AGILE Pre-conference workshop
"Bridging space, time, and semantics in GIScience"
Wageningen, The Netherlands, 9 May 2017

Managing geospatial information and ontologies of historical maps: Empirical evidences from the analysis of Kitchener's survey of Cyprus

Christos Chalkias
Dep. of Geography,
Harokopio University
70, El. Venizelou Str.
Athens, Greece
xalkias@hua.gr

Christoforos Vradis
Dep. of Geography,
Harokopio University
70, El. Venizelou Str.
Athens, Greece
cvradis@hua.gr

Margarita Kokla
School of Rural and
Surveying Engineering,
National Technical
University of Athens,
9, H. Polytechniou Str.,
15780, Athens, Greece
mkokla@survey.ntua.gr

Abstract

Early historical maps are mostly cognitive maps and generally considered a part of our cultural heritage with a mainly aesthetic value due to their low positional accuracy. As maps of the late 17th and 18th century begin to present increased accuracy and precision, their proper documentation becomes a necessity in order to use the historical cartographic information effectively within the context of the semantic web. This short paper summarizes the typical spatial database creation of the 15 sheets from Kitchener's survey of the island of Cyprus

1 Introduction

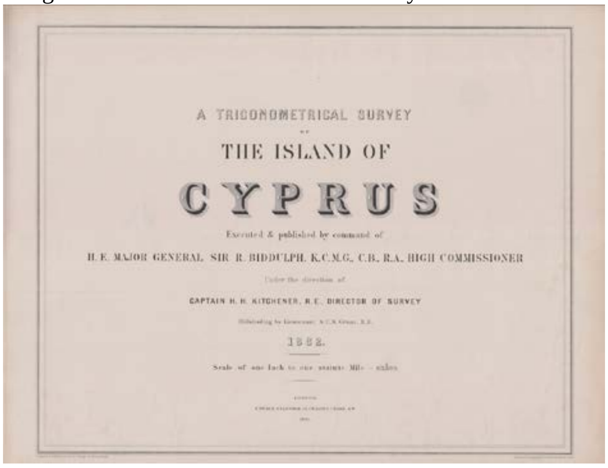
In 1885, Horatio Herbert Kitchener's map (the K-map of Cyprus was published in London by Edward Stanford of 55 Charing Cross Road, S.W. It is considered as the first scientific survey of the island. The K-map, consists of 16 sheets in total (15 plus the title sheet) and immediately became a cartographic model that lasted up until the middle of the 20th century. This surprisingly accurate map contains significant information regarding the geography, history and culture of the island and, even today, requires attentive research by the scientific community. The 15 sheets were georeferenced and selectively vectorised as part of a research project undertaken by the Geography Department of Harokopio University funded by Sylvia Ioannou Foundation. The vector data, initially created in ESRI shapefile format, were categorized in eight thematic layers and later stored in a MySQL database. A minimal web mapping interface was built to enable user exploration of the map's contents.

This standard approach for geospatial information handling and dissemination has already shown its limitations. Research on the semantics and interrelations of generic concepts and properties of objects like time and space, events, persons etc., all within some formal specification like an ontology, as well as the documentation of cultural heritage information using frameworks like the CIDOC Conceptual Reference Model are both quite relevant for the documentation of the content and semantics of historical maps. Finally, we identify potential problems and inherent shortcomings of the model regarding mainly geospatial aspects of the map and its contents.

2 The surveying of Cyprus from H. H. Kitchener

The surveying of the island began as soon as Cyprus became a British protectorate in 1878. It was immediately evident that a modern map was necessary to facilitate the administration of the island. The task was assigned to Horatio Herbert Kitchener, a British army surveyor and Foreign Office delegate. Contrary to his superiors, he envisioned a map that would '*be a model of its kind, and future scholars and archaeologists would be permanently in his debt*'. Although nearly abandoned halfway, the survey was completed by 1883 and the map was published in 1885 by Edward Stanford who promoted the map as follows: "*[…] It shews the Districts and Sub-districts [...] the Roads that have been constructed, and the Telegraph Lines erected [...] the Vineyards and Forests of Fir and other trees; block plans of the Towns and Villages, distinguishing Moslem and Christian Villages; gives the Greek and Turkish names, and the identification of Ancient Sites, in distinct types. Aqueducts, Springs, Wells, Monasteries, Ruins, and other particulars are engraved. Heights of Hills and Mountains, Towns and Villages, above the level of the Sea, are also pretty generally supplied, and Mile-stones, with the mileage written against them, engraved so far as they had been erected at the moment of final revision. [...]*". These key themes were used as the guide for identifying cartographic information. Whether these (and other) claims by Stanford were true, and how close Kitchener actually came to his vision was the main goal of the research project.

Figure 1: Kitchener’s survey title sheet.



Source: Harokopio University, Sylvia Ioannou Foundation collection.

2.1 The spatial database creation process

The 15 sheets were scanned at a high resolution (600 dpi) and georeferenced in a single coordinate system (WGS84). After careful examination and interpretation of each sheet, seven major thematic layers were created (administration, road network, hydrographic network, points of interest, settlements, land cover, various place names[1]) and the cartographic information was digitized on-screen. It should be noted that the legend in K-map of Cyprus is not a complete one as many cartographic symbols of the map are not included in this legend. Thus, to enhance symbol identification, we consulted, among other sources, the legend of the map from the survey of Palestine (that was also conducted by H. H. Kitchener) as it proved to be far more informative from the legend of the map of Cyprus.
Various attributes, annotations (like settlements, rivers etc.) and other place names were stored in appropriate DBF fields. Upon finishing vectorization, a MySQL database was created by importing shapefiles using the GDAL library on a OS X Mac computer.

2.2 Place names database

The various place names possess their own nomenclature, as we identified names for mountains, administrative regions, capes, simple point locations etc. From the fields used to store place names, we created a separate MySQL table holding all the names of geographic attributes along with their source layer and their geometry. The Gazetteer of Cyprus was used to store the modern name of settlements along with their Greek / Turkish appellation as shown on K-map thus enabling searching using the current, modern, names of settlements. This “index” table serves as the single point of reference within the database for faster searching through the map contents by name instead of searching individual tables and fields. The annotation of the K-map has various types of spatial / thematic properties. For example, the annotation of river and mountain names is presented as curved lines and the annotation of villages and towns as text in horizontal direction. Moreover, some text of the K-map corresponds to a thematic description of a point entity. For example with the text “.Ruin” or “R” the cartographer presented ruins on the map.

2.3 Web dissemination

As part of the research project, a minimal web application[2] was developed based on the open-source Leaflet JavaScript library. Through the application, the user may explore the mosaic in various scales, overlay all the vectorised data, and search for any named features.

The mosaic of Kitchener’s sheets was stored as a tiled layer (practically requiring almost no extra server-side software apart from a web server) as 256x256 JPEG files, while the vector data are accessed on demand from MySQL as GeoJSON data by means of a custom PHP script. Searching place names is also supported by means of a custom PHP script that connects to MySQL. This approach results into a small footprint, manageable application, with some basic user tools (map zooming and panning, layer activation/deactivation, feature identification, named features searching).

2.4 Limitations

With the approach described above, many simple and composite questions can be answered by a) examining the mosaic itself, b) by exploring the vector data and c) by means of SQL commands, since most visible attributes of the geographic features are stored in the database and these attributes have defined their grouping into layers. By using the web application, the user can view separate layers of information thus greatly facilitating map interpretation. However, there are some parts of the historical map that could not be treated the same way with the geographic data. The map legend for instance, although used by the research team to identify symbols it has not been part of the database. Additionally, the use of a legend from another similar map (that of Palestine) could not be efficiently documented as part of the final dataset. Another choice we had to make is whether to store individual cartographic symbols that form an areal pattern (as in the land cover layer) instead of creating a vector polygon that can have varying geometry due to possible different interpretation of the pattern. Finally, the transliteration guide embedded in the legend (used to transfer Greek and Turkish names into English), is not a part of the database.
Besides these shortcomings, it is also quite clear that this conventional modelling approach cannot provide a platform that can be used to infer any kind of knowledge or reasoning (spatial, temporal or both). In order to achieve this, we need to involve smarter tools.

[1] Place names in inclined or curved alignment referring to geomorphological data such as rivers, streams, mountains, bays etc.

[2] https://webms.hua.gr/geoprojects/kitchener/ (access requires user registration)

3 The semantic approach

3.1 Geospatial and other information in historical maps

Historical maps, as physical artefacts, are the means to store, depict and communicate historical geographic information. This information can be extracted either directly (by examining the map itself) or indirectly (e.g. by examining digitised / vectorised data). By comparing this with current information, we can deduce changes in geographic entities. Apart from the core geographic content, historical maps contain notes, ancillary text, legends, various ornamental features and more. The geographic content depicted by means of symbols may quite often be ambiguous in interpretation.

3.2 Scope of semantic modelling

Although bibliographic search proves that there are ontologies generally capable of describing artefacts and geographic features, historical maps lie in a rather special domain: they are means of delivering historical, spatial and geometric information as well as objects of art. As Scheider et al (2014) state, few authors have focused on using Semantic Web technology for map descriptions. Bittner and Smith stated that "a map is a specific, simplified and therefore highly efficient representation of the ontology of a certain part of geographical space. It is an ontology because it is an inventory of things that exist in a certain part of the world and of some of the properties and relations between them" (Bittner et al. 2004).

The question we need to answer initially is what kind of formal representation to use and before answering this question, we need to state a number of related questions regarding the map of Kitchener: What are the means to describe map contents involving both named as well as nameless entities? What about named entities (e.g. place names) that are not clearly connected to a specific geographic feature and do not have discernible physical boundaries? What is the most appropriate representation of the map content so that topology relationships can be both modelled and extracted from the geometry if necessary? How can we answer questions that may refer to this map as well as other sources such as *"Which settlements in Cyprus were abandoned in the late 20th century by their Christian / Muslim population?"*. Can all the individual map elements, like the map legend and the transliteration guide be properly documented and then re-used?

3.3 Desribing geospatial information in CIDOC - CRM

ICOM-CIDOC's "Conceptual Reference Model" (CRM) (registered as ISO Standard 21127 "Information and documentation — A reference ontology for the interchange of cultural heritage information") is an ontology designed especially for cultural heritage documentation.

3.4 Limitations of the CIDOC CRM and required extensions

As (Gkadolou 2013) points out, the CIDOC CRM class best describing the concept of a historical map is E84 Information Carrier. The E39 Actor class can generally describe efficiently all involved persons and related actions as creation and publishing can be modelled through the E12 Production class. Other classed involved are as follows:

E55 Type
E54 Dimension
E57 Material
E56 Language
E36 Visual item, used for all ancillary icons, pictures and other map composition elements.
E34 Inscription
E50 Date
E52 Time Span
E31 Document

Stefanakis and Gkadolou (2013) have pointed out that although historical maps, as artefacts, can be well documented in the CIDOC CRM model, there are major map elements (scale of map, spatial reference system, orientation, legend and any notes on metric information etc.) that require some class extensions. Most important is that geographic entities themselves and their relations demand a geographic domain ontology. The integration of geospatial information with CIDOC CRM has been addressed in a few ways (Hiebel et al 2013): the AnnoMAD System, the CLAROS project, the CRM-EH, the WissKI project. We consider of special interest the work to integrate OGC standards within the CIDOC CRM (Hiebel et al 2013).

3.5 Describing Kitchener's map in a CIDOC CRM compatible fashion

By following the ontology shown in figure 3, we can provide a sample semantic documentation of K-map as follows:

"Kitchener's topographic survey of Cyprus" *is a* Historical Map (E84 Information Carrier).
is identified by C.Chalkias (E42 Identifier)
has current owner Harokopio University (E39 Actor)
has current owner Sylvia Ioannou Foundation (E39 Actor)
  has type Foundation (E55 Type)
  has contact point ***** (E45 Address)
  has current location *** (E53 Place)
  falls within **** (E53 Place)
has title "A trigonometrical survey of the island of Cyprus" (E35 Title)
    has language English (E56 Language)
has type Copy of ***** (E55 Type)
has type Original **** (E55 Type)
  has current owner Sylvia Ioannou Foundation (E39 Actor)
has number of parts 16 (E60 Number)
has number of mapsheets 16 (MapSheet)
has dimension (E54 Dimension)
  length 63.5 cm
  width 53.32 cm
consists of Paper (E57 Material)
was produced by MapCreation of *** (MapCreation)
  carried out by Horatio Herbert Kitchener (E39 Actor)
    has type Cartographer (E55 Type)
    has type Director of Survey (E55 Type)
    has type Captain (E55 Type)
    has type Commander (E55 Type)
    has type Surveyor (E55 Type)

carries MapComposition of ***
  has scale 1:63 360
  has Orientation North
  has PrimeMeridian Greenwich
  was created by Topographic Survey (E7 Activity)
  took place at Cyprus (E53 Place)
    has type Country (E55 Type)
    falls within Cyprus (E53 Place)
  carried out by Horatio Herbert Kitchener (E39 Actor)
  was produced by Publication (E12 Production)
    carried out by Edward Stanford (E39 Actor)
      has type Publisher (E55 Type)
    took place at London (E53 Place)
      has type City (E55 Type)
      falls within United Kingdom (E53 Place)
    used specific technique **** (E29 Design or Procedure)
    has time span
      begins at 09/1878 (E50 Date)
      ends at 02/1883 (E50 Date)
depicts Cyprus (E53 Place)
    has type Country (E55 Type)
depicts Nicosia (E53 Place)
  has type City (E55 Type)
  falls within Cyprus (E53 Place)
depicts Larnaca (E53 Place)
  has type Region (E55 Type)
  falls within Cyprus (E53 Place)
is part of the Map collection of the Sylvia Ioannou Foundation (E78 Collection).

4 Conclusion and further research undertaken

The CIDOC-CRM greatly emphasizes information organization around the historical artifact while our needs are central to the contents of the artifact with a strong inclination towards robust geometry handling. A cartographic ontology for historical maps has to be centered on the concepts of map, graphic elements and symbols. The ontology presumes that every representation of a historical map is composed essentially of some graphic symbols (either cartographic symbols or text symbols). The building units for digital mapmaking are the visual variables (colour, opacity, texture, orientation, arrangement, shape, size, focus). The graphic elements and the visual variables are represented by cartographic symbols. Another type of crucial symbols in historical maps are the text symbols (groups of text characters). These symbols may be directly related (or not) to specific geospatial entities. Moreover, the cartographic domain ontology must also handle the complexity of map semiotics. The hierarchy of this ontology is presented in figure 4.

The CIDOC CRM is not designed to handle complex geometric properties and topological relations thus creating the need for an extension. The extension includes the following future steps:

a) Determine the geographic ontology and the required vocabulary that will extent CIDOC CRM. The scope of this step is to document as fully as possible the geographic entities contained in the historical map of Kitchener, their spatial relations that need to be directly modelled as well as those that will be inferred and extracted from their geometry.
b) Define the way that each cartographic ontology is applied in order to be able to document the spatial, thematic and temporal dimension of the entities.
c) Document the semantics of some key elements of the map such as cartographic symbols, legends, notes and guides.
d) Define the relationships between key map elements and cartographic geospatial entities

Regarding the adopted ontology, although fitting the initial purpose, we consider that there can be ways to better express relations between the whole and its parts, e.g. the relation between an individual map sheet and the historical map as a whole.

Acknowledgements

This research supported by Sylvia Ioannou Foundation, which also provided the original copy of the K-map from its historical map collection.

References

Bittner T, Donnelly M, Smith B (2004) *Endurants and perdurants in directly depicting ontologies.* AI Communication 17(4): 247-258

Chalkias, C. et al (2017) *Creation of the cartographic database and development of a web application presenting the geographic information of Kitchener’s map of the island of Cyprus*, Harokopio University, Athens, Greece, 2017

Goerz, G. and Scholz, M. (2010) Semantic Annotation for Medieval Cartography: The Example of the Behaim Globe of 1492. In *Proceedings of the CIDOC 2010 ICOM General Conference*, Shanghai, China, 2010.

Gkadolou, E. and Stefanakis, E. (2013) A formal ontology for historical maps. In: *Proceedings of the 26th International Cartographic Conference, August 25 – 30, 2013*, Dresden, Germany, 2013.

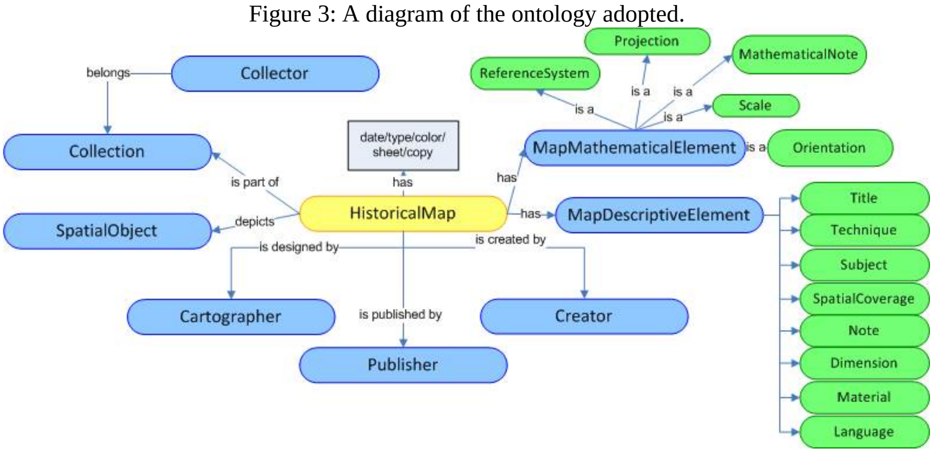
Hiebel, G., Doerr, M., Eide, Oyvind, E. (2013) Integration of CIDOC CRM with OGC Standards to Model Spatial Information. In: *Proceedings of the 41st Conference in Computer Applications and Quantitative Methods in Archaeology*, Perth, Australia, 2013.

Scheider, S., Jones, J., Sánchez, A., Keßler, C. (2014) Encoding and Querying Historic Map Content. In: *Connecting a Digital Europe Through Location and Place.* 251-273

Figure 2: The mosaic of Kitchener's 15 sheets.

Source: Sylvia Ioannou Foundation and Harokopio University.

Figure 3: A diagram of the ontology adopted.



Source: http://ontologydesignpatterns.org/wiki/Image:HistoricalMap.jpg

Fig. 4: Hierarchy of the domain ontology for Historical maps