Wolf-Fritz Riekert
FAW Ulm (Research Institute for Applied Knowledge Processing)
POB 2060, 89010 Ulm, Germany; E-Mail: riekert@faw.uni-ulm.de

Management of Data and Services for Environmental Applications

Abstract: In recent years, systems for processing environmental information have been evolving from research and development systems to practical applications. Many of these systems already support environmental activities of the public sector. Most of these systems, however, were developed as island solutions. The integration of such heterogeneous systems within one distributed environmental information system, is a big challenge of today. It requires new methods to navigate to the services and data offered by the components of such a system. Metadata (such as data and service dictionaries) and codata (such as spatiotemporal references) must be provided in order support the proper usage of these data and services. Systems integration techniques are required in order to overcome the heterogeneity. An approach to fulfill these requirements is shown by examples from recent work at FAW which has been conducted in the context of the Environmental Information System Baden-Württemberg.

1. Environmental Data

Environmental data differ from traditional data in many respects. Often the data are related to a certain location and time interval. Measurement data that are acquired for discrete locations and points in time belong to a special category; these data need to be interpolated if continuous profiles are required (e.g., air pollution data or elevation models). A special case is image data (e.g., satellite sensor data), which typically consist of very large raster data sets. NASA experts often handle terabytes of satellite data; this is more than two orders of magnitude greater than the amount of data managed today in large international financial transaction systems. This requires new storage models for environmental databases, e.g., the use of automatic tape archives and CD ROM jukeboxes as tertiary storage medium.

The most demanding problem concerns the problem of deriving information of interest from existing data. In the ideal case, an environmental database contains all necessary data which are needed to derive the information requested by the user. Apart from the raw environmental data, however, this requires the availability of additional data. These additional data are often refered to as *metainformation* or *metadata* (6). *Codata* is another term which is also used in this context. In this paper, we try to differentiate between metadata and codata. Metadata provide abstract descriptions of the data structures and data formats used in the underlying system, whereas codata include additional instance-specific data about location, time, precision, and revision dates of the data under consideration. Location and time are very important kinds of metadata which are discussed in more detail in the next section.

Metadata and codata are missing today in most existing environmental information processing systems. Often the only data source consists of a magnetic tape containing the raw data (e.g., image data or measurement data) to be analyzed. All of the additional information which is necessary for the correct interpretation of the raw data is implicitly coded in the application programs or must be contributed by the user. Such a tape may become completely useless after a change of the personnel or of the data processing software. Therefore an effective management of metadata and codata is of crucial importance for environmental information processing systems.

2. Spatio-temporal Aspects

Environmental data often describe environmental objects with a spatial and temporal extent. These objects, also known as geographic objects, possess a geometry consisting of point-form (0D), linear (1D), flat (2D), or solid (3D) features. Often geographic objects also have a lifetime which is given by their dates of construction and destruction. During the lifetime of an environmental object, its attributes may change. That is, the values of the attributes are only valid during a certain time interval, and a given attribute (e.g., the population of a city or the land use of a parcel) may possess many possible values during the lifetime of the object. A special case is the change of the geometry of a geographic object such as the growth of a city or the shrinking of a lake or a forest.

Spatio-temporal data are important examples of codata in the environmental domain. The management of spatio-temporal data is a special challenge to applied computer scientists. Abstract data types are required for representing concepts such as the partonomy, the topology, and the spatio-temporal extension of geographic objects as well as the thematic information associated with these objects, such as alpha-numeric attributes and relationships to other geographic objects. An important task is also the (carto-)graphic presentation of these objects. In most cases, non-standard data types are required for representing these kinds of information and, therefore, a trend towards systems which allow the definition of such data types, such as extended relational or object-oriented databases (1), can be recognized in current developments (3).

The special nature of environmental data also requires a query language with special characteristics. Apart from typical SQL-like questions, the forthcoming object-oriented database management systems (2) also allow navigational queries and queries that include user-defined predicates. In particular the latter allows the usage of spatio-temporal predicates in queries. The optimization of spatio-temporal query-processing is a scientifically demanding problem which concerns both storage models and indexing techniques for multi-dimensional data. The integration of these techniques with existing databases is a hard problem which will still require major research activities in the future.

3. Overcoming the Heterogeneity

The harmonization of environmental information at national, European and worldwide levels is of central importance for gaining a reliable description of the environmental situation and, at the same time, is a basic requirement for any reporting system in this context. These requirements, however, are confronted with the existing heterogeneity of hardware and software environments, of database systems, of method bases, of network technology, and of various computer languages.

The task of overcoming heterogeneity also requires the availability of meta-information. Doubtless the development and promotion of standards is of particular importance in this respect. Experience shows, however, that we will still have to cope with competing standards in the future. In addition, technological advances will always produce new heterogeneity problems and will require strategies for migrating the software towards new solutions. Metainformation about data and methods can be used in order to do the necessary translations between different systems. In the public sector, environmental metainformation systems, also known - with a reduced scope - as *environmental data directories* (8), are currently being developed to overcome the prevailing lack of metainformation in existing environmental information systems.

4. Environmental Information Management in the Public Sector

In recent years, environmental protection as an objective of public activities has reached quite a high standard in Europe, particularly in Germany. Communal tasks, for example, include land use planning, approving compliance with environmental standards, management of hazardous waste, and water and energy management for public facilities. Several states and regions of the European Union have already installed effective environmental information systems and powerful sensor networks.

Official systems such as the Environmental Information System (German acronym: UIS), Baden-Württemberg (5) support environmental tasks at various levels: Decision Support Systems are provided for the high-level environmental management, reporting and planning systems are available for middle management, while basic components support the acquisition and management of specific environmental data at the operational level. In addition, interdisciplinary information systems are being used in public environmental management; these systems are not restricted to environmental tasks such as the topographic or cadastral information systems of the surveying offices.

UIS is the organizational, informational, and task-oriented framework for the supply of environmental data and for the processing of both department-specific and interdisciplinary tasks in the environmental domain of the State administration. UIS consists of a large number of components that are implemented on various hardware and software platforms and are operated by various departments at distributed locations. In this context, the INTEGRAL project, which is presented in the following section, aims at a user-friendly and economical way of accessing the functionalities of these distributed system components.

5. INTEGRAL

The INTEGRAL (Integration of Heterogeneous Components of the Environmental Information System of Baden-Württemberg) project has been conducted since 1993 at FAW under commission of the Environment Ministry, Baden Württemberg. The central goal of this project is to continuously increase the integration of UIS components and to improve the ease of use of these components from remote sides (7). The networking concepts available in open systems are a promising option for the necessary integration, because these systems provide highly effective mechanisms for sharing data and functionality within a computer network. Since all commercial operating systems support these communication standards to a large extent, this approach does not severely restrict the applicability of the INTEGRAL concept to dedicated computer systems. In INTEGRAL, the FAW is working on communication interfaces oriented that are oriented towards a client/server model which is typical for modern system architectures. This kind of model is also the target of FAW´s software engineering strategy in general.

At the highest level of abstraction, an interactive interface based on X Window allows the functionality offered by different systems to be shared at the screen level. At a lower level of abstraction, a service interface developed by the Institute for Nuclear Energetics and Energy

Systems (IKE), University of Stuttgart, is provided. This interface makes the functionalities from existing programs available in the form of self-contained services.

The access to these functionalities is based on a hypertext system. The management of these functionalities is supported by meta-information. In this way, preferences of the users may be taken into account. For example, urgent messages may be transported as rapidly as possible, whereas large simulation jobs may be computed at inexpensive rates during the night. Beyond integrating computer systems, INTEGRAL also connects people by empowering them to work cooperatively and to communicate efficiently by means of electronic media.

6. Conclusion

The management of data and services for environmental applications using codata and metadata is of crucial importance for environmental information systems. The INTEGRAL approach shows that the availability of data and services in environmental information systems will be enhanced and standardized by the techniques proposed. The advantage arises from the simplified, uniform, and more user-friendly availability of services offered via the network. Such services can be shared from client workstations all over the network. The “server-side” location of the programs and data that provide the services and information reduces hardware, software, and maintenance requirements.

References

(1) Special Section: Next Generation Database Systems. *Communications of the ACM*. 34(10):30-120, October 1991.

(2) Cattel, R.G.G.: *The Object Database Standard: ODMG-93*. Morgan Kaufmann, 1994.

(3) Günther, O.; Riekert, W.-F.: The Design of GODOT: An Object-Oriented Geographic Information System. *IEEE Data Engineering Bulletin* 16(3), September, 1993.

(4) Jaeschke, A.; Kämpke, T.; Page, B.; Radermacher, F.J.: *Informatik für den Umweltschutz*. Springer-Verlag, Berlin – Heidelberg – New York, 1993.

(5) Mayer-Föll, R.: Das Umweltinformationssystem Baden-Württemberg; Zielsetzung und Stand der Realisierung. In: (4), pp. 313-337.

(6) Radermacher, F.J.: The Importance of Metaknowledge for Environmental Information Systems. In: Günther, O.; Schek, H.-J. (eds.): *Large Spatial Databases*. Proceedings. Lecture Notes in Computer Science 525, pp. 35-44. Springer, Berlin – Heidelberg – New York, 1991.

(7) Riekert, W.-F.; Henning, I.; Schmidt, F.: Integration von heterogenen Komponenten des Umweltinformationssystems (UIS) Baden-Württemberg. In: *2. Workshop „Integration von Umweltdaten”*, KfK 5314, Kernforschungszentrum Karlsruhe, 1994.

(8) Schütz, T.; Lessing, H.: Metainformationen von Umwelt-Datenobjekten – Zum Datenmodell des Umwelt-Datenkataloges Niedersachsen. In: (4), pp. 19-28.