INCOME
Water Care

A.3.1 Establishment and maintenance of database system
Final report

Geological Survey of Slovenia

Katarina Hribernik

Jasna Šinigoj

Martin Podboj

Matija Krivic

Ljubljana, June 2012

Contents

Figures

1 Introduction

The present report presents activities and results of the action A.3.1. - Establishment and maintenance of database system, performed by Geological Survey of Slovenia within the project INCOME (LIFE07 ENV/SLO/000725).

The main objective of the project is long-term effective management of aquifers and preservation of the quality of these water resources for future generations. The project INCOME is co-financed by the European Commission, Municipality of Ljubljana and the Ministry of the Environment and Spatial Planning.

In this action data from different sources and disciplines (physical and socio-economic) were integrated into one consistent database system which enables storing, retrieving, displaying and manipulating all data gathered during the project and other available data relevant for realization of the project. By establishing final application, faster and user-friendly data access with various possibilities of analysis was provided.

The database system consists of two integrated parts:
- relational database, which relates information in tabular way so that rules of relational algebra can be applied,
- geographic information system (GIS), which relates spatial information.

During the project we defined the procedure of database construction, which is called the data modeling and is performed in some logical steps.

As its result the entity-relationship data model (E-R model) is represented.

For developing the database we used MS SQL Server 2008, while physical model is made with program package MS Access.

We also established the internet application, HTML viewer and map based application using ESRI's ArcGIS Server 9.3. The result is fast, user-friendly, stable and easily upgradable application.

2 Database of monitoring points

a. The first step of database construction was the **study of the model of considered problems**. It consists of examination of existing data, collection and preparation of information, agreements with owners and authors of data. The basic components of the new data collection are defined by team work.

b. The next step of data modeling was the definition and description of relations between objects and their functionality. The result of these processes is the **conceptual model**, which represents the simplification, or ideal interpretation of reality with respect to objects and their interrelations.

In this phase all Slovenian partners have started the discussion about the data organization, necessary for the establishment of relation database system.

We developed the conceptual model three basic groups of data: elementary data of monitoring points, average values of chemical measurements for time period 2005-2008 and chemical analyses – time series of the parameters.

Elementary data of monitoring point are:

- Aquifer
- Name of monitoring point
- Type of monitoring point
- Coordinates
- Location
- Environment
- Protection of monitoring point
- Owner of monitoring point
- Performer of monitoring (MP)
- Year of composition
- Depth of MP
- Diameter of MP
- Depth of pump
- Filters
- Material
- Measurement equipment
- Type of measurements
- Way of measurements
- Frequency of measurements
- Typical pollutant
- Profile
- groundwater level
- K

c. The third phase of modeling is the establishment of the **logical data model**, which is based on provable facts. The model consists of logically most acceptable arrangement of tables that are based on data **normalization**.

We received from the project partners their databases structure and list of attributes, which should be included in new database system. They also represented data priority list and give some proposals for visual data presentation.

We noticed several differences between partnership databases. They were designed separately, so they have different structures, data types and media of storage. To transfer data in new system we carried out a thorough analysis, checked duplicated data, added some new fields and tables.

By joining and combining different databases we reached the final **158 monitoring points** which we included in the new data model and tested it.

We also organized several inter-institutional working meetings. We focused on three main subjects, which had to be analyzed:

- determination of common names of objects

We decided how to name and write them and eliminated the duplicates. We also determined their location on the base of available coordinates in comparison with legal network of monitoring points.

During the project we tried to clear out some problematic objects – their location, coordinates and names. We also added some new objects, mainly springs in the database.

We had to clarify the meaning of various usually ambiguous terms and ensure that problems with different interpretations of the terms and concepts cannot occur.

- determination of chemical parameters

We discussed which parameters should be included in the database, regarding some partners' wishes. We also defined the time period for average year values of parameters and how to write them.

Other actions, considering chemical parameters were:

- determination of limit values for each chemical parameter (following the law).

Every value will have its own code.
Here are the codes for values:

0	Only number
1	< or <LOQ=
2	[ ] or <LOD=

- Clarification of ambiguous name of chemical parameters.
- Proper presentation of chemical parameter values taking in account the limit values and transformation of the data in order to satisfy that requirements.
- Construction of several new tables to reach more automatic transfer of institutional data to the common database.

- It was also necessary to check the code list for parameters, we added some new codes and also codes for parameter border levels.
- We also did some transformation of nanograms in micrograms in the database, where needed.

d. establishment of new common data model

We defined new tables, their attributes and relations.

The result of logical modeling is the **relational model** with objects hierarchically arranged. It enables connections and analysis of interactions among the objects, queries and complex insight into the entire system. In each table there always exists an attribute that uniquely identifies the row (identical identifier), and is used as a primary key. By this key connections of various types between tables are established; one to one, one to many, and many to many.
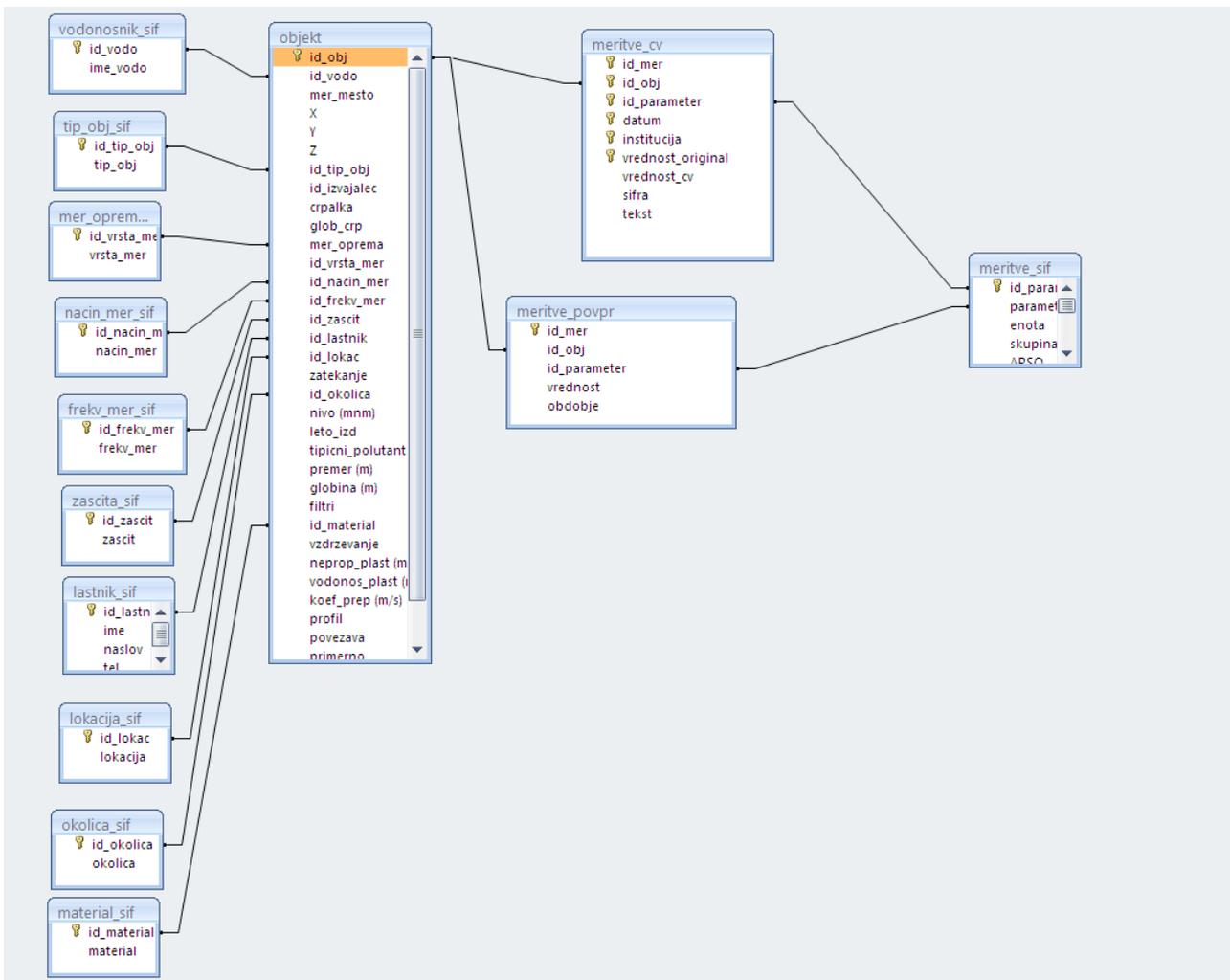
Figure 1: Relational model of the database.

e. Next phase of establishing the database consists of the physical data presentation. The **physical model** shows the actual storage of the data, amounts of the data, access to them and their organization. At this point the decision on the choice of the programming equipment is also taken. In our case, the program package MS Access was selected. The forms enable a simplified input of data and better overview. At the same time the forms serve as tool for checking and restricting of input and for surveying and correcting of records.

Figure 2: Form for elementary data.

3 INCOME GIS

The concept of the INCOME GIS requirements was developed on the base of interviews with the Slovenian partners. They emphasized that INCOME GIS portal should allow visualization of common database and maps.

The INCOME GIS contains a catalogue and a map viewer. The catalogue is used to search for datasets using their metadata and consult the available hydrogeological point data and maps in the viewer application.

3.1 MAP CATALOGUE

The purpose of the map catalogue is to let the user search for geographical layers offering hydrogeological data. The layers will offer hydrogeological point data, hydrogeological maps and reference layers (not necessary hydrogeological maps, but useful to help the user georeference hydrogeological information). The catalogue contains all layers offered by all partners. From a user point of view, the catalogue will start by offering a search page where the user can enter some search criteria for the maps he is looking for.

These criteria are:
- title,
- abstract,
- free text,
- keywords,
- category,
- geographical extent (of the features on the map).

The search criteria keywords and category are controlled lists offered in the Slovenian language. After doing a search, the user is presented a list of all available maps that match the filter criteria. For each item in the list, the information shown is:

- the name of the item (dataset, map,…),
- the abstract describing the item,
- a national flag indicating the origin of the metadata entry,
- if available, a thumbnail picture giving a visual impression of item in the list,
- a link to view the complete metadata for the item,
- a link to add the dataset to the viewer application.

The name and abstract describing the dataset are taken from the datasets metadata. Here, datasets from all partners are offered together and the logo of the institution will be a fast discriminator for the user to decide which dataset he is looking for. The user will always be able to consult the complete and detailed metadata for an item in the list. The metadata that is offered adheres to the INSPIRE metadata standard for discovery metadata. Finally, if the dataset is offered by a geodata service, it can be added as a layer in the map viewer that is part of the INCOME portal.
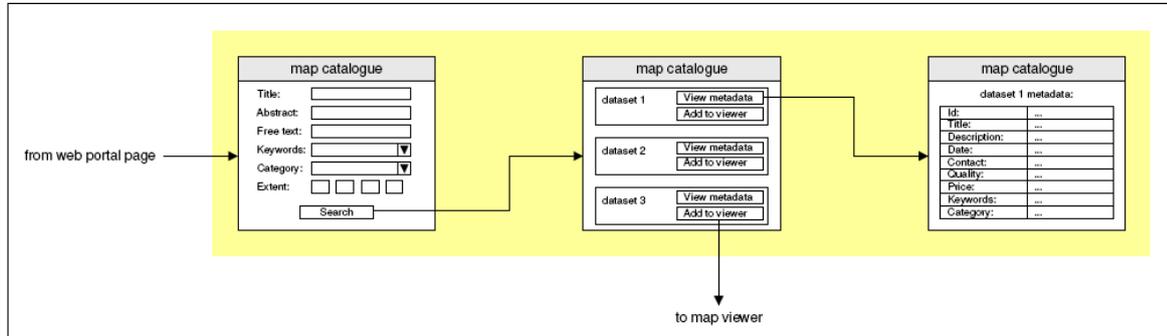
Figure 3: Map catalogue use case.

3.2 MAP VIEWER

The user enters the INCOME portal map viewer from the home page directly or he can select a map layer in the map catalogue.

The map viewer allows the user to:

- change the visible extent of the map (pan, zoom in, zoom out),

- manage the set of layers (visibility, z-order),

- consult the legend for the available map layers,

- perform an identify operation on the selected map layer (only one layer can be set "active").

After identifying a map feature, a dialog shows some attributes of the selected feature on the map.

As the map viewer is capable of retrieving information from any OGC compliant Web Map Service, the user can add layers offered by any WMS he has access to.
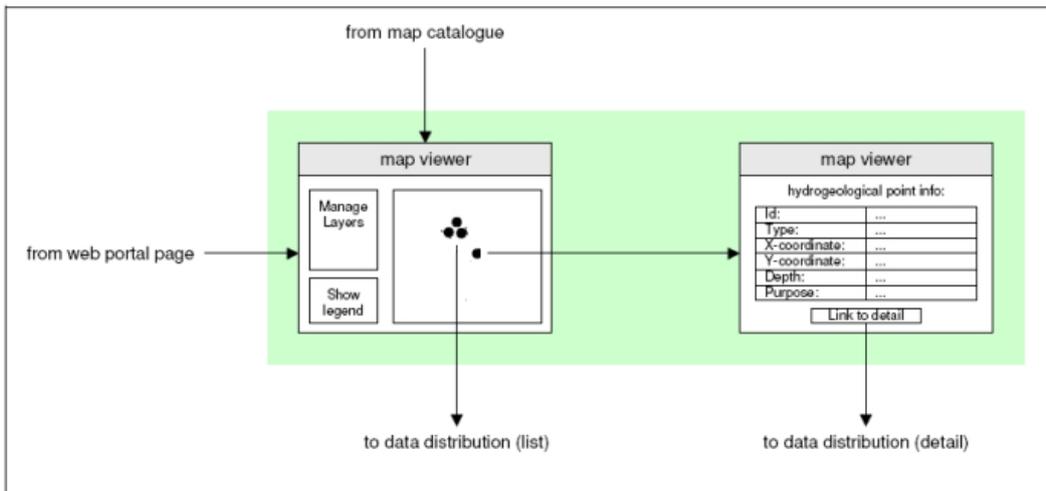
Figure 4: Map viewer use case.

In framework of INCOME GIS system we developed a HTML viewer called LIFE-INCOME. It allows users to browse through published geological maps and to do some simple query/identify functions.

The viewer was developed using ESRI's ArcGIS Server 9.3 software. With it you can:

- Deliver dynamic maps and data via the Web.
- Share data with others to accomplish tasks

Result is fast, user-friendly, stable and easily upgradable application.
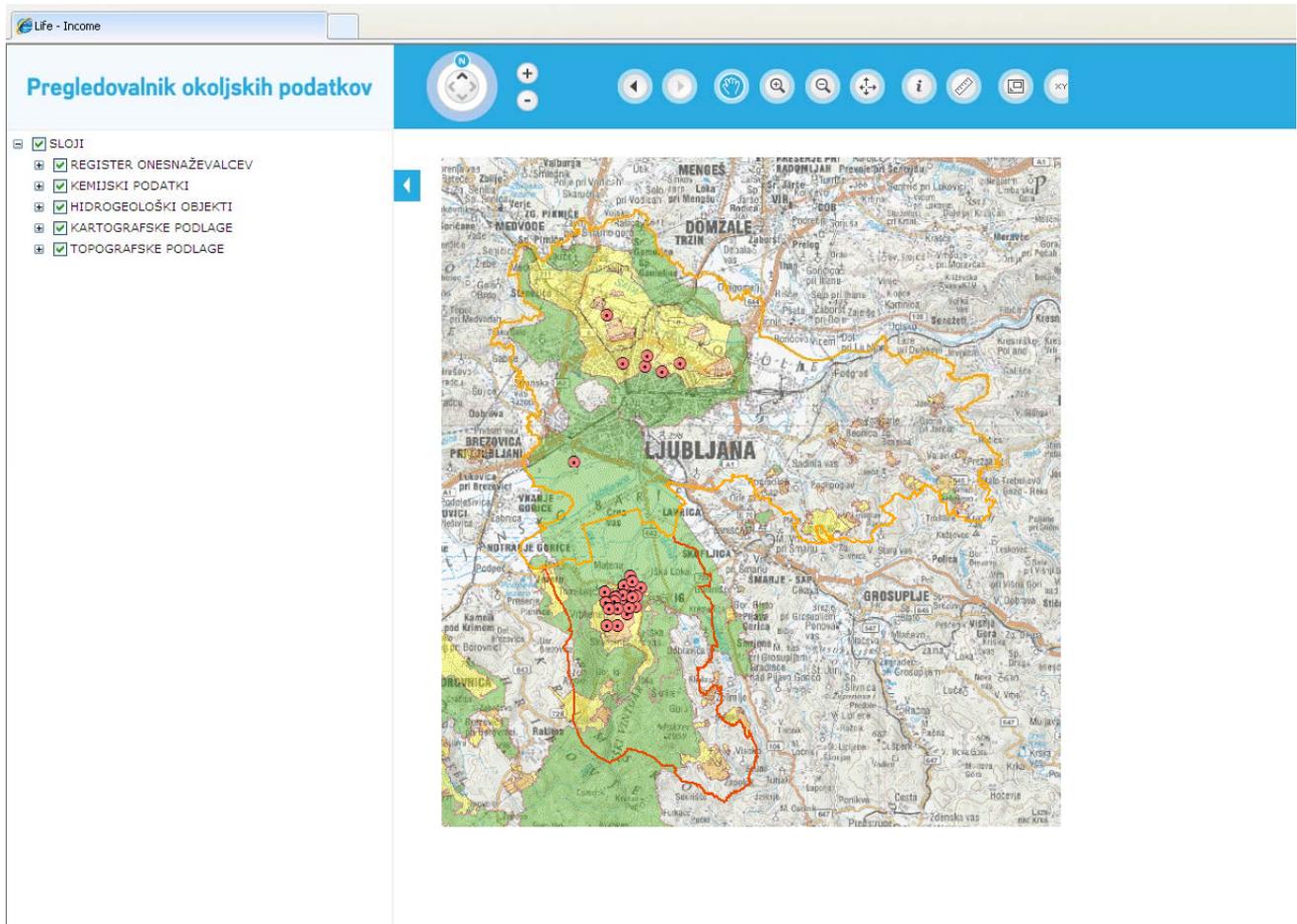
Figure 5: Map Viewer.

TOOLBAR

Toolbar offers basic functions for map manipulation and data query.

Available functions are:

1. Zoom In
2. Zoom Out
3. Full Extent
4. Pan
5. Identify
6. Query
7. Clear Selected
8. Print Map
9. Show Legend
10. Help
11. Report Error
12. Add to Favorites

For testing INCOME GIS system we used selected digital maps:

GROUP	OBJECT CLASS NAME	TITLE	TYPE	GEOMETRY
Popisani objekti	odlagalisca_odpadkov_2006.shp	Odlagališča odpadkov 2006	shapefile	point
	gramoznice.shp	gramoznice	shapefile	point
	ovire.shp	ovire	shapefile	point
	opozorilne_table.shp	opozorilne table	shapefile	point
	viri_tveganja_za_okolje_3_2_06.shp	Industrijska ogroženost 2006	shapefile	point
	vodnjaki.shp	vodnjaki	shapefile	point
Preučevano območje	Proucevano_obmocje.shp	območje terenskega popisa	shapefile	polygon
	meja_MOL.shp	meja MOL	shapefile	polygon
	Meja_IG.shp	meja občine IG	shapefile	polygon
Varstvena območja vodnih virov	Ljubljansko_polje_vvo.shp	Ljubljansko polje	shapefile	polygon
	Iski_vrsaj_vvp.shp	Iški vršaj	shapefile	polygon
	karta_lok_vod.shp	Lokalni vodni viri	shapefile	polygon
OGK100	krpoly; ljpoly; popoly; ripoly	OGK100	cover	polygon
Topografske podlage		Državna pregledna karta 1:250000	tif/tfw	
		Državna topografska karta 1:25000	tif/tfw	
		Digitalni ortofoto	tif/tfw	

On internet application user can also see parameters time series and their diagrams using Visual Studio development system the Web Mapping Application template, which offers basic map display, navigation, and interactive capabilities.

We also used the Web controls directly to create a Web site from scratch and developed our own specialized application in a style that conforms to existing Web site:

- added custom Query attribute task (which enables querying the boreholes by its names)
- added 2 custom search attribute tasks (which enables searching the boreholes by its names or Ids)
- added custom hyperlink tool (which enables seeing a time-line for selected borehole)
-

The INCOME web GIS application is stationed on URL http://akvamarin.geo-zs.si/incomepregledovalnik/Default.aspx.

For the workshop, which took place on 14.10.2010 in Ljubljana the internet application of registry of pollutants was established. It was implemented in already existing INCOME web based application with the same functionalities.

We added required parameters, new objects and new map viewer, especially for the needs of the workshop. Limited access to the application is also arranged with username and password.

On the pictures below the map content, layers and query possibilities are displayed.

Layers are:

1. Active sources of pollution:

   - Agriculture (greenhouses)
   - Mining
   - Industry
   - Energy
   - Dumps
   - Transportation
     …
2. Potential sources of pollution
3. Topographic layers

Figure 6: List of layers.

We continued with working on time-series of chemical parameters and their presentations on web application. We defined some data, necessary for adding pie diagrams and time series diagrams on internet INCOME portal and added them directly in the database.

Border values for chemical parameters are:

Pesticides: 0,1 µg/l
Nitrates: 50 mg/l
Cr 6+: 50 µ g/l
Trichloroethene, Tetrachloroethene : 10 µ g/l

-    Minimum and maximum values for parameters:

Pesticides: 0 - 0,2 µ g/l
Nitrates: 0 - 50 mg/l
Cr 6+: 0 - 50 µ g/l
Trichloroethene, Tetrachloroethene: 0 - 10 µ g/l

Where we have LOQ an LOD values, the average value has to be considered for the diagram.

1. PIE DIAGRAMS

Parameters and 100% value for the circle are:

Nitrates: 50 mg/l
Pesticides: 0,2 µ g/l
Trichloroethene, Tetrachloroethene: 6 µ g/l
Cr 6+: 50 µ g/l

The average value of last year of measurements was also taken into account.
Where we have LOQ an LOD values, the average value has to be considered for the diagram.

On the base of this data we constructed several new queries in the database, which will be constant and will automatically change with the data changes.

One query is the base for the elementary .shp file for the main map on the web applications, with locations of the hydro-geological objects and some elementary data.

Another query is the base for .shp file for representation of pie-diagrams on the web application, with all the necessary data represented above.

We also updated INCOME web application with new functionalities (go to location - by coordinates), pie diagrams for representing the chemical parameters - in results you can select to view graph and pdf files according to their Ids.

We also created new application for graphical presentation of measuring data. The data itself is connected to the SQL database. The user is selecting different data from input form and the graph is drawn on the fly. It can select different data like:

- measurement places (can select one or more to display),
- different parameters to show on graph,
- can chose timeframe by year,
- can select minimum and maximum on the graph to display,
- can export data from graph to Microsoft excel.

Another added features to web application INCOME were:

- New shape with detailed – specific measurement points
- Added new feature for representing graphs by the parameter (show only specific parameters for certain measurement points)
- On graph: now showing default limit for displayed parameters
- On graph: exporting data into Microsoft Excel
- On graph: selecting different periods
- On graph: selecting custom minimum and maximum

Since the application was slow it needed the optimization. We decided to use map caching. This is a very effective way to make web application run faster. When you create a map cache, the server draws the entire map at several different scales, and stores copies of the map images. The server can then distribute these images whenever someone asks for a map. It's much quicker for ArcGIS Server to hand out a cached image than to draw the map each time someone requests it. Another benefit of caching is that the amount of detail in the image doesn't noticeably affect how quickly the server can distribute the copy.

When making a cache we design the map and publish a service from it. Creating the cache take a long time. However, the benefits of caching make up for the time it takes to create the cache.

New feature for representing graphs by the parameter (show only specific parameters for certain measurement points) was incorporated to application.

The custom function for showing specific graphs was established. Wider range of showing multiple graph-lines on graph is now available.

Figure 7: Graphs on internet application.

Exporting data into Microsoft Excel for making further analysis is also possible (new icon and extended function added to the application).

Figure 8: Exporting data to Excel.

The pictures of borehole geological profiles were also added to the application.

Figure 9: Connection to geological profiles.