GIS of Traditional Folk Culture

Jiří Kozel¹, Pavel Bohumel², Petr Kovács³, Andrea Kýnová⁴

¹,²,³,⁴ Masaryk University, Institute of Computer Science

Abstract

This paper describes the process of the construction of the GIS of traditional Czech folk culture between 1750 – 1900. Our project is the first that collects records about all the aspects of folk culture in the Czech Republic systematically. We started with the design of a data model well-suited for heterogeneous folk-culture metadata. At present, we have about 35,000 records in the database, an editing application, and a web map presentation with spatiotemporal filtering written in Google Closure and OpenLayers 3.

Keywords

GIS, folk, culture, data model, web map, filtering

1 Introduction

Traditional Czech folk culture has a long and rich history, especially in the region of Moravia, which covers one third of the Czech Republic. Folk culture was an important vehicle of patriotism during the Austro-Hungarian domination, which took place for four centuries. Many of those customs, songs, crafts, and arts are still alive among ordinary people. Four items from the UNESCO Intangible Cultural Heritage List (314 items in total) are located in the Czech Republic, including the recruit dance "verbuňk" and the "Ride of the Kings" located in Moravia (UNESCO, 2015).

1.1 State of the Art, and Objectives

Before our project started, there had been no systematic collection of all the aspects of the traditional folk culture in the Czech Republic. There exist some web map presentations related only to a single topic of folk culture, e.g. “Hrady a zámky” (http://www.hrady.cz/), related to architectural objects, or the Folk Costumes Map currently being developed at the Czech Institute of Folk Culture. Our goal is to create a systematic collection of metadata records of all the aspects of traditional folk culture in the region of Moravia between 1750 – 1900, and to build a web GIS that would enable editing, presentation, and advanced filtering of the collected data. This unique collection and developed applications enable subsequent research of a given topic, and help to preserve our cultural heritage.

Nowadays the database contains approximately 35,000 records collected by ethnological experts. Each record has attributes that describe place, time, category, and source. Furthermore, we have an editing application and a web map presentation with spatiotemporal filtering written in Google Closure and OL3. This paper focuses especially on the geoinformational aspects of the
building of this collection and GIS.

1.2 Who Stands Behind
This paper forms a part of the project “Geographical information system for traditional folk culture (1750–1900)”, in short “GISTraLiK“. The project is supported by a grant from the Ministry of Culture of the Czech Republic as project No. DF12P01OVV015. Two cooperating partners are involved: the Department of European Ethnology and the Institute of Computer Science, both belonging to Masaryk University in Brno.

2 Data model
The first and the most crucial step was to create a data model well-suited for heterogeneous metadata concerning folk culture. It took hours of discussion among ethnological and GIS experts. We had to deal with a certain level of uncertainty (especially in the case of temporal designation), different levels of geographic localities ranging from municipalities to ethnological regions, and the extremely wide scope of folk culture including customs, products, songs, artworks, etc. Basic concepts of the data model are illustrated in Figure 1.

2.1 Record
The record is the main class of the model. It represents either a mention of a custom, craft, song, social relations, event, etc., or an artwork (e.g. painting, ceramics, or literature).
Each record has several important mandatory properties including keywords, spatial and temporal validity, and the source of the represented mention or artwork. The multiplicity of keywords is 1–n, the multiplicity of other attributes is 1–1. These attributes allow us advanced filtering and record searching in the presentational part of the GIS.
The textual description is an attribute with a free domain, whereas keywords are selected from controlled vocabulary (see the next subsection).
Temporal validity is represented by year range (i.e. integer range) with an optional text. This allows us to store and search for even approximate temporal information, such as “beginning of the 19th century”, “mid 1830s”, or “2nd half of the 19th century”.
Spatial validity is represented by one one locality (see below for details).
A source is either a document that contains a mention about folk culture, or an object that contains an artwork (see below for details).

2.2 Locality
A locality is an abstract class representing a spatial unit; there are five subclasses: a municipality, parish, manor, legal district, and an ethnological region. The first four subclasses are administrative or territorial units that were legal for at least a part of the period 1750–1900. Ethnological regions are used by ethnologists to demarcate areas with similar folk culture.
The spatial extent of each locality is represented by one polygon or multipolygon. It means that we are not able to track the changes of boundaries in time, we just use the widest known delimitation of a given locality in the given period.
The overlapping of localities of the same type is allowed. There are three
reasons for that: the changes of boundaries in time, it is not always possible to delimit a historical boundary exactly, and ethnological regions simply overlap even in reality.

2.3 Keywords
Keywords represent the ethnological classification of records. Keywords form a hierarchical acyclic structure. The highest level contains the most general keywords, whereas the lower levels contain more specific keywords, e.g. Music -> Musical instrument -> Bagpipes. One keyword can appear as a subkeyword of more keywords, e.g. Food -> Egg and Custom -> Easter -> Egg. The connection between a record and a keyword carries information on whether the keyword (e.g. Egg) is used in a general context or whether it relates just to a specific branch of the structure (i.e. to a specific meaning).

2.4 Source
A source is either a document (e.g. a chronicle, book, article, or a map) or an object that contains an artwork (e.g. a museum exhibit). Every source carries complete citation information in order to be identified easily (e.g. authors, year of publication, publisher, magazine, museum, exhibit No., etc.). An attachment (usually a photo or a scan) and a URL that leads to the web page with the electronic version of a given source are very important optional attributes of a source. These attributes enable users to study a record directly without a physical visit to an archive, library, or a museum.
3 Data acquisition

Before the records collection started, a complete index of keywords and complete register of localities had been created.

3.1 Keyword Index

The compilation of the keyword index was done specially by ethnologists because of their expert knowledge of the traditional folk culture phenomena. Attention was paid especially to covering all the phenomena equally to avoid a later expansion of the index. This complex task took about a year to complete. Finally, the index contained 1260 keywords separated into 3 hierarchical levels; the first level contains 28 keywords (Drápala, Doušek, Křížová, Pavlicová, and Válka, 2013). The whole index was also translated from Czech into English.

3.2 Localities Register

The register of localities, with the exception of ethnological regions, was made according to historical documents, such as old cadasters (http://archivnimapy.cuzk.cz/) or maps of Moravian regions by Schenkl, 1841–1845. The numbers of localities by type are available in Table 1. A list of 15 ethnological regions was prepared by ethnologists. The register initially contained only names of localities and their identifiers.
Subsequently, we also filled the geometry attribute specifying the locality area using a polygon or multipolygon. The geometry was derived especially from the vector database of old cadasters made by CUZK.

3.3 Records and Sources
First, ethnologists made a list of about 200 museums, archives, and collections which cover not only all aspects of traditional Moravian folk culture, but also the whole geographic area of Moravia and the whole temporal period (Horáková & Šipöczová, 2013). Similarly, a list of scientific publications and magazines related to the field of interest was made.

Then ethnologists started to collect sources and records. The latter are related to the source, existing keywords, and an existing locality. We had a three-month testing phase to evaluate the data model, and after the evaluation we published the appropriate methodology (Drápala & Malecká, 2013). Today there are about 35,000 records in the database processed by ethnologists according to the methodology.

4 Database and Web Applications

4.1 Database
Prior to data acquisition, a data model had been implemented in the database, and an input application of acquired data had been built. Because of contractual reasons, we used Microsoft SQL Server with “geometry” spatial type as the database engine. Two databases are used: the “source database”, which contains the source data that are managed by the editing application, and the “publication database” with a flattened model used exclusively for the web presentation. The flattened model enables providing and querying the data faster. The publication database is automatically updated once per 24 hours from the source database.

4.2 Web Applications
Since the input application is not public and it is a text-form application without maps, it was omitted from this description.

Parallel to the records and sources acquisition, we started to build a web presentation application with an interactive map and records filtering. It is located at http://gistralik.muni.cz/. We used OpenLayers 3 (http://openlayers.org/) and Google Closure libraries (https://developers.google.com/closure/library/) for client-side interactivity. The server side that serves data from the database to the client is managed only by ArcGIS REST API (http://resources.arcgis.com/en/help/arcgis-rest-api/), used because of contractual reasons. The technical aspects of the client-side solution are described in section 4.3.

The client side has two separate pages: search in the database (vyhledavani-db.html) and search in the map (vyhledavani-mapa.html). Both pages enable to search records by advanced filtering; the former enables to browse found records whereas the latter visualizes the number of found records per locality in a simple choropleth map.

The filtering of records is a very important part of the two pages. Filtering according to locality, time range, keywords, source, and a full-text search is
enabled. The current state of the filter control immediately updates the list of found results or the numbers of found results in the map. Furthermore, the state of the filter is stored in the URL so that everyone can copy the link and go back to the current state anytime later. The web presentation exists in Czech and will be translated into English by June 2015.

4.3 OpenLayers 3 in Web Presentation Application

The ArcGIS REST API is used only for providing vector and attribute data in the JSON format that is similar to GeoJSON. We created OpenLayers 3 class ol.format.EsriJson as a subtype of ol.format.Feature. This class is able to parse ESRI JSON and create ol.Feature instances including geometries (e.g. point, polygon, or multipolygon). We use ol.source.ServerVector class for storing features on client side.

The locality layer is represented by ol.layer.Vector with custom styles (ol.style.Style) and the whole map is rendered to HTML canvas. Even with massive caching of styles and preprocessing (geometric generalization), we reached the limits of the browser reaction time, because too many polygons and vertices are visualized (see following table). Whereas the zoom of the legal districts layer is usually smooth in current browsers, the zoom of the municipalities layer is not.

<table>
<thead>
<tr>
<th>Locality Type</th>
<th>Number of Features</th>
<th>Number of Vertices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal districts</td>
<td>77</td>
<td>14,040</td>
</tr>
<tr>
<td>Manors</td>
<td>415</td>
<td>44,960</td>
</tr>
<tr>
<td>Parishes</td>
<td>873</td>
<td>54,487</td>
</tr>
<tr>
<td>Municipalities</td>
<td>3,163</td>
<td>176,557</td>
</tr>
</tbody>
</table>

*Table 1 - Numbers of Features and Vertices per Feature Type.*

Two further principal extensions of the OL3 library were made. First we created ol.object.Type and ol.feature.Type, which are a simplification of FeatureType of ISO 191xx series (ISO, 2001). This helped us to load and store features and their properties on the client side automatically. Then we extended ol.Object so that every property may have a validator. The validator is a function called anytime an attempt of setting a property value is made. If the validator fails, the new value is not set. This helps us to build robust model classes with business logic separated from the view components.

5 Conclusions

This paper presented the building of the GIS of traditional Czech folk culture in the region of Moravia between 1750 – 1900. The project is realized by GIS and ethnological sections. Our field of study – folk culture – is not a field of study where GIS would be used commonly, so attention was paid especially to the data model and to the ability of web presentation to find and visualize the intensity of folk-culture records according to the filter used by the user easily. The created GIS enables subsequent research of Moravian folk culture and
helps to preserve our cultural heritage. The project is going to continue at least until the end of 2015.

References


