Thematic Maps for Geographical Information Search

(Article begins on next page)

10 April 2024

Thematic Maps for Geographical Information Search

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

Liliana Ardissono
University of Turin
Computer Science Department
Torino, Italy 10149
liliana.ardissono@unito.it

ABSTRACT

The exploration of cultural heritage information is challenged by the fact that most data provided by online resources is fragmented and it is repository or application-centered. In order to address this issue, a data integration approach should be adopted, that makes it possible to generate custom views, focused on the user's information needs, but easily extensible to support the inspection of topically related contents.

In this paper, we present a model supporting the management of thematic maps for information exploration, and their integration with query expansion during the interaction with the user. Our model is based on: (i) an ontological domain knowledge representation for describing the meaning of concepts and their semantic relations; (ii) a semantic interpretation model for identifying the concepts referenced in the user's queries. We are experimenting our model in the OnToMap Participatory GIS, which manages interactive community maps for information sharing and participatory decision-making.

CCS CONCEPTS

•**Information systems → Geographic information systems;** *Ontologies;* **Search interfaces;** *Personalization;*

KEYWORDS

Information search, Thematic maps, Personalization

1 INTRODUCTION

Nowadays, the exploration of cultural heritage takes great advantage of Web-based technologies, which make knowledge available at low cost using standard browsers. However, most information provided by online resources is fragmented because it is repository or application-centered. Thus, it is difficult to provide a unified view of possibly heterogeneous data.

We point out that, in order to make cultural heritage information truly accessible to people, a data integration approach should be adopted, that makes it possible to generate custom views, focused on the user's information needs, but easily extensible to support the inspection of topically related contents. For this purpose, the geographical dimension of a territory should be integrated with a

semantic exploration of content, as described in Semantic GeoSpatial Semantic Web [13], and with a proactive information retrieval support. The idea is not only that of helping the user find the data (s)he explicitly searches for, which is the typical information retrieval problem, but also to support serendipity, by guiding her/him in the exploration of portions of the information space providing extra-helpful information.

Thematic maps, traditionally developed to present specific themes within geographical areas, are an interesting information presentation model because of the abstract view of the territory they provide. However, it is difficult to determine the types of data that should be presented *a priori*: on the one hand, the maps should adapt to users' search behavior in order to reflect changes in the interests of the population. On the other hand, they might not suit the individual user's needs; e.g., (s)he might be interested in a subset of a map, or in a mixture of the content provided by two of them. Therefore, the management of static information layers might not be the best solution to be adopted.

In order to support the exploration of territorial data under different points of view, we propose a model that combines the interpretation of the user's queries with thematic map management to generate dynamic maps that reflect individual information needs. The maps play the role of indexes on information and they support query expansion towards relevant topics that the user might not be aware of. For this purpose, we model the theme of a map at the meta-level, as a cluster of concepts that are usually searched together by people. The visualization of data items on a map comes as a result of selecting one or more concept clusters, or portions of them, to retrieve the corresponding data. The organization of a thematic map as a cluster of concepts enables the user to select the types of information (s)he is interested in and, as a by product, it allows the system to observe what (s)he focuses on. This is the basis to propose personal views on information.

Our thematic map management model is based on: (i) an ontological domain knowledge representation for describing the meaning of concepts and their semantic relations, as well as for handling geographical data as Linked Data [19]; (ii) a semantic query expansion model for identifying the concepts referenced in the user's queries, and thus the types of information relevant to her/his search. Our model is applied in the OnToMap Participatory GIS [3, 18] and currently supports the provision of a set of static thematic maps, among which one about tourism and culture. In the next future, the model will evolve to the management of dynamic concept clusters reflecting general search behavior.

The remainder of this paper is organized as follows: Section 2 provides some background on ICT for cultural heritage exploration. Section 3 presents the main features of the OnToMap application. Section 4 describes the approach we adopted for developing the
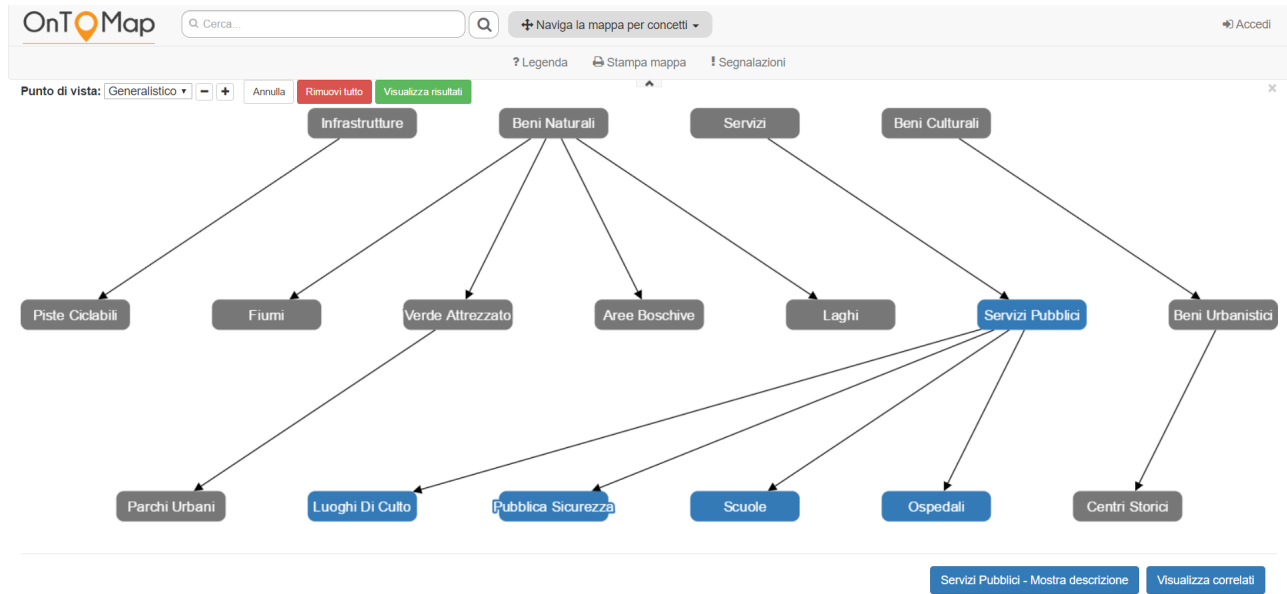
Figure 1: Portion of the concept graph supporting the exploration of the OnToMap domain ontology.

thematic maps and Section 5 presents the management of thematic maps. Section 6 concludes the paper and outlines our future work.

2 BACKGROUND

The Web helps the diffusion of Cultural Heritage information by making available the results of the massive digitalization effort, carried out by cultural organizations and/or crowdsourced as Voluntary Geographic Information [9], by means of online information sources and archives. However, the lack of a framework for semantic data integration, contextualized in a geographical perspective, limits the fruition of this type of information. Specifically:

- Museums, exhibitions and historical sites publish data about physical and digital collections in their own web sites; e.g., see the Tate Online (http://www.tate.org.uk/) and [12]. They present detailed information about art pieces and historical objects. However, they are disconnected from each other and thus provide a partial view on the resources available in the territory.
- Large projects have been developed to support the digitalization and diffusion of artistic and historical content across Europe. However, the resulting services, such as Europeana Collections [7], are managed as catalogs of artworks abstracting from the geographical dimension of information items. These services enable the user to retrieve data according to a fixed number of facets (e.g., by source, color, topic). However, they fail to integrate such search criteria with the geographical aspect, which is a key element to understand the identity of a territory, as demonstrated in the experience of the Parish Maps and of community mapping [16].

- Modern search engines, location-based services (e.g., TripAdvisor [17]) and Web-GIS (e.g., OpenStreetMap [15]) offer graphical interfaces that support information search and a map-based visualization of results. However, they cannot tune the presentation of the territory to the aspects underlying search queries because they lack a semantic approach to their interpretation and to the identification of the relevant types of information.

Some projects proposed the adoption of ontologies for the integration of geographical data as Linked Data, in order to support information retrieval, but focused on data modeling, leaving the management of personalized views on information spaces apart; e.g., OSM Semantic Network [4] and OSMonto [5], and [8].

Thematic maps (e.g., see [11]) are useful to present specific aspects of a territory in an abstract way, by exploiting the geographical dimension to facilitate data visualization and exploration. In other projects, they have been used to describe colors, smells, perception of danger, etc. of a geographical area; e.g., see [1, 6].We use thematic maps to support knowledge discovery, but we aim at focusing them on the information relevant to the user's interests. This is based on the provision of interactive indexes that support the visualization of different types of information belonging to a broad topic, such as culture or tourism.

It should be noticed that other works, such as Google's Knowledge Graph [10], acquire links among individual information items, that people often search together, in order to recommend their inspection, when the user selects one of them. In a complementary way, the clusters we propose are formed by *concepts*, in order to suggest relevant types of information for expanding the user's search. Thus, they support the exploration of the information space
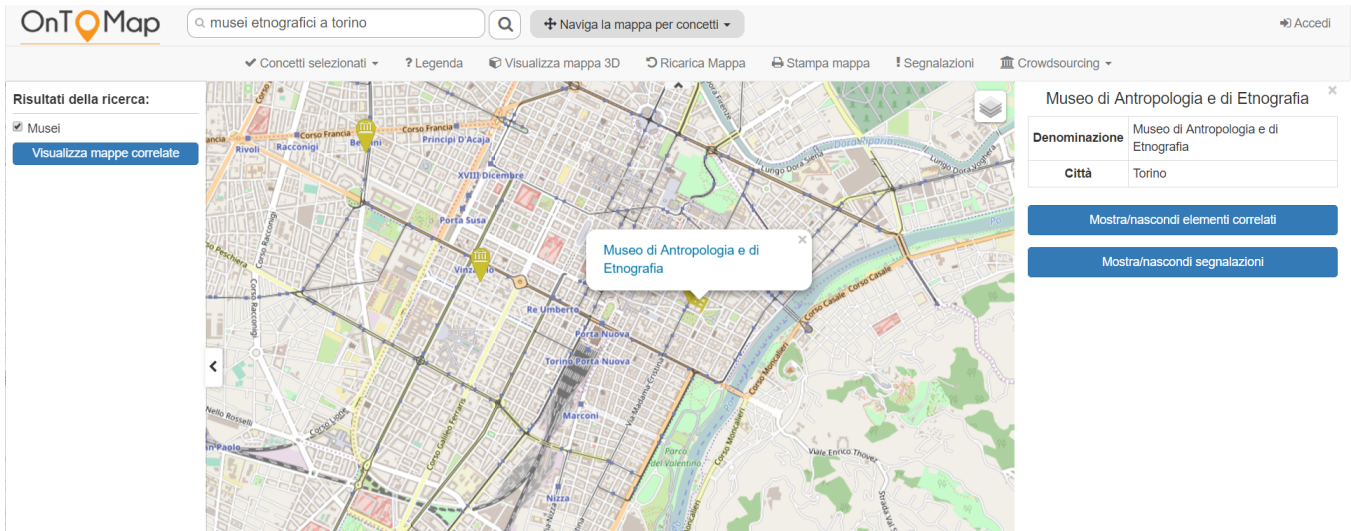
Figure 2: Search results for "musei etnografici (ethnographic museums) a Torino", and visualization of details about an item.

at an abstract level, instead of presenting a possibly large sets of individual information items to analyze.

3 OVERVIEW OF ONTOMAP

The OnToMap Participatory GIS [3, 18] supports the management of interactive community maps for information sharing and participatory decision-making. It supports both the consultation of spatial data and the creation of public and private maps, which reflect individual information needs and can be enriched with crowdsourced content (information items and geo-localized comments) to help project design and group collaboration.

The interaction functions offered by OnToMap are based on an ontological representation of geographical information that allows the abstraction and the categorization of data, and supports multi-faceted data retrieval and presentation. The ontology defines semantic relations among concepts, allowing heterogeneous data integration [8] and the management of heterogeneous data as Linked Data [19]. Moreover, each ontology concept is enriched with linguistic and encyclopaedic knowledge, that makes it possible to use query expansion techniques to interpret user queries, addressing the word sense disambiguation problem; see [2].

The semantic representation of geo-data supports different types of exploration of the information space, thanks to the specification of semantic and subclass relations among concepts, and to the structured representation of information items. Specifically, the application offers two interaction modes, and different granularity levels in the specification of the data to search for:

(1) The user can submit a sequence of free-text queries, or queries referring to multiple concepts, in order to select the types of information to be visualized in the map; e.g., searching for *museums in Torino*. Queries can also be restricted by specifying attributes of the relevant information items, e.g., searching for *ethnographic museums ("musei etnografici")*, as in Figure 2. The system semantically interprets the queries by matching the terms they contain

with the ontology concepts, and filters the results on the basis of the qualifiers specified in the queries. For all the matches, it employs synonyms and linguistic descriptions of concepts to abstract from the terminology used by the user. As a result, it can return data that matches the query by taking the meaning of information into account, instead of only relying on keyword-based matches.

(2) The graph-based navigation of the information space enables the user to explore the concepts defined in the domain ontology by browsing a concept graph that visualizes semantic and subclass relations among concepts. Figure 1 shows the user interface for graph-based navigation. It displays a subset of the ontology concepts and the subclass relations existing among them, in order to help the user understand how information types are related to each other. By clicking on the nodes of the graph, the user can select the types of information that (s)he wants to visualize in the map, at different generality levels. For instance, in the figure the user has selected the public services node, which subsumes several types of services; see the nodes highlighted in blue.

OnToMap displays the search results on a map focused on the geographical area delimited by the query (e.g., in Torino), or by the bounding box selected on the map. However, the user can dynamically change the bounding box (via zoom and drag actions) to view results belonging to different geographical areas. For instance, the map in Figure 2 is focused on the center of the town. While the map employs the OpenStreetMaps layer as a background to provide a general view of the selected geographical area, search results are highlighted by showing them as pointers, or by depicting their geometries, in vivid colors. Results can be clicked to retrieve detailed information about them. For instance, see the sticky note in Figure 2 and the descriptive table displayed in the right portion of the page. From the detailed description of an item, the user can inspect the semantically and geographically related items, following the
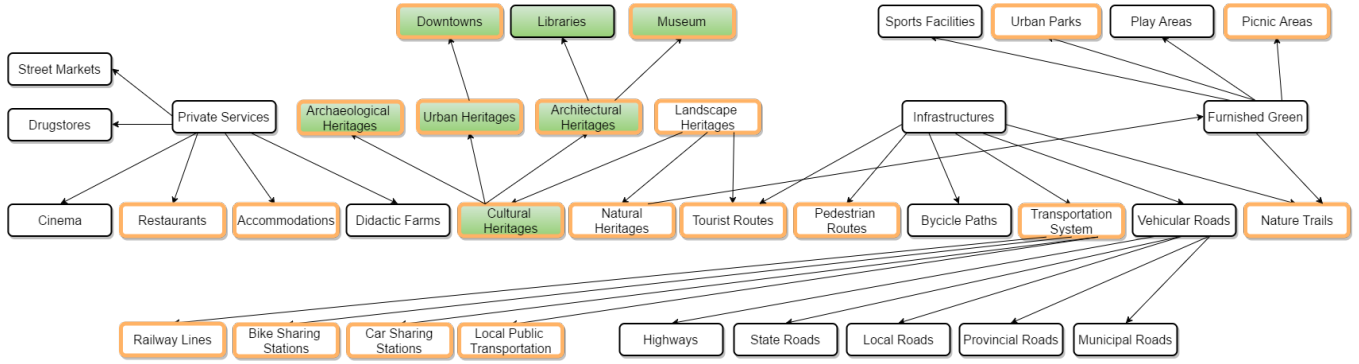
Figure 3: Tourist (orange border) and cultural (green background) thematic maps.

Linked Data model. See button "Mostra/Nascondi elementi correlati" (show/hide related items) under the table. Moreover (s)he can inspect the available comments, added by people via crowdsourcing; see button "Mostra/nascondi segnalazioni" (show/hide reportings).

4 MANAGEMENT OF THE THEMATIC MAPS

Our ultimate goal is the definition of a dynamic set of thematic maps, emerging from users' search behavior, in order to enrich OnToMap with an adaptive suggestion function supporting the exploration of the information space. For that purpose, we are logging the search queries performed by users and we will use that information to identify the sets of concepts that are frequently searched together. However, the collection of this data takes a lot of time. As we aim at shortly testing the usefulness of the support offered by thematic maps, we defined a set of static ones to initialize the application with a preliminary set of concept clusters.

4.1 Static Thematic Maps

In order to develop the first set of thematic maps to be experimented, we spread an online survey in which we asked people to select the concepts that are most relevant to five themes: culture, tourism, nature, family, and services. For each theme, we presented the list of concepts forming the OnToMap ontology and we asked people to select one or more that suited the theme.

We collected 64 questionnaires, which we analyzed to compute the relevance of each concept with respect to the five themes. We assume that the evidence of the relevance of a concept c to a theme T is determined by people agreement; i.e., the more people assigned c to T, the stronger the evidence of association among the two. This enabled us to define the percentages of relevance of each concept with respect to the five thematic maps. We considered a concept as relevant to a theme, and thus belonging to the respective thematic map, if it were selected by at least 50% of people who filled in the survey.

Table 1 reports the percentages of relevance achieved by the concepts selected for the tourism and/or culture themes. Notice that thematic maps can contain rather different numbers of concepts.

Figure 3 shows a portion of the ontology and highlights the concepts relevant to the tourism and culture themes. The concepts relevant to tourism have an orange border and those relevant to

Concepts	Tourist Map	Culture Map
Cultural Heritage	86%	83%
Landscape Heritage	86%	
Archaeological Heritage	84%	76%
Architectural Heritage	84%	75%
Museum	84%	78%
Natural Heritage	79%	
Belvedere Natural	71%	
Natural Monuments	71%	
National Parks	71%	
Tourist Routes	70%	
Downtowns	67%	67%
Restaurants	67%	
Accommodations	67%	
Urban Heritage	65%	54%
Regional Parks	65%	
Urban Parks	65%	
Nature Trails	65%	
Belvedere Artificial	60%	
Railway Stations	60%	
Local Public Transportation	60%	
Railway Lines	59%	
Pedestrian Routes	59%	
Transportation System	57%	
Bike Sharing Stations	56%	
Picnic Areas	52%	
Car Sharing Stations	51%	
Libraries		71%

Table 1: Percentage relevance of concepts to tourism and cultural maps, as collected in the survey.

culture have a green background. Some concepts are relevant to both themes.

As a matter of fact, a larger number of questionnaires should be collected in order to achieve statistically relevant thematic maps. However, the concept clusters we defined so far are a starting point to bootstrap the system.
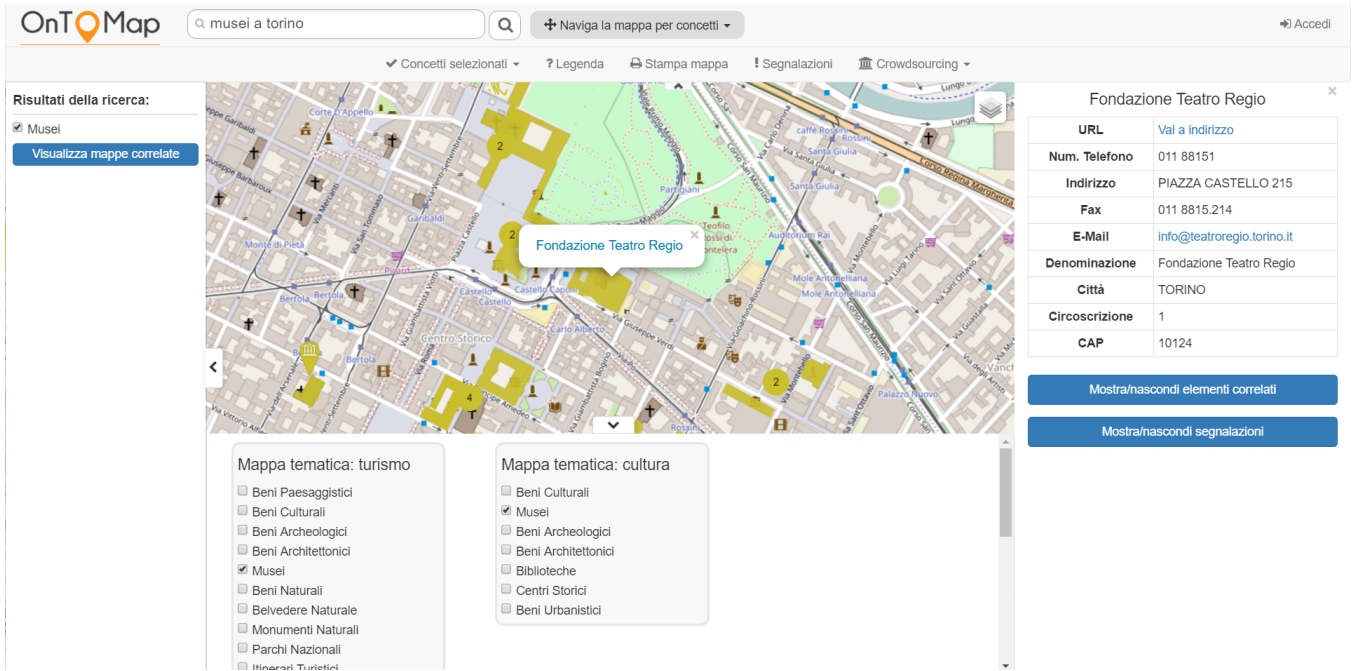
Figure 4: Visualization of concept clusters relevant to search query "musei a Torino" (museums in Torino).

4.2 Dynamic Thematic Maps

Starting from the bootstrapping phase described in the previous section, we aim at creating dynamic thematic maps, by finding new clusters of concepts from users behavior, but also by updating the existing ones. We consider two types of dynamic maps:

- The former is the most general one, and regards concepts that typically co-occur in the search sessions performed by the user population. By analyzing the system's log, we can identify sets of concepts that are jointly searched by users. We think that, if a concept cluster is recurrent, it is relevant, because it reflects the most frequent combined search done by people. In this case, the cluster defines a thematic map whose concepts are weighted by frequency of usage during the interaction with the system in a search task. We aim at automatically generating these clusters, by dynamically updating the occurrences of actions on concepts logged by the system during the interaction with users; e.g., selection of concepts suggested by the system, subsequent searches, and so forth.
- The latter reflects the individual user's search history and is aimed at capturing the maps that (s)he frequently uses. We propose to create ontology-based user models for customizing the system and the information search task, taking the interests of individual users into account; see [14]. The concepts belonging to each user model must be weighted by analyzing the behavior of the specific user. Starting from this, we can deduce which are the most relevant concepts for each user, that we can use to expand the general thematic maps presented above. Indeed, the user models can also be employed to rank the concepts

of the thematics maps, and promote the most interesting ones. This supports a user-adaptive presentation of thematic maps, instead of adopting a single visualization style, the same for everybody.

We aim at automatically updating the static maps created from the survey, the general ones inferred by analyzing the search behavior emerging from the logs of the user population, and those based on the acquisition of individual user models. We hypothesize that the management of dynamic thematic maps, adapting to continuous observations, can be particularly useful to provide users with personalized groups of suggestions, represented as sets of concepts that can be easily explored, without browsing the complete information space. Obviously, our hypothesis has to be verified by testing the model through user experiments.

5 SUPPORTING THE EXPLORATION OF INFORMATION VIA THEMATIC MAPS

During the interaction with the user, OnToMap proposes the thematic maps to the user on the basis of the relevant topics emerging from her/his search behavior. However, as previously specified, we aim at providing a flexible map management model, that enables the user to select the types of information to be visualized, instead of loading complete layers that might contain large amounts of data. We thus integrate thematic maps with query expansion, by presenting the clusters of concepts underlying them, and we leave the user free to choose the interesting ones.

OnToMap dynamically selects the concept clusters to be proposed by matching them to the concepts referenced in the search queries. Given the set of concepts $Q = \{c_{q1}, \ldots, c_{qm}\}$, identified

by interpreting the queries, and the concept clusters $CL_1, \ldots, CL_n$ associated to the thematic maps:

(1) First, the application filters the clusters $CL_i = \{c_{i1}, \ldots, c_{ik}\}$ such that $Q \cap C_i \neq \emptyset$ and at least one concept c_{ij} has relevance $\geq 50\%$ for the thematic map. All these clusters are potentially interesting because they include at least one concept that was referenced in the user's queries.

(2) Then, the application proposes the three best matching clusters, in which the concepts belonging to Q appear with the highest ranks.

The suggestions are activated by clicking on button "Visualizza mappe correlate" (show thematic maps) in the left portion of the page. The result is the visualization of the most relevant concept clusters that have been identified.

Figure 4 shows the suggestions generated for the search query "Musei a Torino" (museums in Torino). OnToMap proposes to select the concepts belonging to two thematic maps:

- The Tourism one ("Mappa tematica turismo") includes concepts related to tourism activities; e.g., cultural heritage, landscape heritage, archaeological heritage, architectural heritage, museums, natural heritage, and so forth.
- The Culture one ("Mappa tematica cultura") includes concepts related to cultural activities: e.g., cultural heritage, museums, archaeological heritage, architectural heritage, libraries, and so forth.

These clusters are suggested because they include concept "Museum", that is referenced in the search query and has relevance $\geq$ 50% in the respective thematic maps.

The concepts of a cluster make the information available in the corresponding thematic map explicit in order to make the user aware of which information could be searched for. They are presented inside check-boxes, so that the user can directly select the types of information (s)he wants to visualize.

Notice that the thematic maps, described as concept clusters, represent topic-oriented views on the domain conceptualization adopted by the system. As such, they could be employed to support the graph-based navigation of the information space described in Section 3, which is currently based on the definition of a set of static subgraphs of the ontology. Indeed, this method might be useful for navigating a large information space, because it presents different views that can be dynamically modified, by reacting to the users' behavior, without the need to manually update the data structures used by the system.

6 CONCLUSIONS AND FUTURE WORK

We presented a thematic map management model supporting a semantic exploration of geographic information spaces in GIS. Our model is based on a semantic representation of geo-data and on a query expansion model based on the recognition of the concepts referenced in the user's search queries.

Currently, our model supports the management of static thematic maps (e.g., a tourism and a culture one) defined by collecting users' feedback by means of an online survey. However, we will extend it to the creation and management of dynamic thematic maps, based on the observation and analysis of users' general and individual search behavior. We hypothesize that this will improve the exploration of large information spaces by making the user aware of the available data, in relation to her/his current search goals, without the need to navigate the complete spaces.

Our model is applied in the OnToMap Participatory GIS, which supports community mapping and information search, and will be evaluated with users in our future work.

REFERENCES

[1] L.M. Aiello, R. Schifanella, D. Quercia, and F. Aletta. 2016. Chatty maps: constructing sound maps of urban areas from social media data. *Royal Society Open Science* 3, 3 (2016), 1–19.

[2] L. Ardissono, M. Lucenteforte, N. Mauro, A. Savoca, A. Voghera, and L. La Riccia. 2016. Exploration of Cultural Heritage Information via Textual Search Queries. In *MobileHCI '16 Proceedings of the 18th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services Adjunct.* ACM, 992–1001.

[3] L. Ardissono, M. Lucenteforte, N. Mauro, A. Savoca, A. Voghera, and L. La Riccia. 2017. OnToMap - Semantic Community Maps for knowledge sharing. In *Proceedings of Hypertext 2017.* ACM.

[4] A. Ballatore, M. Bertolotto, and D.C. Wilson. 2013. Geographic knowledge extraction and semantic similarity in OpenStreetMap. *Knowledge and Information Systems* 37, 1 (2013), 68–81.

[5] M. Codescu, G. Horsinka, and O. Kutz. 2017. OSMonto. http://wiki.openstreetmap.org/wiki/OSMonto.

[6] M. Diaconu. 2011. *Senses and the city : an interdisciplinary approach to urban sensescapes.* Lit Wien ; Berlin.

[7] Europeana. Europeana Collections. www.europeana.eu/portal/it.

[8] F.T. Fonseca. 2002. Using ontologies for geographic information integration. *Transactions in GIS* 3 (2002), 231–257.

[9] B. Goodman and D. Litman. 2007. Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. *International Journal of Spatial Data Infrastructures Research* 2 (2007), 24–32.

[10] Google. 2017. Knowledge Graph. https://www.google.com/intl/it_it/insidesearch/features/search/knowledge.html.

[11] ICSM. Fundamentals of Mapping. http://www.icsm.gov.au/mapping/maps_general.html.

[12] Archives & Museum Informatics. 2017. Museums and the web. http://www.museumsandtheweb.com/.

[13] K. Janowicz, S. Scheider, T. Pehle, and G. Ha. 2012. Geospatial Semantics and Linked Spatiotemporal Data – Past, Present, and Future. *Semantic Web - On linked spatiotemporal data and geo-ontologies* 3, 4 (2012), 321–332.

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

[15] OpenStreetMap Contributors. 2017. OpenStreetMap. https://www.openstreetmap.org.

[16] B. Parker. 2006. Constructing Community Through Maps? Power and Praxis in Community Mapping. *The Professional Geographer* 58, 4 (2006), 470–484.

[17] TripAdvisor. 2017. TripAdvisor. https://www.tripadvisor.it/.

[18] A. Voghera, R. Crivello, L.Ardissono, M. Lucenteforte, A. Savoca, and L. Lariccia. 2016. Production of spatial representations through collaborative mapping. An experiment. In *Proc. of 9th Int. Conf. on Innovation in Urban and Regional Planning (INPUT 2016).* 356–361.

[19] W3C. 2017. Geospatial Semantic Web Community Group. https://www.w3.org/community/geosemweb/.