a large part of the queried ontology back as (unsorted) result set. In contrast, if the search concept is too specific, the result set might be empty.

In case of similarity-based retrieval, the result set is defined as $\mathbb{RS} = \{C_t \mid sim(C_s, C_t) > t\}$; where C_t is the compared-to target concept and t a threshold defined by the user or application [14, 15]. In contrast to subsumption reasoning, the search concept is the concept the user is really searching for (no matter if it is part of the queried ontology or not). As the concepts are compared for overlap (of their definitions or extensions), similarity is more flexible than subsumption reasoning, but it is not guaranteed that the results match all of the user's search criteria. Note that the similarity estimations between search and target concepts are asymmetric which is important for information retrieval. Usually, the results of a

[2] We are focussing on measures which define similarity as degree of conceptual overlap here. Other theories consider similarity as (inverse) distance within a multi-dimensional space, or as a set of transformations [7] from the search to the target concept.

similarity query are presented to the user as an ordered list with descending similarity values.

Summing up, the benefits similarity offers during the retrieval phase, i.e., to deliver a flexible degree of (conceptual) overlap to a searched concept, stands against shortcomings during the usage of the retrieved information, namely that the results not necessarily fit the user's requirements.

To overcome these shortcomings, similarity theories such as SIM-DL combine subsumption and similarity reasoning. In a first step, the context of discourse [11] is defined by introducing a so-called context concept such that $\mathcal{C}_d = \{C_t | C_t \sqsubseteq C_c\}$. Consequently, in the next step only such concepts are compared for similarity which are subconcepts of C_c. This way, the user can specify some minimal characteristics all target concepts need to share. As depicted in figure 3, C_{t_i} and C_{t_j} are compared for similarity to C_s while C_x is not. Note that for reasons of simplification, the figures 2 and 3 show a single hierarchy, while similarity takes all role-filler pairs (e.g., the fact that a river has a spring as its origin $\exists\, hasOrigin.Spring$) into account to compute conceptual overlap.

With regard to semantics-based user interfaces, subsumption reasoning can be applied to include lakes in the result set if the user is searching for (the string) *waterbody*. Mcs and lcs can be used to implement a query-by-example in where the result set is computed out of a set of prototypical individuals (e.g., waterbodies) selected by the user [26, 15]. Similarity measures can be implemented to propose alternative or supplementary results such as reservoirs if the user is searching for (the string) *lake*. However, to propose meaningful alternatives the underlying similarity measure needs to be cognitively plausible, i.e., the rankings returned by such a measure have to correlate with human similarity judgments for the same set of compared individuals or concepts based on the same definitions[3]. This has been shown for several measures such as MDSM [19] and SIM-DL [14, 13]. As these similarity measures compare DL concepts defined within ontologies, the question what is similar to *Lake* depends on the formal definition (and does not necessarily match the mental conceptualization of a particular human user).

SEMANTICS-BASED USER INTERFACES

While there are specifications on how to display RDF triples to the user [4, 20] and various semantic Web browsers such as *PowerMagpie*[4], there is little work on how to integrate the reasoning capabilities of the semantic Web within end user interfaces for information retrieval. Schraefel and Karger [22] argue against the assumption that data on the semantic Web should be displayed to end users as graphs simply because graphs are used for the computational representation (as data model). As in contrast to HTML, XML (and therefore RDF and OWL) separates design from content, the visualization should depend on the task to solve instead of the content or the type of data. Van Kleek and colleagues [24] point to the obstacles users face when diving into the semantic Web and introduce an agenda on how to improve knowledge creation and access for end users. Van Ossenbruggen et al. argue why it is difficult to evaluate semantics-based user interfaces [25]. Basically, one needs to distinguish between the quality of the underlying data, the involved reasoning engine, and the user interface as such. In this work, we focus on the user interface. Additional information on the underlying data and ontologies is discussed in [12], while the similarity reasoning is described in [14].

The Atom interface [21] realizes a circular layout to support users in navigating through semantic Web data. One example for a combination of Web 2.0 interaction and semantic Web techniques is given by Heath and Motta [8], who describe how non-experts can publish and consume RDF data by means of tags in a reviewing Website. Whereas their work deals with tagging data and proposing related tags, we present an approach to interact with data classified in a taxonomy (of feature types) and focus on similarity and subsumption instead of relatedness. A promising approach to integrate semantic autocompletion for geographic locations is described by Hildebrand et al. [9]. They propose an interface which allows for searching geographic places by name where the suggestions are grouped by country or feature type. Instead of focusing on a search by name, in this paper, we implement an additional feature type parameter, allowing for a more restrictive search of places but also demanding for an intuitive feature type selection process. An approach to ease concept selection is presented by Sinkkilä et al. [23]. They present an interface that combines semantic autocompletion with cross-ontological context navigation. Whereas Sinkkilä et al. chose to visualize the concept hierarchy as well as related concepts, we suggest to hide the taxonomic complexity from the user.

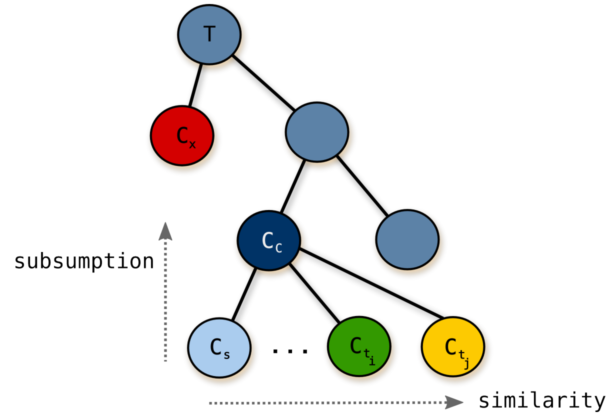

Figure 3. Combining subsumption and similarity reasoning for information retrieval. C_s is the search concept, C_{t_i} and C_{t_j} are the compared-to target concepts, while C_x is not a subsumee of the context concept C_c, and hence is not compared for similarity to C_s.

[3]This is a weaker requirement than cognitive adequacy which would imply that the computational similarity theory reproduces the process of human similarity reasoning.

[4]http://powermagpie.open.ac.uk/

IMPLEMENTATION

In order to demonstrate and evaluate how semantics-based information retrieval can be integrated in Web interfaces for end users, we have implemented a prototypical interface for the ADL gazetteer. The interface can access the records stored in the gazetteer via the ADL API. To implement the search-while-you-type form as well as geographic feature type recommendations, the prototype makes use of AJAX (Asynchronous JavaScript and XML) which allows to update certain parts of the Web interface without the need to reload the whole page.

As depicted in figure 4, the gazetteer Web interface allows for specifying three search parameters: place name, geographic feature type, and location of the searched place. As argued above, while specifying the name and location of a searched place is straightforward, choosing the correct feature type is more difficult, especially in combination with the other parameters. An example is shown in figure 4; note that most of the places categorized as reservoirs are named lake or pond, hence they would not appear in the result list if the user would search for geographic features of type *Lake*.

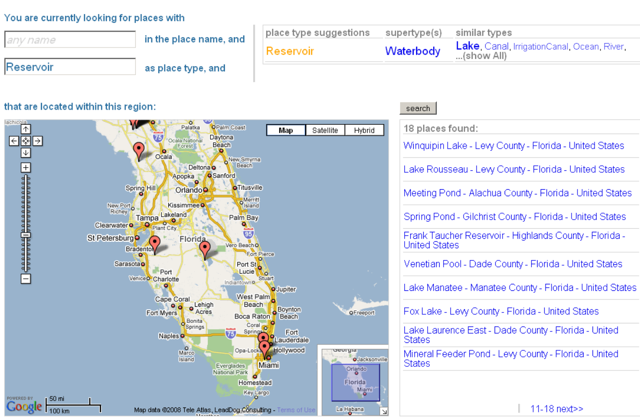

Figure 4. The semantics-based Web interface for the ADL gazetteer showing the search result for reservoirs in southern Florida.

When the users begin to enter characters in the feature type input field a table appears showing suggested types, starting with the characters already entered in the input field. Therefore, users can immediately see which feature types are supported by the application without browsing the feature type thesaurus (or ontology) manually. In addition, the table provides information about similar types and super types (which are computed by the SIM-DL similarity server[5] at query time). These recommendations support the user in browsing the underlying feature type ontology (vertically and horizontally; see figure 3).

To display the similarity rankings in the Web interface font size and opacity are used. This is known from tag clouds, as used by several blogs or Web 2.0 recommendation portals, and should therefore be familiar for the user. The font size and opacity visualize the ranked order of the similarity values, where a big font size and high opacity indicate a high similarity. Each feature type entry in the table is a hyperlink. If a user clicks on one of the suggested feature types, it is selected as search parameter and the similar and super types are computed thereupon. The color of the chosen feature type changes to orange to indicate that this type is part of the new query.

The features resulting from a specific query are displayed at the right side of the interface and on the map. Clicking on these links or map markers opens a text box which contains the information from the ADL gazetteer (e.g., names, footprints, broader administrative units).

Figure 4 shows the result of an exemplary search for places of type *Reservoir* located in southern Florida (without name restriction). Using the interface, the search could be broadened by clicking on the supertype *Waterbody*, leading to a larger result set (including the results for *Reservoir*). In case the search for reservoirs did not yield the desired result, the user could also decide to search for similar geographic feature types such as *Lake* or *Canal*.

To point out why a semantics-based interfaces is valuable for end users, imagine the following example. One of the search results shown in figure 4 is *Lake Manatee*. Due to its name, a user searching for this specific place might be tempted to deduce that it is of type *Lake*. Specifying a query as shown in figure 6 seems obvious, but will not return *Lake Manatee*, since it is of type *Reservoir*. However, facing an empty result set, the user could decide to either broaden the search by selecting the supertype *Waterbody*, or try out a similar type to *Lake*. In the latter case, searching for *Reservoir*, which is most similar to *Lake*, yields the desired result.

EVALUATION

In this section we discuss first results and lessons learned from a human participants test carried out using the new interface.

Our hypothesis was that an interface combining subsumption and similarity to support the user in selecting appropriate feature types would be more effective than interfaces solely based on either subsumption or similarity reasoning. We did not compare our interface to the original ADL interface. This is for the following reasons:

- The original ADL interface makes use of the ADL Feature Type Thesaurus while the new interface is based on a feature type ontology which contains only hydrographic feature types so far [12]. Additionally, we do not distinguish between preferred and non-preferred concepts and have one root node instead of six.
- The new interface supports search-while-you-type, while the ADL interface does not. Consequently, the participants would have to scroll through the whole ADL feature type list – while only some of these types are relevant for the test.

[5]The SIM-DL similarity server as well as the interface are free and open-source software and can be downloaded at http://sim-dl.sourceforge.net/.

- The ADL interface allows for multiple comparison operators for the place names (*has any words*, *has all words*, *has phrase*, *equals*, *matches pattern*). To reduce complexity for end users, the new interface only supports *has all words* for comparison.

- The ADL interface map does not support map markers as used in the Google Maps™mashup in the new interface to narrow down the search to a particular map extent.

To test the hypothesis we developed three versions of the semantics-based interface, the combined, a similarity-based and a subsumption-based version. While the subsumption-based interface shows super and sub types, the combined interface only lists the super types (see figure 5). The underlying assumption was that users will select specific types rather than general ones and can use the super types to broaden their search. Strictly speaking, this could be another hypothesis for further testing. Showing similar, super, and sub types may overload the interface and overstrain the user.

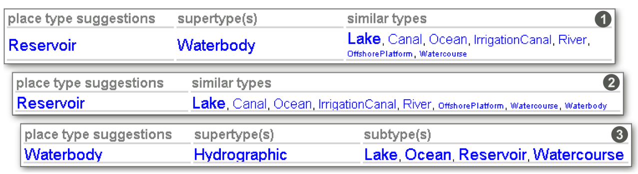

Figure 5. The geographic feature type recommendation part of the interfaces: (1) combined interface, (2) similarity-based interface, (3) subsumption-based interface.

To verify the hypothesis we investigated how many tasks were successfully solved per interface and how many user interactions (e.g., clicks) were necessary. As the interface acts as front-end for the ADL gazetteer (whose response times differ), the time needed to solve the tasks cannot be taken as criterion. A total of 30 participants performed the test[6], i.e., each interface was tested by 10 participants. To ensure that the participants understand how the interfaces work, a video was shown which stepwise presents how to solve an introductory task:

Ems-Task

- *Find the river Ems in Germany. Make it the only result shown in the 'found places' list. Try using just the 'place name' field. Try using both fields, 'place name' AND 'place type'.*

- *Find canals connected to the Ems. Try using just the 'place name' field. Try using both fields, 'place name' AND 'place type'. (Note that the number of 'found places' differs. Do you see why?)*

Next, the participants were asked to solve the task on their own. Afterwards, four additional tasks (with sub tasks) were given to the participants. They were asked to speak out loud while filling in the test and each test was recorded on video. To make sure that the tasks are not biased towards subsumption or similarity reasoning, the questionnaire was designed in a way that it involves both kinds of reasoning as well as interaction with the map and the place name field. Finally, the participants could give general comments on whether they liked the interface and how it could be improved.

As depicted in figure 4, all interfaces were designed in a way that the search parameters form a sentence of the type

> *You are currently looking for places with … in the place name, and … as place type, and that are located within this region: ….*

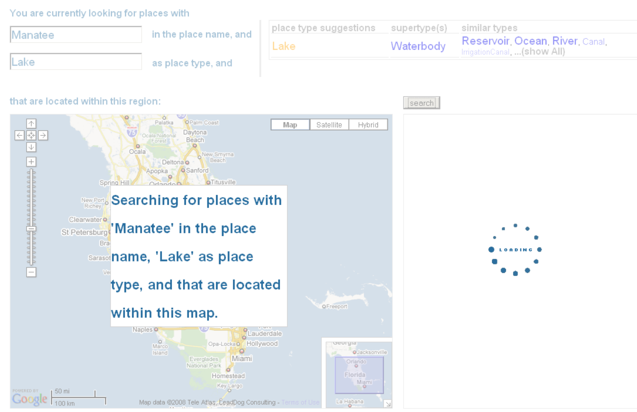

Figure 6. The interface shows the sentence resulting from the user's search parameters while querying.

During a series of pre-tests it turned out that many participants are so accustomed to single keyword search fields (as used in most common Web search engines[7]) that they put the place name as well as the feature type into the place name field. This was still the case when they were asked to speak out the query before submitting it. Therefore, the interfaces were redesigned to display the sentence resulting from the entered search parameters in the map area while loading (and the participants were still asked to speak out the query). This reduced the number of wrong queries.

Analyzing the first results from our human participants test, it turns out that from a total of 40 tasks per interface, 72.5% were solved completely (i.e., including all sub tasks) using the combined interface. The participants using the similarity-based interface were able to solve 65% of the tasks, while 62.5% of the tasks were solved using the subsumption-only interface.

As depicted in figure 7, there is no clear difference in the total number of user interactions necessary to solve the tasks. However, the inter-quartile range differs clearly. It increases from the combined interface over the subsumption-based to the similarity-based interface.

[6]With an age ranging from 21 to 30 years, 9 were female and 21 male (most of them students of either geoinformatics or computer science). The questionnaire and the introductory video can be downloaded at http://sim-dl.sourceforge.net/downloads/.

[7]In contrast to upcoming semantics-based search engines such as Powerset (http://www.powerset.com/).

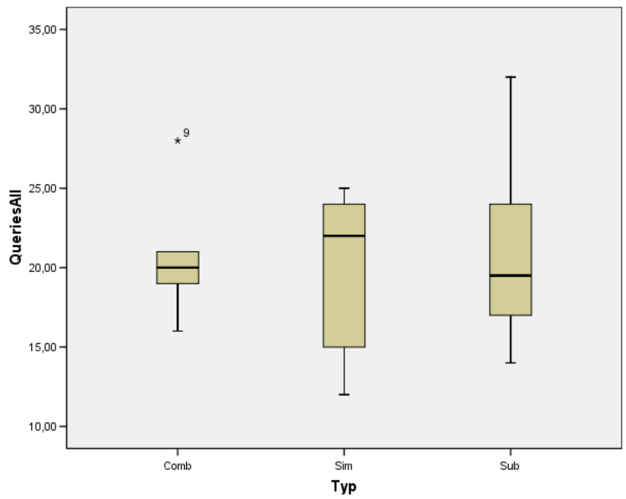

Figure 7. The box plot shows the median as well as the 0.25 and 0.75 quartiles for the number of total user interactions (clicking, typing in words, following suggestions, interacting with the map) per interface type. *Comb* is the combined interface using subsumption and similarity, while *Sim* only shows similar types as suggestions and *Sub* only sub and super types.

While these first results support our hypothesis, the test also points so some aspects which let us rethink the design of the Web interface as well as the number of user interactions as quality measure. Comparing manually entered versus clicked feature types, intuitively one would expect that the more types are suggested, the more will be clicked instead of being manually entered. Hence, the number of manually entered feature types should be lower in the combined interface compared to the other ones. This turns out to be a wrong assumption. In fact, the ratio between clicked and manually entered types was about one to one in case of the combined and the similarity interface, while being higher for the subsumption-based interface. To understand these results, we reviewed the videos taken from the participants during the test. It turns out that several participants had entered the type manually after seeing it in the suggestion table. This points towards two difficulties. First, we have to investigate whether this was caused by the interface design, i.e., by the positioning of the forms and the table (for instance using eye tracking). Second, this makes it more difficult to understand to which degree the suggested types were helpful to the user in selecting the next type to be displayed.

After completing the tasks, the participants were asked to answer additional questions about the design and functionality of the interface. While the participants liked the interfaces in general and rated the type suggestion table to be useful (giving both the second best median rank on a 5-level Likert scale), it turns out that some of the participants still had problems dealing with two input fields. A given suggestion for improvement was to have just one input field and automatically identify keywords representing feature types. Some participants recommended to cache previous queries, allowing for an easy comparison of several result sets using the browser's navigation (back and forward) buttons.

In the current version of the interface, only ten entries of the result set and their corresponding map markers are shown at a time. Although it is possible to go forth (or back) to see the next (or previous) ten entries, several participants stated that they almost overlooked the fact that there are more than ten result entries. As a consequence, they proposed to show all markers on the map, where those currently not listed in the sidebar could be greyed out. Some participants also requested that features of similar types should be directly displayed on the map using colored markers to visualize their similarity instead of the suggestion table.

Finally, a few participants mentioned that upon sending a query to the server and reading the query summary, they realized a mistake in one of the search parameters, and hence missed a button to cancel the request. As this was not possible, the wrong query was counted as user interaction during the test.

CONCLUSIONS AND FURTHER WORK

In this paper we discussed results from testing the integration of subsumption and similarity reasoning into end user interfaces to support horizontal and vertical retrieval. While a combination of both seems to be the most promising approach, it is not surprising that there is a gap between being able to implement more *intelligent* user interfaces and making them intuitively useable [25]. One solution would be to hide even more of the complexity (e.g., the suggestions table) from the user by directly displaying features of very similar types. However, this may be interpreted as paternalism and reduce the user's trust in such interfaces. For instance, while reservoirs and lakes might be similar enough for most tasks, some application areas such as water management may require a clear distinction between both. The extended retrieval capabilities of the semantic Web should be used to offer support, while the final decision should still be up to the user. This leads to the question of how to incorporate contextual information in the user interfaces and to which degree context influences the resulting suggestions [1, 16, 11]. From the viewpoint of the social Web, one could also think of allowing human users to add similar types by hand to improve the suggestions.

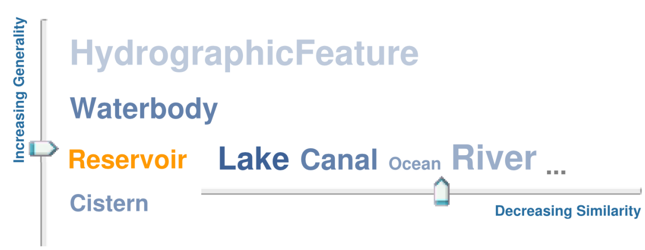

Figure 8. Conceptual design for the feature type suggestions allowing for vertical and horizontal navigation.

Based on our experiences and the test results, we designed a new navigation component which will be integrated in the next version of the Web interface (figure 8). This component will replace the feature type suggestion table in the upper right part of the interface. It will appear as soon as a feature type is selected or typed in by the user. The suggestion table itself will appear just below the feature type input field as usually implemented in autosuggestion forms. The new component visualizes explicitly the difference between horizontal and vertical navigation. A vertical slider can be used to specify more general or more specific feature types than the searched type. The result set only contains places that are of a type under the slider. The horizontal slider allows for including similar types in the result set. This integrates the idea of a similarity threshold which was discussed in the semantics-based information retrieval section. Here, only those places are included in the result set that are to the left of the slider. As in the presented user interface, less opacity indicates decreasing similarity. The use of font size, however, differs. Whereas previously the font size also represented the similarity value, within the conceptual navigation component the font size indicates the number of features found within the map extent for that specific type; where a larger font size points to a higher number of found places.

In figure 8, the immediate super type of *Reservoir* is *Waterbody*, although another immediate super type of *Reservoir* specified in the underlying ontology is *Manmade*. Allowing for generalizing also towards *Manmade* now yields two problems. First, the complexity of the navigation component would increase, ending up in a graph-like representation of the ontology (which is the opposite of our design goal). Second, generalizing towards multiple super types will result in a large and diverse result set. That is why we propose to only generalize along one inheritance path. Deciding which path of inheritance is shown in the navigation component depends on the context concept C_c whose integration in the interface is an open issue.

ACKNOWLEDGEMENTS

The comments from three anonymous reviewers provided useful suggestions to improve the content and clarity of the paper. This work is funded by the *Semantic Similarity Measurement for Role-Governed Geospatial Categories (SimCat)* project granted by the German Research Foundation (DFG Ra1062/2-1), as well as the International Research Training Group on *Semantic Integration of Geospatial Information* (DFG GRK 1498).

REFERENCES

1. R. Albertoni and M. D. Martino. Semantic Similarity of Ontology Instances Tailored on the Application Context. In A. K. Dey, B. Kokinov, D. Leake, and R. Turner, editors, *ODBASE - OTM Conferences*, volume 4275 of *Lecture Notes in Computer Science*, pages 1020–1038. Springer, 2006.
2. ANSI/NISO. *ANSI/NISO Z39.19 - 2005 Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. NISO Press, 2005.
3. R. Araújo and H. S. Pinto. Towards Semantics-Based Ontology Similarity. In P. Shvaiko, J. Euzenat, F. Giunchiglia, and B. He, editors, *Proceedings of the Workshop on Ontology Matching (OM2007) at ISWC/ASWC2007, Busan, South Korea*, 2007.
4. C. Bizer, R. Lee, and E. Pietriga. Fresnel - Display Vocabulary for RDF. Technical report, The World Wide Web Consortium (W3C), June 2005.
5. A. Borgida, T. Walsh, and H. Hirsh. Towards Measuring Similarity in Description Logics. In *Proceedings of the 2005 International Workshop on Description Logics (DL2005)*, volume 147 of *CEUR Workshop Proceedings*. CEUR, Edinburgh, Scotland, UK, 2005.
6. C. d'Amato, N. Fanizzi, and F. Esposito. A Semantic Similarity Measure for expressive Description Logics. In *CILC 2005, Convegno Italiano di Logica Computazionale*, Rome, Italy, 2005.
7. R. L. Goldstone and J. Son. Similarity. In K. Holyoak and R. Morrison, editors, *Cambridge Handbook of Thinking and Reasoning*. Cambridge University Press, 2005.
8. T. Heath, E. Motta, T. Heath, and E. Motta. Ease of Interaction Plus Ease of Integration: Combining Web 2. 0 and the Semantic Web in a Reviewing Site. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1):76–83, 2008. Semantic Web and Web 2.0.
9. M. Hildebrand, J. van Ossenbruggen, A. Amin, L. Aroyo, J. Wielemaker, and L. Hardman. The Design Space of a Configurable Autocompletion Component. Technical report, CWI technical report INS-E0708, November 2007.
10. L. L. Hill. *Georeferencing: The Geographic Associations of Information (Digital Libraries and Electronic Publishing)*. The MIT Press, 2006.
11. K. Janowicz. Kinds of Contexts and their Impact on Semantic Similarity Measurement. In *5th IEEE Workshop on Context Modeling and Reasoning (CoMoRea'08) at the 6th IEEE International Conference on Pervasive Computing and Communication (PerCom'08)*, pages 441–446, Hong Kong, March 2008. IEEE Computer Society.
12. K. Janowicz and C. Keßler. The Role of Ontology in Improving Gazetteer Interaction. *International Journal of Geographical Information Science*, 10(22):1129–1157, 2008.
13. K. Janowicz, C. Keßler, I. Panov, M. Wilkes, M. Espeter, and M. Schwarz. A Study on the Cognitive Plausibility of SIM-DL Similarity Rankings for Geographic Feature Types. In L. Bernard, A. Friis-Christensen, and H. Pundt, editors, *11th AGILE International Conference on Geographic Information Science (AGILE 2008)*, Lecture Notes in Geoinformation and Cartography, pages 115–133, Girona, Spain, 5-8. May 2008. Springer.

14. K. Janowicz, C. Keßler, M. Schwarz, M. Wilkes, I. Panov, M. Espeter, and B. Baeumer. Algorithm, Implementation and Application of the SIM-DL Similarity Server. In *Second International Conference on GeoSpatial Semantics (GeoS 2007)*, number 4853 in Lecture Notes in Computer Science, pages 128–145. Springer, 2007.

15. K. Janowicz, M. Wilkes, and M. Lutz. Similarity-based Information Retrieval and its Role within Spatial Data Infrastructures. In *5th International Conference on Geographic Information Science (GIScience 2008)*, Park City, Utah, USA, September 2008; forthcoming.

16. C. Keßler. Similarity Measurement in Context. In B. M. Kokinov, D. C. Richardson, T. Roth-Berghofer, and L. Vieu, editors, *6th International and Interdisciplinary Conference, CONTEXT 2007*, volume 4635, pages 277–290. Springer, Copenhagen, Denmark, August 2007.

17. R. Küsters. *Non-Standard Inferences in Description Logics*, volume 2100 of *Lecture Notes in Artificial Intelligence*. Springer, 2001.

18. R. Möller, V. Haarslev, and B. Neumann. Semantics-Based Information Retrieval. In *Proc. IT&KNOWS-98: International Conference on Information Technology and Knowledge Systems, 31. August- 4. September, Vienna, Budapest*, pages 49–56, 1998.

19. A. Rodríguez and M. Egenhofer. Comparing Geospatial Entity Classes: An Asymmetric and Context-dependent Similarity Measure. *International Journal of Geographical Information Science*, 18(3):229–256, 2004.

20. L. Rutledge, J. van Ossenbruggen, and L. Hardman. Making RDF Presentable: Integrated Global and Local Semantic Web Browsing. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 199–206, New York, NY, USA, 2005. ACM.

21. K. Samp, C. Basca, B. McDaniel, and S. Decker. Atom Interface - A Novel Interface for Exploring and Browsing Semantic Space. In *Semantic Web User Interaction at CHI 2008: Exploring HCI Challenges*, 2008.

22. M. Schraefel and D. Karger. The Pathetic Fallacy of RDF. In *International Workshop on the Semantic Web and User Interaction (SWUI) 2006*, Athens, Georgia, USA, November 2006.

23. R. Sinkkilä, E. Mäkelä, E. Hyvönen, and T. Kauppinen. Combining Context Navigation with Semantic Autocompletion to Solve Problems in Concept Selection. In K. Belhajjame, M. d'Aquin, P. Haase, and P. Missier, editors, *SeMMA*, volume 346 of *CEUR Workshop Proceedings*, pages 61–68. CEUR-WS.org, 2008.

24. M. Van Kleek, M. Bernstein, P. André, M. Perttunen, D. Karger, and M. Schraefel. Simplifying Knowledge Creation and Access for End-Users on the SW. In *CHI 2008 SWUI Workshop*, April 2008.

25. J. van Ossenbruggen, A. Amin, and M. Hildebrand. Why Evaluating Semantic Web Applications is Difficult. In *Semantic Web User Interaction at CHI 2008: Exploring HCI Challenges*, 2008.

26. M. Wilkes and K. Janowicz. A Graph-based Alignment Approach to Similarity between Climbing Routes. In *First International Workshop on Information Semantics and its Implications for Geographic Analysis (ISGA '08) held in conjunction with GIScience 2008, the 5th International Conference on Geographic Information Science*, Park City, Utah, USA, September 23 2008.