Information system for maintaining a database of geobotanical descriptions while studying a landscape

K Shchukova[1]

Russia, Tomsk, 634050, National Research Tomsk Polytechnic University

E-mail: embrre@yandex.ru

Abstract. Nowadays, there is a need in developing a common information space for providing a centralized data storage of geobotanical descriptions while studying a landscape and custom software for collecting, storing and processing semi-structured data. In the paper, the results of information system development for maintaining a geobotanical database are discussed. The database was developed via the bottom-up design approach and ER-model (Entity-Relationship). The analysis results of geobotanical data obtained from field studies and the database structure of geobotanical descriptions while studying a landscape are considered. At the moment, the information system for maintaining the geobotanical database is applied by the Institute of Steppe, Ural Branch of Russian Academy of Sciences.

1. Introduction

A xylium and phytocoenosis inventory process is one of the most important scientific tasks because the results of studies provide to analyze biological diversity and perform ecological vegetation monitoring of the areas being explored.

The growing popularity of computer-based activities has encouraged not IT-specialists in Vegetation Science to use a wide range of various software tools for solving their scientific fundamental and applied problems. In the field studies of sample areas, a large amount of geobotanical data is constantly being accumulated. In order to provide a central data storage, it is necessary to ensure a common information space. The custom software has to have a user-friendly interface and sufficient functional capabilities for data processing and agile reporting.

Our analysis of similar information systems focused on geobotanical and ecological data processing has shown that the best-known products are TurboVEG, Juice and IBIS. The systems are intended for maintaining a phytocoenotic database. They also allow to significantly automate analysis and classification processes of different plant communities.

However, the software products are not fully satisfied by the user requirements. Unfortunately, there is no opportunity to modify these systems because they are proprietary or based on obsolete technologies.

Thus, the main aim of the study is to create a geobotanical database (DB) while studying a landscape and an information system (IS) for its using intended for collecting, storing and processing tremendous volumes of geobotanical data. The functional capabilities of the IS were discussed in cooperation with the Institute of Steppe, Ural Branch of Russian Academy of Sciences (UB RAS).

---

[1] To whom any correspondence should be addressed.

2. Geobotanical data analysis and database development

Nowadays, most specialists in Geobotany apply a rich variety of theoretic and practical methods to perform geobotanical descriptions studies.

In the field investigations while studying a landscape, they fill out special forms of geobotanical descriptions, create documents with the results of the chemical soil analysis in Microsoft Word and take pictures of a landscape being studied.

The use of common forms provides to standardize and unify a traditional procedure of geobotanical description handling. The geobotanical data obtained from the field studies are primarily used for analyzing and comparing phytocoenotic data.

The information about ecotope features and geographical coordinates of sample areas is gained while studying a landscape. In order to determine exact geographical coordinates and an altitude of areas being studied, GPS devices are primarily used.

While studying a phytocoenosis, all vegetation species are classified according to the stratification of the plants. A vegetation layer is a structural part of a vertical phytocoenosis building.

The woodland structure includes the following layers: Canopy/Tree Layer (A), Understorey/Shrub Layer (B), Field Layer (C) and Ground Layer (D).

In order to describe various vegetation species, the following quantitative characteristics are used depending on a particular layer: abundance evaluation according to the scale of O. Drude, a plant cover, landscape aspects and a common plant cover.

The characteristics obtained from field studies are recorded in special forms of geobotanical descriptions for each layer and sublayer. A name of vegetation species is identified according to dominate plant species of each layer.

On the basis of described forest and herbaceous vegetation characteristics, all the plants are classified into the following categories: a life form, an ecological group, a coenotic group, latitudinal and longitudinal groups, a particular value. In addition to the vegetation classification, a statistical analysis of species, communities and families is carried out for various geobotanical descriptions.

In the process of database design, the following information sources provided by the Institute of Steppe, UB RAS, were used: geobotanical descriptive forms created in MS Word, documents of chemical soil analysis created in MS Word and Excel, summary tables of phytocoenotic data generated in MS Excel, landscape images of studied areas, the results of biomorphological and phytocoenotic vegetation analysis described in documents of MS Word and Excel, GPS tracks including a point name, geographical coordinates, an altitude and etc.

In general, these forms contain the following characteristics of forest and herbaceous communities: a number of a geobotanical description, a value of a sample area, descriptions of vegetation layers, names of plants, geographical coordinates of a studied area, a description of relief features, a description of soil properties, the results of chemical soil analysis, human and animal impacts, a total plant cover; vegetation species, plant heights, abundance evaluation according to the scale of O. Drude, phonological stages, aspects, location features and etc.

Then, a taxonomic, biomorphological and phytocoenotic analysis is performed by using the obtained geobotanical data [1].

At the stage of conceptual database scheme design, semantics, typical structures of phytocoenotic descriptions, and relationships between geobotanical objects were identified. In the process of geobotanical data analysis, information inaccuracy and incompleteness have been founded out in all kinds of documents provided by the Institute of Steppe, UB RAS. Semi-structured data representation has led to the conflict with data unification and formalization requirements of IT.

The infological analysis has revealed that the most common type of relationships between objects of geobotanical descriptions is the many-to-many relationship. For instance, various vegetation species may live in the same area, and vice versa. A single abundance evaluation according to the scale of O. Drude may be characterized by several descriptions of woody species, and vice versa.

Geobotanical objects are also related by the one-to-many relationship. A plan community may include a lot of plant families, but each family consists of various species. Vegetation layers may involve a large number of sublayers that may be related to a lot of vegetation descriptions.

While designing the physical data model, the database of geobotanical descriptions has been successfully developed. This database structure allows to expand a scope of work without any information system modification. A database framework was developed via the bottom-up design approach and entity-relationship (ER) diagram. The developed database consists of 22 tables in total including 10 lookup tables.

In order to distinguish different types of tables, the terms 'lookup table' and 'domain object' were introduced. A lookup table is a table storing reference data. These tables are used for storing data frequently used by users. For example, lookup tables of vegetation layers and phytocenosis types allow to eliminate a manual data input. Due to the use of lookup tables, the time for data search can be significantly reduced.

The developed database includes the following lookup tables: phytocenosis types, vegetation layers and sublayers, plant families, communities and etc.

A 'domain object' is a table containing useful data of the field being studied with a more sophisticated framework in contrast with a lookup table.

The database also involves the following domain objects: a geobotanical description for a specific sample area, a description of a vegetation layer and sublayer, and etc.

In order to store chemical soil analysis files created in MS Word and Excel and obtained images, special tables in the database were generated.

In all database tables, surrogate keys were introduced to ensure uniformity across the database framework. The alternative keys have been generated for table attributes which are candidate keys [2].

3. Development features and functional capabilities of the information system

In order to describe a functional mechanism of IS in considerable details, the IDEF0 Functional Modeling method was applied for effective visualizing of IS functional capabilities [3].

The top level of the IDEF0 model is shown in Figures 1.
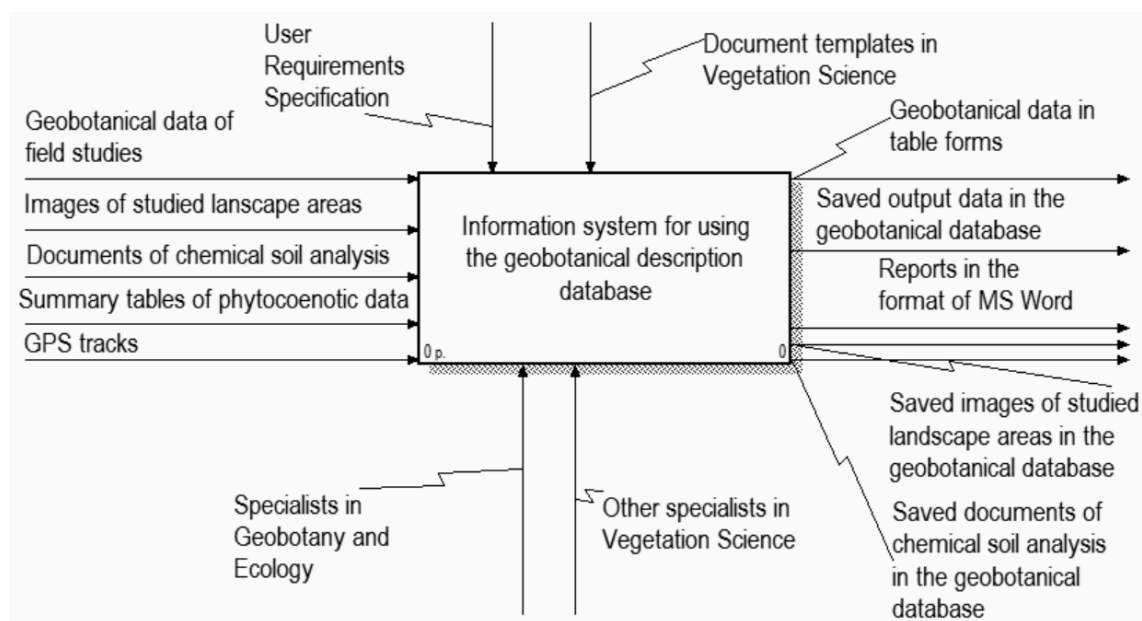
Figure 1. The top level of the IDEF0 model.

The system consists of four subsystems (Figure 2):
1.   Data handling subsystem. It is intended for transforming input data into a database format.

2.   Data storing subsystem. This subsystem provides a centralized data storage in the database of geobotanical descriptions.

3.   Reporting subsystem. The subsystem allows to generate a report in a necessary format.

4.   Data visualization subsystem. It is used for visualizing various geobotanical data, for example, in table forms.
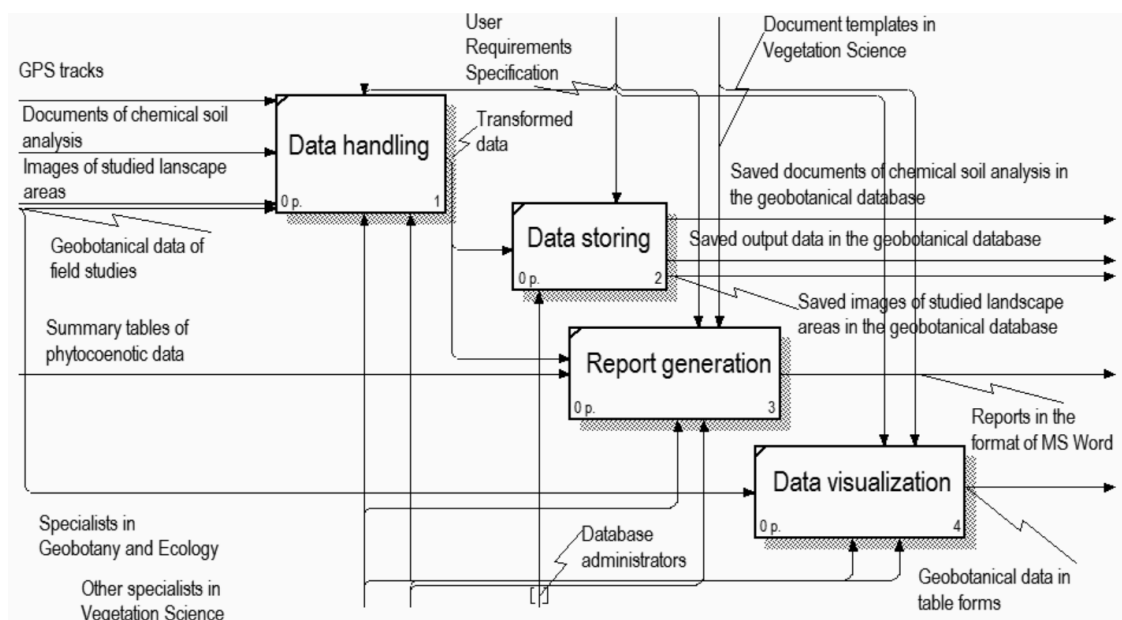
Figure 2. The decomposition of the IDEF0 model.

The information system for geobotanical database maintenance has been implemented in Microsoft Visual Studio 2013 by the C # language. MS SQL Server 2008 R2 Express was used as a database management system (DBMS). The use of ADO.NET helps to accelerate the process of data access functionality encapsulation and eliminate a lot of mistakes caused by a human factor.

In order to connect with MS Word, the Microsoft.Office.Interop.Word and Microsoft.CSharp libraries were added in a project. The interface of the information system includes various drop-down lists and lookup tables providing a quick data access. All input user's data is controlled by special software tools developed for checking input data and spelling. The IS has a user-friendly interface and offers efficient tools for processing, manipulating and extracting geobotanical data. The system operation does not require specific knowledge and skills in IT. Hence, it may be applied by any specialists in Geobotany and Vegetation Science. The IS is intended for intensifying a routine work with tremendous volumes of the customer data.

4. Conclusion

Due to the use of enabling technologies, the information system for geobotanical database maintenance with sufficient functional capabilities has been successfully developed. At the moment, this system is used by the Institute of Steppe, UB RAS. In the future, the DB framework and IS functional capabilities are being planned to expand in order to carry out a biomorphological and phytocoenotic vegetation analysis and set up a connection with GPS devices and Google Earth.

References

[1]   Korchagin A A 2012 *The Field Geobotany. Vegetation Building* vol 5 (Moscow: Book On Demand) pp 200-320

[2]   Date C J 2006 *An introduction to Database Systems* (Book On Demand) p 1328

[3]   Jesse R and Cohn R 2015 *IDEF0* (Book On Demand) p 105