Intelligent and Personalized Community Maps

Noemi Mauro
University of Turin
Torino, 10149 Italy
noemi.mauro@unito.it

IUI'17 Companion, March 13–16, 2017, Limassol, Cyprus
ACM ISBN 978-1-4503-4893-5/17/03.
http://dx.doi.org/10.1145/3030024.3038282

Abstract
My PhD project focuses on Participatory GIS (PGIS). In the project I analyze two methodologies to offer personalized search results in community maps and a natural interaction with the system. The first consists of automatically gathering the terms according to which the users express their information needs, in order to enrich the domain conceptualization of a PGIS, giving common definitions for places. The second concerns the creation of ontology-based user models that reflect the interests, lexicon and modality of expression adopted by each person, mapped to the domain ontology adopted by the PGIS. In the project I also analyze how these techniques may be jointly used during the query expansion process to retrieve more accurate and relevant search results.

Author Keywords
Semantic search; personalization; ontology-based user model; participatory GIS; linked data; ontologies.

ACM Classification Keywords
H.3.3 [Information Search and Retrieval]; H.3.5 [Online Information Services]; H.5.1 [Multimedia Information Systems]; H.5.3 [Group and Organization Interfaces]

Introduction

With the digitization of the information and the development of crowdsourcing a large amount of data concerning the territory is available. Geographical maps have proven to support an intuitive and expressive representation of the territory, leveraging the geographical position to help users orientate themselves during the exploration of the information space [2]. Specifically, PGISs (Participatory Geographic Information Systems) support the management and analysis of spatial data, offering a visual representation of geographical information based on dynamic map generation and map sharing within and between stakeholder groups. However, dynamic maps challenge users with an information overload issue, that has to be solved to make information really usable to people. For this purpose, an efficient and accurate method is required for building a search engine that correctly answers the user's information needs, possibly tailoring results for improving their fruition. The main goal is to enable the user to specify her/his information needs in a natural way. In the related research, this is done by using query expansion techniques, a natural language model for interpreting the queries, and personalization techniques for showing on the map tailored suggestions. For this purpose, the semantic representation of geographic information is exploited, following the Geospatial Semantic Web's approach [3].

In order to improve the fruition of information, in this PhD project I will analyze an interaction method based on natural language processing that interprets the search queries submitted by the user and offers different visualizations of the results, based on the her/his interests and behavior. The idea is that, depending on the user's behavior, different aspects of information are relevant, and an information system should take this into account to optimize the presentation of data. Moreover, understanding how people specify their information needs (i.e., which terms are closer to the users' lexicon for representing a specific concept) enables the system to understand individual expression modalities.

OnToMap Project

The PhD project is framed in the context of "OnToMap - Mappe di Comunità 3.0" (https://ontomap.ontomap.eu), that aims at supporting citizens' participation in public policy design, as well as knowledge sharing about territorial information. The main result is the development of a web application which can be used to consult spatial data, create custom community maps, report to local administrators critical issues or new proposals in order to give a local representation of the territory. The main aspects of OnToMap are (i) the use of an ontological layer supporting the integration of heterogeneous data (managed as Linked Data) and describing semantic relations to enhance the exploration of the information space, (ii) the support to the management of personalized and persistent community maps for project development, and (iii) the crowdsourcing of new items related to the concepts of the ontology in order to enrich the knowledge base of the application. With project MIMOSA ("progetto di Ateneo Torino_call2014_L2_157", 2015-17), OnToMap has been extended with information search facilities for improving the users' experience in browsing and searching data, developing a semantic-based search engine [2].

Common-sense knowledge about places

The first assumption of my work is that, by learning how an individual user formulates the information queries (her/his vocabulary), the system could improve its capability of understanding her/his information needs. In this perspective, one of the aspects of my PhD project

concerns the gathering of linguistic knowledge starting from an analysis of the users' search behavior. With project MIMOSA the concepts of the OnToMap ontology were enriched with linguistic and encyclopedic knowledge to address the word sense disambiguation problem (i.e., synonyms, textual description, keywords). In order to offer a more natural interaction with the system, it is important to define the common-sense knowledge about places that represent how people refer to the ontology concepts: this could be different from the knowledge extracted from official sources (dictionaries, encyclopedias, etc.) because each person expresses her/himself in a different way and often uses a colloquial language instead of a formal one. By gathering the terms directly from the user, the system can automatically expand the linguistic knowledge associated to the ontology concepts in order to make it more similar to the users' modalities of expression.

The idea is to annotate every concept with the terms used by the user for their search task and to weight terms on the basis of the frequency of usage, and recentness. In that way, the system can learn how users express their information needs, analyzing which terms are most popular for a concept, locally to the individual user, but also globally, considering general user interaction with the application. Indeed, every time a term t is related to a concept c, the relation between them will acquire importance. When it exceeds a threshold, t will become relevant for the concept c. These terms can be put beside the linguistic and encyclopedic knowledge and can be used to personalize query expansion techniques.

Ontology-based user models

Another major aspect of my PhD project concerns the creation of ontology-based user profiles to customize the system and the information search task taking individual interests and knowledge needs into account [4]. The user profiles can be combined with the domain knowledge with the purpose of enriching the search queries and increasing the likelihood that none of the potentially interesting data is missing. In order to better understand the user's behaviour, it could also be useful to analyse her/his query reformulation process. Web log analysis is, in fact, one of the most well-known research methods to capture user interactions with a Web search engine [5].

I propose to associate each user to a user model that reflects the structure of the ontology. The system gives a weight at each concept contained in the user model, considering the user's interest gathered analyzing her/his behaviour. Furthermore, each relation between concepts has a weight, in order to collect the relevant semantic relations for every user. For instance, if we consider a mother interested in the *school* and *park* concepts, it is likely that the weight of the relation between these two concepts in her model is high. These relations can be used to produce a set of association rules, useful for generating personalized suggestions (e.g., "You might also be interested in..."). In each user model there are also all the terms that define the common-sense knowledge about places, but in this case they are weighted with the usage frequencies and recentness for the specific user. In this sense, the user model represents the lexicon and modality of expression adopted by a person. The proposed model is inspired by the mechanisms guiding the declarative memory module of the cognitive architecture ACT–R, in which "the weight of a concept is a sum of a base-level activation, reflecting its general usefulness in the past, and an associative activation, reflecting its relevance to the current context" [1].

Personalized search results

The previously described methods may be combined for two purposes: (i) to retrieve correct search results by analyzing the terms used by the user; (ii) to visualize on community maps both results and suggestions, tailored to the user and constructed by observing her/his behavior. Both approaches can be used for query expansion; the result will be a personalized expanded search query that can be used to search for matching ontology concepts and to retrieve more relevant and accurate results. Indeed, the two ontology layers, the user model and the one representing the enriched ontology, allow to increase the ability to understand what people mean when they submit a search query. For instance, if a person uses a term that isn't in her/his model, but is present in the enriched ontology, the correct concept could be retrieved and suggested. Instead, if the user profile contains the term, the search results could be shown on the map because they probably represent what the user is searching for.

Conclusions and discussion

The described PhD project aims at improving the interaction between the users and intelligent systems and at making it more natural, leveraging learning mechanisms and user modeling techniques, in order to automatically generate an internal representation of the user profiles. The current natural language model adopted in OnToMap will be refined for retrieving better search results in terms of precision and recall, as well as user satisfaction. The developed methods will be tested by applying traditional HCI evaluation techniques. A relevant social impact of this project is related to knowledge diffusion and education. Thanks to a flexible information search model, people will be guided in the exploration of a large amount of data about the territory that, otherwise, would be hardly usable. Another impact is the support to public engagement in policy making, which presupposes the existence of an efficient knowledge sharing and information search model raising general awareness about the territory, as well as inspection of particular projections of information. Concerning the search engine's accuracy, it will be important to offer useful and personalized results and reduce the noise of the search results. In order to improve this capability, the model will not only analyse punctual information: it will also consider the context in which the items are placed, the users' interests, and the common-sense knowledge of places learned by analyzing user behaviour.

Acknowledgements

I would like to thank my supervisor Prof. Liliana Ardissono from Computer Science Department of University of Turin.

References

[1] Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S. A., Lebiere, C., and Qin, Y. An integrated theory of the mind. *Psychological Review 111*, 4 (2004).

[2] Ardissono, L., Lucenteforte, M., Mauro, N., Savoca, A., Voghera, A., and Lariccia, L. Exploration of cultural heritage information via textual search queries. In *MobileHCI '16 Proceedings Adjunct*, ACM (2016).

[3] Janowicz, K., Scheider, S., Pehle, T., and Hart, G. Geospatial semantics and linked spatiotemporal data – past, present, and future. *Semantic web 3*, 4 (2012).

[4] Jiang, X., and Tan, A.-H. Learning and inferencing in user ontology for personalized semantic web search. *Information Sciences 179*, 16 (2009), 2794–2808.

[5] Rieh, S. Y., and Xie, H. Analysis of multiple query reformulations on the web: The interactive information retrieval context. *Information Processing and Management 42*, 3 (2006), 751–768.