COLOGNE'S CUTTING-

DIETMAR HERMSDOERFER and ALFRED WALGENBACH explain how a spatially referenced information system has reformed one German city's management practices—and taken local government into the business of data warehousing.

The city of Cologne's department for statistics, residency and European affairs recognised the need for a strategic information system (SIS) some years ago. The idea for a municipal information management system had been under discussion since 1988, long before the subject of a data warehouse was mentioned. In 1990, these discussions led Cologne to develop an information system which took account of management, planning departments and political decision-making bodies' needs. The SIS was conceived as a joint project, in which a number of other large towns and the statistical departments of state authorities also participated, under the city of Cologne's management.

The objective of the SIS was to make relevant data available swiftly and in a form that could be readily interrogated by a range of different users including strategic planners, administrators, marketing and sales departments, purchasing departments and political advisors. It had to combine differently structured data from diverse sources. Information which had been collected for quite different reasons had to be turned into precise, spatially referenced data through a process of selection, aggregation, transformation and combination. This is how the system works:

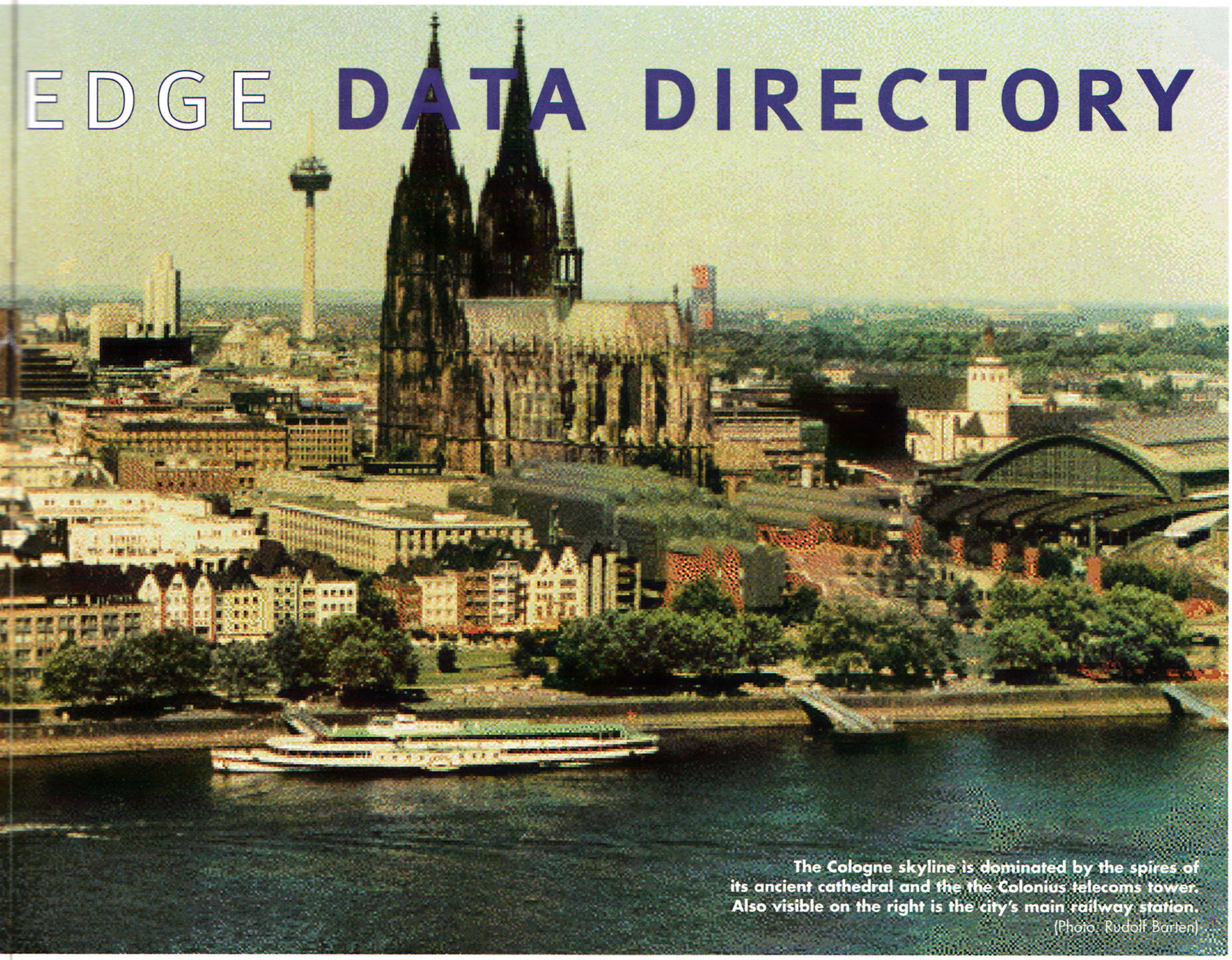
- The SIS reads in raw data from operational applications (such as data extracted from automated administrative and measurement procedures, as well as information from surveys and market research) and links it with data from other sources (for example, official statistics or external data sources). The information, which is all differently organised, has then to be combined into a homogenous data structure.
- For reasons of performance and efficiency, the base data, which is often very extensive and tends to be updated at different intervals, must be combined in the SIS into multi-dimensional parcels of information.
- These parcels of information are used to organise the data along thematic, spatial and temporal axes. Based on the requirements of the analysis, users can construct any number of combinations of data out of the parcels. Flexible transformation, selection, aggregation and linkage processes can then be performed.

The multi-dimensional SIS takes into account the different data structures of the individual operational processes. Not only the data model, but also the range of functions offered by the SIS is completely open. The idea is to make complex individual database enquiries possible without programming effort.

Building on success

Right from the start, the city of Cologne realised that spatial information was an essential component of a central information management system. Nearly all business and administrative sectors, as well as the everyday activities of living, working, shopping, relaxation and education have a spatial dimension. The ways in which these activities are distributed within a city are very complex but rarely come about by chance. Understanding this distribution and the processes affecting spatially related decisions is

EDGE DATA DIRECTORY



The Cologne skyline is dominated by the spires of its ancient cathedral and the the Colonius telecoms tower. Also visible on the right is the city's main railway station. (Photo: Rudolf Barten)

of great importance, not only to local authority departments responsible for town planning and the provision of schools and social services, but also to the strategic planning departments of private companies—marketing and sales managers in particular.

Over the past few years, a statistical RBS (*Raumbezugsystem*; a spatially referenced system) has been developed to make this spatial information available within a central information management system. The object of RBS was to classify the entire area of the city of Cologne and the surrounding region at various levels of detail according to the municipal structure plan (see right). The entire street network is included in the database, as are the addresses of the 141,000 or so buildings which make up the town, their attributes and relationships to other objects. The geo-database now contains about 250,000 objects. The benefit of SIS in relation to RBS has not only been to automate previously manual administrative tasks, but to make new modes of analysis possible. This has been possible because there is almost no limit to the ways in which the geo-objects, and relevant data about their relationships, can be used to select and aggregate attribute data within the SIS data warehouse.

The wider picture

The key to SIS is the system's metadata control mechanisms, which structure, standardise and describe all the data held. They build a framework into which data from external sources is loaded, and enable valuable information to be processed transparently.

The metadata file allows users to search records and retrieve data according to subject, data source, or particular characteristics. Metadata describes and structures the information, so users can navigate their way through the data model and obtain the required information without having to write a program.

Large municipalities often have a wide variety of different hardware and software systems installed. SIS must therefore offer not only the possibility of linking various data sources, but also provide logical and physical support of heterogeneous hardware networks consisting of Unix servers, work-

Municipal structure plans

The municipal structure plan (*Kommunale Gebietsgliederung*) is the mechanism by which many German cities subdivide the town into regions, districts and wards for administrative, statistical and planning purposes. It also forms the basis of an RBS. The most important element is the address, as most other information is related to this. Closely associated with the address is the street network.

The city is also divided into smaller blocks at various levels of detail. These blocks provide subdivisions at the level of postcode, electoral district, school catchment areas and traffic cells, for example.

Germany

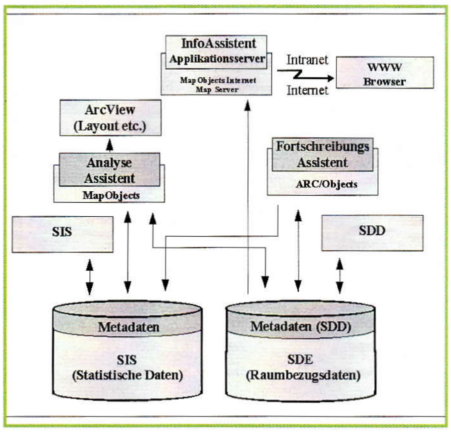


Figure 1. Integrating GIS into the SIS ensures data consistency between the two halves of the system.

stations and PCs. A range of different types of database system must be supported, and an open systems environment offered for future developments. A number of interfaces are provided to typical GIS applications such as traffic and environmental planning, and other specialised applications. Standard tools such as statistical analysis software, spreadsheets, business graphics or cartographic software can then be used to analyse and visualise the data.

Data warehouse

The concept of the data warehouse is borrowed from automatic warehousing technology, in which goods (data) are stored in a structured way. Specific items can be called down from a catalogue (metadata file) and if necessary made up into complete orders (parcels of information) for a customer.

The originator of this idea is considered to be W. Inmon, who defined a data warehouse as a collection of thematically oriented, integrated, permanent, time-based data that satisfies the information needs of managers.

The kernel of a data warehouse is a multidimensional data structure that can adjust to different sorts of management questions and can be used by people without extensive knowledge of data processing. In the SIS, these data structures are known as parcels of information. Apart from the actual data (values) they also describe metadata. Metadata is high-level data which describes the structure of the basic data. Its role is to navigate through the data warehouse and make flexible analyses possible.

Cologne achieved reciprocal access by integrating GIS into the SIS. Its Department of Statistics, Residency and European affairs created spatially related structures for the SIS, which were updated using RBS (see figure 1). Via linkages to the attribute data, these structures were then used to carry out spatially referenced analyses. All updates to the geometry of geo-objects, their attributes and relationships, were made first of all in the RBS and then carried over to the SIS. In this way, data consistency was guaranteed between the two systems.

Three steps to migration

The prototype RBS was developed using the GRADIS GIS, made by the now-defunct Strässle company. When the system had to be transferred to a new software platform, Cologne chose ESRI as a new development partner. ESRI proposed a three-stage migration plan, which safeguarded the future of RBS and allowed further development of the spatially-referenced data warehouse.

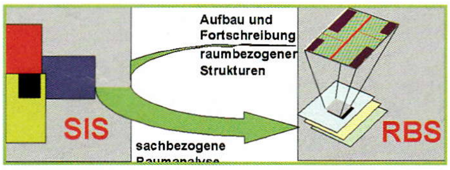


Figure 2. Improving data access over the Internet is one of the roles of the "geo-assistants".

At the same time, the system allowed ESRI products in use by utilities companies and other municipal departments to be easily embedded. The further development of an integrated geoinformation system will be based on the SIS, together with ESRI's Spatial Database Engine (SDE) and various "geo-assistants", which are controlled using metadata (see figure 2).

The first step in the migration involves transferring the geodata and the modelling functions of the current RBS, including its linkages to the SIS, to ARC/INFO. The second step provides access (via metadata) to both geodata and attribute data via the SDE. In the third step, the "geo-assistants" are generated, to make access to data over the Internet/Intranet quick and easy.

The Semantic Data Dictionary (SDD) forms a descriptive layer on top of the SDE. It not only documents the geodata, but also stores various amounts, enquiries and map-based presentation material for further use. The Update assistant, based on ARC/Objects, uses metadata not only to support the SDE geodata but also to keep all of the spatially referenced information in

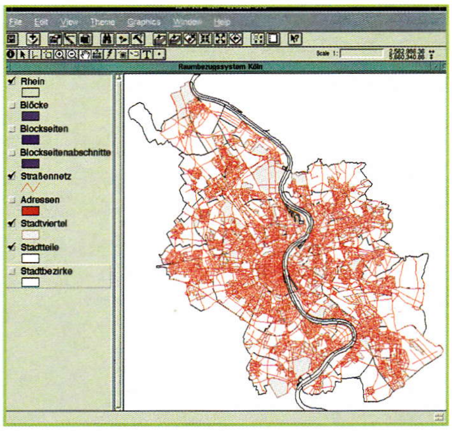


Figure 3. ESRI's ArcView GIS is used to generate finished maps of Cologne's *Raumbezugsystem*.

the SIS up to date. It can also take account of changes in the data model. The Analysis assistant, based on ESRI's MapObjects, enables SDE and SIS data to be processed. Relationship links can be used to aggregate attribute data according to any desired criteria. Enquiry definitions can be stored, edited and re-used. Data and presentation material can be processed in ArcView, for example, to generate a finished map (see figure 3).

European aid

Cologne's Strategic Information System has benefited from a number of EU information technology projects. Developing the SDD as the basis for all the geo-assistants, for instance, is being undertaken within the EU project ENTRANCE. The development of the Update assistant and the Analysis assistant is part of the EU project EUROSCOPE, and the creation of the Info-assistants forms the basis of the EU project ENTIRE. Since making information available over the Internet/intranet has a high priority, another project proposal under the name GALA is being submitted to speed up the work.

The data warehousing system which began in the Department for Statistics, Residency and European Affairs has since been implemented not only throughout the council but also in other large municipalities and state administration departments. It has even been installed in several large private companies, demonstrating that the open concept on which it is based can be applied across a range of public and private organisations.

ALFRED WALGENBACH can be contacted at the city of Cologne's Department of Statistics, Residency and European affairs, Athener Ring 4, 50765 Köln, Germany. Tel: +49 221 221 1855; Fax: +49 221 221 1705. DIETMAR HERMSDOERFER can be contacted at the same address. Tel: +49 221 221 1910; Fax: +49 221 221 1705