A Framework for Adaptive Context and User-Related Management of Multimedia Contents

M.G. Fugini [1], J. Finocchi [1], E. Rossi [2]

[1] *DEIB - Politecnico di Milano, Milano, Italy*
[2] *DIG - Politecnico di Milano, Milano, Italy*

Abstract

This paper presents the specification of a software platform where *knowledge* is associated to *maps*, through the organization and presentation of geo- and chronological-referenced *multimedia contents*, in such as a way that this association is automatically *adaptive* to different contexts (application areas) and to different *typologies of users*.

The platform adaptivity has the purpose of *reducing* the *information overload* of the maps and of offering the user targeted navigation and search functions, tailored to his/her typology (e.g., domain expert, citizen, public administration operator, decision maker, stakeholder, etc.) and dependent on the loaded contents. To achieve this, we focus on the problem of *selecting and aggregating* the multimedia contents of the map, in a way that *depends on the context and on the user*. This reduction is obtained as a dynamic process, using a metamodel in the form of ontology of tags.

Keywords

Georeferenced Data, Data Discovery, Map Enrichment, Context Adaptivity, Multimedia Content, Multimedia Tagging, Metadata Analysis.

1. Introduction

The typical aim of geographic-based multimedia applications is to enrich a map with georeferenced contents, allowing one to explore selected areas and run thematic analyses at various levels of detail, and with different aims and focus, so that is will be possible to *perform analytics* on multimedia map contents. Such analyses usually aim at exploring various levels of interest on the map contents, and at obtaining zoom levels on the represented realm.

In this paper, we propose an adaptive framework that enriches maps with knowledge using an *ontology of tags*, which describe the map multimedia contents, the geo references, the time references, the context, and the user behaviour, so that the contents can adapt to: 1) different contexts; 2) different users.

Proceedings Name, Month XX–XX, YYYY, City, Country
EMAIL: mariagrazia.fugini@polimi.it (M.G.Fugini); jacopo.finocchi@ mail.polimi.it (J.Finocchi); elisa5.rossi@mail.polimi.it (E.Rossi)
ORCID: 0000-0002-0692-0153 (M.G.Fugini)

The set of tags associated to multimedia data are modelled in an ontology with:

- sets of concepts shared among various contexts (e.g., “geo-location” tag, “time”, “user”, “media_type”, etc.) and
- sets of concepts typical of a context (e.g. “year_of_discovery” and “host_museum” tags in the Cultural Heritage context, “approval_document” tag in the Urban Planning context, and “level_of_robotization” and “efficiency_formula” tags in the Industry context).

Domain experts develop and specialize the ontology for each context.

The tags guide the dynamic generation of *thematic map layers*, as defined in [1], which are the basis for content navigation. In fact, the navigation interface is based on a set of layers, some recommended by the system, others customized by the user. Each layer contains the points of interest belonging to one or more categories, i.e., connected to contents associated with those tags. Each tag identifies a category of contents. The choice of how to split tags among the various layers is taken dynamically considering which nodes of the ontology are actually instantiated by the contents present in the context.

The paper is organized as follows. In Section 2, we review related work. In Section 3, we present the overall approach of associating knowledge to maps. In Section 4, we discuss about navigating tags for adaptivity and draw the conclusions.

2. Related work

Nowadays, geographic-based applications are a widespread object of study. Developers are implementing many different solutions that combine geo-referenced information related to a map, ending up with maps that can be used either as general-purpose exploration tools, or as professional tools for specific application contexts.

Some of these tools combine geographic and multimedia information to improve content search, e.g., in [2], where visual features of pictures are integrated with geographic data, based on the points of interest concept.

The research in [3] tackles the semantic gap between semantic needs of users and the visualization of multimedia content. The proposed solution is focused on how to extract information from the context of multimedia documents and their metadata, rather than from media raw content itself.

Also [4] enhances the semantic annotation of multimedia contents by leveraging their web context and the user comments. Videos are indexed via a large set of labels, and are linked with related contents.

The solution presented in [5] implements a semantic tag recommendation technique for image tagging, relying on a graph of relationships among words.

The development of an adaptive multimedia content navigation is a main goal of adaptive hypermedia systems, mainly based on the user profiles, as discussed in [6].

The problem of adaptivity in accessing multimedia contents is often addressed in the field of e-learning systems. For example in [7], ML techniques are applied to adapt the difficulty level of the presentation to the learner style and to student assessment.

In [8], the authors adopt an ontology-based approach to integrate the relevant knowledge in the content presentation. Their ontology not only models the domain knowledge but also the teaching activity tasks, to support personalized flipped learning.

3. Associating knowledge to maps: issues and our approach

In our platform, knowledge associated to geographical maps links geographical elements with *geo- and temporal-referenced multimedia contents.* The framework is being developed to be *adaptable* to:

- various contexts, such as history, cultural heritage, urban planning, digital twinning of devices and artefacts, and citizen journalism.
- different user typologies.

The key innovative features of the proposed framework are as follows.

1. *Handling of multimedia content:* the framework focuses on the organization and navigation of georeferenced multimedia, such as pictures, video, audio tracks, point clouds, unstructured text or documents.
2. *Adaptivity* to different contexts and users: the framework comprises a meta-structure (semantic tags ontology), which allows instantiation onto different contexts, via self-adaption to loaded contents.
3. *Temporal dimension:* chronological metadata are treated separately, so that specific navigation modes can be provided along the time dimension.

Contents are linked to the map through the definition of a number of *Points Of Interest (POI)*, which are selected and clustered according to our self-adaptive semantic tag ontology. Content classification adapts to the context-specific ontology, by matching the content metadata with the most relevant ontology node.

In the case of massive contents loaded from an existing data source, the classification of multimedia under our semantic tags is performed by a *supervised ML* classifier, starting from the metadata natively associated to each content item in the data sources, along with geo and temporal tags.

3.1. Modeling

For the representation of geo and temporal information, we propose a *flexible data structure* that joins existing cartographic formats to enable linking pieces of information one to another and to map.

The approach consists in adding the following issues to geographical objects:

- *unstructured multimedia content*, together with source of information attributes as well as domain-specific attributes, and
- *temporal dimension*, to reach the concept of 4-dimensional map.

The prototype we are developing will be grounded on an existing open cartographic system (precisely, OpenStreetMap) with various added layers of data, representing the context-specific elements. Compared to a classic map, this model allows browsing the data as a 4D map, enriched with multimedia content.

Tags representing the temporal dimension allow associating a point of interest to a set of *events* that occurred in its location or describing the evolution of the associated contents along time.

The knowledge associated with the map through the metadata of the multimedia contents is organized along three information dimensions: the *geographic* dimension, the *temporal* dimension, and the *thematic* dimension. The first can be navigated with the interactive tools typical of maps, i.e., panning and zooming. The second is navigable through chronological selection filters. The third is navigable through the definition of content layers, with the support of an ontology of semantic classes.

The first two dimensions are organized independently of the context, while the thematic dimension is strongly linked to the context of application.

We started from the definition of a data model, aimed at representing and managing the *semantic contents* associated to the map. The data model allows integrating the cartographic data with heterogeneous structured data (coming from different sources) and with unstructured data (multimedia), including the chronological dimension.

The hinge between geographic elements and contextual content is the definition of POI. Since knowledge is often related to a geographical element rather than a single point, a POI can also correspond to an area, accompanied by related context information.

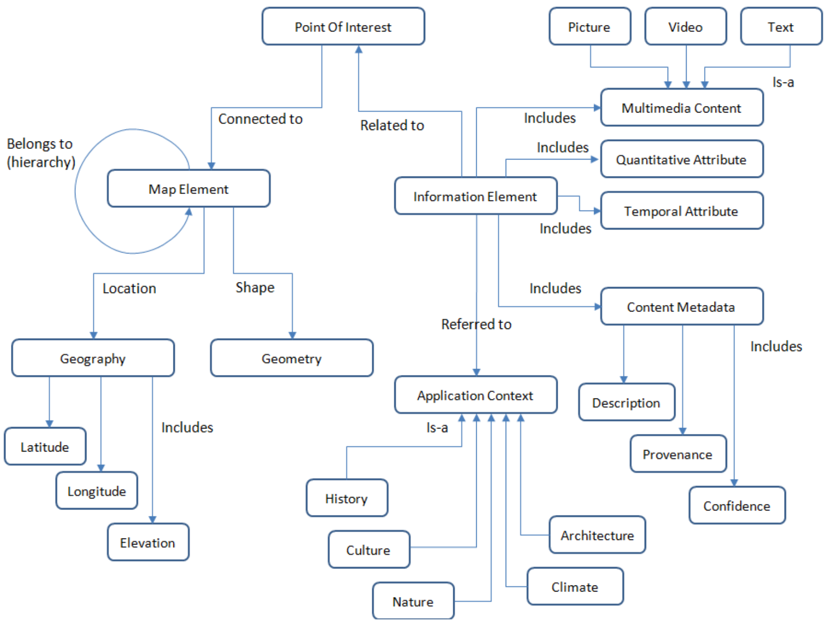

Figure 1: The ontology of tags

Figure 1 shows an outline of the ontology we are developing. A *map element* is defined with geographic coordinates and can represent a single point (like a building), a line (like a street) or a polygon (an area).

On the right-hand part of Figure 1, we represent the content associated to the map, organized into different contexts and including the multimedia content and its related *metadata*.

Deriving the content classification from a semantic ontology allows for a language-independent consultation, since each node of the ontology describing the contents can be expressed with a specific tag for each language managed by the system. This is different from textual search in metadata, which is instead dependent on the language used to specify the metadata.

3.1. Functional architecture

The functional architecture of the proposed framework is shown in Figure 2.

The *Storage* area contains a spatial database for geographical and temporal data connected to the map, together with an additional database storing information about contexts and multimedia.

In the *Analytics* area, the *Adaptation* module implements the adaptivity of the framework, while transversal modules deal with *Data Quality and Validation*.

In the topmost area of our architecture, the *Interface* layer allows users to interact with the analysis modules, to enter new data and to visualise the map, dealing with *Security and Privacy* issues by filtering accesses according to the different user privileges. Here, some modules aim at validating the provenance of the data sources, and to model the additional data according to the target representation, which is visualised in a personalised way according to the rights of the different users.

The framework is designed to be used by different categories of users: Public Administrations, professional users such as architects or urban designers, and individual citizens. Each category is granted access to different contents and functionalities according to the needs and roles.

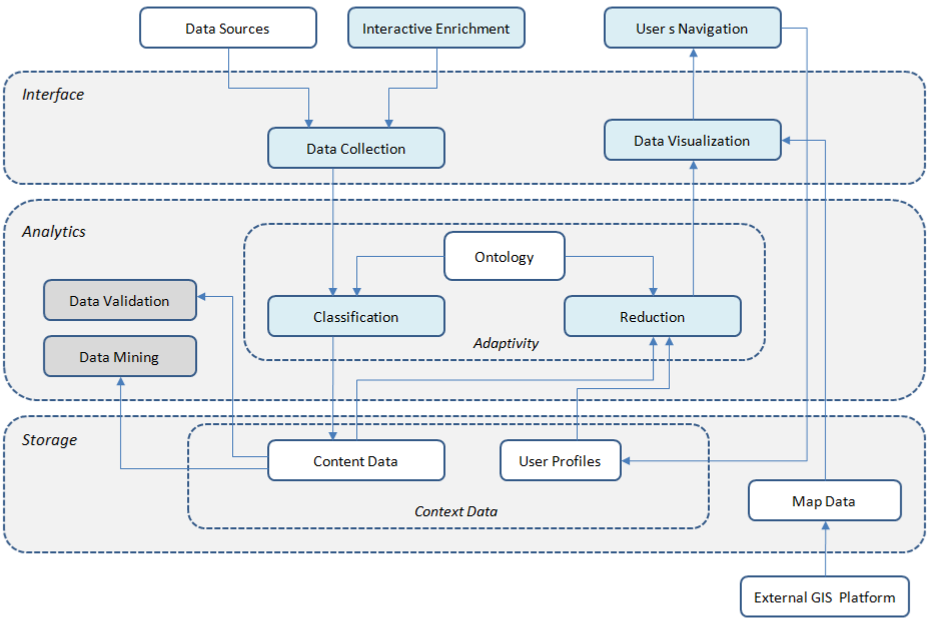

Figure 2: Framework overview

3.2. Enriching a Map: Semantic Annotations

After the definition of the general framework, the following step is to design an interactive tool aimed at enriching the map layer with content provided by *external and voluntary sources*, according to *crowdsourcing* models, since, for testing purposes, we are considering importing a series of publicly available georeferenced multimedia contents.

In the content acquisition phase, the interface layer assigns to data a set of attributes derived from metadata and from source registration information, such as privacy, licensing, confidence, and so on) and to content, to enable its thematic classification.

Considering *classification of the content*, either we provide a tree of predefined tags for homogeneous and language-independent classification, or we leave the contributors free to enter user-defined tags in the tree.

Concerning *who can upload content*, the framework is open to ordinary users, voluntary citizens and qualified sources with credentials, such as public administrations, universities, associations, or professionals.

Regarding *who can read data*, we envisage public, private, password-protected information or payment-based access. The type of access depends on the type of licence the system grants: data can be *public domain*, or can be used under creative-common licences, or provided under the authorisation of the owner. Some types of users can add data, which they can decide to share with other users (a community – selective access, all users – public access, no users – restricted access) and combine them with the existing data. The combination is guided by the ML algorithms, that help classifying data according to what defined in the ontology.

4. Discussion about Navigating Tags for Adaptivity and Conclusions

The presented framework aims at offering a query and analysis environment that supports navigation along tags and enables analyses, e.g., for decision making, by intersecting geo/temporal referenced data coming from different sources and with different formats and informative content, including multimedia data.

The navigation and query interface would *adapt to the context*. The way to achieve adaptation is, for example, the selection of different query filters or the dynamic generation of a layer containing context-dependent elements.

As an alternative, we envision the dynamic generation of the POI and the clusters of multimedia content associated with them. The idea is that not only the content of the clusters can be dynamically navigated (which is somehow obvious), but also their structure should adapt to the context. Tags adapt to the uploaded content or vary based on user queries. We plan to have *three navigation modes* to reduce information overload in a way that is adaptable to the context, where the context is defined by the set of entered contents and by the user preferences:

- *Selection*, namely filtering, i.e., showing only some contents and filtering out irrelevant contents.
- *Grouping*, namely clustering, i.e., showing a single object that summarizes various different contents.
- *Ranking*, namely proposing only the most relevant contents according to a given order of relevance.

The framework could adapt to the data entered into the archive (multimedia data, metadata, and tags). This should occur not simply by returning different data (otherwise it would just be a normal data-driven system) but by *modifying the navigation* and *presentation methods* depending on the type of data. For example, the modification can be made dependent on: a) whether we are dealing with images or videos or texts etc., or b) on the value of the tags.

If applied to multimedia elements, some ML techniques can be used to generate and better structure the tags associated to the different media contents, in order to run further analyses of clustering or aggregations.

Our proposed *navigation and query mode* is conceived to display information on the map, considering the different granularity of data, depending on the displayed area. The higher the geographical zoom level, the higher the temporal detail and the semantic detail of the returned content information. In this way, information on the map is presented at different levels of granularity, automatically linked to each other.

The *zooming* function can operate onto the three dimensions: *spatial*, *temporal* and *thematic*. The first is the classic geographical zoom; the second groups events into macro-events or chronological periods; the third is a *semantic* zoom, where one can navigate a taxonomy dynamically derived from the context-specific ontology.

It is also interesting to *adapt* the user *analyses* to the context; this is a matter of current study.

5. Acknowledgements

This paper is partially supported by the Projects *BigData* and *Seamless* of Regione Lombardia, Italy, and by the EU Horizon 2020 research and innovation programme under grant agreement No. 826232, project *WorkingAge* (Smart Working environments for all Ages).

6. References

[1] H. Karimi, H. Zeinivand, Integrating runoff map of a spatially distributed model and thematic layers for identifying potential rainwater harvesting suitability sites using GIS techniques, Geocarto International, (2021), pp. 320-339.

[2] E. Purificato, A. M. Rinaldi, Multimedia and geographic data integration for cultural heritage information retrieval, Multimedia Tools and Applications, 77(20), 2018.

[3] T. Bracamonte, B. Bustos, B. Poblete, T. Schreck, Extracting semantic knowledge from web context for multimedia IR: a taxonomy, survey and challenges, Multimedia tools and applications, 77(11), 2018.

[4] D. Fernández, D. Varas, J. Espadaler, I. Masuda, J. Ferreira, A. Woodward, E. Bou, Vits: video tagging system from massive web multimedia collections, in: Proceedings of the IEEE International Conference on Computer Vision Workshops, 2017, pp. 337-346.

[5] H. K. Hong, G. W. Kim, D. H. Lee, Semantic tag recommendation based on associated words exploiting the interwiki links of Wikipedia, Journal of Information Science, 44(3), 2018, pp. 298-313.

[6] A. C. T. Klock, I. Gasparini, M. S. Pimenta, J. P. M. De Oliveira, Adaptive hypermedia systems, in: Advanced Methodologies and Technologies in Media and Communications, IGI Global, 2019, pp. 217-228.

[7] H. Anantharaman, A. Mubarak, B. T. Shobana, Modelling an adaptive e-Learning system using LSTM and Random Forest classification, in: 2018 IEEE Conference on e-Learning, e-Management and e-Services (IC3e), 2018, pp. 29-34.

[8] Y. L. Chi, T. Y. Chen, C. Hung, Learning adaptivity in support of flipped learning: an ontological problem-solving approach. Expert Systems, 35(3), 2018.