Chapter IV

Using Environmental Information Efficiently: Sharing Data and Knowledge from Heterogeneous Sources

Ubbo Visser, Heiner Stuckenschmidt, Holger Wache and Thomas Vögele
University of Bremen, Germany

INTRODUCTION

Environmental information systems have gained more importance both in the public administration and industry since the beginning of 1990. For example, in public administration, every state in the Federal Republic of Germany has developed a type of environmental information system. National and European legislation demanding far reaching transparency in the state of the environment encouraged this development. In industry on the other hand, environmental information systems are used for cost- and product-specific recording of waste flows. These are used to point out weak points within the companies’ processes.

Both application areas share the need to store and process large amounts of diverse data, which is often geographically distributed. Most environmental information systems use specific data models and databases for this purpose. This implies that making new data available to the system requires, that the data be transferred, into the system’s specific data format. This is a process, which is very time consuming and tedious. Data acquisition, automatically or semiautomatically, often makes large-scale investment in technical infrastructure and/or manpower inevitable. These obstacles are some of the reasons behind the concept of information sharing. The solution of information sharing applies because existing information can be accessed by remote systems in order to supplement their own databases. The advantages of successful information sharing is thus obvious for many reasons:

• Quality improvement of data due to the availability of large and complete data.
• Improvement of existing analyses and application of the new analyses.

Cost reduction resulting from multiple use of the existing information sources.

Avoidance of redundant data and conflicts that can arise from redundancy.

However, in order to establish efficient information sharing, difficulties arising from organizational and competence questions and many other technical problems have to be solved. Firstly, a suitable information source must be located which contains the data needed for a given task. Once the information source has been found, access to the data therein has to be provided. Furthermore, access has to be provided on a technical and informational level. In short, information sharing not only needs to provide full accessibility to the data, it also requires that the accessed data be interpreted by the remote system. While the problem of providing access to information has been largely solved by the invention of large-scale computer networks, the problem of processing and interpreting retrieved information remains an important research topic. This paper will address three of the problems mentioned above:

- finding suitable information sources,
- enabling a remote system to process accessed data,
- solutions for helping the remote system to interpret accessed data intended by its source.

In addressing these questions we will explore: meta information systems, data exchange standards, converter and mediator systems and techniques in the area of intelligent information integration, always bearing in mind the special needs of environmental information.

META INFORMATION SYSTEMS

Meta information systems are tools to organize and document the large and ever growing resources of environmental and geo-spatial data compiled by government agencies, academia, nongovernmental organizations, and industry. Their main purpose is to provide an overview of the available information sources and to support manual and semiautomatic searches for environmental information.

Good examples for national and international meta information systems are the German/Austrian Umweltdatenkatalog (UDK), and the European Catalogue of Data Sources (CDS). The UDK was created in the nineties to improve access to environmental and geospatial data sources held mainly by government agencies in Germany and Austria (Günther, 1998). (Legat et al., 1999) describes their five years experiences with the Austrian UDK (the most advanced UDK) and concludes, that with the help of the system, the process of finding and administering data has drastically improved.

The CDS (ETC/CDS, 2000) is a locator system for environmental data in Europe. It collects and maintains environmental meta information of all 18 countries participating in the European Environment Agency (EEA) and, in the future, the 13 PHARE countries (PHARE is an European Commission Program, the main channel for the European Union’s financial and technical cooperation with the countries of central and eastern Europe). The CDS is maintained by the Topic Centre on Catalogue of Data Sources (ETC/CDS), based in Hanover, Germany.