

SHORT COMMUNICATION

DOI: 10.1515/ffp-2016-0020

Open access to research data in the data warehouse of the Forest Research Institute

Dorota Grygoruk

Forest Research Institute, Department of Forest Ecology, Sękocin Stary, Braci Leśnej 3, 05-090 Raszyn, Poland, phone: +48 22 7150414, e-mail: farfald@ibles.waw.pl

The concept of open science has a long history. Its origin dates back to the 17th century and is connected with the release of the first scientific journal in Europe: *Journal des Savants* and *Philosophical Transactions of the Royal Society*. Today worldwide, there are more than fifty four thousand scientific journals, of which more than a million scientific articles are published annually. The contemporary concept of open science includes not only an open access to the research results, that is an open access to scientific articles, but also more and more open models for other areas of scientific work like access to research data or conducting research with ‘open notebook’. The dissemination of open science concept is a global trend, which is connected with the development of information and communication technologies.

In Poland, the policy guidelines on open access to publications and research results are consistent with the Commission Recommendations of 17 July 2012 on the access to and the preservation of scientific information (2012/417/EU). Among the scientific institutions conducting studies on forest ecosystems in Poland, the Forest Research Institute was the first to use advanced information technology for open access to scientific data. The project ‘Forest Information Centre – information platform of environmental monitoring in Poland’ (The Innovative Economy Programme 2010–2014) was developed as a model of the scientific data warehouse. The data warehouse designed based on the concept of *Business Intelligence* uses the following tools: SQL Server Integration Services, SQL Server Analysis Services, SQL Server Reporting Services and SQL Server Report Builder. The adopted

technological solutions enable integration and standardization of data from distributed sources, loading multidimensional data structures (OLAP cubes), and perform multidimensional data analysis (Fig. 1). The process of integrating the data warehouse may include archival data of the Institute, data from operational bases in current projects, as well as data from external sources like the Information System of State Forests (SILP). The structures of the multidimensional warehouse (OLAP cubes) are designed regardless of the source databases. There are two basic categories of data in the warehouse, namely facts and dimensions. Facts are the analysed data. For example, in case of the Institute, facts relate to the growth of trees, the occur-

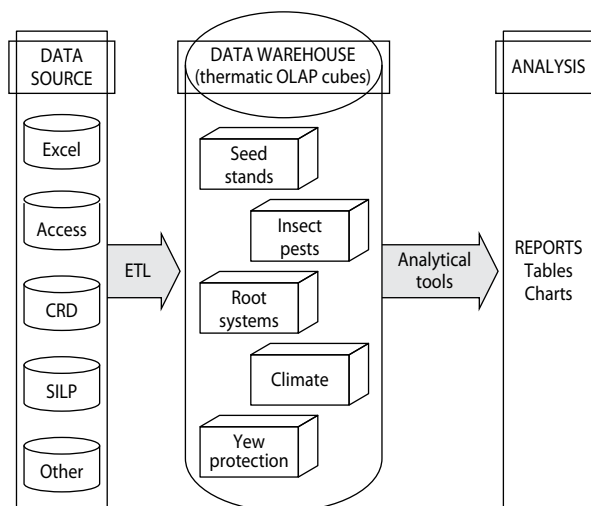


Figure 1. Scheme of data warehouse architecture of the Forest Research Institute

rence of insect pests, the development of root system of forest trees, etc. Facts are described by measures such as the height of trees, the number of insects per 1 ha of stands, root biomass, etc. Dimensions specify the condition of analysis; in case of the Institute, dimensions relate to the time (year, month) and location (geographical coordinates or forest address consisting of the codes, such as Regional Directorate of State Forests, forest districts, precincts forest, forestry, division, sub-division, etc.). Scientific data selected for integration is subjected to a process ETL, which is responsible for the selection and downloading of data from the source databases, data transformation, cleaning, conversion, sorting, standardization, and then loading of data to the structures warehouse. The integrated scientific data through the use of modern computer tools create new sources of valuable scientific information, such as, on the environmental changes in an era of global climate change. The results of the analysis can

be presented in the report as graphs or tables or they can also be exported to CSV files, PDF, Excel, Word or XML. The results of the analysis are stored in thematic catalogues (e.g., seed stands, root systems, insect pests, climate, yew protection, etc.) and they are available in the Repository Reports of Data Warehouse (Fig. 2). In the last decade, a number of scientific institutions in Poland have undergone modernization, and have expanded their information systems based on modern technologies. Modern database systems guarantee data security, fast search of information and provide authorized access to data. Data warehouse enables the storage of historical data and operational data at one place and at the same time, allows various kinds of analyses like identifying trends, directions of change, development forecasts, etc. The possibility of data mining changes the natural environment in Poland and gives significant support for the development of forest research and scientific cooperation.



Figure 2. Distribution of thematic data warehouse of the Forest Research Institute

REFERENCES

- Andrzejewski W., Królikowski Z., Morzy T. 2010. Databases and information systems and their impact on the development of science in Poland. In: The Polish and world achievements of science: technical sciences. Fundacja im. Wojciecha Świątosławskiego na Rzecz Wspierania Nauki i Rozwoju Potencjału Naukowego w Polsce, Gliwice, 345–388.
- Dudycz H. 2003. Data warehouse is veritable component of Business Intelligence. Wydawnictwo Akademii Ekonomicznej w Katowicach.
- Rączka K., Kowalski M., Gąsiorek S. 2007. Business intelligence systems. *Inżynieria Rolnicza*, 6 (94), 205–212.
- Todman Ch. 2011. Design a data warehouse. Support customer relationship management. Helion, Gliwice.
- Urbańska M., Gierszal H., Nowacki M. 2010. Decision support system for domesticated plants protection. *Inżynieria Rolnicza*, 7 (125), 223–228.
- Wyskwarski M. 2012. Data analysis in ERP system – use of business intelligence concept. *Zeszyty Naukowe Politechniki Śląskiej*, 61, 311–318.
- Zaskórki P. 2012. Integration IT services in the business organization. *Zeszyty Naukowe WWSI*, 7, 203–227. <http://web.pages.wwsi.pwr.edu.pl>