

**S&T Slovenia - Hermes Plus d.d.**

Oracle Business Unit, Ljubljana

# **LOGICAL DATA MODEL FOR DISSEMINATION OF SPATIAL DATA**

**by Joze Senegacnik**

**Chief Technology Officer**

e-mail: [joze.senegacnik@hermes-plus.si](mailto:joze.senegacnik@hermes-plus.si)

Ljubljana, August 2002

## Preface

This paper describes a logical model of the web based system for data dissemination that is used to support the dissemination process in the Slovenian government.

S&T with other partners is currently working on two ongoing projects concerning data dissemination. One project is for the Surveying and Mapping Authority of the Republic of Slovenia and is described in this paper; the other one is for the Statistical Office of Slovenia where we support the dissemination process of statistical data. Both projects have similar objectives and also we are encountering similar problems. The main difference is that in the dissemination of statistical data we don't work with personal data.

In this paper we will try to shortly describe the logical model required to support all necessary activities in electronic data dissemination. This logical model was prepared in a preceding theoretical project and is a good entry point in preparation of physical implementation of data dissemination process in the Surveying and Mapping Authority of the Republic of Slovenia.

## The Key Success Factors for Modern E-dissemination Systems

Web based e-systems have a different structure in comparison to classic information systems. The main difference is in the number of participants – users. This fact demands bigger automation and efficiency. This has a consequence that the demanded response time is shorter and the level of end-users satisfaction is higher.

The base participants in web-based systems are:

1. **E-users.** These users access data through network. We can split these users in different group:
  - a group of users which exactly know their demands and also have knowledge how to satisfy their demands.
  - a group of users which don't know their exact demands and behave like customers in a big shop.
2. **E-partners:** e-partners provide certain e-services like internet service providers. e-payment, analysis and logistic.
3. **E-Network:** The entry point of the system is based on WEB server, FTP server, WAP and different formats.

4. **Physical level** consists of network, supporting lookup tables, services for users authentication and data security, order entry systems, data preparation and data dissemination.

## **Logical Model of the Dissemination Process**

The logical model is based on the idea of using web services, which can be accessed through unified portal. This model tries to:

- offer data dissemination and e-government services to all interested users
- data accessibility of 24x7
- offer cheaper and more efficient services.

The classical paper orders method will be also supported for all those users that can't access the unified portal through internet.

The logical concept of the whole system is based on next levels:

- **level 1:** one entry point with all required and needed services
- **level 2:** there is one (or several) dissemination database in the background providing all necessary services
- **level 3:** there are system services available in the background to support security and system administration and system support. (data transformation, recording of all data orders, payment system)

For Internet users the dissemination system is seen as one entry point – internet address called Spatial Portal. This address contain links to different services of the portal (ordering, data browsing, data acquiring, general and special information, etc...). This spatial portal will also use internal governmental users, which will use different access points according to their requirements:

### **1. Data dissemination for external users**

In this case the spatial portal is behaving in same way as for the Internet users. The access point in this case is the spatial portal.

### **2. Other activities related to data dissemination**

These activities are indirectly related to the data dissemination process and can be partly manual procedures. The users also in this case use as the entry point the Spatial Portal to access the internal services.

### **3. Data dissemination for maintaining databases**

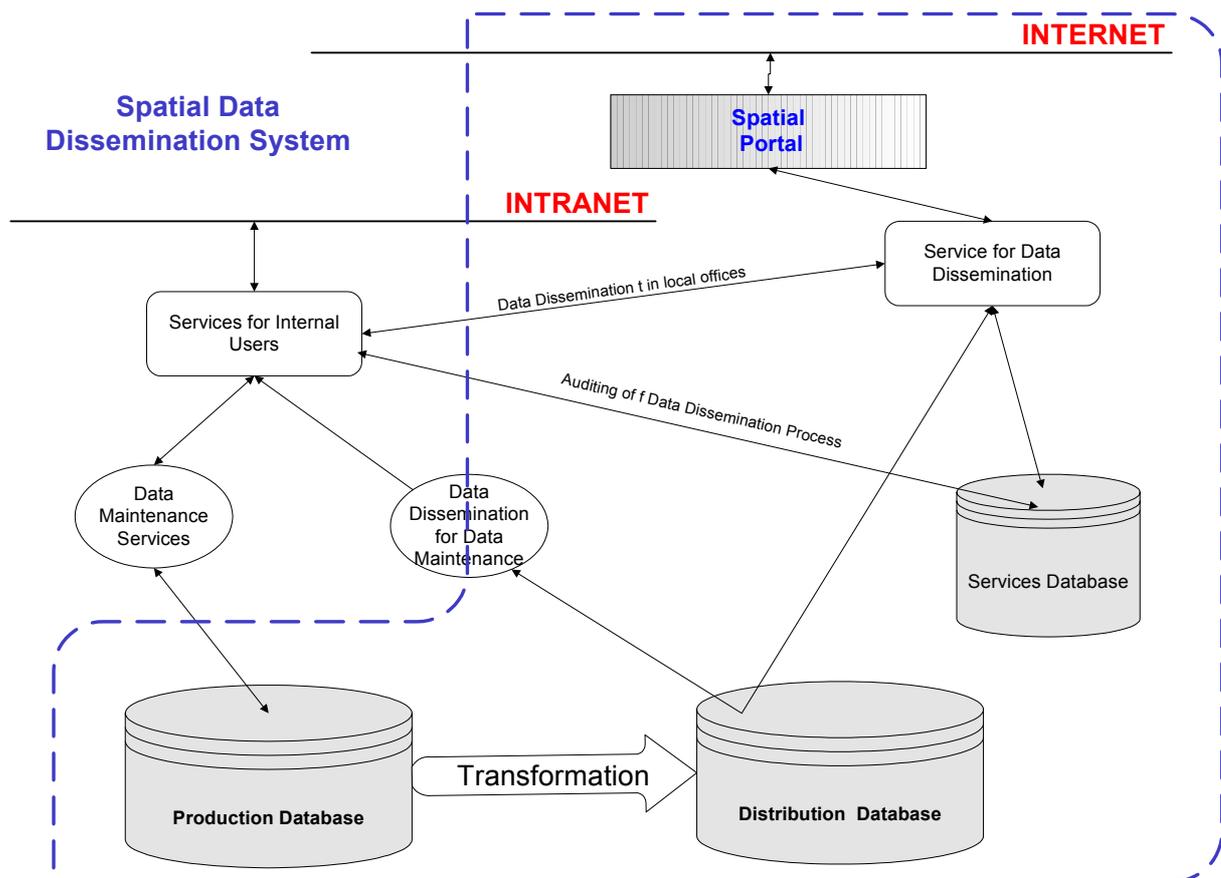
For regular maintenance of the own databases the issuer has to provide data to special organizations – contractors, which are in charge

of executing the maintenance procedures defined in the legislation. We can consider these organizations also as all other users although they have a special contract for data maintenance. They also access the data from the dissemination database rather than production database, which they maintain through predefined procedures. Only in some special cases the data dissemination process can use as a source the production database.

#### 4. Internal database maintenance

The internal governmental users access data in production database in the process of regular data maintenance. We must distinguish this case of data accessing from the data dissemination process because this is not a data dissemination process. The access to the data is provided through the application software.

Now let's take a closer look to the important parts of the spatial portal.



**Logical Spatial Data Dissemination Model**

## The key processes and their relations

From the above figure we can see the logical model of our system with the main data flows and processes. On the highest level the logical model is presented with the data flow diagram Spatial Portal that consists of:

- global system processes
- data warehouses
- data flows among data warehouses and processes

**Data warehouses** in the system are:

1. **Dissemination database** – is the main data source of the dissemination system
2. **Production databases** – these databases are not a part of the dissemination system but are the main source of the data flow for new or updated data that are through transformation process entered in the dissemination system.
3. **The metadata database** - in this database we keep the metadata about our data records.
4. **The overview layers database** – in this database we keep additional metadata attributes at the level of data maintenance, all planned data maintenance or data acquisition activities and also the constraints about the minimum amount of data that can be acquired.
5. **The Statistics or Services database** – keeps all necessary data for overview and managing the dissemination process.
6. **The system database** – contains the system setup data and all other data used for controlling the dissemination process.
7. **The users database** – contains the data about the registered users, their attributes and rights that influence the data dissemination process and also the price of the service.
8. **The price list database** – keeps the records for different services and data types.

**Processes** in the Spatial Portal are:

### 1.1 Registration

This process provides the authentication of portal users. All new users that want to access services with restricted access must first register and they are automatically redirected to the registration process.

### 1.2 Data dissemination processes

The main part of the system are next processes for data dissemination:

- data browsing services for browsing the spatial data, metadata, overview layers etc...

- data ordering services,
- special services

The users that are not registered can access some of these services also.

### **1.3 Services for the internal users**

The Spatial Portal also provides the services for activities that are not directly related with the dissemination process but are part of the internal procedures of the data providers.

### **1.4 Common e-governmental services**

The Spatial Portal is connected to the common e-governmental services in the process of:

- User authentication
- Payment process
- Electronic data interchange of ordered data with the digital signature

### **1.5 Billing and Payment system**

This is a general service used for preparing the necessary documents for the billing process.

### **1.6 Statistics**

The system keeps records about all necessary statistics about all accesses to the system and data requests.

### **1.7 Auditing system**

The purpose of the auditing system is to keep the track off all accesses to the data especially to the data with defined attribute of privacy. According to the Slovenian legislation the personal data has a high degree of privacy and thus special procedures of providing privacy of this data are required.

In the auditing process we should implement next rules:

- Due to the potential high load of the system all the auditing records should be kept in such manner that they don't degrade the performance of the system
- The audit records should be periodically archived and proper analysis of these records should be made in regular time intervals. The audit records are in place only for the data with certain degree of privacy.
- For each type of data the audit process should be carefully planned.

### **1.8 Data transformation**

Data transformation process is used for transformation of the production data into the dissemination data and properly updating all the overview layers and metadata database.

### **1.9 Help system**

The purpose of the help system is in providing context and technical help to the end users.

## Implementation of the data dissemination system

The whole dissemination system will be implemented in web technology. The database system is ORACLE RDBMS that is a standard in the governmental environment.

### Different Types of Database Organization

The organization of **production database** supports:

- The normalized structure of all data types that is easy to maintain (update); each attribute appears only once; the de-normalized structure is an exception in the case of serious performance problems.
- All data changes are recorded in the history of changes. The audit record contains also the information who made the change.
- All business rules how to maintain data are implemented at the database level (triggers, database procedures) or are provided by the application code.
- In the application data browsing is supported for data maintenance function and not for the data dissemination process.

The organization of **dissemination database** supports:

- The data model is optimized for querying data regardless the point in time
- The number of business rules implemented in dissemination process is smaller than in production database.
- The data organization can be also optimized for supporting certain standard formats.
- For all personal data with certain degree of privacy special rules are in order. The user rights to access data are checked differently than in production systems.

The above considerations lead us to a conclusion that dissemination database should be organized for dissemination process. This means that the physical structure of the dissemination database can be quite different from that in production environment.

The process of de-normalization of production data is one the key success factors for successful dissemination model. The advantage of the de-normalized data model is:

- The preparation of standard and non-standard queries is a simple process.

- The rules for organizing data warehouse are used in the process of the de-normalization.
- The process of optimizing the physical data structure reduces the overall load of the system and provides good response time thus minimizing the end user irritation.

### **Bibliography:**

1. The Model and Implementation Plan of the Dissemination of the Geodetic Data - University of Maribor - Faculty for the electronics, Igea d.o.o. and Aster d.o.o, Ljubljana, 1991

### **About the author:**

Joze Senegacnik has fourteen years experience in working in GIS and with Oracle RDBMS. At the beginning of his IT career he was an analyst/programmer/DBA for City of Ljubljana ( capital city of Republic of Slovenia) where, among other things, he was in charge for the implementation of the Information system for City of Ljubljana, which also included the geographical information system (GIS). From 1992 on he is an independent researcher in IT field and works as a permanent associate of Hermes Plus that is part of the S&T Corporation. For some time he was working in Oracle Technical Support and later on in development of GIS application systems for Surveying and Mapping Authority of Republic of Slovenia, Agricultural information system for the Ministry of Agriculture of Republic of Slovenia and also in the development of applications for other governmental institutions all based on Oracle database. In 1994 he decided to put all graphical data in Oracle database what was a lot before Oracle started to develop Oracle Spatial Database Option (Oracle Spatial).

He is regular speaker on all Slovenian Oracle Users Group Conferences (SIOUG) and also IOUG-A and European Oracle conferences. His contact address is: [joze.senegacnik@hermes-plus.si](mailto:joze.senegacnik@hermes-plus.si).