

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676060

LONGPOP

**Methodologies and Data mining techniques for the analysis
of Big Data based on Longitudinal Population and
Epidemiological Registers**

***Tool to visualize indicators of
environmental exposures***

Deliverable n. 2.4

Disclaimer: This publication reflects only the author's view and the Research Executive Agency is not responsible for any use that may be made of the information it contains.

WP2: 3-ULUND Deliverable 2.4 – Tool to visualize indicators of environmental exposures

1. Background

Centre for Economic Demography, Lund University (ULUND), has built up a demographic database SEDD in recent decades (Bengtsson et al., 2014). SEDD contains demographic information on all persons who have lived in five rural parishes in southern Sweden from the 1700s to today. In addition, we have digitized historical maps covering these five parishes and transformed them into digital map datasets containing object-lifeline representations of geographic objects. Thereafter we have geocoded the population in the SEDD (for the period 1813-1914) at the micro-level; in total, 53,000 individuals have been linked to the property units where they lived (c.f. Hedefalk et al. (2017)). These database has been utilized to finalize the deliverable for ULUND, which is described in detail in the following text.

Related to the work within WP2 in the LONGPOP project, we have developed a map service/tool for a historical population in five Swedish rural parishes (Deliverable 2.4). The map service contains historical maps and geographical data for the period 1757–1914, as well as geocoded demographic data. Historical objects such as property units, buildings, communications and wetlands, as well as modern geographic data such as elevation and soil type, are included in the map service. By combining the geocoded demographic data with the historical and modern geographic data, it is possible to visualize indicators of environmental exposure such as soil conditions and the natural landscape, population density, and distance to wetlands. The tool is targeted both to researchers and to the public.

The demographic data are taken from the SEDD database established by the Centre for Economic Demography (CED), Lund University, in collaboration with the Regional Archives, Lund (Bengtsson et al., 2014). SEDD has been used extensively in research during the last decades. SEDD contains demographic and economic information about all the people that lived in five parishes, located in southern Sweden, named Hög, Kävlinge, Kågeröd, Sireköpinge and Halmstad, from the 17th century onwards. The primary sources for SEDD are vital registers, annual poll-tax registers, catechetical examination registers/parish registers and land surveyor acts. The database also includes extensive contextual information. In this tool the user can view an extract of the SEDD database containing around 53,000 individuals from 1813 to 1914.

The historical geographic data are mainly based on the historical maps from four map series described in Hedefalk et al. (2017). The historical maps have been scanned, geocoded and digitized. Approximately 1150 property units, 4,800 buildings as well as a substantial amount of roads, railways, streams and wetlands have been digitized. In this map service the user can view both the scanned and digitized maps.

Lastly, the individuals have been linked to property units (Swedish: Hemman). Approximately 53,000 individuals who lived in the five rural parishes are linked to the property units where they lived (Hedefalk et al., 2017). In the tool the user can view how an individual has moved during the course of his or her life as well as explorative information on indicators of environmental exposures.

2. Overview of the Map Service

The start view of the tool is shown in Figure 1, which shows the borders of the five parishes. The map service contains a base map as background. The user can select which base map that is shown (section 3.1). On top of this base map the user can view historical map information (section 3.2), historical exposure indicators (section 3.5) as well as statistics and animations in time (section 4.2). The map service can be used for several types of studies. In sections 4.1 and 4.2 we provide two examples of use cases that can be performed.

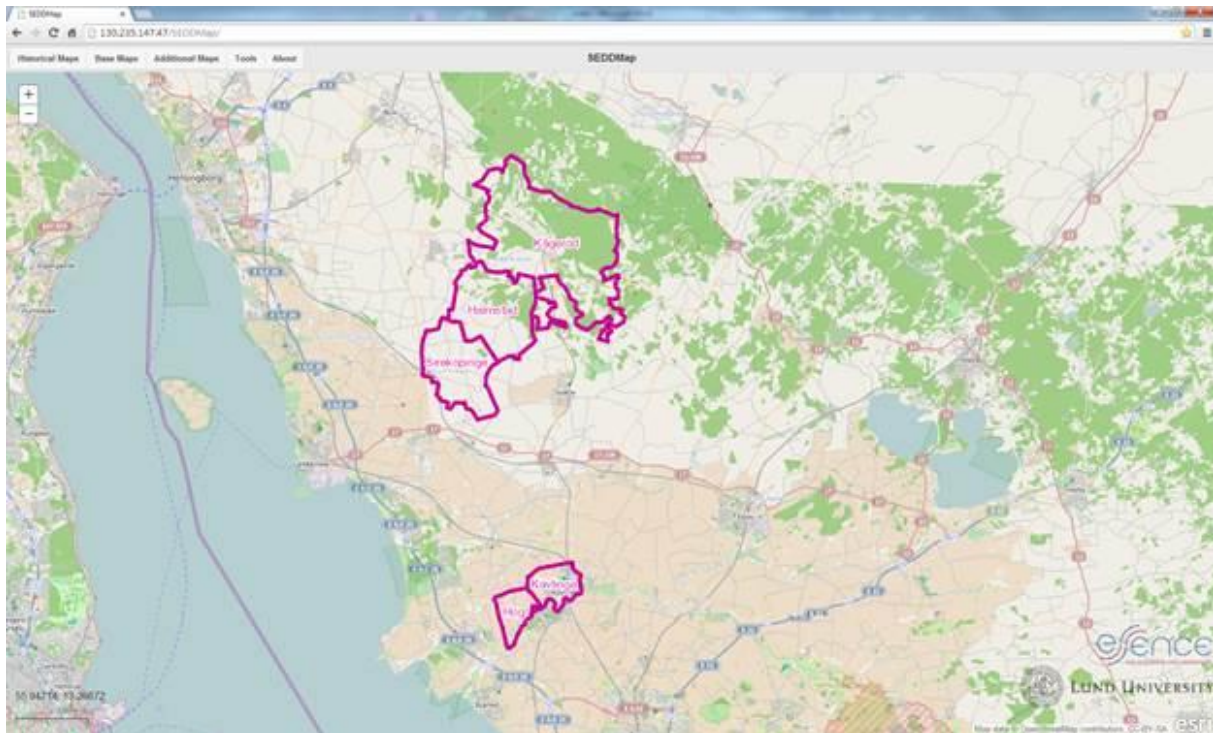


Figure 1: Overview of the map service/tool. The parishes Sireköpinge, Halmstad and Kågeröd in the north, and Hög and Kävlinge in the south.

3. Maps and geographic/demographic data in the tool

The geographic data available in the tool are: (1) base maps (modern and historical); (2) scanned historical maps and historical geographic vector data; and (3) exposure indicators. The demographic data are based on individual-level information which has been linked to the property units. Below follows a description of these data.

3.1 Base Maps

It is possible to select several base maps on which the historical scanned maps and digitized geographic vector data can be overlaid. These base maps are in raster format (and often retrieved using external web services). The base maps are of two types: maps and images. Most of the maps show modern geographic data, with some exceptions as the 1940 aerial photography. The following base maps are available:

- OpenStreetMap: Volunteered topographic data (see <http://OpenStreetMap.org>)
- Bing - Roads: Topographic data from Bing maps (Microsoft)
- Bing - Aerial: Aerial photos (orthophotos) from Bing maps (Microsoft)
- Bing - Aerial with Labels: Aerial photos (orthophotos) from Bing maps with additional labels added on the map (Microsoft)
- 1940 Aerial Photography - Aerial photographs (orthophotos) from the 1940s from Lantmäteriet and the GIS Centre, Lund University (see <http://www.giscentrum.lu.se/english/geodataArielPhoto.htm>)
- Topography (Hillshade): A hill shade map created from the national height model of Sweden (2 meter ground resolution) created by Lantmäteriet. The hill shade map has the sun placed in north-west direction. Be aware that the national height model is a ground model where buildings, vegetation, etc. are removed.
- Lantmäteriet's Topowebb. A detailed topographic map from Lantmäteriet. Be aware that you currently need an account associated with Lund University to view this map. Furthermore your browser must allow pop-ups from this page.

- Lantmäteriet's Cadastre. Lanmäteriet's Topowebb map with added property unit borders. Be aware that you currently need an account associated with Lund University to view this map. Furthermore your browser must allow pop-ups from this page.

3.2 Historical Scanned Maps

The following georeferenced historical maps are available in the tool.

- Economic maps from 1910-1915 (EM)
- Topographic maps from 1860-1865 (TM)
- Military topographic survey map of Scania from 1812-1820 (MTSM)
- Land surveyor maps from 1757-1863 (LSM)

These maps has been obtained in digital format from the Lantmäteriet (the Swedish mapping, cadastral, and land registration authority) (Lantmäteriet, 2018). The LSM maps primarily focused on documenting juridical borders, such as property units, but also documented buildings and land use, communications and the natural landscape. The MTSM and TM were constructed for military purposes and thus, focused on the natural landscape, land use, topography, communications and physical objects, such as buildings. The EMs documented property unit borders and the natural landscape, land use, communications and buildings.

3.3 Digitized Geographic data

The geographic information includes the following feature types: property units, wetlands, buildings, roads and railroads. The property units and wetlands are stored in object-lifeline time representations (information about creation, changes and ends of objects are recorded in time), whereas the other feature types are stored as snapshots in time. The user can access these geographic objects by clicking on a drop-down list to the left of the web map (Figure 2). For example, in figure 2, houses and roads have been selected from the economic and topographic maps as well as the parish borders from the economic map. The digitized objects are visualized in middle (parish level) or detailed (property unit level) scales. The map tool is not designed to visualize the historical data in an overview map.

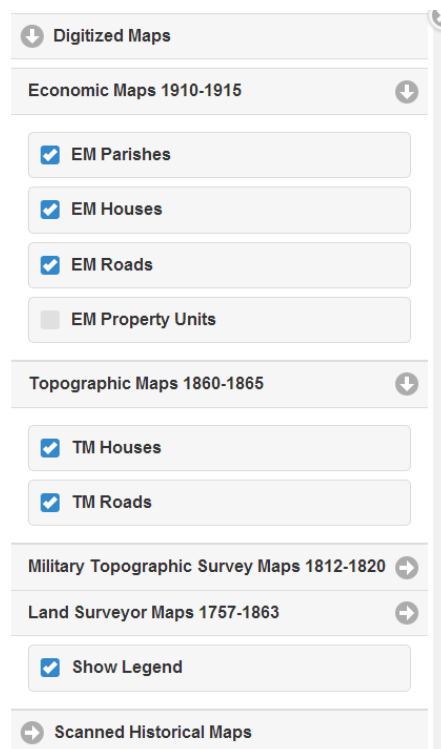


Figure 2: Drop-down list of the geographic data digitized from each historical map series.

3.4 Linked individual demographic data

Approximately 53,000 individuals who lived in the five rural parishes are linked to the property units where they lived. In the map service the user can view how an individual has moved during the course of his or her life (see section 4.2). This information can thereafter be combined with information on the exposure indicators.

3.5 Geographic and demographic exposure indicators

The exposure indicators contain information that is useful for exploratory analysis of the historical data. Because of the historical nature of the data used in the tool, these exposure indicators are limited to the historical data available, except for some modern data that are thought to be static in time (e.g., soil type and elevation). The current environmental and demographic exposure indicators in the tool are:

- Elevation contours (1 or 5 m) (static in time): Elevation contours based on the national height model of Sweden. In detailed maps (i.e., when the user has zoomed in to a small region) the contour level is 1 meter while in other maps the contour level is 5 meters (Obtained from Lantmäteriet).
- Soil type (>1: 50 000) (static in time): Soil data from the Swedish Geological Survey (SGU) (SGU, 2018). To see the most detailed classification of the soil data, the user must zoom in to the map (to only a few property units). When the user zoom out to view information about more than one parish, only the main soil type classes will be shown. An explanation of the soil types in the map is shown if the user click in the bottom of the menu on Soil Map Legend.
- Wetlands as well as distance to wetlands in metres (both dynamic and static in time)
- Population – total population in each property unit (the 30 June for the particular year) (dynamic in time)
- Population density (people/ha) – population per hectare (ha) (dynamic in time). One hectare equals 10,000 m². The area here is the total area of the property unit; that is, it does not represent the cultivated area.
- Population density (people/mantal) – population divided by the taxation value of the property unit (dynamic in time).
- Mantal - a measure of the productivity of a property unit based on the quality and size of the fields. Mantal was also used as a basis for taxation (dynamic in time).
- Property rights – states the type of property rights (crown, church, freehold or noble) (mostly static in time).

4 Main functionalities

4.1 Visualize exposure indicators and other data through time

The map service displays temporal information about property units and individuals (that are linked to these property units). This visualization is based on a database that contains information about when each property unit was created, changed and deleted. As mentioned previously, to the property units we connect other information such as which individuals that lived there and e.g. property rights (crown, church, etc.). This database enables visualization of the property units (and related information) for each year in an animation. The user can use a time slider to move forward or backwards in time.

To visualize some specific factors, the user can first select the factor of interest and thereafter select the year of interest using the time slider. In this case the base map is of less interest since the added geographic information will cover the information in the base map. The first step is to add information about property units on top of the base map. This can be done by clicking on the *Visualize Through Time* in the Tools menu. Thereafter the exposure indicators can be selected. The user can start an

animation by pressing the start button to the left of the time slider or jump single steps by using the buttons to the right of the time slider. It is also possible to change the year by clicking (or dragging) on the time line in the time slider. For example, in figure 3, the user have selected the taxation values of the property units for Hög parish. The user can thereafter use the time slider in the bottom of the figure to animate time and how the taxation values as well as the property unit borders are changing in time. It is also possible to select any of the other exposure indicators such as wetlands, population density etc. Lastly, the users can zoom in to each property unit to study in detail the geographic aspects of them; for example, by adding the detailed information on elevation.

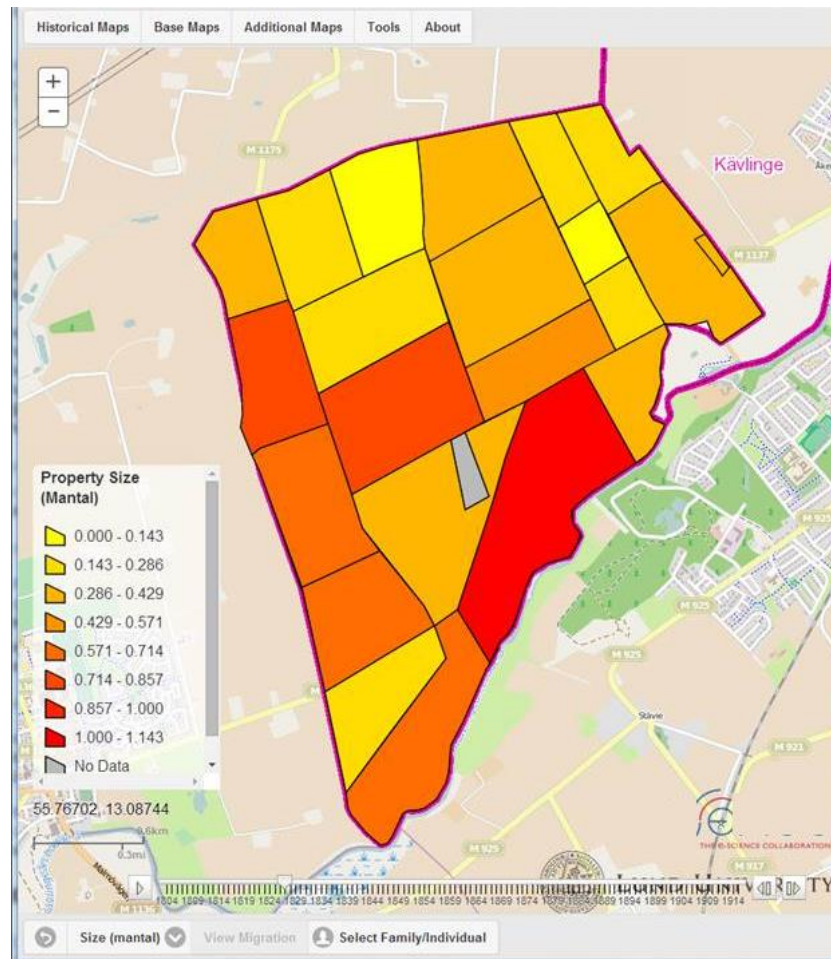


Figure 3: Taxation value and property unit borders of Hög parish for a specific year. The time slider in the bottom of the figure can be used to animate time.

4.1 Visualize Individual Demographic Data through time

The user can start by selecting the background information which in this case is the base map and some digitized historical data. If they are interested in the living conditions for a historical person the best base map is 1940 Aerial Photography. This is selected in the Base Maps menu. Another option would be to have a scanned historical map in the background (selected in the Historical Maps menu under Scanned Maps).

On top of this base map, the user can select appropriate historical vector information. Which information to select depends on when the people lived. If, for example, the users are interested in a person that lived early in the 19th century they could select LSM Houses, MTSM Roads, MTSM Streams, MTSM Lakes and MTSM Wetlands. All of these selections are done in the Historical Maps menu under Digitized Maps.

Thereafter, to study information about individuals and families the user can press on the button Select Family/individual in the tool. The users then gets a new menu where it is possible to specify the name, birth year, sex and/or location of the person they are interested in. The location can be selected either by name or by clicking in the map. Thereafter a list will be provided with all persons that satisfy the search condition. The users can select the person, or family, they are interested in by clicking the corresponding row in the list. The user will then be presented with a chronological list of events as well as a map view of the events (Figure 4).

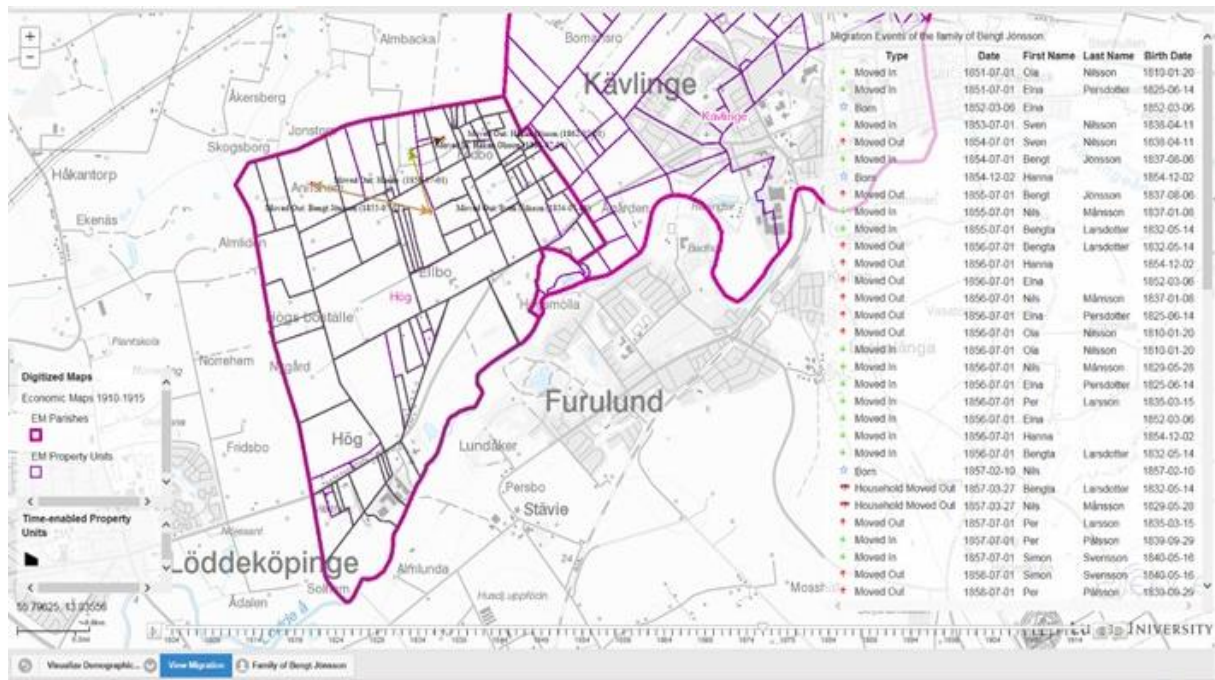


Figure 4: Chronological list events (migration, birth, death, marriage) for individuals belonging to a specific family (to the right in the figure), which are shown in the map dynamically as the time slider is used.

5. Technical information and installation

5.1 Accessing the tool

A zip-file of the web map application can be downloaded from the following [link](#).

To install the web application, see sections 5.2 and 5.3 for technical information.

5.2 System architecture

The architecture used by the web application consists of three separate servers with different roles, and a workstation for managing the system. These are:

- A database server for storage of data. Most of the vector-based geographic data and some demographic data are stored in an external database. The raster-based geographic data are stored locally on a server.
- ArcGIS for Server to publish and manage web map tool.
- A standard web server, developed in in Microsofts ASP.NET Web API, which enables server-side requests to the database.

Thus, when utilizing the web application on the user's own data after extracting the zip-file, the users need to set up their own architecture of how to access and store their data. This could be in an external

database (as mentioned above), on a local server, or directly within the ArcGIS environment. Moreover, since the application uses the ArcGIS for Server, the users need to have access to either ArcGIS Server or ArcGIS Online (some modifications needed) to utilize the application in its current form. However, several parts of the code for the tools can be used in other environments as well.

5.3 Application structure

The application has been developed using HTML, CSS and JavaScript. The main JavaScript library used was ESRI's ArcGIS API for JavaScript (v. 3.8). The ArcGIS API for JavaScript is built on the Dojo library, but jQuery (1.9.1) were also used for code simplification and browser compatibility. The Dojo version used by the ArcGIS API for JavaScript uses the Asynchronous Module Definition (AMD) coding format, which is modular, portable and handles dependencies well among other advantages.

The application is run on four tiers; on the client, on the web server, on the separate GIS- and database servers and on external web services. When the application is used, a client web browser starts by making an HTTP connection to the web browser and downloads the index.html file. It defines the structure of the web page and includes references to style sheets (CSS), static content such as images and to JavaScript libraries that should be loaded.

For maintainability and well structured code, styles and logic (scripts) are kept separated from the structure of the web page in the HTML file. It should be noted, however, that the jQuery Mobile JavaScript library, when loaded, reads and reconstructs the Document Object Model (DOM) in index.html based on given options.

When the client web browser loads index.html, it also loads d6_4.js, which is the file containing the main logic of the application. This script, in turn, creates the map object which handles the displaying of base maps and geographic data to the user. It then adds the map services from the GIS server for the map object to display. It listens to user interaction with button objects on the web page and, when required, it loads additional modules to handle specific tools and functionality (most of which are loaded as the application is initialized).

Through the map service objects created in d6_4.js, map images are fetched from the GIS server using http requests. The GIS server makes SQL requests to the database server for the vector data and reads the raster data locally.

The ChartTool.js and TimeTool.js modules accesses non-geographic data in the database by making HTTP requests to an ASP.NET Web API web service which is run on the web server. It makes SQL queries to the database and returns the results as JSON to the client web browser.

Lastly, d6_4.js and its modules (e.g. BaseMaps.js) also access external web services, such as WMS services from Microsoft's Bing Maps etc.

6 References

- Bengtsson T., Dribe M., Quaranta L., & Svensson P. (2014). The Scanian Economic Demographic Database, Version 4.0 (Machine-readable database), C.f.E.D. Lund University, Lund.
- Hedefalk, F., Svensson, P., & Harrie, L. 2017a. Spatiotemporal historical datasets at micro-level for geocoded individuals in five Swedish parishes, 1813-1914. *Nature - Scientific Data* 4 (170046).
- Lantmäteriet. 2018. Historical Maps. Available at <http://historiskakartor.lantmateriet.se/arken/s/search.html>
- Swedish Geological Survey (SGU). 2018. PRODUKT: JORDARTER 1:25 000-1:100 000. Available at <http://resource.sgu.se/dokument/produkter/jordarter-25-100000-beskrivning.pdf>