GEO^{2}DIS evaluation report

Document Reference: GEO2DIS/MCS/SR/WP540/01/V1.1

Version: V1.1

Date: 24/03/97

Author(s): B. Bouvier, R. Havas, P Hayot, R. Della Maggiore, A. Musone, P.Mogorovich, S. Riedhammer, U. Rugani, M. Trevisani

Distribution: Steering committee, Exploitation board, Project Manager, Site managers, Project Secretariat

Abstract: This document summarises all the operational results of the GEO2DIS project

Keywords: Products, market, dissemination, geographical data server, catalogue, object oriented, spatial index

References: see chapter 2

Approved by: R. Havas, U. Rugani, H. Schuette, M. Trevisani

Amendment History		
Date	Issue	Description
September 16, 1996	0.1	First draft for comments
November 29 , 1996	0.2	First release
March 10, 1997	0.3	Release for comments from the partners which takes in account comments from the reviewers after the final review
March 14, 1997	1.0	Release for review by the CEC
March 24, 1997	1.1	Change of the name of the company MATRA CAP SYSTEMES into MATRA Systèmes & Information

ABSTRACT

The geographical data, both raster (satellite images, aerial photographs, scanned maps, ...) and vector (DCW, IGN products like BDTopo and BDCarto, ...) are growing up exponentially in amount, complexity and variety. Need of geographical data are growing in the same way among governmental and local autorithies and commercial and industrial operators for various applications like urban planning, environment monitoring, civil engineering project impact studies and communication networks design .

Therefore, potential users have a ever increasing need of services provided by responsive organisations (data distributors) using efficient tools for acquiring geographical data, i.e. identifying, locating, selecting, purchasing and retrieving it.

The ESPRIT III project n°8017 GEographical Object Oriented heterogeneous Data Information Server (GEO2DIS) was born out of this need for efficient tools for geographical data acquisition. It was entrusted to a Consortium whose partners are MATRA SYSTEMES & INFORMATION from France. DORNIER from Germany and INTECS from Italy, with associated partners : IGN, O2 Technology, CNUCE-CNR, Region Toscana and the municipality of Pisa.

Project goals were the definition of consolidated user requirements at the European level, the definition of a common architecture of the geographical data server and the validation of the concept and of the architecture through three applications.

The project has designed, implemented and validated a system based on two types of stations the data distributor station and the data user station, connected through LAN or WAN. The data user station allows users to submit queries to the catalogue and to define and issue a request for data and to receive them using powerful MMI, where the footprint of the data are located on a Reference Background. This MMI is customisable to the user's needs. The data distributor station answers to the requests of the data user station and offers the services of managing the catalogue and the data. managing the

For the description of the heterogeneous geographical data sets a common data model has been defined as an object class. For GDS (and GDS descriptions) request, the Geographical Object Query Language (GeOQL, extension of the OQL (Object Query Language), which is part of the standard ODMG 93.

This architecture allows to build the system on various GIS or DBMS, according to the operational requirements of the data distributor. It was actually implemented on the GIS Arc-Info and the OODBMS O2 which has been fitted with a spatial indexing module giving responses times divided by ten compared with previous tools.

The system has been validated by three users using it for acquiring data for three very different applications : Assessment of the Visual Impact of a sport and leisure centre in Deulemont (France), monitoring of the land use, suitability analysis for an waste disposal site near Pisa.

The innovation and benefits of the system are that, even now, no other tool than GEO2DIS offers all the features needed for acquiring geographical data : the capability to manage large amount of data of both vector (including 3D objects) and raster types and the geographical continuity, use by potential non IT specialists users, efficient response time and acquisition methods.

As the needs of the users and the state of the technology continuously progress, some improvements have to be planned for the system. The most important of them is the use of Web clients.

Partners are marketing the system. Multilingual commercial leaflet and technical description have been produced. Presentation have been made to potential customers.

Contents

1. Introduction

1.1 Purpose of the document

This document gives overall picture and evaluation of the ESPRIT III project 8017 GEographical Object Oriented heterogeneous Distributed Information Server (GEO^{2}DIS). It summarises all the relevant information about the project.

1.2 Content of the document

The initial objectives of the project are first recalled in the fourth chapter of the document.

The achievements of the project against the objectives are described in the fifth chapter of the document.

Innovation and benefits of the system are then outlined in the sixth chapter.

Recommendations for the improvement of the products are given in the seventh chapter.

The results of the project are then described in detail in the eighth chapter : common European requirements for the system, state of the art at the begin of the project and its evolution during the project, technical description of the implemented system, technological advances needed for this achievement, products to be developed from the prototypes and marketed, evaluation of the performances of the implemented system.

The report ends with an overview of the exploitation of the results of the project by the partners.

2. Applicable and reference documents

2.1 Applicable documents

Following documents have an prescriptive value :

AD1 ESPRIT PROJECT 8017 : Geographical object oriented heterogeneous data information server, GEO^{2}DIS contract

AD2 Consortium agreement related to ESPRIT III Project N° 8017 "Geographical object oriented heterogeneous data information server" ("GEO^{2}DIS")

AD3 Quality assurance plan (GEO2DIS/MCS/PM/WP000/01)

AD4 Project Glossary (GEO2DIS/MCS/PM/WP000/03)

AD5 Project Plan (GEO2DIS/MCS/PM/WP000/02/V1.0)

AD6 GEO2DIS workplan update (GEO2DIS/MCS/PM/WP000/06/V2.2

2.2 Reference documents

Documents in the following paragraphs have an informative value :

2.2.1 Documents issued by the project

RD1 User's requirements (GEO2DIS/RTO/RD/WP110/01)

RD2 Specification of the generic model (GEO2DIS/DOR/SD/WP120/01)

RD3 Geographic query language (GEO2DIS/DOR/SD/WP120/02)

RD4 State of the art of geographical data formats (GEO2DIS/CNR/SR/WP131/01)

RD5 Synthesis of available tools (GEO2DIS/INT/SR/WP132/01)

RD6 O2DBMS design schemes specification document (GEO2DIS/MCS/SD/WP140/01)

RD7 Specification of the update solution (GEO2DIS/INT/RD/WP150/01)

RD8 Specification document of the catalogue content (GEO2DIS/MCS/SD/WP160/01)

RD9 Client part specification (GEO2DIS/INT/SD/WP161/01)

RD10 Requirements for MCS's application (GEO2DIS/MCS/RD/WP170/01)

RD11 Requirements for DORNIER's application (GEO2DIS/DOR/RD/WP170/01)

RD12 Requirements for INTECS's application (GEO2DIS/INT/RD/WP170/01)

RD13 System requirement (GEO2DIS/MCS/RD/WP180/01)

RD14 System specification (GEO2DIS/MCS/RD/WP180/01)

RD15 MCS list and catalogue of acquired data (GEO2DIS/MCS/DD/WP210/01/V1.0)

RD16 DORNIER list and catalogue of acquired data (GEO2DIS/DOR/DD/WP220/01)

RD17 INTECS list and catalogue of acquired data (GEO2DIS/DOR/DD/WP220/01)

RD18 O2 engine for spatial DB detailed design (GEO2DIS/O2/DD/WP320/01)

RD19 O2 engine for spatial DB release 1 reception report (GEO2DIS/MCS/AF/WP320/01)

RD20 Implementation report on data clustering (GEO2DIS/O2/DD/WP330/01)

RD21 Specification of the Server part requests manager (GEO2DIS/MCS/SD/WP340/01)

RD22 Software test plan of the Server. part requests manager (GEO2DIS/MCS/TD/WP340/01)

RD23 Software test report of the Server. part requests manager. (GEO2DIS/MCS/TD/WP340/02)

RD24 Design of the Client part implementation on PC (GEO2DIS/INT/DD/WP350/01)

RD25 Client part user manual (GEO2DIS/INT/SD/WP350/01)

RD26 Design of the Client part implementation. on UNIX Station (GEO2DIS/INT/DD/WP350/02)

RD27 Specification of the Transformer for Terra-Logic (GEO2DIS/DOR/SD/WP360/01)

RD28 Specification of the Transformer for Arc-Info (GEO2DIS/INT/SD/WP360/03)

RD29 Integration report of the Transformer for Arc-Info (GEO2DIS/INT/TD/WP360/02)

RD30 Scenario of the MCS's application (GEO2DIS/MCS/TD/WP381/01)

RD31 Design of the MCS's application (GEO2DIS/MCS/DD/WP381/01)

RD32 Scenario of the DORNIER's application (GEO2DIS/DOR/TD/WP382/01)

RD33 Design of the DORNIER's application (GEO2DIS/DOR/DD/WP382/01)

RD34 Scenario of the INTECS's application (GEO2DIS/INT/TD/WP383/01)

RD35 Design of the INTECS's application (GEO2DIS/INT/DD/WP383/01)

RD36 Integration report of the server in MCS site (GEO2DIS/MCS/TD/WP410/02/V1.0)

RD38 Integration report of the server in INTECS site (GEO2DIS/INT/TD/WP430/02)

RD39 Validation plan (GEO2DIS/RT/TD/WP510/V1.0)

RD40 Exploitation plan (GEO2DIS/DOR/SR/WP600/01)

RD41 MATRA SYSTEMES & INFORMATION Exploitation plan (GEO2DIS/MCS/SR/WP600/01)

RD42 DORNIER Exploitation plan (GEO2DIS/DOR/SR/WP600/01)

RD43 INTECS Exploitation plan (GEO2DIS/INT/SR/WP600/01)

RD44 GEO2DIS test results (GEO2DIS/RT/TD/WP520)

2.2.2 Other documents

None

3. Terminology

3.1 Glossary

Vector

See AD4 for the terms used in the project.

Terms specific to this document are the following :

- **CAD objects**
  Objects provided by a Computer Aided Design (CAD) system

3.2 Abbreviations

See AD4 for the acronyms used in the project.

Acronyms specific to this document or very often used are the following :

CAD	Computer Aided Design
CEC	Commission of the European Community
DD	Data Distributor
DOR	DORNIER GmbH
DP	Data Provider
DSS	DORNIER GmbH SatellitenSysteme
DU	Data User
EM	Exploitation Manager
IGN	Institut Géographique National (France)
MCS	MATRA Systèmes & Information
SSU	Selection Support Unit

4. Objectives of the project

4.1 Rationale

The GEographical Object Oriented heterogeneous Data Information Server (GEO^{2}DIS) project was born out of a growing awareness about the following trends :

- One is of technology push type and resides in the availability of large amount of data like maps, satellite images, aerial photographs, the progress of advanced data base management systems, man machine communication modules,....

- The second is of market pull type and resides in the needs for managing aspects like industrial pollution, transportation infrastructure, hydrography, forestry by local or regional authorities.

In this contexts, the aim of GEO^{2}DIS was to develop an information server platform which allows the users (professional mostly) to manipulate (store, retrieve, transform, annotate, correlate) any type of data relevant to the management of geographically referenced objects.

Existing systems have significant shortcomings. They hardly can manage large amount of data and geographical continuity is not easy to obtain. They are basically restricted to vector type information and their exploitation is more suitable for Information Technology specialists than potential non IT users. Most of them are limited to a two dimensional (2D) view of the world and cannot manipulate multi-dimensional objects as needed, for example, in construction work impact studies. The raster/vector integration is poor and the manipulation of complex objects is awkward. Their response time is long and they do not efficiently deal with geo-coding and acquisition methods.

The emerging object oriented (OO) technology appeared as a promising solution to meet the above requirements because of its capacity to manipulate richer data and processing structures. The GEO^{2}DIS team thus includes a leading provider of object oriented solutions and planned to build on this expertise to research ways to best benefit from OO potentialities in a geographic context.

Building a new system based upon state of the art technology imply some commercial risks. To prevent users from rejecting the new system, the consortium needed to provide friendly user interfaces with existing commercial systems, so that users can build and preserve their initial investment. In this way GEO^{2}DIS was aimed to add new efficient management functions to existing data sources and a smooth up-grade to state-of-the-art OO solutions.

Improve the competitive position of the European industry and the co-operation between the European commercial and industrial companies, universities and public organisations from all the countries of the EEC are major objectives of the ESPRIT programmes. Therefore, the geo2dis project has also for objectives the improvement of the competitive position of the partners and of the co-operation between the partners of the project.

4.2 Description of the objectives

The following objectives were set up for the project according to the rationale defined in the previous paragraph :

O1 : To implement an geographical information server.

O2 : To demonstrate the capabilities of the geographical information server.

O3 : To assess the use of the Object Oriented Technologies for the management of geographical data.

O4 : To improve the co-operation between the partners.

O5 : To improve the competitive position of the partners

4.2.1 O1 : to implement an geographical information server.

The main features of the server to be implemented were as follow :

F1 : to allow the users to pool all types of geo-data in a single server

F2 : to allow the users to manage efficiently cartographic and geo-data

F3 : to allow the users to manage easily large amount of data

F4 : to allow the users to implement efficiently applications

F5 : to offer user friendly MMI

F6 : to be implemented on COST components, O2, Arc-Info and Terra-Logic

The objective is meet as far the implemented system has the planned features.

4.2.2 O2 : To demonstrate the capabilities of the system.

To evaluate the server performance, three applications were planned to be defined together with the end user and implemented upon the server. These applications had to demonstrate the server ability to manipulate all kind of data easily. They had to manage large amount of data, to fit operational aspects and their interfaces had to reflect the real end user wishes.

The objective is meet as far as the system pass successfully the tests of an validation plan defined with the end users.

4.2.3 O3 : To assess the use of the Object Oriented Technologies for the management of geographical data.

The use of an OODBMS compliant with the ODMG-93 standard (O_2) for the management of geographical data was to be assessed, following a first experience made by the IGN (Project GEO2).

4.2.4 O4 : To improve the co-operation between the partners.

As the co-operation between partners can be in many fields, this objective has to be analysed according to these fields. The co-operation between the partners of the geo2dis project had to cover the following :

- to have a common understanding of the requirements of the system
- to define a common architecture for the system
- to share the implementation of the components of the system
- to define a common exploitation of the results of the project

4.2.5 O5 : To improve the competitive position of the partners

This objective have to be split into the following sub-objectives

O51 : To attract potential end users for this kind of server.

O52 : To define the objectives and the means of the exploitation of the results of the project by the partners.

O53 : To demonstrate the know-how of the partners in the field of the processing of the geographical information.

O54 : To provide to the partners components for building systems embedding geographical information.

5. Achievements

The following paragraphs describe the achievements of the project against each of the objectives defined in the previous chapter.

5.1.1 O1 : to implement an geographical information server.

In order to assess the achievement of the project against this objective, we have to examine if the system has the planned main features. That is done in the following sub-paragraphs for each feature.

5.1.1.1 F1 : to allow the users to pool all types of geo-data in a single server

This feature was implemented and demonstrated. Heterogeneous data has been loaded in servers based on both technologies (O2 and Arc-Info).and related entry has been created in the catalogue. Loaded data are listed in RD15 to RD17.

5.1.1.2 F2 : to allow the users to manage efficiently cartographic and geo-data

This feature was implemented and demonstrated on both technologies.

5.1.1.3 F3 : to allow the users to manage easily large amount of data

This feature was implemented and demonstrated on both technologies. The spatial index integrated to the O2 OODBMS worked effectively.

5.1.1.4 F4 : to allow the users to implement efficiently applications

This feature was not demonstrated by the planned applications.

5.1.1.5 F5 : to offer user friendly MMI

Both the client part and the ATLAS application provides user-friendly interface with powerful graphical tools for the definition or the area of interest for the selection of the data. That was demonstrated during the final demonstrations.

5.1.1.6 F6 : to be implemented on COST components, O2, Arc-Info and Terra-Logic

The server part of the system was implemented only on O2 and Arc-Info. The implementation of the server on Terra-Logic was cancelled because of the evolution of the market. The client part was implemented on the Arc-View product from the ESRI, using the AVENUE language.

5.1.2 O2 : To demonstrate the capabilities of the system.

The tests defined in the validation plan defined in RD40 were applied to the system. They were mainly successful. The loading and retrieval of 3D data for the Visual Impact application was not done because the CAD interface of geo2dis is now based on the OPEN INVENTOR format (subset of the VRML format), because of market considerations. The previous interface was not supported in C++ 4.0, in order to spare efforts. The application, based on the EVER product, still use the old interface. The import-export of data in the VRML format for the EVER product was still to be integrated at the time of the demonstrations. Only the most demanding interworking test (every data provided by server of every technology) was successful : Arc-Info coverages were loaded in the server based on O_2 and retrieved for the urban planning application. The other tests were not undertaken, partly because of the delays caused by the fact that the Arc-Info export in VPF do not comply with the standard VPF.

The Visual Impact and the AES demonstrations failed to demonstrate the capabilities of the system, although impressive in itself.

The urban planning demonstration illustrated the power of the GEO2DIS architecture : data administration catalogue browsing, order and fetch of heterogeneous data both from an Arc/Info server and a O2 server, spatial filtering and management of Reference Backgrounds.

The ATLAS application by MATRA SYSTEMES & INFORMATION demonstrated the power of an O2 based solution using the spatial index. The performance was extremely impressive.

5.1.3 O3 : To assess the use of the Object Oriented Technologies for the management of geographical data.

The use of an OODBMS compliant with the ODMG-93 standard (O_2) for the management of geographical data was demonstrated, mainly by the ATLAS application. A database of more than 1 000 000 of geographical objects was loaded into the server implemented on O2 and performances at least ten times better than for solutions based on RDBMS were obtained for the response time of geographical queries or complex queries.

The OMT analysis and design method (Rumbaugh & alii) was used for the analysis and the design of the system (RD14), of the client part (RD9 and RD24) and of the server part request manager (server kernel, RD21). It provided a powerful common language to the consortium. The main limitation is the lack of support for the management of many levels of subsystems.

The Avenue language (ESRI) was used for the implementation of the client part.

5.1.4 O4 : To improve the co-operation between the partners.

The co-operation between the partners of the geo2dis project had the following results :

- a common understanding of the requirements of the system was obtained ; it is described in the documents RD1 and RD7 to RD9 ; a common glossary was defined in AD4 ;
- a common architecture for the system was defined ; it is described in RD14 ;
- the implementation of the components of the system was shared between the partners according to the project plan described in AD5. After the withdrawal of DORNIER from the implementation of the server on TerraLogic and of the client part on UNIX station, the update of the project plan was defined in AD6.
- commercial agreements will be signed for a common exploitation of the results of the project between MATRA SYSTEMES & INFORMATION and INTECS on one side, INTECS and the CNUCE/CNR and MATRA SYSTEMES & INFORMATION and O2 Technology on the other side.

5.1.5 O5 : To improve the competitive position of the partners

This objective was globally reached. Each sub-objective was reached as described in the following subparagraphs :

5.1.6 O51 : To attract potential end users for this kind of server

Efforts in direction of the Region of Toscana, the municipality of Pisa, the Urban Community of Lille (CUDL), EDF (Electricité de France), DORNIER SatellitenSysteme did not succeed, not resulting in a operational use of the system by these potential end users. Further efforts are undertaken against most of them, and against other potential end users. Commercial leaflet and technical description were produced in order to attract further end users.

5.1.7 O52 : To define the objectives and the means of the exploitation of the results

An general exploitation plan was defined in RD40. Specific exploitation plan were also defined by MATRA SYSTEMES & INFORMATION (RD41) and INTECS (RD43). A exploitation plan specific to O2 Technology is annexed to RD40. DORNIER withdrew from the exploitation.

5.1.8 O53 : To demonstrate the geographical information know-how of the partners

The client part , the Urban Planning demonstration and the ATLAS demonstration give an impressive picture of the know how of the partners in the field of the geographical information.

5.1.9 O54 : To provide to the partners components for building systems

As GEO^{2}DIS has been defined as a modularised system, the marketable products described in the paragraph 8.5 « Marketable products » can be integrated in more important systems. The Geographical Objects Libraries implemented by MATRA SYSTEMES & INFORMATION are planned to be integrated in Control and Command Systems, Photo Interpretation tools and Satellite Ground Segment Systems. These systems belong to main business area of MATRA SYSTEMES & INFORMATION.

6. Innovation, benefits and added value of the system

The novelty of GEO^{2}DIS, built on top of existing commercial systems, resides in its ability to :

- manage easily large amount of data,
- pool comprehensively and exhaustively all types of geodata in a single server (users have no longer to be individually connected to each data producers),
- efficiently manage cartographic and geodata without being an IT, a DB or even a GIS specialist,
- deliver geographically continuous geographical data sets,
- -overlay heterogeneous raster and vector geographical data sets.

Therefore geo^{2}disGeO^{2}DIS provides the following benefits :

- A broader knowledge on geographic data.
- Wider access to geographic data.
- Quicker response times in accessing and processing geographic data.
- Avoidance of data collection duplications.
- Providing of a homogeneous interface to heterogeneous data.
- Usage by people who are neither GIS nor database experts by user friendly interface
- Optimised data selection by interactive catalogue help with search and preview of data (e.g. to check cloud cover).
- Cost efficiency improvements
- Cost reduction by delivery of partial datasets, extracts of raster images, layers or coverages of vector images.

The novelty of the C++ geographical object libraries is that these libraries of C++ classes offer all services required for the management and the use of geographical data, using a powerful spatial indexing module (M4) for data retrieval.

The benefits of the components for the implementation of systems are :

- an easy integration of geographical objects in the applications, using well-defined methods,
- an efficient production of new type of geographical objects, using inheritance,
- performances in data retrieval, using the spatial indexing,
- huge gains in productivity, because the complexity of the file structure of the geographical data is masked to the developers.

7. Recommendations for the improvement of the products

Following improvements have to be implemented in the system :

a) encapsulating the Server (DDT) within an http server (thus improving security , reliability, integrability with existing ISs, etc.);

b) moving ahead in the realisation of a set of Server Administration Services, exploiting all what realised already; scenario and design outlines could be further to a re-visit of the INTECS unsolicited study reports on the subject;

c) experiment the integration of a third GIS (e.g. Spatial Database Engine by ESRI, MGE by Intergraph);

d) comply the Catalogue content to the CEN TC/287 standard (for what is applicable, according to the various types of geodata);

e) enhance the ergonomy of the Client MMI;

f) Enhance the Data User Tool (DUT) for using the http protocol.

g) Implement the security classification of the data.

h) Implement the management of the archives.

An industrial plan was set up for implementing these improvements in the next two years.

8. The results of the project

This chapter details how the achievments outlined in the chapter 5 have been reached.

This chapter describes first the common European requirements for the system, then the state of the art at the begin of the project and its evolution during the project. It follows with the technical description of the implemented system and technological advances needed for this achievement. The performances of the system are then assessed. The chapter concludes with the definition of the products to be developed from the prototypes and marketed.

8.1 Requirements

The requirements were specified in the following way : first the general user's requirements were gathered from the « Region de Haute Normandie » experience, from the « Region Toscana » and from and enquiry at potential and actual data users in the Region de Haute Normandie. This user's requirement were formalised. Then, more detailed specifications were derived for update solution, catalogue content and client part, using the experience of the industrial partners. These detailed specifications were validated by the Region of Toscana. Catalogue content specifications were also derived from the users enquiry in the Region de Haute Normandie and validated by the Region of Toscana.

8.1.1 User's requirements

This paragraph summarises the user's requirements as specified from the experience of the end user's and from an enquiry at potential and actual data users in the Region de Haute Normandie. The context is first recalled, then the operational scenario in which the system have to be operated is outlined. After that, the general requirements related to the data are outlined, and then the needed organisation is defined. The detailed requirements related to the data and functionalities are defined in annexes 1 and 2. Analytical requirements are specified in RD1.

8.1.1.1 General Considerations

The GIS market is one of the most active in the area of information sciences, and is related to software, data automation and services. However, the development of applications and the market itself are affected by data-related problems, the most important of which are: the existence, the availability, the cost, the quality and the date of the last update of the data.

This has led to the need for information systems that can document geographical data. Such systems should allow users to know what geodata is available, which is its type, and which is its geographic domain. The system must be accessible on-line via teleprocessing and provide information on the data itself.

This is something like an ‘intellectual bet’ based on :

- the skill of the end users, their capacity to understand the value of this new service and their willingness to pay for it ;
- the truth of those who produce geodata in this tool, as an instrument to sell more data;
- the ability of this tool to be an accelerating factor in the GIS market.

On the other hand, this approach has to deal with a series of technical and organisational problems.

This project involves the same problems that other working groups find when standardising GIS data at a European level. This is due to the high heterogeneity in Europe between different countries and within a country itself. From this point of view, the project has also a role of homogenisation at an EU level.

There are various technical points to focus on :

- the user will not have any specialised computer skills, therefore the User Interface is of prime importance;
- the heterogeneity of geodata in a wide-ranging scale from various thematic contexts;
- the constraints due to the existing platforms, which we cannot obviously change as far as communication technology is concerned.

As the trend in documentation systems is to group together as many services as possible (new geodata, new platforms to link), the system should be open to new developments.

Finally, there are legal and copyright problems.

As there is no consolidated experience in this area, some requirements cannot be considered technical requirements but rather user expectations.

Most end user requirements has been accepted as system requirements to be implemented in the scope of the project, but not all, due to the size of the problem and the constraints of the project. System requirements are defined in the paragraph 8.1.5.

8.1.1.2 Scenario

8.1.1.2.1 Introduction

Geographical Data is used for managing those phenomena that relate to a particular geographical area. Public bodies are mainly responsible for this, but also to a lesser extent, private organisations.

The context for the use of geographical data is very complex and is analysed below in various stages.

8.1.1.2.2 A simplified first ideal scenario

Every state is organised via Authorities (AUs) at various levels that manage the territory (and therefore use geographical data) in a hierarchical structure.

A simplified model of this organisation could be based on three levels:

- Governmental Authorities (GAs) (e.g. Ministries, Central Agencies, etc.)
- Intermediate Level Local Authorities (ILLAs) (e.g., Regions, Provinces, etc.)
- Peripheral Local Authorities (PLAs) (typically Municipalities)

Each AU has different responsibilities, but the geographical area is obviously the same, and so we might expect that the geodata needed for the various Authorities be the same, though it may differ in scale, detail of information and organisation.(see Figure 8-1).

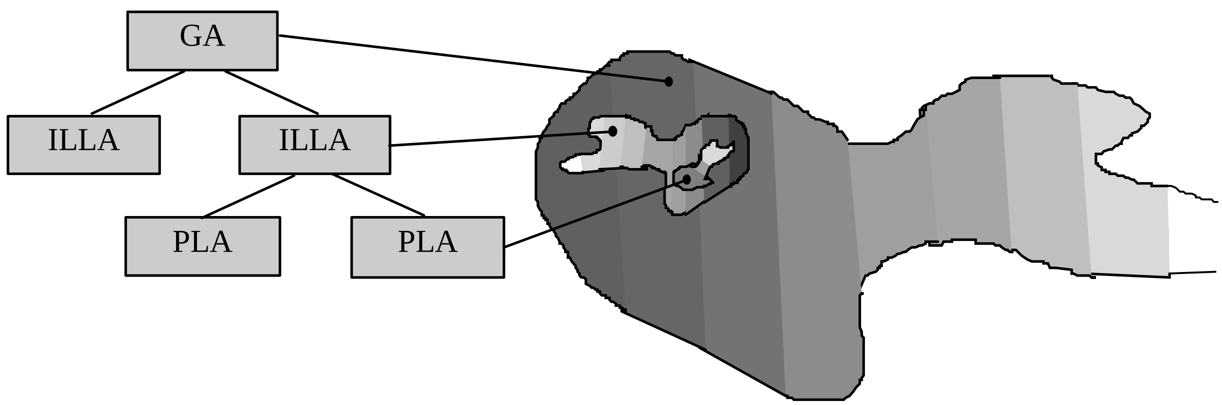


Figure 8-1 Data needs for the different levels of Authorities

Each AU produces some data and uses other data. The data used outweigh those produced, and generally come from other AUs or organisations that produce data. Clearly, a continuous data exchange between the various AUs is needed.

The data in question involves various levels of abstraction in the real world and consequently has quite different characteristics (detail and geometrical structure, number of attributes, historical series etc.). Moving from the PLAs to the GAs the geometrical identity is gradually lost and statistical significance comes into play.

The circulation of the data in this scenario is based on a hierarchical model, which we will call "vertical".

8.1.1.2.3 Second scenario

The scenario models that follow derive from those in the first scenario, they are more complex but more realistic.

Each AU consists of subsets (Departments or Offices), each of which has its own specific responsibilities, and which produces other data. However, besides using the data it produces, it also uses a lot of other support data. This is particularly true for the ILLAs, as that can be seen in .

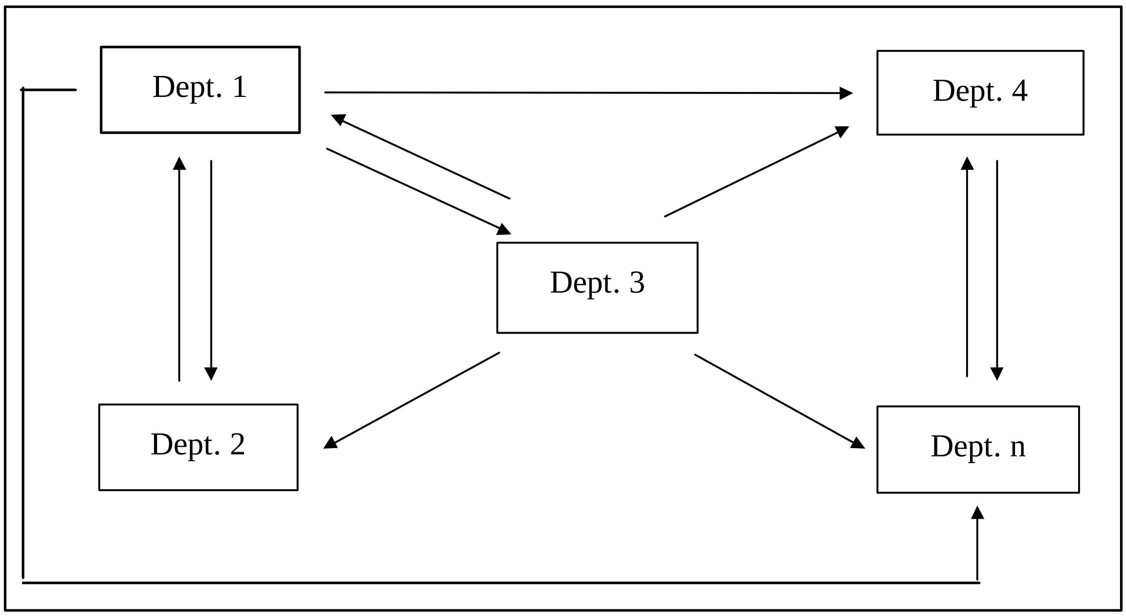


Figure 8-2 : Data exchanges between subsets of Authorithies.

In this case the scale of the data processed is generally the same. The data is structurally similar, but the applications are different and therefore the data models are also different. Nevertheless, the data has to circulate within the AU. We call this "internal circulation".

8.1.1.2.4 Third scenario

A more detailed sketch of the hierarchical structure outlined in n in

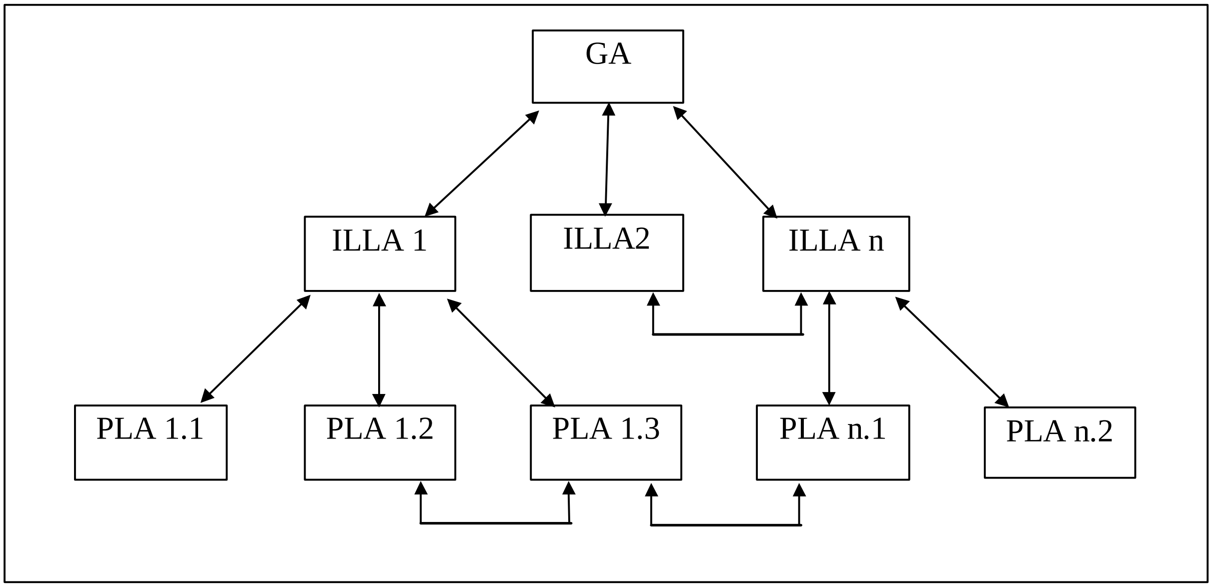


Figure 8-3 General geodata exchange between various Authorithies

Many territorial management problems relate to phenomenon that are not geographically circumscribed by an individual AU (e.g. road management, forest conservation, atmospheric pollution). There must thus be a circulation of data between AUs on the same hierarchical level, generally contiguous.

This data circulation is shown by dashed lines; we will call that "horizontal circulation".

In this case the structure of the data that circulates is similar, but there are difficulties due to the lower capacity for co-ordination, leading to differences in standards, updates, etc.

8.1.1.2.5 Fourth scenario

Due to the long history of our continent, the situation is actually more complex due to the presence of other intermediary AUs that are interposed in the hierarchy. These give rise to new levels of data circulation, which is still in any case hierarchical.

Finally, there are several AUs that act at the same hierarchical level, though with different tasks. They exchange data in what we shall call a "transversal" logic.

At this point the scenario consists of many AUs, within a mixed structure so that each AU has other AUs that act at a higher, lower or similar level, and data exchange needs to be kept active between them.

8.1.1.2.6 General scenario

In the last level of complexity there are other factors that affect this data exchange in an even more different way.

Each country has some AUs whose purpose is to create geographical data. These AUs are not generally data users, but only producers, and the data they create is normally on a medium or small scale (greater than or equal to 1:25000).

These AUs are inserted into the data network outside any hierarchical, horizontal or transversal structure, and in fact supply their products to all those who request them.

Private enterprises, too, both producers and users, are in the network at each level.

The general scenario is thus more articulated, and cannot be easily represented in a diagram. Broadly speaking, each Authority or private enterprise has to read data from and provide data to any other Authority or private enterprise. This relationship may be modified by the technical know-how and organisational expertise of the various AUs.

8.1.1.2.7 Future scenarios

In the scenario described, the circulation of data between the various AUs is of prime importance for at least three reasons :

- the opportunity to exploit economies of scale since a lot of data can certainly be used by many AUs ;
- the need for an AU to always have available the most up-to-date data from other AUs ;;
- the need to avoid duplications due to acquisition of the same data by several AUs.

All this must be possible when managing the territory, where several possibly conflicting interests are involved, such as industrial development, infrastructural development, social problems, problems relating to the quality of life, and preservation of the environment. These activities affect virtually all the AUs in a country.

A scenario thus needs to be hypothesised in which all data on a particular geographical area is accessible to everyone, though naturally there may be limitations due to national security, confidentiality, and cost.

Equipped with suitable HW/SW, users should be able to access a network within which they should be able to

- see what data is available along with its characteristics;
- take this data (perhaps on payment) and then use it as seen fit.

8.1.1.3 Actors

There are two type of actors in this scenario:

a) actors that furnish services

b) actors that use services.

The former provide data related documentation, and a series of services that allow rapid and efficient use of such information. We shall call these suppliers "DDs" (Data Distributors). The latter can carry out the operations described in 8.1.1.2.7 and we shall call them "DUs" (Data Users). Most DDs will in fact also be DUs.

DDs and DUs are not the only actors in our hypothetical structure. We also need to consider data producers, i.e. those who for economic or institutional reasons create information that can then be sold and distributed; we shall call them "DPs" (Data Providers). IGNs and corresponding institutes in other countries are examples of DPs.

The DPs, the DDs and the DUs play different roles as far as data is concerned; the DPs provide data, the DDs offer a service mainly concerning data distribution, and the DUs use the data. Despite the different roles, all the actors can be considered users of this complex organisation whose context is the market in geographical data.

8.1.1.3.1 Problems

To operate a structure based on one or more DDs to which various DUs are linked is an organisational task. This could be carried out either via private investment or more likely through an AU, assisted by a private organisation.

The basic problems are :

a) the geodata : here the difficulty is in standardising the geodata on a technical level (geometric compatibility) and semantic (homogeneity of meanings). Further problems relate to the geodata themselves, and their quality and completeness ;;

b) the tools : there must be quick access to data (or at least to corresponding documentation). This entails developing sophisticated software tools supported by existing services (networks) and consolidated and widespread technologies ;;

c) the users : there are many DUs, likely with a heterogeneous background, who are experts in their fields (e.g. geology, urban planning), but many of whom are not experts in computers, and if they do use them at work are reluctant to change their tools.

8.1.1.3.2 Discussion of possible general architectures

There are many DUs, and many DPs too and as a consequence the system must offer a lot of heterogeneous data.

There are many possible architectures for the system, with various levels of complexity. They are discussed in annex 1.

We believe that the solution 2b is the most suitable for the project; some of the requirements given in annexes 2 and 3 and in RD1 are given taking into account this hypothesis. According to this solution there are many DDs, and each DD has (and therefore can supply to DUs) only their own specific data which is chosen following various criteria (e.g. by theme, territorial responsibility, institutional responsibility). However each DD also has the documentation (CAT) of the data of all the other DDs. DUs can connect with the most convenient DD and check for the data they need; afterwards, if the data required is available they can get them directly from the same DD, otherwise they have to make a link with the DD where the data is. This method has the advantage of gaining access to the DUs and doesn't lead to geodata being duplicated, but it does entail some organisation (Figure 8-4).

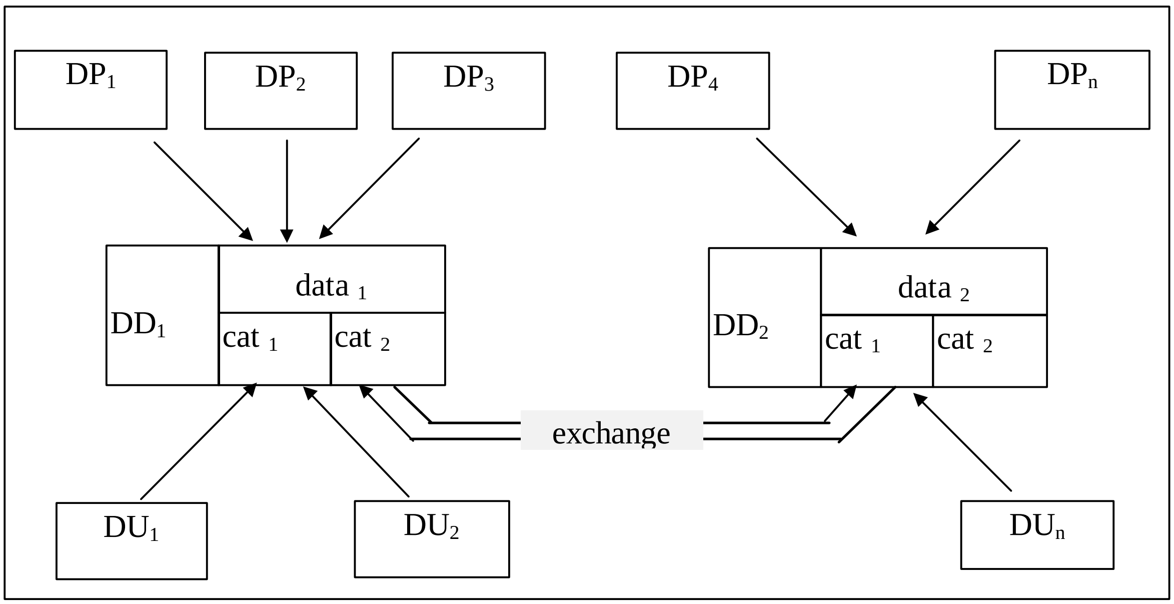


Figure 8-4 : Level 2b implementation

8.1.1.4 The Data

The data is what a DD sells via geo2dis; it is therefore a critical element in the system.

Producing geodata and digitising already existing geodata is not the task of a DD. However a DD wants to carry out operations on the data he receives from a DP so as to offer a service that is more complete than that offered by the DPs. In fact a DD offers data coming from several DPs, a check on the quality of the data, a comprehensive catalogue and distribution facilities.

The data managed by geo2dis relates to the territory, and therefore has a geographical component. A lot of data is conceived without the geographical component being explicit (a lot of statistical, legal and normative data), but since it is applied to territorial entities it has a geographical value. If we take this idea to the extreme then all the information we manage is geo information. geo2dis must be able to manage the kinds of data described in the following sections, which cover most of geo data types, but will also be open to new types of data.

User's requirements related to the data are detailed in Annex 2.

8.1.1.5 Functionalities

geo2dis is based on two types of physical station: the DDís station and the DU's station. On the DDís station the data is prepared for distribution, the information in the catalogue is organised and managed, and communications with the DUs are kept active. The DU's station has to allow users to submit queries to the catalogue, and to define and issue a request for data and receive it, all this within one session alone. Subsequent processing by the DU is independent of GEo2dis and must make use of suitable software, whose owner is the DU.

The DD's station has to operate on a workstation with a UNIX operating system; the DU's station must be able to operate on:

- a workstation with a UNIX operating system
- PCs with DOS 6.2 operating system with WINDOWS 3.1 or a later version.
- PCs with WINDOWS-NT
- Macintosh systems.

The DD's station must be designed so as to be easily adaptable to new types of data from the DPs, to new formats, and to new GIS software that may appear on the market.

The combined use of the server and the DU's stations must enable the DU to:

- identify the data needed for a specific application
- select from available data
- see a generic sample of the selected data
- select a subset of data related to a specific geographical area, typically a polygon area
- select a subset of geographical elements from those available (e.g. select only the main roads from all the roads)
- select a subset of attributes from those available and sort them freely
- preview the data thus selected, that means. allow subsampled vector or pixel information to be displayed on the DU’s screen so that the DU can evaluate the real data he will receive if he purchases it.
- transform the geographical projection and the reference system
- convert the data into a required data transfer format
- activate data delivery from the DD to the DU.

Functionalities required by the users are detailed in Annex 3.

8.1.1.6 Organisation

8.1.1.6.1 Data Distributor

The DD, from an organisational point of view, will be a public organisation created for this goal with respect to a mixed pool of private and public organisations, depending on the political and administrative situation in each country.

The DD will deliver data to a set of DUs within a non-profit making approach. This does not mean necessarily that the service is free, but something may be paid to compensate for the investment. In the future a profit making approach may be planned.

In any case the DD should promote Geographic Information and its use through Geographic Information Systems.

The DP is the primary element responsible for the quality of the data; whereas the DD is responsible for the global service, i.e. data documentation, access to data and any other value added service. The DU perceives the DD as the one who provides the data, thus giving the DD some kind of global responsibility: this should be borne in mind during operational activity.

The DD will provide the DU with a software running on a set of platforms (hardware and operating system) that will enable the DU to use the facilities of the DD. Optionally the DD may provide the hardware too. It is important to maintain active a technological link between the DD and the DU as, in many cases, the DU is unaware of technological developments.

8.1.1.6.2 Organisation Of The Data Distributor

The DD may be considered as an organisation whose main items are hardware, software, data and skilled people

The first three items concern technology; people are concerned with organisational aspects.

As far as this special point is concerned, the DD will activate some special offices (one or more) whose main jobs are to:

- promote the services offered by the DD itself
- authorise users to use the system
- define a profile of the user
- inform users of new data or new products
- monitor user activity
- prepare bills for users.

The commercial strategy of the DD may differ according to the environment it is operating in. With a closed group of users within the same organisation a financial account may not be necessary, or it may only be useful for monitoring the DU's activity. In other cases an exact account of the DU's activity must be kept.

In each DD installation there will be a DBA (Data Base Administrator) whose job is to :

- insert the data furnished by the DP and update it
- create the catalogue
- manage the whole system from a technical point of view
- monitor the system
- offer technical assistance to the DUs.

8.1.1.6.3 DATA USER

DUs belong to many categories: public organisations, private companies and professional individuals.

At present only a few dozen DUs are active, equally balanced between public and private users. In a few years they will increase, in some cases up to some hundreds. In particular, the number of private users will increase. These figures come from the implementation areas planned in the project, thus being only a small part of the whole market.

The number of DU varies considerably according to country.

DU access to the system also depends on the various situations. At most each user will query the catalogue once a day, but only a few queries (from 2% to 10%) will be followed by a request for data. The maximum load of the system will be 40 queries to the catalogue with one request for data per day. This figure will rise to 100-200 and the number of requests for data will go up to five.

In some cases requests for data involve delivering a map on paper or other means, instead of a telematic transmission of information.

As can be seen, accessing catalogues is rather heavy, and if an off-line catalogue has not yet been implemented, concurrent access may exist. This is also subject to the duration of a session which may be from some minutes up to one hour depending on the DU's skill and the performance of the system.

8.1.2 Specification of the update solution

8.1.2.1 Introduction

The update solution, outlined in its general terms in the related specification document, has been implemented accordingly and compliant with the restrictions of the System requirements (RD13). In particular it has been confirmed that updating the GEO2DIS Server means a combined and consistent update of the GDS (Geo Data Set) in the Server (on each specific GIS), the related access structure and the Catalogue. This gave rise also to the concept of 'long transaction', to emphasise this three steps process. The proper functioning of the system, in operations, is assured only when there is consistency among Catalogue, access structure and GIS database.

Formerly it was outlined a two-fold consistent update only: GDS and catalogue. The third element (GIS specific access structure) has been deemed necessary in the implementation phase. In fact, with the current implementation, metadata records in a catalogue are unique according to the following three attributes: hostname, Geo-db and GDS-ID. These triplets are necessary unambiguous on each Catalogue instance. With the hostname it is located the Server on the Internet, with the Geo-db it is located the hosting GIS on said Server, and lastly with the GDS-ID it is located unambiguously a given dataset. Access structures (one for each GIS on each Server) map GDS-IDs with geodata file system locations.

While the updates of GDS and its access structure affect the GIS and its specific administration facilities (each GIS has own tools for self administering), the latter update (catalogue) is the common item above the heterogeneous environments. Focus of the implementation has been such Catalogue administration tool, which is the key-tool necessary to any customer wishing to maintain GEO2DIS.

8.1.2.2 Updating

The updating strategy in GEO2DIS was outlined early 1994, and things have evolved in the years. In the recent years there has been a significant effort towards a standardisation of content of digital spatial metadata. It is worth mentioning the recent US FGDC activities, and the much more recent ISO and CEN TC/287 ones. These initiatives, being further to world- or Europe-wide standardisation committees, have been considered worth for comparison with the GEO2DIS approach. Though happened later, these authoritative initiatives have uncovered the soundness of the basic considerations of GEO2DIS.

As a common standard content of digital spatial metadata it has been recognised both the "capture date" (date at which geodata and ground are consistent) as well as the "last update date", while making a bit redundant the "date of insertion into the catalogue". This latter, in fact, can be simply replaced by a notification technique of a Data Distributor (email, Web pages, etc.), therefore manageable with larger flexibility. Among these, there is also the already outlined whole Catalogue versioning, rather than each singular metadata record.

As expected, though GISs have evolved largely in this respect, it has been confirmed that automation of catalogue updating is unfeasible: GDSs do not contain all the metadata needed in the catalogue. several attributes still require manual entry (e.g. dates, equivalent scale, theme). Large effort is in progress in GIS industry and Governmental Committees in order to combine the two. The emerging US SDTS (Spatial Data Transfer Standard) standard, for instance, is expected to carry both in the same file, while ESRI with their DOCUMENT tool is enhancing the semi-automatic extraction of metadata from coverages.

8.1.2.3 The implementation

An interactive environment has been defined which let administrators define new metadata records corresponding in content and structure to new holdings. This tool, called DAT (Data Administration Tool), is highly interactive and graphic. As usual, it runs on Unix, Windows and Windows '95.

The tool produces new metadata records in the agreed exchange format. Since the format has a formal grammar behind, the tool is not a free graphical editor but instead is compliant with said (customisable) grammar and then acts as a sort of Syntax-Driven Editor (that means prevents the definition of wrong metadata).

Upon completion of the new metadata, the Client allows their import onto the catalogue. Again further checks are carried out in this careful operation, in order to prevent possible data corruption.

Combined with this tool facilities, the administrator has to invoke the specific GIS facilities for completing the missing two thirds of the transaction. These must allow a fair and consistent copy/move of GDSs and related access structure updates.

The same Client allows the export of any metadata record from the catalogue. This facilitates their exchange on the Internet. In addition it facilitates the update of a record. With the Client the metadata record to be updated is exported from the Catalogue. It is then read by the metadata editor and changed. The new saved metadata record is finally imported back in the Catalogue. At this stage it can either replace the former or represent a new one.

8.1.2.4 Towards more advanced solutions

The problem of maintaining an Internet-accessed heterogeneous Data Server of geographical data, through an uniform toolset lying above the heterogeneous GISs, is an objective which was not affordable by GEO2DIS. In the frame of the three years duration the problem has seen several studies conducted and papers presented world-wide, but no one commercial or prototype solution. It is not by chance that, in the technical description of one of the very few systems competing with GEO2DIS (i.e. Alexandria) it is written literally "..metadata management is the Achille's heel of a digital spatial library".

It is in the frame of this state-of-the-art that INTECS produced two unsolicited study reports ("Scenario of the Administration Services", "Design of the Administration Services"), where both the scenario and the architectural solutions of a comprehensive administration software solution are outlined. The reports considered also the emerging WWW technology as well as standards on metadata (FGDC and CEN TC/287).

The results of such studies are confirmed and what produced in GEO2DIS is coherent with the medium-long term enhancements to GEO2DIS.

8.1.3 Specification of the catalogue content

8.1.3.1 Introduction

The content of the catalogue consists in the set of data, which describes the data available within the data server, organised in such a way the user can easily travel through.

To define this set of data , it was necessary to previously analyse the needs of the different actors of the data server. In the following paragraphs the requirements of DU, DP and DD are summarised. Detailed requirements are specified in RD8.

8.1.3.2 DU's point of view

8.1.3.2.1 Overall

When the DU connects to the server, its purpose is to get information's about the data the server is supposed to have and which complies with his needs (his application). The end user's needs may be classified into technical and financial.

The information needed by the end user in order to assess the technical suitability of the data for the application which he need are the geographical area of interest(SSU), the geographical data available within it and the classes of geographical objects. The geographical data have to have a content (information), a structure (vector, pixel, text) and a format compatible with and understandable by, the application software.

The end user also have specific requirements on the data as age, quality, coherency when different data are used simultaneously.

Over the technical constraints, listed before, the end user may have financial ones : price, delay for delivery,...

If we assume that GEO2DIS has to perform on-line delivery, it implies that the end user needs an exact quotation before purchasing the data.

8.1.3.2.2 Selection of the SSU

There are different kinds of thematical or specialised maps of the same territory. These maps, in their multiplicity, can be considered as geographical areas. Despite the difficulties, the interest is to have a synoptic vision of the maps which analyse the different scales of the same territory.

Each discipline takes into account its own phenomena. Each one has spatial characteristics, which allow the user to bring out the spatial sets and to plot their outlines on a map. Unfortunately, the spatial sets of each discipline do not coincide and most often the phenomena are inclusive, exclusive or the spatial sets are intersecting. However, it is possible to use a portion of these spatial sets as referred SSUs.

The user or client will have the choice between 4 types of SSUs :

- administrative SSU (hierarchical organisation of spatial sets which have a common link and a perfect inclusion. Ex : regions, departments, districts, communities,...),

- technical SSU (involving the large spatial sets of the user disciplines. Ex : river basin, employment area, climatic zone),

- user's SSU (client's own type of spatial division),

- "user predefined SSU".

To be useful, the first two ones must be stored into the server and made available to all the users. They are said "server predefined SSUs"

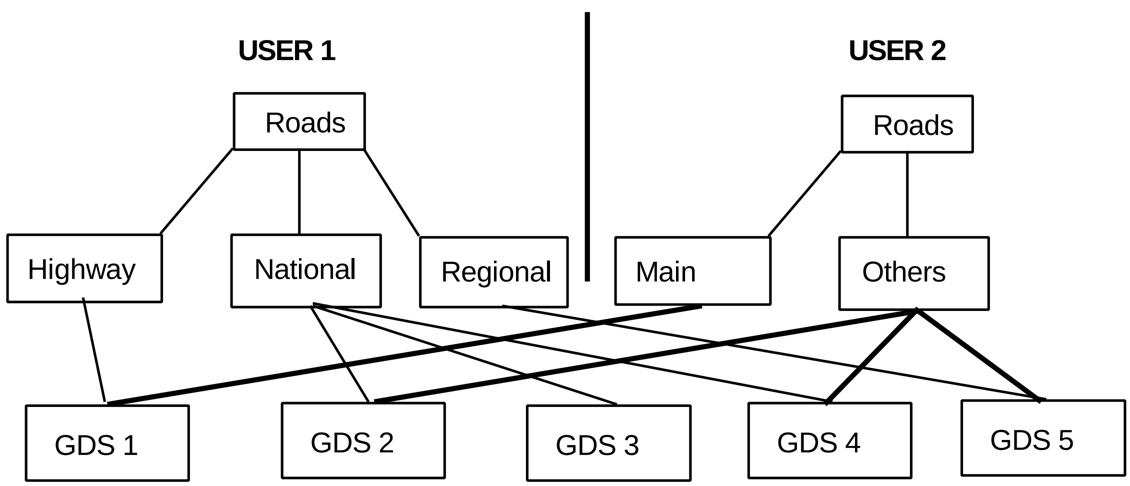
The third one is a bounding box defined interactively by the data user (polygon interactively defined, circle centred on a particular point, buffer, predefined and named area).

The fourth one is a graphic area which the user has defined and named to use it later and several times.

The catalogue must offer functions and data allowing the definition and the selection of this SSU. The most important data for that is the « Reference background », which is an map displaying the main features of the area, like main towns, roads and rivers and administrative limits)

Depending on his technical field, a DU has his own view of the geographical space . Let us consider the "roads" geographical class for example, a urban planner is interested by a detailed information and a hydrographer may only needs to know where are the roads.

This view may be expressed using a tree (said User Tree) structuring the thematic classification with three levels : theme, sub-theme, geo-data.



8.1.3.3 DP's POINT OF VIEW

8.1.3.3.1 Overview

As an general rule, DPs are very interested by the support which a geographic data server may be for the development of their market . That was clearly identified during talks which we had with different national or local DPs.

8.1.3.3.2 Information / Promotion

For the DPs, a server would be an efficient tool to promote their products and advise the customers (DU) of news like the future availability of data or the evolution of existing products.

They also plan to use the server as an information vector to explain to the DU the way in which the GDS was produced and the limitations of their usage. As an example, INSEE said us that they sometimes have troubles because some of their customers compute the total population of a department as the sum of populations of the cities. For INSEE, a server would give the opportunity to inform the DU that such an usage is non coherent with the content of the information and the way in which it was produced.

8.1.3.3.3 Limitations of use

In Europe, the geographic information is generally subject to copyright. Because of that, the DP needs to know who buys and for what. These constraints must be taken into account by the DD who must be able to deliver, on DP's request the list of clients who bought GDS subject to copyright.

For given geographic data, the DP may impose restriction concerning the volume of data to be delivered. As an example, IGN, in France, does not accept to deliver the entire 75 meters DTM over France. This delivery must be subject to a specific authorisation which doesn't depend from IGN only.

On the other hand, the DP may impose a minimum volume for delivery.

As a consequence of this, it may be necessary to have, at the catalogue level, for each GDS, the limits, min and max, of the deliverable data volume.

8.1.3.4 DD'S POINT OF VIEW

The DD level is where the catalogue is managed (creation, updating).

The DD, as well as the DU, may need to navigate into the catalogue (for example : to validate the operations of creation and updating).

8.1.3.4.1 Creating the catalogue

The function "creation" has no impact on the content of the catalogue

8.1.3.4.2 Updating the catalogue

Update of the catalogue must occur at following times :

- arrival of new GDS not already known by the server,
- arrival of a new release of a GDS.

8.1.3.4.3 DU's subscription

To use the data server, the DUs may contract a subscription . This subscription lists, for a given DU, the GDSs he can access and, for each one, describes the legal, financial and technical conditions in which the DU uses the data server.

The legal conditions describe, for a DU, and for the GDSs which he can buy, the legal constraints as listed in chapter 6 of RD8.

The financial conditions specify , for a DU, and for the GDSs which he can buy, the price conditions. These conditions may depends on different factors like the status of the DU (administration, private company, partner, ...) or the commitment of the DU concerning the volume of data which he plans to buy over a given period.

The technical conditions which depend on the way the DU is connected to the data server are conditions like the use of a dedicated hardware or not ; if yes, hardware pertaining to the DU or to the DD or previewing or not.

Two solutions may be proposed for the management of subscriptions. First one supposes that any DU may be strictly different from the other ones, so that it may occur as many types of subscriptions as DUs. Second one consider that subscriptions may be classified within a finite list of types (10 to 20).

Referring to existing data server in other fields of applications, we consider the second solution as sufficient for a geographical data server. At DD level, a user is characterised, by his subscription contact number, and his subscription has a type chosen within the pre-definite ones.

With this assumption, for a given user, the legal, financial and technical conditions are described as a function of his subscription type.

8.1.4 Specification of the client part

8.1.4.1 General system view

The GEO2DIS system is an heterogeneous Data Information Server based on an Object-Oriented approach. It consists of a Server acting as a GDS (Geo-Data-Set) provider, which responds to Client requests of said GDSs. Client and Server run on a LAN or WAN configuration

The main "actors" of the whole system are Data Users (DUs), Data Distributors (DDs) eachone with their own Administrator, and Data Providers (DPs). In brief DPs actually prepare GDSs and related descriptions, DDs distribute them and DUs buy and use them. Such a view has been commonly agreed between various Partners, despite of some National peculiarities slightly differing each other. An Administrator function is associated to each DD.

The Client software must promote geodata access and somewhat “GIS software interoperability” by supporting the capability of (*Open Geodata Access* paradigm):

- locate the data (whether on a local network or halfaway around the world);
- determine if such data (through their metadata) meets the user needs;
- retrieve the data;
- ensure the data are in the format required by the user software;

In addition it must provide support to DPs and DPs in order to publish data and metadata easily and consistently, from their heterogeneous GISs. Such functionalities will be realized through the cooperation of the Client software, namely the Data User Tool (DUT) and the Data Administration Tool (DAT).

8.1.4.2 Problem statement

What do User Requirements ask definitely for what concern the "Client software"? The User Requirements ask for appropriate functionalities in support of the Data User and the Data Administrator. Data Users wish to browse the Catalogue in order to see which GDSs are available on a given geographic area, which are their characteristics, in which format they can be retrieved, and so on. Their demand is for querying such a Catalogue, in order to see the GDSs availability, but having also a clear reference to the territory. Queries are therefore driven by a Reference Background (RB) map of the concerned territory (usually at lower scale of the GDSs ones). The DU can see which "portions" of the territory are actually covered by each GDS. Queries must involve the selection of specific features on the RB (e.g. a Municipality boundary) and the application of spatial operators on them (e.g. which GDSs intersect such a boundary?). In summary the queries belong to two classes:

- where are these data?
- which data are here?

On the basis of the GDSs availability, DUs wish to definitely import one or few of them, by applying also additional qualifiers. Among them it is worth mentioning clipping, filtering on attributes values (e.g. all rivers whose width > 50 mt.), attribute filtering (e.g. removal of columns), coordinate transformation (e.g. UTM to Gauss-Boaga) and format transformation (e.g. from original Geo-db format to DXF). All of these give rise to the concept of GeoData SubSet, just to enforce the global filtering applied to the original GDS. However, from the conceptual point of view, Geo Data SubSets are GDS themselves.

The Data Administrator must have facilities in order to update and administrate the Catalogue and setting the access rights to the Geo-dbs (and related accounting). In this project more relevance is given to the former support (Catalog administration) rather then the latter (Geo-db access rights).

The problem statement is completed by the demand of having the Client part running on both Unix WorkStations (WSs) and PC,s exploiting at the best the basic software available on the respective platforms (X11 with OSF/Motif and Windows 3.11). During the project also the Windows95 platform has been inserted. Unix workstations are LAN and WAN connected to the Server, while Pcs may be also connectionless. In such a case the functionalities will be limited to Catalogue queries only.

8.1.4.3 Analysis

As outlined in the General Architecture, up to three different software components must be implemented in GEO2DIS:

- DU Client;
- Data Administrator Client;
- Server;

The Client software has to be envisaged both connected and not connected to the Server. When connected it will be a LAN connection, "open" to WAN. As stated previously, the connectionless option on PC is expected to enhance the promotion of GEO2DIS, without the constraint of Unix-LAN sites. Therefore there are several possible sites where to host the whole GEO2DIS or the Client part only. The architecture must encompass all of them in a flexible way. In fact in a Client/Server architecture on a LAN, the resources can be almost arbitrarily distributed and shared.

Given the various needs of the Customers, eachone having slightly different habits, it results that a "winning" architecture, common to three different Countries and fulfilling eachone's Customers demands, must not bind itself to any physical allocation of the functionalities on the nodes of the LAN (and optionally of the WAN). On the extreme, in an Unix-LAN architecture all the Client and Server parts could be running on the very same machine. On the other, we can have a standalone PC with all the Client software running there ("open" to connect remote Servers on the WAN).

As hardware allocation, the only constraint is given by the Server part and the related Geo-dbs, which must necessarily run on an Unix WS. In fact both O2 and ARC/INFO run on Unix. The flexibility of the architecture must also accommodate the peculiarities of the hosting DBMS for the Catalogue, which is accessed and logically comprised in the Client part (see also later).

The Administrator Client software is a "super" DU one. In fact the difference stays in the Administrator capability to create new Catalogue entries, and consequently update the Catalogue, not possible for the DU. The DU Client, instead, has no rights to change the Catalog. The Catalog updating is carried out by collecting and sorting Update Requests. These in turn represent "raw" GDS descriptions, to be completed in order to be inserted into the Catalogue. GDS descriptions (alias Catalogue entries) can either refer to owned GDSs or elseone's. In any case they are the result of Update Requests issued by other DDs and DPs (or DUs temporarily acting as Dps).

The Administrator defines the Reference Backgrounds (RBs). Those are separated from the Catalogue. There can be several RBs, allowing an additional degree of freedom for the DU. For instance a default RB for a German user can be the entire Germany, but a specific DU may desire to work mainly on a restricted area (say the Sachsen-Anhalt). Such an option is set by the DU in his User Profile. This contains as many options as possible that can speed up the DU session (e.g. default query template, default display of previously retrieved GDS descriptions).

The Catalogue evolves due to new GDS descriptions, in turn further to incoming Update Requests. At any given time a "master" Catalogue is available to DUs. The Administrator directly updates said master Catalog. In fact in a Client/Server architecture on a LAN, the "master" Catalogue is unique and directly accessed by DUs, in readonly mode.

Note that DUs are allowed to have local copies of the Catalogue. Despite of the obvious data redundancy and of the (potential) arising consistency problems, such an option must be envisaged due to one or more of the following reasons:

- faster and easier access to a disk local to the Client station (no network overloading);
- copy of limited Catalog portions of specific interest of the DU (especially when the whole Catalogue becomes huge);
- export to a "sibling" DBMS on a connected PC, such DBMS in turn becoming the "master" for said standalone PC platform;

If there is no connected PC, the latter export is carried out with traditional means (e.g. tape or CD-ROM). This is also called Catalogue "downloading". In principle, and as previously stated, copies of the master Catalogue create consistency problems upon any change (and next delivery) of said updated master. These can be avoided, basically by means of an unique Catalogue ID plus the adoption of the attribute "date of insertion into the Catalogue”.

8.1.5 System requirements

In order to be compatible with the budget and time frames, the implementation of some requirements specified in RD1 to RD12 was limited.

Limitation related to the user's requirement were defined in RD13.

The generic data model as defined in RD2 had to be fully implemented.

GeOQL had to be used for geo data subset requests. Its implementation had to conform to RD3 and be limited to the extend needed for that purpose.

Geographical data format among those described in RD4 which had to be loaded/output by each implementation of GEO2DIS were defined.

Implementation	**O2**	**Arc-Info**	**Terra-Logic**
DXF	IO	IO	IO
IGES	N	IO	IO
DLG	N	IO	N
EDIGEO	IO	N	N
VISILOG	IO	N	N
ERDAS	IO	IO	IO

LEGEND : I for input, O for output N for none

Available tools were analysed in RD5. Following tools were selected for the implementation of GEO2DIS :

- MATRA SYSTEMES & INFORMATION implementation: O2 OODBMS, SOLARIS facilities
- DORNIER implementation: Terra-Logic, ORACLE, Tel-Use
- INTECS implementation: Arc-Info, PC DBMS (DBASE family), ArcView2

The O2 DBMS design schemes as defined in RD6 had to be fully implemented in MATRA SYSTEMES & INFORMATION implementation only.

A simplified update policy had to be implemented in DORNIER implementation.

Full provisions had to be made in the MATRA SYSTEMES & INFORMATION implementation for the update solution.

Limits of the implementation of catalogue content and functions were defined in RD8.

Applications as defined in RD10 to RD12 had to be fully implemented.

8.2 State of the art

8.2.1 At the start of the project

8.2.1.1 Geographical data formats

At the begin of the project (early 1994) there was neither a consolidated nor an emerging international standard for exchange format of GIS data. The CEN TC/287 was working on the definition of a standard to exchange GIS data, but unfortunately the results were scheduled for the end of 1996.

The CEN TC/287 Interchange Format should have been be the best choice for the project, but as GEO2DIS was not synchronised with CEN TC/287 activity we were forced to choose market formats whilst remaining open to international standard like CEN TC/287. In any case the adoption of CEN TC/287 Interchange Format, whenever available, was strongly recommended for the GEO2DIS products family.

For vector data the DXF format was chosen as it seemed to be the most appropriate for the project purpose though it has not the possibility to transfer the attribute information attached to vectors in GIS systems. To fix this problem an extension for attribute shall be defined. Such extension will also fulfil the requirements concerning the delivery of tabular data.

The most suitable formats for raster data exchange are ERDAS and TIFF.

TIFF format has compression facility, but it does not allow to georeference the raster data and in addition it has some slight limitations concerning the format of data.

The ERDAS LAN format allows the georeferencing of data through the specification of the co-ordinates (direct or indirect) of each pixel, the referential system and the projection adopted, the only limitation being the lack of the characteristics of the ellipsoid adopted for the projection. At the moment ERDAS LAN did not support data compression.

The best solution seemed to use the ERDAS LAN format, which from the technical point of view is more suitable to transfer GIS raster data and in addition allows a self consistent description of the data, without the use of an external file for the georeference specification. A compression algorithm will subsequently be applied to ERDAS LAN files.

8.2.1.2 Available tools

8.2.1.2.1 General

The GEO2DIS project aims at developing a geographical O-O heterogeneous data information server, comprising also, as an essential "ingredient", currently available commercial Geographical Information Systems (e.g. ARC/INFO, Terra.logic). Therefore heterogeneity addresses not only the data type (e.g. raster, vector), but also the data format (i.e. the way such data types are actually stored by each GIS, often independently each other).

In the Technical Annex it had already been preliminarily outlined the rationale behind the GEO2DIS objectives. That addressed current lacks of the commercial Geographical Information Systems (GIS), possibly still persisting in the near-term future. That represents the "background" and "rationale" of the project. Therefore this Report intends to refine said basic background of the GEO2DIS project, by analyzing in more detail some features of current GISs, especially those features which can better insulate the novelty and the added value of GEO2DIS. It is therefore implicit that this Report is not intended to provide an exhaustive and omni-comprehensive description of the current GISs, since that could be too much dispersive. Focus will be every time given only to GISs possible limitations and constraints. In turn those limitations are likely to become the next generation GISs' challenges.

8.2.1.2.2 Current GISs: potentials and limits

A GIS can be defined as a "..set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world and for a particular purpose" As an extension to said concept, it could be conveniently added also the "not spatial" data (alias attributes).

The GIS technology is quite consolidated, and that is confirmed by the relatively high number of commercial products.

GIS have seen a growing interest in the last years. Reasons for that are the rapid development of hardware and software combined with the spreading of knowledge of both the current possibilities and the potentials of GISs to new application areas. In fact, besides the handling of information which is strongly related with geography (like land-use, environment control, agriculture planning, etc.), there also is a significant potential in geo-referencing an enormous lot of data which are still largely handled on alphanumeric DBMSs only (e.g. census data, registry office data).

The GEO2DIS rationale of an "heterogeneous Data Information Server" is coherent with the assumption that users nowadays have their own (commercial) GIS. Lacks can be identified in the fact that, regardless of the advanced technology, each GIS still has its own "world". Standardization, and easy inter-operability, largely claimed, isn't yet fully reached.

Moreover, though large steps have been done in this field, many GISs are still either vector or raster oriented. The use of raster images (satellite but aerial also) is increasing, together with special services such as ortophotos mosaic and superimposition with vector information. Traditional vector-oriented GIS do not perform image processing, and they usually rely on third-party software packages (e.g. ERDAS-ARC/INFO). However ARC/INFO has an integrated raster-GIS tool (GRID).

Therefore there is "room" for investigating how to build a "bridge" between different GISs (and DBMSs) and let users access and import heterogeneous data, with an O-O approach.

Returning the point already outlined in the Technical Annex with the "push and pull" paradigm (i.e. push of data and technology availability and pull of potential market demands), it is possible to point out the following items where current GISs show some deficiencies (and which might represent the challenges for the next years):

- Client-Server Approach;
- Object-Oriented Technology;
- Data Exchange;

Those topics are analyzed in turn, trying to outline synthetically the current state-of-the-art and expected trends. The technologu evolves quickly: therefore what stated here might require significant revision in short time.

8.2.1.2.3 The Client-Server Approach

Once high performance and relatively low cost computer networks have been available, a sort of change of habits begun in the architecture of Information Systems (ISs). This was mainly driven by the possibility to share processing resources among various computers.

As a first result, the Client-Server architecture has progressively replaced the traditional centralized ones, in many IS fields. In fact centralised databases (such as GIS ones) when accessed by remote nodes, may unacceptably load the communication network, slowing down the performance. The solution is to split up large data volumes to reside at different sites and to split up applications to run on different processors, at user will. This is only possible with a Client-Server architecture.

A Server is a process, running on a particular machine, and waiting for being connected to one (or more) Clients, running on the same or other machines connected in the network, and which ask for various services from the Server. The architecture provides a great flexibility, since the processing power is split in two (i.e. Client and Server) and it allows for the most convenient allocation of the processes according to the processing power distribution in a given computer network.

With database systems, Clients are usually that part of the system conceived for handling the Graphical User Interface (GUI), querying, and displaying, while Servers take care of the database, receiving and analyzing Clients requests and serving them.

Note that, according to this preliminary outline, the capabilities of the X11 Window System (in turn based on the Client-Server approach) offer an additional degree of freedom: splitting the X server process (necessarily running on the physical Client Workstation) from the Client application code, which can run somewhere else in the network.

The Client-Server approach allows also for a more efficient check of data consistency when the same data are accessed by multiple users. This because the data are managed by a single program (the Server) which can therefore implement any sort of safeguarding policy (e.g. record locking, deny access).

Apart of the above benefits, there isn't yet any "declared", commercial GIS built over the Client-Server architecture. In fact the demand for separation of processing and fragmentation of data is emerging in big public companies and corporate environments, but it is not yet completely established. It is not by chance that in [23] said Client-Server architecture is mentioned, together with others items, as a possible solution to the "GIS challenges for the year 2000".

GISs are being installed throughout the world, but still mainly on a local basis. However soon the need for data sharing will raise among sibling "institutions" (e.g. Districts and Municipalities, Gas, Water and Electricity Companies, etc.), and that will inevitably lead to give more emphasis to key-terms like distribution, heterogeneity and Client-Server.

In any case Industry is moving. ESRI, has presented ARC/INFO version 7 together with ArcStorm [24]. It is a new product, built on the Client-Server architecture, which manages the whole set of geographic information, ensuring its integrity and security also in a multiple users environment (e.g. features locks, transaction recovery, historical versions of database). ArcStorm will take care of the whole ARC/INFO "public" database, which users can browse, navigate and query for subsequent import on local workspaces. It acts as Server of several potential Clients (e.g. ARC/INFO, ArcView2), by means of an ArcStorm Application Programming Interface (API).

Smallworld Systems Ltd. (UK) too have announced a major new enhancement to their product (Smallworld GIS), for handling very large databases in a distributed environment. The new functionality, called Database Cache, provides greatly improved performance for very large systems in which hundreds of GIS workstations simultaneously access one GIS database server.

8.2.1.2.4 The Object-Oriented Technology

Preliminaries”

Object-orientation applied to GIS is not a complete novelty in the IT world. It is a number of years that such a topic is encountered in specialized papers. However its definition is still subject to different interpretations. Therefore, due to the multiple facets of the O-O approach, [3] suggests that it is more appropriate "to ask in what way a GIS is object-oriented, instead of asking if it is or not". Therefore a preliminary attempt to clarify how things are is deemed necessary.

One of the most common features of GISs is their capability to handle geographic (alias geometric) data and attribute data. It is suggested to view the data managed by GIS at the following different levels of abstractions:

- reality, the real geographical world;
- (conceptual) data model, an abstraction of the real world which contains properties and relations which are thought to be relevant for the best definition of the characteristics of the geographical phenomena; of course, each GIS has its own data model;
- data structure, the logical translation of the model in terms of tables and diagrams;
- file structure, usually the most hidden and representing the physical storage and organization of data on a computer;

The features above, but the real world and the file structure, are then described in the following.

.The Data Model

The key-role is played by the data model, which represents the first bridge between the real world and the presentation, manipulation and analysis functionalities which are to be performed on the data by GISs. These models have to define the locational aspects (coordinates, topology) as well as the spatial characteristics (areas, perimeters, lengths and other attributes). GEO2DIS will have to define a Data Model of its own.

However some clarifications are necessary in order to define the most appropriate scope of the Data Model for GEO2DIS. In fact each GIS has its own data model: this was historically the problem of the many existing GISs. Now each GIS handles its own geo-data which reflect its own data model, which in turn represents an abstraction of the real world.

The GIS data models are usually based on the concept of point, line and polygon, grouped into "layers" (alias tiles, maps). Points, lines, and polygons have a geographic and attribute part. Note that the real feature which characterizes GISs data models is the topology, i.e. the capability to manage the continuity and contiguity of the various geographic features.

In GEO2DIS there is to model existing geo-data, which are structured according to their own data models, with an object-oriented approach. That leads to the concept of geo-object. This represents the most convenient abstraction of (heterogeneous) geo-data.

Therefore instead of interfering with the intimate structure of the GIS's geo-data (GIS's specific data model), we have to move to an higher abstraction level in order to "merge" and "model" heterogeneous geo-data

The Data Structure

In terms of data structure, GIS can be divide in two types: "hybrid" and "integrated".

Hybrid systems (based on the known "coupling technique" or "dual architecture") store geographic information (not attributes) in a set of independently-defined operating system files for direct, high speed access. Each file, also called map layer or map tile ("coverage" in ARC/INFO), represents a selected set of closely associated geographic entities (e.g. roads, buildings, streams). Layers could also be organized according to an homogeneous geometric type (e.g. points, lines and polygons), but that is not a constraint.

The attributes, viceversa, are stored in standard (commercial, relational) DBMSs such as ORACLE, INFORMIX and linked to the spatial data by links (keys). This approach is also known as the geo-relational model [9]. For instance hybrid systems comprise ARC/INFO (by ESRI, USA), MGE (Modular GIS Environment, by Intergraph Corp., USA), Atlas*GIS (by Strategic Mapping Inc., USA), GRASS (USA public domain), MapInfo (by MapInfo Corp., USA) and terra.logic (by DASA/Dornier, Germany).

Note that it is usually a deliberate policy of hybrid systems to encourage the use of "external" DBMS's for attributes, just for saving the highly probable Customer previous investments on such technology. In fact, one of the main novelties of GISs is their capability to provide a spatial reference to "traditional" textual information. But in order to do that, such textual information can stay where it has always been. That's the winning feature of hybrid systems.

It is worth noticing that, differently from other "hybrid" systems, ARC/INFO does run also without an external DBMS, for storing the attributes part. In fact ARC/INFO provides an independent "INFO" part which is a tabular, flat file manager for handling attributes.

Inconveniences for hybrid solutions stays in the risk of inconsistencies between spatial and not spatial data. The query languages are usually different. In addition hybrid systems are single user, since the geometry is single user .

There is a rationale behind hybrid systems: they match the requirements of all those Customers who have already an Information System, based on relational DBMSs. In most cases, GISs have to be complementary to these pre-existing applications.

In the "integrated" types of GIS, both geographical and attribute information are stored in a (commercial) relational DBMS. These systems do not define independent layers or tiles, but just one seamless map (the "continuous map" model). Note however that also many hybrid systems, though having the tile concept, can handle the continuity of the Relational tables hold all the coordinates, the topological information and the attributes. Authoritative examples of this category are Smallworld GIS (by SmallWorld Systems Ltd., UK), System9 (by Computervision Corp., USA) and GEO++.

With the integrated approach, being spatial and not-spatial data stored together, consistency and concurrency controls on both of them are easier.

Integrated solutions overcome hybrid solutions inconveniences, but these systems lack of flexibility and openness (no use of standard products but a comprehensive "turn-key" system). Moreover stored data must submit to relational data base constraints (normalization), therefore requiring costly join operations when answering queries. In fact attributes, but not coordinates, find their best place in commercial DBMSs. Moreover the DBMSs have a poor data typing and the adopted query languages (e.g. SQL) are not complete, usually requiring to be merged with traditional programming languages.

Concerning performances there is still a debate about if and to which extent hybrid systems are better than integrated ones.

8.2.1.2.5 Object-Orientation and GIS

Given the previous preliminaries, there is to converge to the point of how and to which extent object orientation can be applied to GISs. It is obviously out of the scope of this Report to define the principles of object orientation. However it is worth while mentioning that it is a model which attempts to be as closely as possible to the real-world applications, and to achieve software reusability and extensibility. The most important aspects of object-orientation are abstract data types, information hiding, inheritance (single, multiple), object identity and polymorphism.

How GIS and "object-orientation" can be combined

- object-oriented programming languages;
- object-oriented data models;
- object-oriented databases;

Those are described in the following.

O-O programming languages

In the programming languages field, further to the first launch of Simula67, others have appeared, like Ada, APL, Smalltalk and C++. Taking into account the commonalities they offer (e.g. data abstraction, data hiding, inheritance), they can be considered superior w.r.t. other procedural languages, especially if considering reusability and maintainability aspects . For instance, once a set of geographically-related C++ classes have been defined accurately and comprehensively, any new development is easier and faster.

In addition O-O programming languages facilitate both the cooperative work (parallel developments in geographically-distributed teams) and the individual testing

Only one commercial GIS is known having an object-oriented programming interface: it is Smallworld GIS with Magik, which is an interactive object-oriented programming language and environment (it supports both O-O and procedural methodologies).

ESRI is launching ArcView2, an easy-to-use desktop GIS, which is also supported by a development environment comprising a C++ like scripting language (Avenue), which allows the users a great flexibility in order to develop their applications. However ArcView2 has lesser functionalities w.r.t. ARC/INFO.

O-O data models

Object-oriented data models, differently, are used to model the reality as a variety of interrelated objects, which can be viewed at different levels of abstraction . Objects can be complex (alias composite) or primitive. Typical geometric primitive objects are points, lines and polygons. Objects are characterized by their structure and behaviour. Objects having the same structure and behaviour are said to belong to the same object class. Objects are therefore organized in a tree structure, with the following possible relations: gen-spec (alias "is a"), aggregation (alias "part of") and association

There is a great consensus that O-O data modeling is quite appropriate for handling geographic data, especially due to the nature of its data types: spatial (geometric and topological) and not spatial .

Note that just some O-O designers state that looking for for an O-O data model, general enough to encompass a large variety of geographical applications, is a difficult enterprise. All of that is likely to represent a challenge for GEO2DIS, where the common data model will have to accommodate different application domains, GISs and DBMSs (Geo-dbs).

It is important to outline that an O-O data model is independent from an O-O database. The former can be implemented also over a relational one. However difficulties arise in the implementation of the behavioral aspects, by packaging methods with the data structure. Note that System9 claims that they provide an object-oriented data model of its geographic database, though being strictly implemented with a relational technology.

O-O DBMSs

Lastly there are the O-O DBMSs. It is out of doubt that there is a significant trend towards the extensions of the object-orientation concepts to database systems. Though receiving a growing interest, O-O DBMSs are still rare and there is a persisting sort of "reluctancy" towards their introduction, due to a consolidated presence of the traditional (mainly relational) technology. However this is not a novelty, since what is "new" has always to overcome some difficulties related to the old user's habits.

This sort of resistance is also true for current, commercial and claimed "object-oriented" GISs, which still use the relational DBMSs for storage (i.e. Smallworld GIS). The use of an O-O DBMS, selectively for a few GIS functionalities, could represent one of the expected novelties of GEO2DIS.

In the literature it is claimed that GISs based on an O-O DBMS can overcome the single-user limit of "hybrid" systems. In such systems in fact, being the geometric part implemented directly over the file system, rarely any sort of access control, record locking, etc., is implemented. However the ArcStorm and ArcView2 experiences, just to make a few examples, seem to show that "traditional" Industry is moving.

An hypothetical GIS, based on a O-O DBMS, should be seen as "integrated", according to the distinction which was made previously (i.e. uniform and coherent storage of spatial and not spatial data in the same DBMS), but without having the inconveniences of the "integrated" solutions with relational DBMSs. As such, however, in case that a Customer already has a commercial DBMS containing possible attributes, the O-O DBMS solution would inherit the nasty problem of either loading (and duplicating) such data or interfacing the DBMS, by building links with its tables. In this case, however, it is the O-O world to move, where an SQL access is declared to be under development.

8.2.1.2.6 Data Exchange

A comprehensive discussion finalized to the choice of the adopted exchange format(s), further to a survey of the existing commercial, international or de-facto standards, is part of WP-131 objectives, and it is discarded here.

Some relevant European and US initiatives, concerning the geographic information standardization, are described in the following. These simply confirm the large importance that the geographic information has nowadays, and will have in the future, in many human activities.

In Europe there are nowadays several geographic exchange formats, adopted in various Countries. For instance in UK there is Ntf (National transfer format), in Germany Atkis, in The Netherlands Nen 1878, in Denmark Dsfl, in Norway Sost, in Italy Ntf (Cadastral Office) and Digest (military) and in France Edigeo (derived from Digest). Digest was conceived by a set of NATO countries.

It is worth mentioning that the CEC has charged the DG XIII (Telecommunications, Information Industry and Innovation) to create, within the CEN (Committee for European Normalization), a specific Technical Committee for coordinating the cartographic standardization. On 1991 the CEN/TC "Geographic Information" 287 is established. This begun to work on mid 1992, once National delegates had been nominated.

On June 1992, in Paris, the working policy of the Technical Committe is defined: critical and comparative analysis of what already exists, in terms of standards and/or formats, either "de jure" or "de facto", and avoiding to conceive a new specific one (or a family of). Four working groups are then formally created:

- WG1: General structure - reference model for data;
- WG2: Conceptual model - data description languages;
- WG3: Geographic Information Transfer and Exchange;
- WG4: Reference models - geo-referencing;

Being the geographic standardization a so critical matter, the Technical Committee decided, on December 1992 in Brussels, to establish official "liasons" with other authoritative organization, such as:

- Cerco (Comite' Europeen des Responsables de la Cartographie Officielle) ;
- Dgwig (Digital Geographic Information Working Group) ;
- I.h.o (International Hydrographic Organization).

Due dates for the definition of a preliminary european standard (pre-standard) are end of 1996 - mid 1997. Definite standard not earlier than 1998 -1999.

Therefore there are authoritative initiatives aiming at the definition of a comprehensive, Europe-wide conceptual data model (and related exchange format), having the level of detail of the elementary geographic objects (say, point, line and polygon), and being "standard" across different commercial GISs.

8.2.1.2.7 Conclusions

The study on the available tools had the objectives to highlight the challenges for present GISs, namely in the Client/Server approach, the O-O technology and the Data Exchange issue. As a first conclusion of this study, it was confirmed that the above topics are really up-to-date and quite "hot" in the GIS world. This successfully confirmed the general scenario and trends of GISs, as initially outlined in the Technical Annex. Therefore the focus of the project is still a leading-edge one, as it must be for an ESPRIT III project.

The study tried to evaluate the "distance" between expectations and actuals in such topics, taking into account both the current commercial GISs as well as the prototype products.

The report output by the study served as valuable input to analysts and designers in GEO2DIS, in order to make the appropriate choice for the project itself, namely:

- the most opportune/effective tasks, nowadays and looking at a 30-months time frame, of a geographical, O-O, and heterogeneous Data Server;

- the most effective exploitation of O-O technologies, in terms of Data Model, Programming Language and Data Base;
- the most opportune exchange format, since no comprehensive and universally-accepted format is nowadays available, though many authoritative initiatives are in progress;

8.2.2 Evolution and trends

8.2.2.1 Geographical data formats

From the beginning of the GEo2dis project, the most important events in the field of the geographical data formats is the release of several standards, addressing either metadata (e.g. US FGDC and European CEN TC/287), or spatial query languages (e.g. spatial extension of SQL3-MM), or common GIS API interfaces (e.g. those under definition by Open GIS), and transfer formats (e.g. SDTS). Those have been considered for compliance and impact, but not implemented in the system. An suggested improvement of GEo2dis is the compliance with the European CEN TC/287.

8.2.2.2 Tools

During the three years of the project, there was no dramatic change in the field of the GIS. The main products remain, with continuous improvements with the releases. Main improvements are related to the MMI, the integration of raster data and the response time.

There is two important and recent events :

- the recent appearance of MAP OBJECT from ESRI, which is a library of OLE components offering basically the services of the Arc-View product with an OLE interface. This product is by definition « restricted » to the PC world and is aimed to the visualisation and not to the management of geographical data.
- the ESRI has also developed and markets jointly with the ORACLE company since the end of 1995 the Spatial Database Engine (SDE) which permit to manage, store and retrieve spatial data within an ORACLE 7 DBMS. SDE offers an API which permit to develop applications in C, C++ or Visual Basic. SDE remains based on the relational model.

8.3 Technical description

8.3.1 Services

The project has designed and implemented a system based on two types of stations: The server and the client which can be connected by LAN and WAN. On the server station, which hosts the geographical DB system, the data are prepared for distribution, the information which fully describes data is organised and managed in a catalogue, and communication with the client station is maintained active. The client has to allow users to submit queries to the catalogue and to define and to issue a request for data to receive them.

The combined use of the server and the client stations enables a generic user to fulfil the suitable functions to extract data from the geographical DB system, mainly :

- to identify the data needed for a specific application,
- to select a specific layer from available data,
- to see a generic sample of selected data,
- to elect a subset of data related to a specific geographical area (e.g. a circular area around a point of interest);
- to select a subset of geographical elements from those available (e.g. select only the main roads from all roads);
- to select a subset of attributes from those available;
- to preview selected data;
- to transform the geographical projection and the reference system;
- to convert the data into a required transfer data format;
- to activate the data delivery from the server station into his own one.

GEO^{2}DIS manages structured and non structured vector data, tabular data to be associated with vector data, raster data, maps, photos and free texts that are directly and indirectly linked to geographic entities.

Special attention is given to the user interface.

The interaction techniques, the interaction language and the degree of detail of the information in the catalogue are presented to the user in a customised form, according to the skill of the user.

Fundamental elements in the user interface are the tools, offered to the user to select the desired data. The system presents to the user a representation of the territory (reference background), so that the user can define the area to operate on through the identification of the main reference geographical entities.

The Reference Background is managed within a logic of a geographical continuum and contains physical data (rives, lakes, coastlines, etc.),infrastructural (roads, railways, urban areas, etc.) administrative (administrative borders etc.) and a reduced toponymy.

The user can select entities of any type on the Reference Background or by choosing one or more existing entities (a river, a certain number of contiguous administrative entities, etc.) or by defining new entities (a polygon, a line, a point...). Such entities are used to select geographically the data to be extracted from the server on the basis of geographic operators: If the entity is an area, the operation may be: The data within that area. If the entity is a line, the operation may be: The data that intersect that line. The point type and line type entities can be transformed into areas by buffering, thus allowing a vast number of selection criteria.

An important feature is that the server can create an off-line support (e.g. a CD-ROM) containing the catalogue and a query software; this support will grow up the number of those who are aware of the data available, as it does not require any kind of telematic connection, but only a PC.

Services offered to integrated systems by the reusable components developed in the frame of the GEo^{2}dis project are basic operations on geographical objects (distortion, inclusion, intersection), geographical co-ordinates translation, C++ interface for vector data and raster data, layering of vector and raster data, geographical objects persistence, GeOQL requests and the spatial indexing module (M4).

8.3.2 System architecture

The GEO^{2}DIS system architecture is based on a client-server structure. It is shown in **fig.1.** Many clients can be connected at the same time to the same server part.

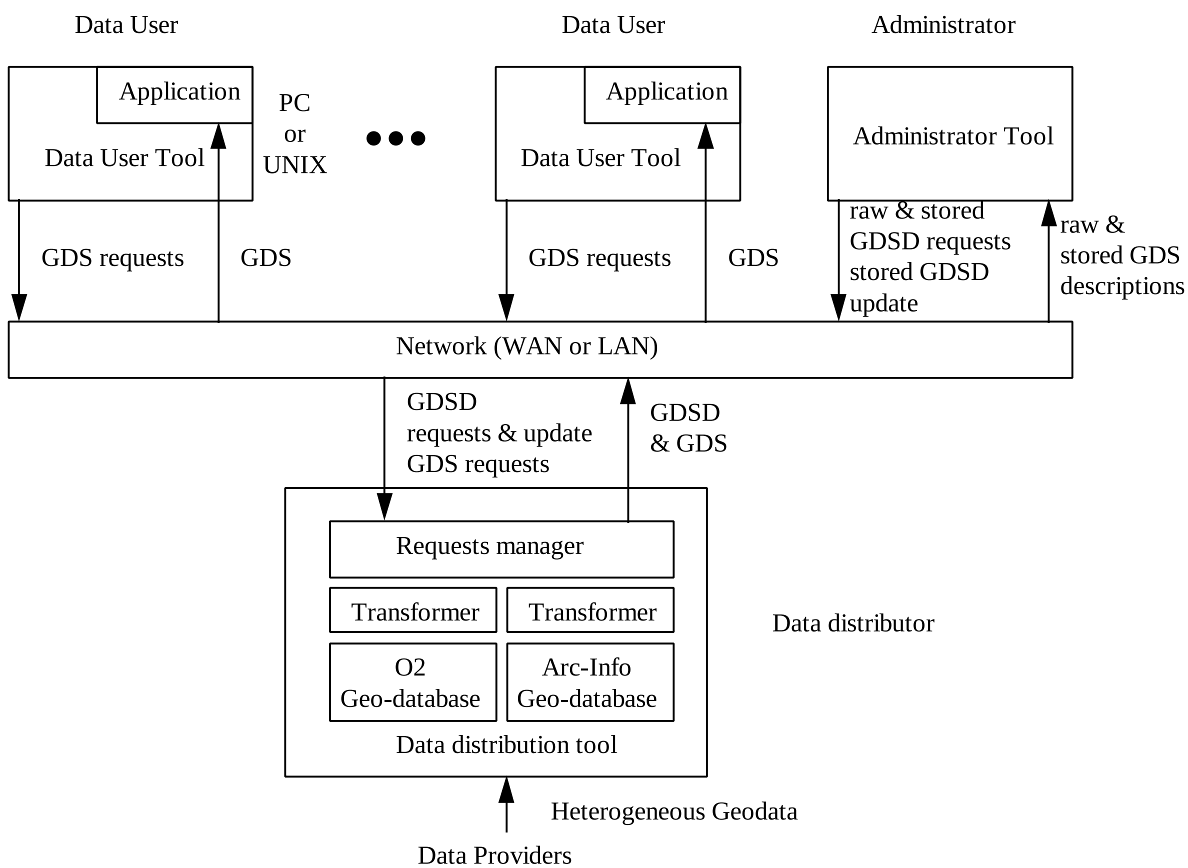


Fig. 1: GEO2DIS System architecture

The client of the geographical data server allows the data user to build requests of Geographical Data Sets (GDS) by selecting wanted GDS in a GDS catalogue. The catalogue purpose is the presentation of GDS descriptions. GDS requests contain filters on geographical objects contained in GDS, allowing data user to retrieve subsets of GDS tailored to his needs.

An extended version of the client part will allow the data administrator to manage data and catalogue.

The server part will send back GDS subsets corresponding to the in a format readable by he applications, among those supported by the geographical data server (IGES, EDIGEO, ERDAS, etc.)

The server part is already implemented on two software products: The O2 OODBMS and Arc-Info. Both implementations have a common part, the "request manager", which manages the request stack, redirects requests to the geo data base and sends back GDS to the right client part.

A common generic GDS description data model has been defined, which is an object class.

For GDS (and GDS descriptions) requests, an object oriented geographical request language has been developed. This is GeOQL ("Geographical Object Query Language"). This language is a spatial extension of the OQL ("Object Query Language"), which is part of the emerging ODMG 93 standard.

A spatial indexing module integrated in the core of the O2 OODBMS has been designed and implemented, in order to improve request response times. Those are to long in current systems, implemented in hierarchical or relational data bases.

C++ geographical object libraries were developed for the implementation of the server on the O2 OODBMS.

8.4 Technological breakthroughs

In order to build the flexible architecture previously described and reach the performance objectives, some technological advances have been performed.

8.4.1 Generic data model

8.4.1.1 Introduction

The generic data model provides a generic way to describe the structure of each geo-data set which could be exchanged by the means of the GEO2DIS server. The geo-data sets are stored in heterogeneous Geographical Data Banks (geo-DBs), thus they are organised with different data models, but they will be presented to the users in an homogeneous and object oriented way.

Different proposals have been made about which was the most opportune abstraction level at which to model the heterogeneous data stored in different geo-DBs. It has been decided to model data at an high abstraction level, that is to take into account the geo-data sets, regardless of how the single geo-objects are represented inside the geo-data sets. This decision has been taken on the basis of two consideration:

- the DU will access GEO2DIS in order to purchase geo-data sets (or subsets of them) and not to purchase the single geo-object representation. Even if the information layer is composed of geo-object representations (e.g. the hydrography layer is composed of the rivers representation), in GEO2DIS this is not relevant, because the DUs want to import the hydrography, even though restricted to a particular zone, and not a single river).

- to model at a lower abstraction level means to provide a common model for all the geo-data in GEO2DIS. This is a task of the same complexity of defining a new interchange format between the involved GISs and it risks overlapping the work of some organisations created for this purpose.

With this choice the generic data model mainly becomes the frame for the GEO2DIS catalogue data model.

DU in GEO2DIS will be provided with a catalogue which contains information on geographical data sets. By browsing the catalogue the DUs can identify which geo-data sets are interesting and decide to purchase them (or part of them). A request for these geo-data sets is issued to the server, and geo-data sets are received back by the DU.

Then the DUs can use the geo-data set with their own GISs, and they can access the fine-grain geo-data within the geo-data set as they wish.

Via the catalogue the users will choose geo-data sets which they want to purchase and to import in their local Data Banks. Thus users need as much information as possible about the geo-data sets structure and content to know if the geo-data sets they want to import are really interesting for them.

Moreover GEO2DIS must support the users when ordering the geo-data sets, that is to prepare a request for the geo-data. This request depends on the identification of the geo-data set (e.g. the name of the geo-data set) and further information to know where the geo-data set is stored, in which geo-DB, and so on.

Thus the generic data model has threefold scope:

- to provide the users with an uniform object oriented presentation of the geo-data sets handled by GEO2DIS,
- to provide the more exhaustive description of the geo-data sets structure and content,
- to provide the information about how to access the geo-data sets.

8.4.1.2 Rationale

GEO2DIS is a server working with GISs, and GISs work with geo-data, strictly speaking with geo-data-sets. The GEO2DIS will use geo-data-set-descriptions, that is, data which describes the geo-data resident in the GISs. This is often referred to as metadata.

GEO2DIS models the geo-data-sets with a generic data model.

A GIS is used to manipulate numeric abstractions of real world entities. The real world entities are the ones we are walking on (the earth), driving on (roads), living in (buildings), or crossing (rivers). The numeric abstractions are built according to a geographic model. Each GIS or data source has its own models, one for each real world entity. These numeric abstractions are called geo-data in the following. For an end-user, the only way to access these geo-data is to use the corresponding GIS or data access software, as these alone know how the geo-data are stored, organised and indexed.

The scope of GEO2DIS is not, of course, to become another GIS or data source. The GEO2DIS is a data server, and must give the user access both to the geo-data-sets and to its meta-information, through a common interface, whichever GIS or data source is providing the geo-data.

This does not mean that GEO2DIS has to define its own geo-data by doing type intersection or type union. For example, if roads are modelled with 4 attributes (width, name, category, and quality state) and a line by GIS A, and with 2 attributes (number of lanes, name) and a double line (a set of parallel lines) by GIS B, GEO2DIS should not define a type "road" that would cover these two types (for example 5 attributes : width, name, use category, quality state and number of lanes). This kind of solution does not allow access to other, different data sources. It would be easy to imagine a geo-data type which would not fit into this "super" type. Also, there are European initiatives and agencies like CEN, already trying to define such "complete super" types.

The generic data model used by GEO2DIS is not intended as a conversion bridge between GISs or data sources; it is not an exchange format. The generic data model has to allow a unified view of the geo-data information. The purpose of the GEO2DIS server is to allow a user or an application to access the information contained in these two kinds of geo-data-sets (roads A and roads B), and their associated meta-information, through the same object-oriented interface. This means that the GEO2DIS server provides objects that correspond to the geo-data-sets (an object for roads A geo-data-set, an object for roads B geo-data).

It is possible to provide a uniform object-oriented interface for most (or all) geo-data-sets from differing GISs or data sources because they are all models of geographic entities. So they have a descriptive part (optional) and a geometric or geographical part (must be present). The geographical part can be used with a common interface, because it has specific data and behaviour (for example, the co-ordinates). The purpose of this document is to specify a generic data model, which has the structure and behaviour needed to access all the different types of geo-data-sets. This structure also enables storage of meta-information about the geo-data-sets.

8.4.1.3 What is geo-data?

In the GIS world and among the cartographic community there are many definitions; each system has its own geo-data types (a road is not modelled the same way by Arc-Info and by DCW). A very generic definition could be that geo-data is the association of the description of a phenomenon and its location. Roughly, it is a relation between something and somewhere. So geo-data has two parts:

8.4.1.3.1 Descriptive Part

The descriptive part of geo-data can also be defined as the semantic or textual information. It is optional, and is usually achieved by a list of attributes and their values. For example, to describe cities, one can use attributes name, population, mayor, number of districts and so on. It is clear that no two geo-data from different producers, even when representing the same physical phenomenon, have the same list of attributes. Even when attributes have the same name, they may not represent the same physical thing. The list of all possible types of attribute value has to be known beforehand.

8.4.1.3.2 Geographical Part

The geographical part is sometimes referred to as the location of the phenomenon or its geometric part. The location is usually made of a list of co-ordinates (along with the system in which they have been calculated, called the referential) and a shape or form in which they have to be considered (point, line or polygon). Usually the geographic object is mingled with its location. A geo-object can have any geometric shape (at least one of the more common in the geographic or cartographic community) and has the basic geometric behaviour of these concepts: intersection with a rectangular region, change of co-ordinate system (called a referential change). The complete behaviour has to be determined along with the query language allowed on geo-objects. The GEO2DIS server will not provide specific geographic actions like "extract roads between Paris and Lyon", or spatial joins (query like "extract all cities close to a forest"), but it has to provide fast and easy access to the geo-objects through spatial query (query like "extract all objects intersecting this rectangular region").

8.4.1.3.3 Use of geo-data

Geo-data can be used and manipulated in several ways:

- As whole objects when performing operations common to all objects (store, copy, create, delete ...)

- Through their descriptive parts only (statistical analysis).

- Through their geometric shape only (display).

Through relations (how many cities in this rectangular region?). This part is really specific to geographic objects. This kind of question is called a geographic (or spatial) query, because the first criterion of selection of objects is location. In general, they are the queries allowable on spatial indexes.

8.4.1.3.4 Geo-data-sets

GEO2DIS is primarily concerned with the large-grain data, the geo-data-sets.

For many data sources, the geo-data are manipulated through geo-data-sets. The fine-grain geo-data is structured into large-grain geo-data-sets according to the organisation of the GIS in which they reside : Arc-Info has coverages, IGN EDIGEO has packages, VPF DIGEST has libraries, TerraLogic has layers. The underlying concept is expressed by the geo-data-set. A geo-data-set is a homogeneous composite object: "homogeneous" meaning sharing some common characteristics, in the sense defined by the data source. For example, in VPF, a library can be represented as a geo-data-set: all objects inside share the same referential, scale, and some meta-information. The geo-data-set is the minimum grain of information which is visible from the server. However, the server can ask the GISs to extract from a geo-data-set a sub-set (which is also a geo-data-set).

8.4.1.4 The geo-data-set description

This is an object-oriented description of the geo-data-sets, as defined above.

The server must handle geo-data-sets from various sources or producers. Thus the administrator must be provided with a way of describing the geo-data-sets, so that the server can use and manipulate the data through a common interface. The generic data model has to be as flexible as possible, to allow most existing geo-databases to use the server.

The Figure 8-5 below shows a simplified representation of the structure of data in GEO2DIS. The geo-data-set-description can be divided into three sections: global description, geography or geometrical description, and content description.

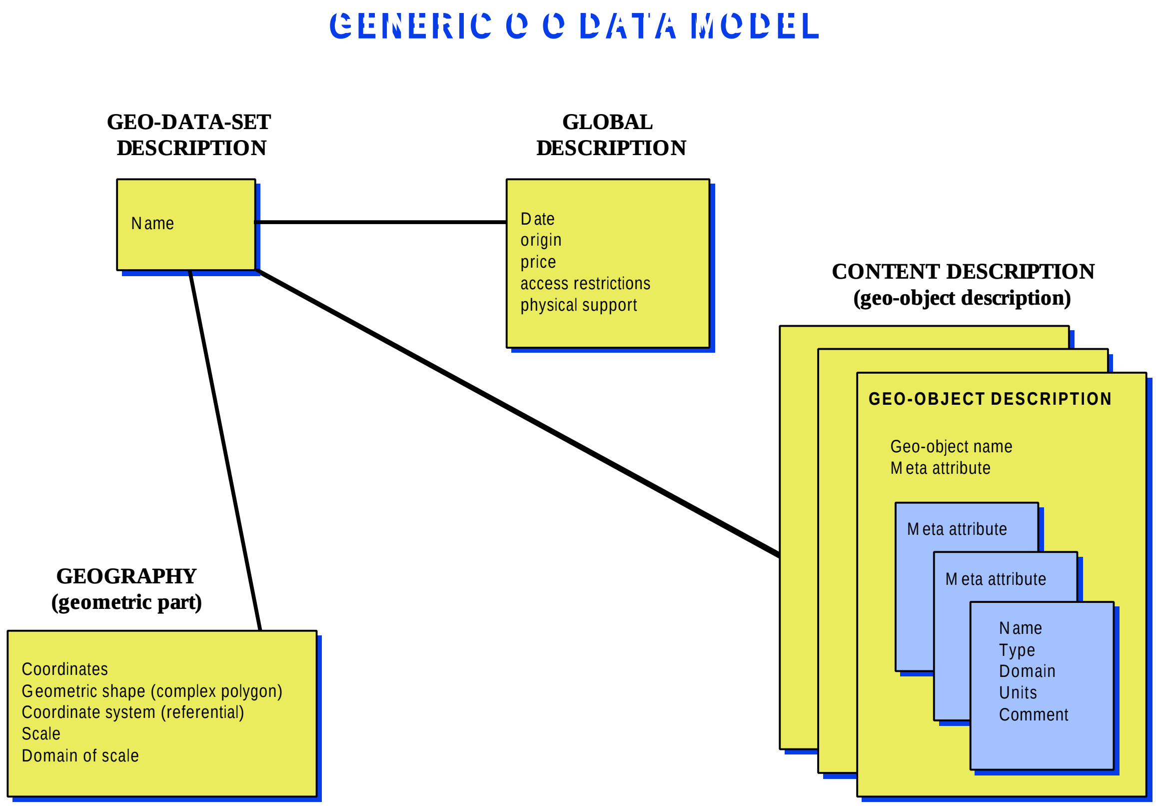


Figure 8-5 : Generic OO data model

8.4.1.4.1 Global Description

The global description covers general non-technical characteristics of the geo-data-set considered as a whole, like the producer (origin), date of production, date of entry in database, scale, format, price, quality, access restrictions, and the physical support. The physical support contains at least the GIS name and the node name, in general it contains all information needed to access the geo-data-sets required.

8.4.1.4.2 Geography

This is the geometric part; everything about the location. It includes the co-ordinates, the system in which they have been calculated (called a referential, for example. UTM, WGS84, Lambert, ...), the shape of the geo-data-set (in this case always a complex polygon), and the domain of scale (usually geo-data-sets are produced from maps, which have a scale. Since numeric representations allow any scale, it may not be significant to use some geo-data-sets above or below certain scales).

The shape is here defined as a complex polygon, and thus covers most possible shapes, including extension and intension shapes. Other shapes can be added later, if it becomes necessary.

GDSs handled by GEO2DIS server may be three dimensional, but GDS descriptions shall be two-dimensional, as their geography is the geographical domain of the GDS.

8.4.1.4.3 Content Description

This is a description of the contents of the geo-data-sets. The content description or geo-object description contains a list of meta attributes for each geo-object described. A meta-attribute has four attributes: name, type, domain, and units.

The meta-attributes are defined more precisely in Work Package T160, Catalogue Contents. The attributes and operations of the object classes are described briefly in Appendix A. (see also Glossary.)

8.4.1.4.4 Composite objects

A geo-data-set-description can be composite, that is, made up of several other geo-data-set-descriptions; for example a hydrography layer can be considered as being composed of many parts. Of course, this may make some parts of a composite geo-data-set-description meaningless; the referential of a composite geo-data-set-description does not exist, unless all parts share the same one. Thus when the result of a query is a composite geo-data-set whose components have different referentials, GEO2DIS must transform some components so they all have the same referential.

8.4.2 Geographical data classes

The Geographical data base on O2 has been implemented as C++ classes. This approach ease the extension of the server to new formats, using the inheritance mechanisms. It conforms to the generic data model described in the previous paragraph.

8.4.2.1 Class Geo-Data Set Description

The Geo-Data Set Description class is defined in the previous paragraph. It Describes a set of logically grouped Geo-Data.

8.4.2.2 Class Geo-Data

The Geo-Data class is the base class for all geographical data (that is geo-referenced data). It provides with a uniform interface for raster and vector data. It is composed of a Description and a GeoGraphy. The Description stores the semantic information (that is attribute values). The Geography class stores the location (Shape class), visual information (Palette Class), scale (Scale Class), layer (Layer class), and geographical datum (Referential class).

A Geo-Data can be composite (that is made of several others Geo-Data).Composition of Geo Data is implemented by the Composite Geo-Data Class.

Geo-Data can be related to each other : the Geo-Data Relation class stores all pertinent information.

Each Geo-Data knows its Geo-Data Set Description. Some of the Geo-Data parts (like Referential, Layer, Scale) can be factored in their Geo-Data Set Description.

The Geo-Data class has two attributes, the type and the Bounding Box.

The type is the name of the type of data which the GeoData is representing : this name is chosen by the administrator importing the data source.

The Bounding Box is the minimum enclosing rectangular box of the Geo-Data, expressed in the same referential as the Geo-Data. It is made of two 3D co-ordinates: the minimum point and the maximum point (its edges are parallel to the axis of the referential in which it is calculated).

The storage of semantic information is implemented through AVAs. The AVA class is virtual: the method value has to be redefined by the subclasses. There could be as many subclasses as atomic types.

8.4.2.3 Class Vector data

The Vector Geo-Data class is the base class of all vector data : all subclasses will inherit its interface and data. This class is able to create instances of subclasses using only their names. All Vector data have to inherit from it.

The Attribute Classname is the name of the Class, which allow the base class Vector Geo-Data to create instances of sub-classes using only the name of the Subclass. All SubClasses have to have their own name (Vector Geo-Data verifies consistency of new names).

The storage of the co-ordinates is assured by the various sub-classes of Shape They are in charge of implementing the geometric methods: intersect, contain, deform and Bounding Box. A very important point for performance is the persistent storage of co-ordinate. A Co-ordinate is made of three float (or double). Since they are always manipulated by groups, there is no need to have a persistent Co-ordinate class. Instead there is a CoordinateHeap class which implement the storage of and access to a group of co-ordinates, as well as efficient copy and browse methods. All Shapes (that is geometric primitives) like Point, Line, 2DSurface, 3DSurface and Volume use this CoordinateHeap for inner storage of their co-ordinates.

The choice between double and float for the co-ordinate is a compromise between the available amount of disk space and the precision required by the Referential in which the co-ordinates are calculated.

In most GIS or Geographic data-bank, the topology is stored and implemented through physical sharing of geometric primitives. Here the Topology is stored through relationship between geometric primitives: Node, Edge, Winged Edge, 2DFace. This does not prevent to share primitives between geo-data: it depends on the organisation of the imported data; for example, in DCW, the topology is implemented through physical sharing: it is simpler to have the corresponding Geo-Data classes share their primitives (it facilitates the loader development).

For each imported data source, a loader is developed: this software is in charge of reading the data source in its organisation, create instance of the corresponding Geo2dis type and to fill them with the information in the appropriate order. The Loader uses the information given by the administrator to select which data should be read and where to copy it. The loader will use the class method create_instance of Vector Geo-Data to get instances of Geo2dis objects. The loader has knowledge of the external data source structure and knowledge of the interface of Geo2dis objects and server.

The Geo-Data may have to be written into a specific format. Instead of having each class to know every needed format, a writer is developed for each format. This piece of software both know the Geo-Data structure and the organisation of the data in the desired format: it is the inverse of the Loader.

8.4.2.3.1 VPF Classes

The Attributes of VPF classes correspond to the structure of VPF elements. They are not stored using AVAs, because they are specific to VPF: a query on these attributes will apply only to VPF classes.

The VpfDictionnary is a private VPF class that allow the transcription of codes into their string description. It is as close as possible to the original data structure of VPF database.

The method fill_specific_info has to be redefined by the subclasses of a particular VPF database (for example DCW). These five classes are the basis that manages all information common to all VPF database. Subclasses of these have to be created to manage information specific to a VPF database (see below).

8.4.2.3.2 Classes for raster data

The storage of and the access to the pixels is implemented by the RFile (RasterFile) template class. Given the name of the file and the type of the pixels (unsigned char, short, float, complex,...) the Raster Geo-Data can create an instance of the corresponding type. The method read of the RFile fills a memory area with the radiometries of the pixels, taking as input a rectangular area, expressed in pixels. The RFile class takes care of resampling if needed, through the method deform.

The attributes of the classes for raster data are the name, the width and the height. The name is the name of the image if any. The width and height are the dimensions of the image in pixels (integers).

8.4.2.3.3 3D Object

This is a special subclass of Shape which stores information for CAD objects that will be stored in the server. For more details see RD31.

8.4.2.3.4 Connection with O2

Since O2 has a C++ interface, the defined C++ classes constitute the schema of the O2 database. The idea is to make persistent the objects that have to be stored by the server. The chosen classes will then be imported into O2 using the import tool of O2. This tool will generate equivalent persistent C++ classes that will be able to persist in the O2 database. For the persistent classes, the methods that have to be triggered from the query language also have to be imported (see chapter 4 of RD6).GeOQL language

8.4.3 The GeOQL query language

The GeOQL language allows to make the selection of the Geographical data according to the generic OO data model.

This language is an extension of the OQL language, part of the ODMG 93 standard.

A GeOQL query allows to formulate a geo-data-set request, under geographical and attribute conditions. This geo-data-set is delivered inside a geo-data-file, into a required format and in a required referential. Below is an example of query showing some of the capabilities of the language :

define interest area **as** "RegioneToscana";

define geodataset **as flatten**(

select gds._set **from** gds **in** GeoDataSets

where gds._name = "DfadItaly");

clip

select (GeoData)area **from** area in DfadAreas

where geo_intersect (area, interestArea)

and area **in** geodataset

and area._tree > 3 **and** area._roof < 5

on interestArea

project Mercator

content DfadArea(_roof)

The elements of this <geOQL> request mean the following , words in bold characters being key words :

« define interest area **as** "RegioneToscana" » :

this element is self explanatory

« define geodataset **as flatten**(

select gds._set **from** gds **in** GeoDataSets

where gds._name = "DfadItaly"); » :

this element specify the name of the GDS from which the requested objects must be extracted.

« clip

...

on interestArea »

this element makes the geometrical shape of the selected objects clipped on the area of interest.

select (GeoData)area **from** area in DfadAreas

this element specify that the data have to be searched in the set of the objects of the type « DfadArea ».

« **where** geo_intersect (area, interestArea) »

this element specify that the selected objects must intersect the area of interest.

and area **in** geodataset

this element specify that the selected objects must belong to the specified geodataset.

and area._tree > 3 **and** area._roof < 5

this element specify that the selected objects must have a value of the attribute « tree » > 3 and a value of the attribute« roof » < 5.

« project Mercator » :

this element specify the required referential : - Mercator

content DfadArea(_surface_type) :

this element specify that the values of the attribute « surface_type » for the selected objects must be extracted, ordered and written in the result file.

8.4.4 Spatial Indexing

The Spatial Index Module (SIM in the following) is a specific module, integrated in the O2 OODBMS.

It manages spatio temporal indices based on an R Quadtree algorithm (Figure 8-6). Spatio temporal indices deal with time as a third or fourth dimension and optimise response time of spatio temporal queries which are the most frequently used in the geographical field

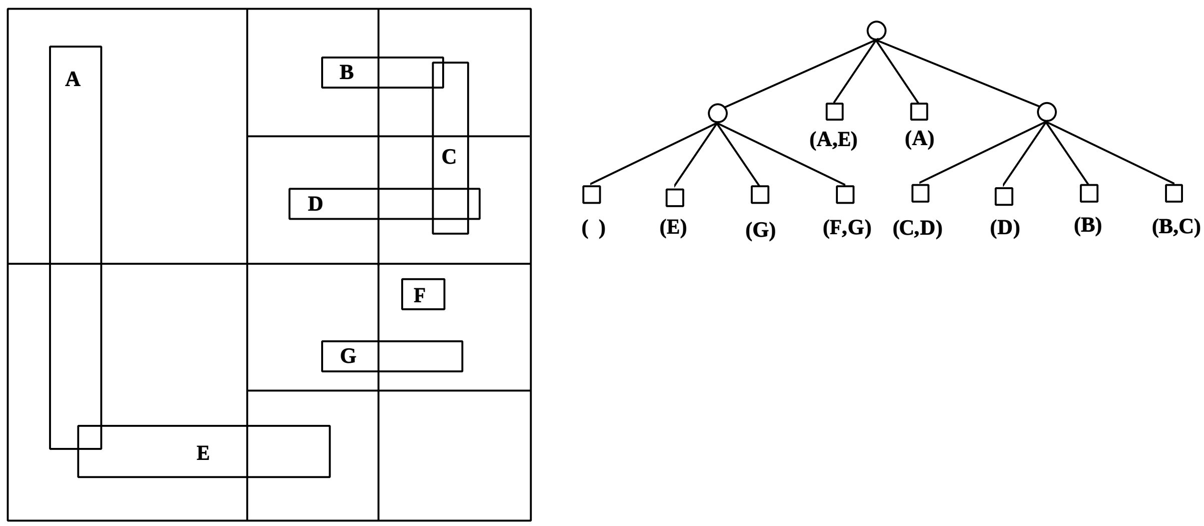


Figure 8-6 : R Quad-Tree algorithm

The R Quad Tree algorithm consists in recursively dividing the n dimensional space in m^n equal cells (n = m = 2 for the example shown in the figure). The cells are organised in a tree, each cell being the daughter of the cell which contains spatially it.

The R Quad Tree is build in dividing a cell each time that the number of objects of which the spatial coverage intersects or is contained in it exceed a threshold (3 for the figure, up to 100 or more for an operational system.

Each object is linked to each leaf cell which intersects or contains it.

The SIM was developed by MATRA Systèmes & Information and O2 Technology in |co-operation in the course of the overall GEO2DIS development. It is integrated in the low level layers of the O2 OODBMS. This solution relies on the virtue of the O2 OODBMS open architecture and allows maximal performances.

Among the main requirements for the system, the SIM yields significant advantages in performance, flexibility and ease of use.

A performance improvement of one magnitude of time is achieved for the retrieval of heterogeneous geographical data.

The dimension of the space of indexation and the type of keys (length in bits) can be defined by the user.

The SIM is a C++ library with a well defined and easy-to-use API.

As a conclusion, the user of the O2 OODBMS fitted with the SIM extension will take advantage of a powerful OO data model and of outstanding response time for geographical application development

8.4.5 O2 Engine for spatial databases

8.4.5.1 Goals and principles

The Geo2dis geographical data server basically consists of three major components :

- a query interpreter for a geographical query language,
- a database engine which supports the logical representation of objects defined by a geographical data,
- an advanced mechanism for optimising the evaluation of queries.

The request manager of the Geo2dis server is the front end for user access to the geographical data. It to analyse and perform queries expressed in the GeOQL language. The GeOQL language is based on query language, and extends OQL by defining several geographical operators for building predicates location.

The database engine supports the data model defined for the Geo2dis geographical data server. It provides a persistent and recoverable storage for geographical information and a concurrent access to the geographical data.

The goal of O2 Technology in the Geo2dis server kernel is to integrate an advanced mechanism for optimising geographical query evaluation in its database engine.

The technique used for optimising geographical queries is to provide a spatial index that can be collection of geo-referenced objects.

Due to the richness of the object-oriented data model, it is possible to implement a spatial index as a component of the database schema. However, the drawback of such approach is that one lose the physical data since indexed data should be explicitly identified by the user.

Furthermore, when a spatial index is manipulated as a database object, it is quiet?? difficult to control activity required to access the index node. Since the number of page faults during an index traversal is a performance factor, we believe a better solution is to integrate a spatial index as a storage structure of the which has the capability to manage the buffering of disk pages, and the location of index nodes in pages.

The solution for integrating a spatial index in the O2 database system consists in extending the current index mechanism of the O2 Engine, in order to transparently manipulate spatial indexes as other existing indexes.

The spatial index can then be seen as an independent module. The constraint on the O2 database engine and on the spatial index module is only to have a common functional interface to interact.

8.4.5.2 Architecture of the O2 based Geo2dis server kernel

The architecture of the O_2 based Geo2dis server kernel described by the Figure 8-7 below, shows the different module which participate in the processing of geographical queries on the Geo2dis server.

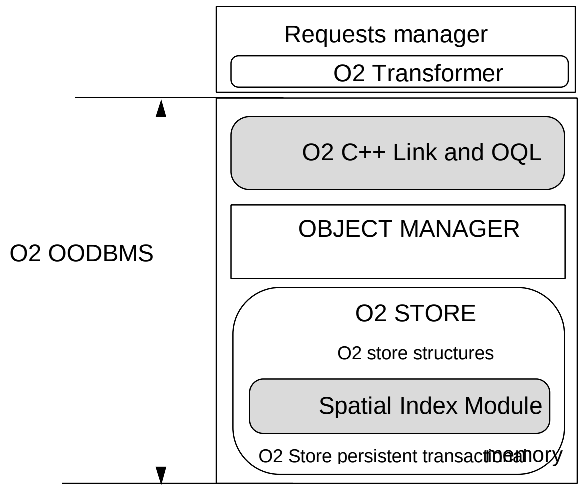


Figure 8-7 : Architecture of the O_2 based Geo2dis server kernel

The O2 transformer parses GeOQL queries in order to translate it in a pure OQL query. In it is in charge of the optimisation of queries by translating the geographical predicates into call to the spatial index modules.

The OQL interpreter embedded in the O_2 C++ Link is in charge of the evaluation of the query passed by the GeOQL interpreter.

As a object-oriented query language, it has the capability to perform regular SQL selections and ods calls. The method call capability is used to allow the call to the set of C++ functions of the O_2 transformer modules from OQL. It requires to be able to import in O_2, C++ member functions of persistent classes which will be called by the OQL interpreter.

The C++ functions of the O_2 transformer module have access to the object manager which is aware of the spatial indexes associated to the collections of objects. From the O_2 Engine API, the object manager transmits the operation on the spatial index through the O_2 store in the spatial index module.

8.4.5.3 Geo2dis spatial index definition

A spatial index is physically structured as a tree in the database system. But various techniques for the building of the tree are available, depending on which access method, pointing or windowing, is the more often used.

The choose of a spatial index algorithm was critical for the performance of the Geo2dis kernel query implementation and it was a very important feature that the spatial index module of the Geo2dis server works with several implementations of the spatial index in order to test the different algorithms, in particular the Rtree, R+tree, and NQuadTree. The R Quad Tree was eventually chosen (see the related paragraph).

Thus, the functional interface of the spatial index is independent from the algorithm chosen. Only the initialisation of a spatial index requires specific initialisation parameters to notify which implementation is used when the index is created. These parameters known at the application level can be passed by the O_2 Engine to the spatial index module as opaque information.

8.4.5.4 Integration of a spatial index module in the O_2 DB engine

8.4.5.4.1 Goals

The integration of a spatial index module in the O_2 database engine was guided by the three following goals:

- transparency,
- performance,
- independence.

8.4.5.4.1.1 Transparency

The spatial index is seen by users as other indexes used with the O_2 database engine. It is thus created and used from the O_2 Engine API index interface, callable from the C++ binding. However it cannot benefit from the agement mechanism provided by O2 Engine because the geodata bounding box which serves as the spatial index key is not only stored as an object attribute but can also be computed by an external function.

8.4.5.4.1.2 Performance

In order to achieve the requested performances, the spatial index module is integrated into the O2STORE index level. Thus, the physical representation of the spatial index is performed by using the data abstraction the O2STORE level, namely, files and pages, and the implementation can manage the placement of index pages and the residency of index pages on cache.

8.4.5.4.1.3 Independence

Since the O_2 database engine and the spatial index module are separate components plugged together, each component have to be as independent as possible from the other.

From the spatial index module, the O_2 database engine provides a persistent memory, and concurrent recoverable and transactional accesses. Thus, the implementation of the algorithms should be independent from the database aspects.

From the O_2 database engine, the spatial index module provides a common index interface which hides the geographical concepts. Thus, the picking and windowing functions should be translated into ordinary functions of the index module, and the structure of a bounding box should not appear in the definition of a spatial index key when it is transferred by the database engine to the index module.

8.4.5.4.2 Principles

The figure below describes the principle of the integration between the O_2 database engine and the spatial index module.

O2STORE provides a spatial index internal interface which is an extension of its current index interface. Furthermore, it gives access to its IO, buffer, locking and logging managers in order to support the integration of the persistent memory allocator of the spatial index module.

The spatial index module implements the spatial index interface published by O2STORE by calls to the spatial index functions in the intermediate layer called "o2store to spatial index interface". It is also a memory allocator which allows to allocate index nodes in the O2STORE pages. Thus, the implementation of the spatial index algorithms is independent of O2STORE and only depends on a common memory allocator interface.

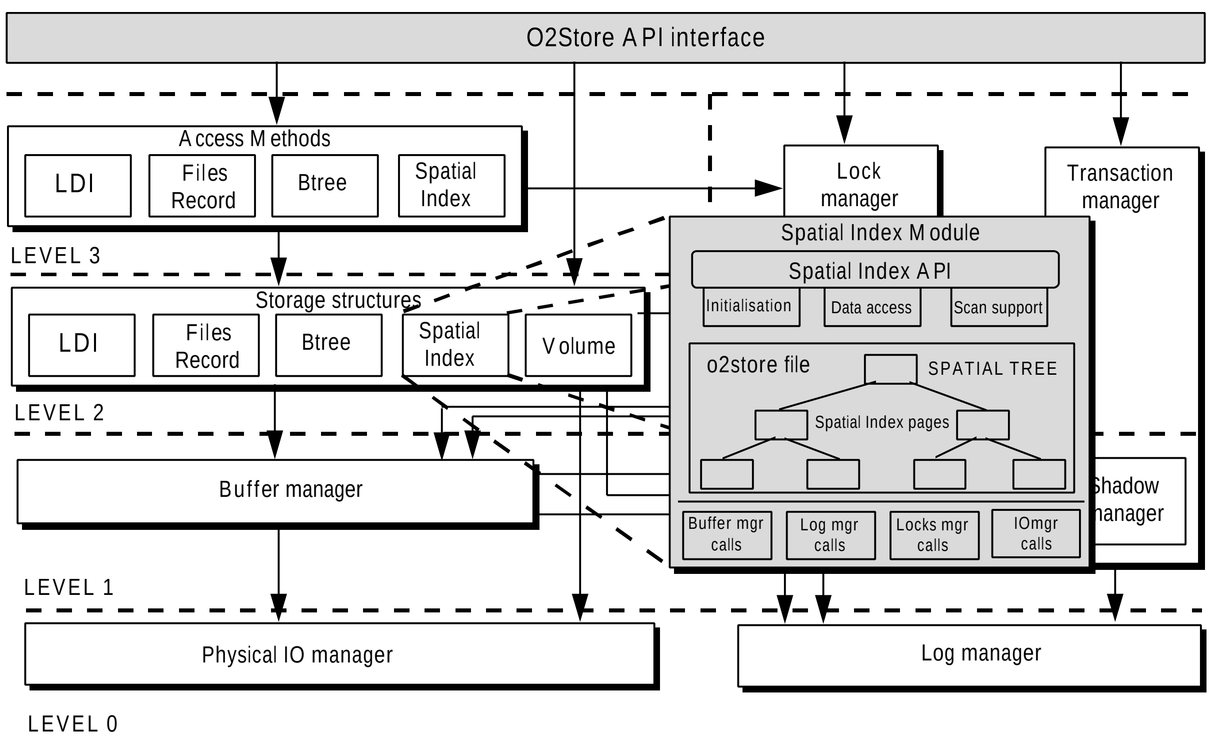


Figure 8-8 : Integration of the Spatial Index Module into the O_2 OODBMS

8.4.6 ARC-INFO Transformer

ARC-INFO, by ESRI, Inc. (US), is the leading GIS software in the world. It runs on any Unix platform and WindowsNT. For more details contact your local ESRI distributor. A comprehensive description can be seen at *http://www.esri.com/*.

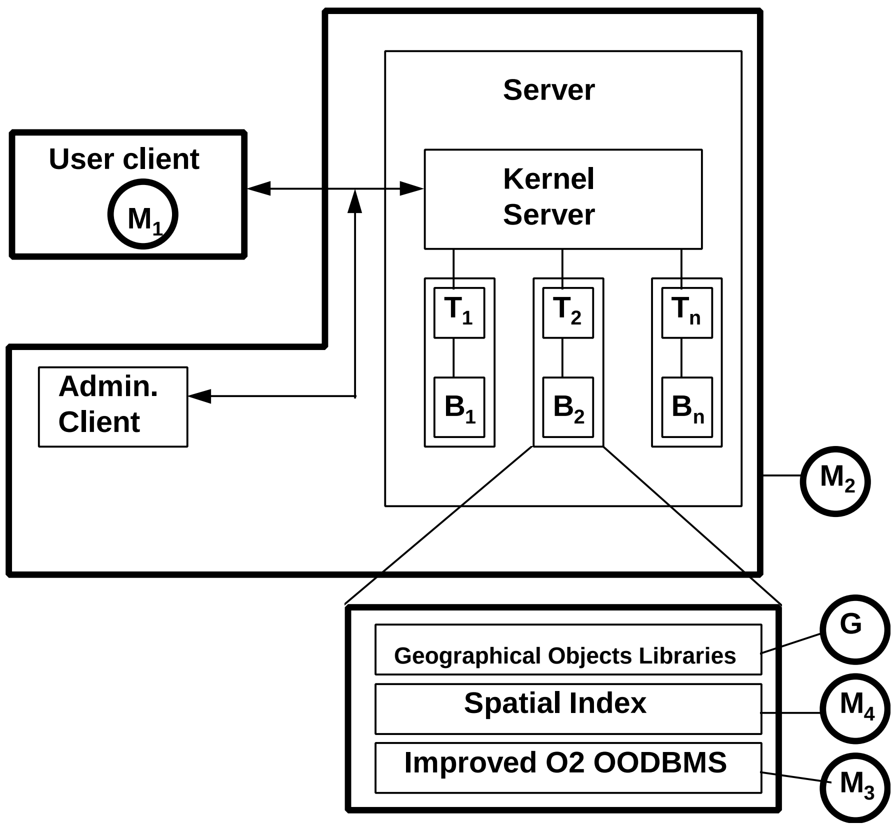
Of particular relevance is the architecture of the ARC-INFO transformer. In turn this has been realised with a **Client-Server** approach. This has been made possible by exploiting the **IAC** (**Inter-Application Communication**) functionality, available since Arc-Info release 7.0.3.

IAC enables software applications on local or remote machines to communicate with each other. In particular IAC provides a way for external applications to request services to an Arc-Info process when it is in "server mode". IAC makes it possible for Arc-Info to react to external events, such the receipt of a specific requests from other applications.

In GEO2DIS, therefore, Arc-Info will be operated in "server mode", and it is invoked only for processing an incoming request. This has the tangible advantage of preserving the Arc-Info licence taken for any other use (e.g. interactive session). In fact IAC has a queuing facility and does not deny a concurrent Arc-Info interactive session: Whenever this latter is idle or waiting user input, Arc-Info switches to "serve" possible queued requests and processes them until completion. After completion, control is returned to the user interactive session. Therefore the license is shared between the GEO2DIS Server and any other use.

8.5 Marketable products

GEO²DIS has been defined as a modularised system, made from marketable modules, based on the technical architecture previously defined. These modules have been defined according to the needs of the various potential customers. These products are shown below.



Legend: M = GEO²DIS Exploitable Modules

T = Transformers

B = Geographical databases (GIS or DBMS)

M1 = Data User Tool (client part for data users)

M2 = Data Distribution Tool (server part) +

Data Administration Tool (client part for data administrator)

M_3 = Improved O_2 DBMS (See the O_2 separate exploitation plan in AD5)

M_4 = Spatial Index Module (SIM)

G = Geographical Object libraries

That means for the market that :

1) **Data distributors** have to use a server (M_2), in order to provide the required data to interested customers. The server is composed of the following elements :

- A kernel server, which takes care of the data handling activities.
- Various translators (T_i), which handle the communication with the corresponding GIS or DBMS. Every GIS or DBMS (like Arc Info, Intergraph, TerraLogic. O_2, ORACLE; etc.) requires a specially designed translator.
- An administration client, which handles the internal server administration.

2) **Data users** have to use a user client (M_1) in order to have access to the corresponding data distributor servers.

3) **System integrators** will use the reusable components, that means the modules M_3, M_4 and G which are included into the server on one side and are marketed as stand alone products (See chapter 5 and 6 of AD5). M3, the improved O_2 OODBMS is described in the separate O_2 exploitation plan, included in AD5. M4, the spatial indexing module (SIM), is integrated in the OODBMS O2 for optimal performances (see Annex 1 for further description), The geographical object libraries G are :

- a C++ geographical objects library, providing basic methods on geographical objects : distortion, inclusion, intersection,
- a module for the translation of geographical co-ordinates, based on the GeOLIB library, providing translation between main geographical co-ordinates systems,
- a geographical objects libraries providing a C++ interface to data in the format VPF, DFAD or EDIGEO ; these libraries provide huge gains in productivity, because the complexity of the file structure of the geographical data is masked to the developers ; these libraries provide the same basic methods as the geographical objects library,
- a geographical objects library providing a C++ interface to raster data in the format VIFF, TIFF, Sunraster or vlog ; this library provide all the operations on the pixels useful for the surveillance applications ; this library process all the pixel formats defined in these formats and multichannel images,
- a display and layering module for raster and vector data,
- a geographical objects persistence module, allowing the storage of geographical data in every OODBMS ODMG compliant ; an option of this module offers the GeOQL request facility ,
- a spatial indexing module (SIM), integrated in the OODBMS O2 for optimal performances (see Annex 1 for further description),

- an ATLAS application, which displays geographical objects and which was used as « browser » of the database, for testing purposes in the frame of the project and is used as MMI for the geographical applications embedded in the systems produced by MATRA SYSTEMES & INFORMATION.

8.6 Evaluation of the performances

8.6.1 Evaluation methodology

The principle of the validation of the system is to use it for providing geodata for three very different applications, covering a large range of the possible applications. The Geo database on O2 was validated using an geographical data visualisation tool previously implemented by MATRA SYSTEMES & INFORMATION and adapted to the Geo database on O2 during the project, ATLAS. The validation tests were conducted according to a validation plan.

8.6.1.1 Validation plan

The validation of the system was done according to a validation plan. The main features of this validation plan were as follow.

There will be eight test sets for the validation of the system. The names and main features of these tests are defined in the Table 1 below :

Name	**Aim**	**Responsibility**	**Begin-End**	**Site**
MMI	MMI validation	RTO	10/95-02/96	INTECS
O2 Performances	O2 Performances	MATRA SYSTEMES & INFORMATION	09/96	MATRA SYSTEMES & INFORMATION
CAD Interface	CAD Interface	MATRA SYSTEMES & INFORMATION	10/96	MATRA SYSTEMES & INFORMATION
Visual Impact	Visual Impact Application	MATRA SYSTEMES & INFORMATION	10/96	MATRA SYSTEMES & INFORMATION
AES	Land use monit. Applicat.	DSS	10/96	DORNIER
Urban Planning	Urban Planning Applicat.	RTO	10/96	INTECS
Interworking	Interworking	MATRA SYSTEMES & INFORMATION	10-11/96	MATRA SYSTEMES & INFORMATION
WAN	Validation on LAN	INT	10-11/96	RTO INTECS

Table 1 : Validation tests definition

The validation tests will be done on five sites, in order to share the workload of the validation between the partners and to optimise the travels of the end users. The names and main features of these demonstrators are defined in the table below :

Site	Validations	Demonstrations	Components	responsibility
MATRA SYSTEMES & INFORMATION	O2 Performances Interworking & data heterogeneity	Visual Impact appli. O2 Performances CAD Interface Interworking & data heterogeneity AES Application Urban plan Appli.	Client on PC Client on UNIX Server kernel GeODB on O2 GeODB on Arc-Info CAD Interface Visual impact Appli. AES Application Urban plan. Appli	MATRA SYSTEMES & INFORMATION
MATRA SYSTEMES & INFORMATION	Visual Impact appli. CAD Interface	none	Client on PC Server kernel GeODB on O2 CAD Interface Visual impact Appli.	MATRA SYSTEMES & INFORMATION
DORNIER	Land use monitoring application (AES)	none	Client on PC Client on UNIX Server kernel GeODB on Arc-Info. AES Application	DORNIER
INTECS	Urban planning Appli. Client part WAN	Client part	Client on PC, Client on UNIX Server kernel, GeODB Arc-Info, Urban plan. Appli.	INTECS
RTO	WAN	none	Client on PC, Client on UNIX Server kernel, GeODB Arc-Info, Urban plan. Appli.	INTECS

Table 2 : Sites definition

8.6.1.2 Applications description

8.6.1.2.1 Visual Impact Application

The Visual Impact application has been defined in collaboration by MATRA SYSTEMES & INFORMATION and the CUDL. The main features of this application are the following :

The EVER product, developed under a partnership of MATRA SYSTEMES & INFORMATION and EDF brings heterogeneous GIS, CAD and photographs data together to deliver a highly realistic visual rendering of civil engineering construction into the real environment.

The MATRA SYSTEMES & INFORMATION application builds on the initial success of the EVER product and on the performances of the GEO2DIs server.

The EVER product software is adapted for being connected to the GEO2DIs server and thus exploiting the server facilities.

The case study aimed to demonstrate the effectiveness of the complete MATRA SYSTEMES & INFORMATION GEO2DIs tool (server+application) was the impact study of the sport and leisure centre in Deulemont.

The « Visual impact » application was aimed to demonstrate that the GEO2DIs server meets the initial objectives of the project in the following way.

The geographical data server had to handle large amounts of data fast for the visual impact demonstrator. A typical data set for an impact study is about 1,5 Gigabytes. For managing several studies, the server has to provide efficient access to large data bases.

The geographical data server had to handle heterogeneous data and sources for the visual impact application. This application requires very heterogeneous geo-referenced data like raster, 2D and 3D vector and CAD data of multiple sources. Therefore it was able to demonstrate the ability of the GEO2DIs server to manage coherently these data.

The MATRA SYSTEMES & INFORMATION application is dedicated to land and infrastructure planners. The application users are not specialist of GeoGraphy or information technology, and they will have the possibility to efficiently manage an heterogeneous set of geographic data through the GEO2DIs server.

8.6.1.2.2 AES Application

After the reform of the common agricultural policy in 1992, the emphasis for support of farmers was moved from price control to direct payments based on the area of crops planted. Procedures were laid down for administering and monitoring these grants. Farmers can apply for grants for the production of cereal, leguminous crops, oilseed, linseed and for set-aside and disadvantaged areas (in fact, for nearly everything except intensive fruit and vegetable farming). As sometimes profiled in the media, this system can be misused, so a monitoring system has been introduced. The body responsible for monitoring land use can choose between manual inspection of the fields and the use of remotely sensed data.

When using remotely sensed data in the past, the process of image evaluation has been carried out manually. Around 5% of farms were selected at random in a suitable area. All parcels for the farms selected were evaluated by visual inspection of pseudo-coloured SPOT images (multispectral and multitemporal), compared with a manually calibrated colour table. This was then compared with the data given in the grant application form, again manually, and a report prepared. The report stated whether the form was correct, or whether further action was required, such as visual inspection of the land.

Project AES (Agricultural Monitoring System) is a demonstrator for the automatic evaluation of remotely sensed images for land-use monitoring. It uses GEO2DIS client/server components for data access.

The use of GEO2DIS enables quick access to satellite images, to cadastral maps, and to the grant application forms themselves. These data are in raster format, in vector format, in tabular form and in ordinary ASCII files. The catalogue browser gives a uniform feel to the distributed data sources, and the user-friendly interface simplifies searching for and ordering data. Raster images can be requested for preview, to check cloud cover and weather conditions.

The work can be performed in stages, and the interim results stored in GEO2DIS until the next stage. Thus the work can be performed on different machines and by different people with a minimum of administrative overhead. Paper copies can be kept to a minimum.

The AES application consists of two subsystems related to different evaluation phases. The subsystem for crop identification is Unix based. The subsystem for grant assessment is an ArcView application written in ArcView's programming language Avenue. It can be run on a PC or on an ArcView distribution for Unix as well.

Crop identification is achieved by image evaluation. Images are evaluated by classification. The classifier used is of the supervised pixel-based neural net type. The classifier was trained using ground truth data. The figure below shows the classifier in action on an example SPOT image. The trained neural net is seen on the right. The small coloured polygons at the top of the picture are training data. After classification crop codes are attached to individual pixels of the image.

During **pixel to crop conversion** the classified image is analysed in combination with cadastral data to produce a table showing the crop grown in each field. First the classified pixel co-ordinates of the raster image are converted to a vector file. The vector data is combined with the cadastral map, also in vector format, to obtain the area of each crop in each field. Then a table is produced, showing field identifier (parcel id) and area of crop in square metre.

Grant assessment combines the results of crop identification with the grant forms to state whether a form is correct or not. These are mainly tabular operations. Specific parcels and grant forms can also be selected by coloured maps.

First **plausibility checks** are carried out on the forms - is the form submitted in the proper manner, are the crops eligible for grants and so on. Only the forms that pass all plausibility checks are candidates for further evaluation. Their stated crop areas will be compared with the actual crop areas given by crop identification results. Implemented decision rules help to identify problematic parcels which need further investigations.

For problematic parcels **maps and plans for visual inspection** can be produced. These maps are used to identify the fields in question. They are accompanied by tables to state the problem and to enter the results of visual inspection. The results are then entered into the system.

Inspection results are incorporated and forms are re-evaluated to produce **final reports** in the format required. Besides the investigation of grant forms the system also provides tools for statistical analysis of grant data and classification results. Statistics will be displayed using business graphic diagrams (bar charts, pie charts) or coloured maps.

8.6.1.2.3 Urban Planning Application

The suitability analysis application demonstrate that the GEO2DIs server meets the initial objectives of the project in the following way.

The geographical data server will handle large amounts of data fast for this demonstrator.

The geographical data server will handle heterogeneous data and sources for the suitability analysis application : slope maps, floods risk map, hydro-geological vulnerability map, soil use, hydrography map, infrastructure maps, provided both by Region Toscana and Pisa Province.

8.6.2 Results of the validation tests

General results of the validation tests have been outlined in the description of the achievement of the objective O2 « To demonstrate the capabilities of the system ».

Detailed results of the tests are recorded in RD 44.

8.7 Performances of the system

The system fulfils all its requirements but few points : historical data are not managed, decompression algorithms are not integrated in the client part and the selection by intersection with a line is not supported. Weak and strong points of the system have been analysed.

8.7.1 Weak points

a) WAN security of the DDT (Server). Though the basic protocol stack has been successfully chosen (TCP/IP), the used services are nowadays not totally appropriate for an extensive use of the Server on the WAN. Security is scarcely compatible with the use of rsh. This can work safely in LAN (Intranet) not in WAN (Internet).

b) Exploitation of http technology. Compared with the GEO2DIS schedule, the technology has become workable too late. Its use could have been twofold: 1) being http servers so widely available nowadays and world-wide supported, the GEO2DIS Client-Server protocol would have inherited all its features (e.g. reliability, security, open-endness, easy integration in Customer Information Servers, etc.); 2) on the Web it could have been possible to outline an Index of Catalogues, singular Catalogues or also singular metadata entries, and those searched, browsed and downloaded at will, thus enforcing the organisational infrastructure of GEO2DIS;

c) a set of Server administration services, comprising GIS dependent and independent components, is only partially available. This was due to limitations in budget (e.g. missing extract metadata service from heterogeneous GISs);

d) Though conceived open and flexible, it could have been more sound if the system had comprised a third, commercial GIS.

e) Late compared with the GEO2DIS schedule have emerged several standards, addressing either metadata (e.g. US FGDC and European CEN TC/287), or spatial query languages (e.g. spatial extension of SQL3-MM), or common GIS API interfaces (e.g. those under definition by Open GIS), and transfer formats (e.g. SDTS). Those have been considered for compliance and impact, but not implemented.

f) Ergonomy of the Client Man-Machine Interface could have been improved. There are some limits in the (free) text and form-oriented display of metadata.

g) User have still to define their requests according to the various data model of the geographical data sets. That means that queries are expressed explicitly in term of the specific object types and attributes of the data models and that it is not possible to request roads with an width greater than 10 meter from datasets in both VPF and EDIGEO formats by a simple query. This point is not easy to solve and can be solved only by a standardisation process of the data model or by translation between data models. Interoperability studies made mainly in the military field have shown that it is an hard problem with only partial solutions. From a practical point of view, the user is greatly helped by the MMI which allow him to define request with self-documented menus and can define request without prior knowledge of the data model.

h) The security classification of the data is not managed.

i) The archives (historical data) are not managed.

8.7.2 Strong points

a) the fundamental 5 requirements addressing the "Open Geodata Access" paradigm are implemented: 1) locate the data, spatially and attributive; 2) determine if the data meets user needs; 3) retrieve the data; 4) ensure the data are in the format required by the user software; 5) maintain data & metadata. Several "sibling" experiments are available, world-wide, sometimes providing, on singular requirements, either more functionalities, or better performances, or more user-friendly interfaces than GEO2DIS. No one such of them, however, covers all the requirements above.

b) the whole Client-Server architecture is sound and open to comply with the GIS-related emerging standards and technologies (e.g. as from US NSDI, EGII, EUROGI, Open GIS, etc.), which are revealing as the "2000 year challenge".

c) The catalogue data model is general enough to accommodate and document various geodata. In addition, with a limited effort, it can also comply with the emerging CEN TC/287 standard content.

d) the metadata exchange format is simple and general enough either to be customisable in order to comply with possible emerging standards (say html, sgml), or for being extended (free text parts).

e) the ArcView technology, necessary on the Client, is revealing commercially and functionally successful; Tcl/Tk is confirming as a successful free development kit, which is platform and window system -independent and it is now available with compilers too;

f) Catalogue query performances are excellent, compared with the actual limitations of some "experiments" on the Web. The Web technology well applies only to search and browse, and not to spatial queries.

g) The choice of the TCP/IP protocol stack which is widely used and allows an migration to the http protocol and the Web, when performance issues will be solved (see the previous point).

h) The architecture of the system allows easy use of further GIS or DBMS software products, the use of a new tool being possible by the implementation of a new transformer.

i) The server on O2 can easily manage geographical data sets with new data models, with the implementation of couple of loader and writer, the extension of the internal data model by inheritance being easy.

9. Exploitation of the results of the project by the partners

9.1 Introduction

The partners of the geo2dis project will market the products of the project, defined in the paragraph 8.5.

The market for each product has been qualified in term of extension, homogeneity, growth rate, level of competition, purchasing decision process, differentiation and standardisation process.

Marketing rules have also been defined for each product.

Requirements for each market have been specified, and market structure outlined.

The following paragraphs outline main features of this exploitation. See RD40 to RD43 for further details.

9.2 Responsibilities for the exploitation

The exploitation responsibility for the products of the project would be organised as follows :

Partner➔ ➔Product	MATRA SYSTEMES & INFORMATION	DORNIER	INTECS	O_2	RTO	CNUCE	IGN
M_1+M_2	X	X	X				
M_3				X			
M_4	X (1)			X			
G	X						

Legend: (1) = With APIC and O2

- M_1 and M_2: The whole GEO2DIS heterogeneous geographical data server system (see paragraph 8.3.2).
- M_3 :The improved O_2 OODBMS (see the paragraph 8.4.4.).
- M_4: The spatial index module (See the paragraph 8.4.5.**)**
- G : geographical objects Libraries

9.3 *General policy for the exploitation*

9.3.1 GEO²DIS Positioning

GEO²DIS provides an increasing attractiveness with an increasing data volume and complexity. Prime candidates for GEO²DIS applications are therefore :

- Environmental Information Systems
- Land use monitoring
- Automated Mapping
- Facility management
- Regional and environmental planning
- Catastrophe Prevention

Data Type					
Raster					
* Satellite				1,2,3,6,7	
* Aircraft			1,3,6,7		
* Other		1			
Vector					
* Cadastral		2,3	6		
* Thematic Maps	4	3,1	6,7		
* Aero Photographic Data			1,3,6,7		
Attributive (Alphanumeric)					
* Statistics		6	4		
* Thematic Contents of Vector Data		1,2,3	6,7		
* Measurement Data				1	
* Other		7	2,5		
CAD	5,6				
Texting	5,6				
	Small	**Medium**	**High**	**Very High**	**Data Volume**
	(shaded)	**Applications,attractive for GEO2DIS**			

Legend: **1 = Environmental Information Systems**

2 = Land Use Monitoring

3 = Automated Mapping

4 = Demography, Marketing

5 = Facility Management

6 = Regional and Environmental Planning

7 = Catastrophe Prevention

Figure 9-1 : Positioning of GEO²DIS

9.3.2 Customer Profile

In the potential GEO²DIS markets the following customers have been identified as the most promising:

- Public bodies:
  - Municipalities and communities
  - Surveyors.
  - Cadastral offices.
  - Planning offices.
  - Environmental monitoring authorities.
  - Others.
- Universities and institutions (for research purposes)
- Large companies with needs for geographical data like
  - Electrical power distributors (e.g.Electricité de la France)
  - Civil engineering companies
  - Large facility operators
- Private institutions
  - Network operators
  - Consultants
  - Architects
  - Data service companies
  - Insurance companies (for example flood insurance)
- Public service bodies
  - Fireguards
  - Catastrophe prevention forces
  - Military

9.3.3 Market Requirements for Success

9.3.3.1 Basics

The following basic requirements apply fully to GEO^{2}DIS:

- The applied technology has to be sound and consolidated.
- It must be guaranteed that a potential customer if he decides to install our product has a protection of his earlier investment.
- Application of standards
- GEO2DIS must be connectable to other well introduced GIS on the market.
- The consortium should come up as soon as possible with a first convincing and functioning version of GEO2DIS. In this case the timeliness of availability is more important than the complete market maturity of the product (time to market).

9.3.3.2 Drivers:

In a highly competitive environment, like in the geographical data market, the main driver for a sustainable market success certainly will be the full spectrum of customer benefits, as detailed in the chapter 7.

First successful GEO^{2}DIS applications will accelerate further GEO^{2}DIS demands, if intensively communicated to the market.

9.3.3.3 Barriers:

A smooth and quick GEO^{2}DIS market introduction will be hindered by some barriers, which have to be considered for a successful market introduction. Those are :

Potential customers can only take full advantage of the GEO^{2}DIS benefits, if their counterparts have an operative GEO^{2}DIS capability, too, as shown in Figure 9-2.

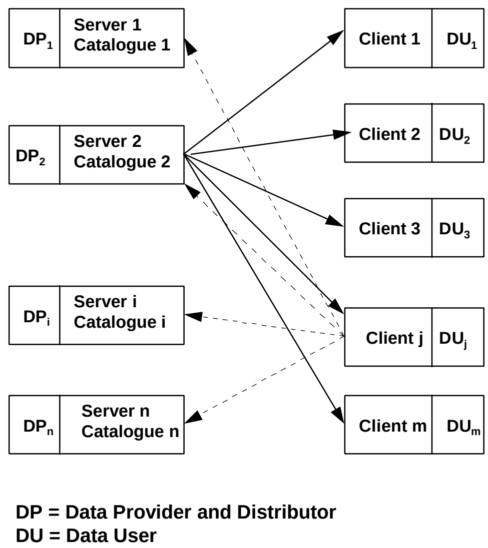


Figure 9-2 Increasing GEO²DIS attractiveness with increasing pairwise installations

- A data user may only benefit from a GEO²DIS investment, if his relevant data provider and/or data distributor has installed a fully operational data server and if corresponding comprehensive data catalogues are available.
- A data provider and/or distributor may only benefit from a GEO²DIS data server and catalogue investment, if a sufficient number of his clients have installed a GEO²DIS client.
- benefits are not clearly visible at once. They have to be intensively communicated to the market.

9.3.4 Potential Exploitation Channels (Exploitation Philosophy)

9.3.4.1 Regional Introduction Sequence

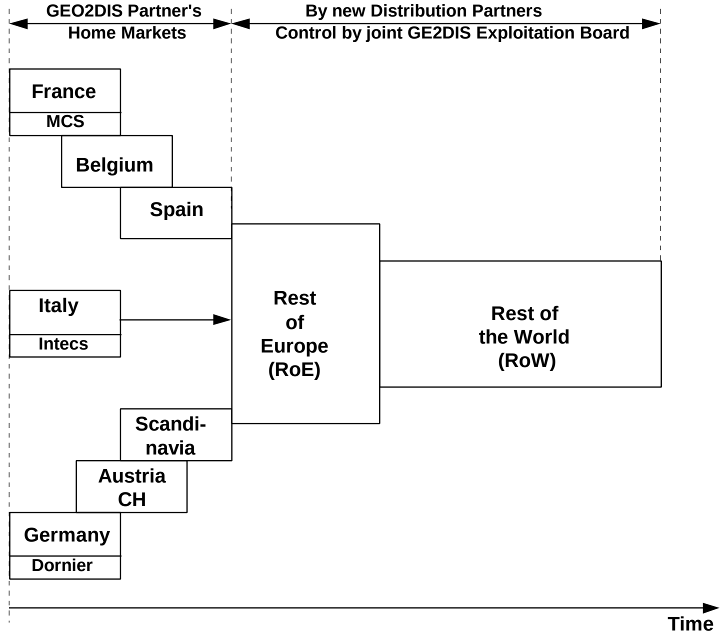


Figure 9-3 : Qualitative example of the anticipated GEO²DIS Market introduction policy

It is recommended to apply a stepwise GEO²DIS market introduction. That means :

- All GEO²DIS partners start with marketing activities in their home markets
- Stepwise marketing activities will be extended to surrounding countries
- RoE and RoW markets will be served by distribution partners

This approach reduces the risks of warranty problems. However, the speed of market development has to be adjusted to possible competitive developments. Here, further market research is necessary.

9.3.5 Synchronisation of Development and Exploitation Activities

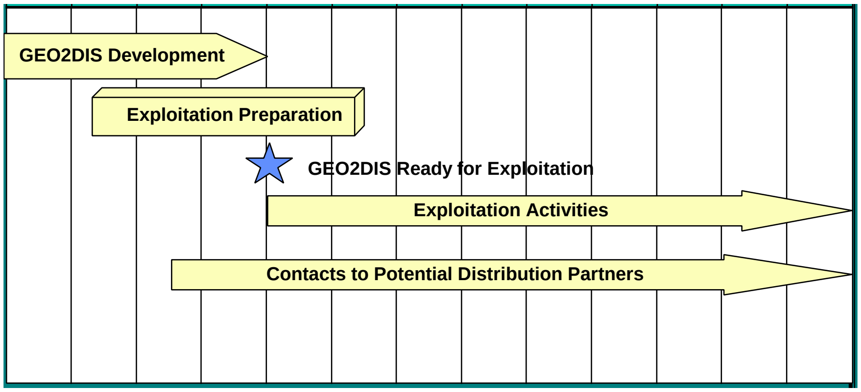


Figure 9-4 : Co-ordination of GEO^{2}DIS development- and exploitation activities

Synchronisation of development- and exploitation activities will be achieved by a co-ordinated timing as shown in . A joint exploitation board, established by all GEO^{2}DIS partners will decide on start of all corresponding activities.

9.3.6 Treatment of Market Introduction Barriers

In order to counterbalance the market introduction barriers, described in paragraph 9.3.3.3, it is mandatory to create some well functioning references as soon as possible. Therefore he following policy objectives are suggested :

- For an early market stimulation some leading data providers and/or distributors have to be equipped with GEO^{2}DIS servers. Most probably for this purpose special sales conditions are necessary like discounts, improved payment conditions, test installations with payment only in case of success etc.

- Some prominent data users, preferably with multiplier functions like universities, public bodies with a high public attention etc., have to be equipped with GEO^{2}DIS clients under similar conditions, as mentioned above.

- All GEO^{2}DIS project partners have to offer relevant consulting activities to those potential promotion partners

- All corresponding exploitation start up activities have to be accompanied by broad and intense communication activities. Partnership relations are recommended with one or more of the leading GIS manufacturers like Arc Info or Intergraph and/or hardware suppliers like e. g. IBM or Hewlett Packard in order to benefit from their world-wide distribution channels. In case, this is not possible, alternatively partnerships with specialised software distribution companies are possible.

9.3.7 Product Policy

GEO^{2}DIS product policy will be ruled by the following highlights:

- A high degree of modularisation of the GEO^{2}DIS system in order to achieve a maximum of differentiation in respect to customer requirements and ,by that, in order to facilitate a successful market penetration
- Continuous development of **Translators** for important GISs.
- Systematic utilisation of **all GEO2DIS developments** by all partners **in all markets**
- **Continuous improvement** of the GEO^{2}DIS performance in order to maintain the GEO^{2}DIS benefits as long as possible.

9.3.8 Pricing

Application of a pricing policy with the following objectives :

- **Benefit provision** to the customer.
- **Full cost coverage** to all GEO^{2}DIS partners
- Generation of sufficient resources to **support further GEO^{2}DIS developments**, as required in chapter 7.
- **Avoiding** of potential **competitive developments**
- Application of a wide spread **discount policy** in respect to
  - Key accounts
  - Quantities
  - Multipliers (like universities, scientific institutions etc.)

9.3.9 Training

Training will be realised by the training departments of the partners. There will be sessions for DD and DU.

9.3.10 Customer Service and Support

All GEO^{2}DIS partners have long lasting relations and experiences in the relevant markets. Correspondingly well established training, service and support infrastructures and experiences are available.

9.3.11 Competitive Analysis

The ESRI has developed and markets jointly with the ORACLE company since the end of 1995 the Spatial Database Engine (SDE) which permit to manage, store and retrieve spatial data within an ORACLE 7 DBMS. SDE offers an API which permit to develop applications in C, C++ or Visual Basic. SDE remains based on the relational model.

9.3.12 GEO2DIS Exploitation Actions

9.3.12.1 Dissemination

9.3.12.1.1 Dissemination material

- Production of GEO^{2}DIS data leaflets
- Production of GEO^{2}DIS technical descriptions with:
    - Problem description
    - Objective of GEO^{2}DIS
    - Technical solution
    - Systems environment
    - Benefits
    - Application examples
    - Systems information
- Production of contributions to
    - technical newspapers
    - non technical newspapers
- Production of presentation material, like
    - overhead foils
    - a PC based demonstration (for example by MS Powerpoint)
    - a GEO^{2}DIS demonstrator
- Data Catalogues
    - Data catalogue production adviser
    - Production of standardised data catalogues

9.3.12.1.2 Dissemination Actions

- Direct Mailing
- Information workshops
- Fairs participation (direct and indirect, like EGIS)
- WWW contribution
- Incentive actions with
  - Key accounts
  - Scientific institutions
- Closure of consultancy contracts with multipliers (like professors etc.)
- Advertising campaign in specialised magazines
- PR actions, like
  - press conference at program start
  - publications in specialised magazines
  - congresses participation (with active GEO^{2}DIS presentations)
  - selected key account contacts
- STCs (selected short term consulting actions)

9.3.12.2 Sales activities

- Establishment of a joint exploitation board
- Establishment of a sales organisation.
- Acquisition of selected beta customers.
- Start of sales activities.

9.3.12.3 After Sales Service (ASS)

- Program documentation (Handbook and other).
- Elaboration of a GEO^{2}DIS training course.
- Installation of a hot line.

- Installation of a problem solution service.
- Organisation of a GEO^{2}DIS user board with
  - workshops,
  - written experience exchange,
  - incentive actions,
  - Others .

9.4 Exploitation of the GEO2DIS project by MATRA SYSTEMES & INFORMATION

MATRA SYSTEMES & INFORMATION shares the overall GEO^{2}DIS objective to provide :

- quick and cost effective identification, selection and delivery of geographical data to any customer
- use of a user friendly catalogue,
- digital transmission and
- object oriented technologies.

Rationale for this objective is given in AD5.

MATRA SYSTEMES & INFORMATION has also the objective to develop technology, know-how and components in order to provide to its customers a fully integrated, up-to-date use of geographical data in Control and Command systems (C3I) and spaceborne and airborne surveillance systems. These components are needed for the implementation of the server on the O2 OODBMS and can be reused in other systems.

geo2dis technology and components are used in a project of photo interpretation station demonstrator for the UEO, for a demonstrator installed in the French Armed Forces, for a study for the photo interpretation station for the future (French Defence) and for some tenders for photo interpretation stations.

9.5 Exploitation of the geo2dis project by DORNIER

Remarks to the objectives of DORNIER:

- Initial objectives:

The goal of the project GEO2DIS is a client/server system for geo-data, to be developed and exploited by a consortium of three partners. Each partner will produce an application. The DORNIER application Agrar-Erfassungs-System AES (translation: Agricultural Monitoring System) is subtitled "Use of satellite data to identify land use and green areas at single land parcel level". This research project is interesting for the Ministry for Agriculture in a Land of the German federation. The results of the project shall include an operational system suitable for government use.

- Intermediate modifications:

In 1994/95 DORNIER as part of the Daimler-Benz Aerospace DASA has been restructured.

For the Profit Center VIM Reconnaissance and C³I-Systems a new policy was decided, having as content that all future activities will be restricted to the military market. So all commercial projects have to be finished without additional company funding for future marketing and product support.

- Actual objectives:

Based on this the goal of DORNIER is to enable a regular end of the project GEO2DIS in January 1997. Therefore the relevant exploitation actions/results of DORNIER are:

- General GEO2DIS Exploitation Plan: see AD5
- Exploitation Actions:
  - Sales Promotion Material (generation and co-ordination):
    - Leaflet (7000 units)
    - Technical description
    - Presentation material (view graphs)
  - Sales Promotion Actions:
    - **Consulting activities to obtain an agreement with DORNIER SatellitenSysteme (DSS) to take over the activities for the commercial market**
    - Presentation activities for the German military market related to the client/server system or/and reusable components of GEO2DIS.

As DSS did not choose the AES application for the commercial market, DORNIER will not undertake further exploitation action.

9.6 *Exploitation of the Geo2dis project by INTECS*

9.6.1 INTECS Sistemi market and market trend

Originally the Company business was only in the Space and Defence fields. There it played (and still plays) a relevant role as technology supplier in many software programs funded at European level (CEC, ESA, WEO, NATO, etc.). It is worth mentioning the ESA Space Programs like Columbus and Hermes, and the Defence HELIOS initiative. The Company involvement was in the requirements, design and development of medium-large software systems, exploiting all the leading-edge technologies: software engineering, process engineering, distributed systems, client-server applications, multimedia systems, hard real-time systems, etc. All such initiatives have also lead the Company to build an internal Quality System which was positively audited by several big Companies (MATRA MARCONI SPACE, Aerospatiale, Alenia), certified by ESA and, since last January 1995, compliant to ISO 9000 requirements too (as certified by DNV Italia, accredited Certifying Company).

A few years ago the Company, on the basis of an estimated negative trend of the "traditional" demand of complex systems in the Space and Military fields, decided an early strategic investment in the "new" field of the Information Systems for the Public Administration. The rationale was that the large and valuable experience in leading edge technologies, as gained in said Space and Defence fields, could have been positively exploited in favour of the Public Administration, whose "de-facto" technologies were on average, and still are, a bit behind.

Thus an analysis followed, aiming at identifying which internal technologies would have better fit the new demand and which additional technologies it could have been worth invest on. The result was that, beyond the generic field of the Information Systems, the development (integration, customisation, etc.) of
the specific Geographic Information Systems appeared as the most promising and fitting the Company background. Other fields, identified as most promising, were Information Systems for the Environment Control, Multimedia Technologies, Multimedia Networking and Process Engineering.

Again the analysis lead to the conclusion that, in order to approach a stable, medium-long term success, the winning principle was to invest overt rue quality products with tangible added-value. That could have been possible through two combined, parallel initiatives:

- heavy investments in R&D of product and applications, related to the above technologies;
- distribution of commercial products, based on a sound technology and easily recognised as "leaders" on the market;

That has lead on the one side to the participation to a set of Research Projects, and on the other to the promotion of commercial activities and projects with important Local Public Administrations like Regione Abruzzo, a Consortium of 19 Municipalities of Benevento Province, the Province of Pisa, the AEM (the Milan Electric Utility Company) and others. In parallel the Company obtained the distribution of FMS products (by Facility Mapping Systems, USA) for the PC, and became distributor of ESRI products too (by Environmental System Research Institute, USA), for both PC and Unix workstations.

Very soon, beside the traditional business related to the development of stand alone GIS applications, a new emerging market trend, linked to the growing availability of telematic infrastructures and to an increasing demand of GIS interoperability from the User's side, was identified as strategic for INTECS. The market for an European Company operating in the GIS field is, as a matter of fact, mainly related to GIS applications development being the GIS tools technology mainly owned by USA GIS producers. GIS Interoperability, instead, open a new potential market to European Companies addressing the development of middle-ware software components resolving the problem to make heterogeneous GIS and Geographical data interoperable.

The technology offer on this side is still weak and, in spite of the international efforts involving standardisation bodies and organisations to define a common approach to GIS Interoperability, a concrete alternative solution to GEO2DIS has not yet been reached in this field.

INTECS considers this market strategic and GEO2DIS results will allow the Company to compete on this new emerging market by a competitive position.

9.6.2 INTECS Sistemi exploitable results

Given the above preliminaries, it is quite tangible that GEO2DIS, in its objectives and technological choices, contains added-value and innovative solutions for creating the organisational framework that is asked by each LA.

The scenario has been described in short, but it remains complex in any case. Politics, power and responsibility conflicts are involved as well, as it can be easily guessed. INTECS goal in GEO2DIS, is to move towards a solution to these problems involving the different actors of this scenario. Though GEO2DIS cannot provide a turn-key solution to said scenario, we wish that it moves in a direction coherent with said objectives, at least. It's a challenge that the Company wishes to face with.
The baseline design decisions and technological choices help in building the GEO2DIS system compliant with the requirements arising from the scenario.

GEO2DIS is an heterogeneous Geographic Data Server and its implementation is largely based on a sound technology, at the same time consolidated, reliable but leading-edge too. It is worth pointing out that innovation and soundness, related to a software technology, don't always come along together in an easy way. Should any of the two slightly exceed the line of an implicit balance, there is the risk either to propose something robust (and, as such, usually accepted by users), but scarcely innovative and with poor stimulus of new demand, or to venture into some fruitless experiments, with doubtful and riskful results and no returns.
The Company is quite convinced that the technology adopted for the Server implementation in INTECS site embraces both innovation and soundness, in a good balance. That is expected to fit the ultimate purpose of the CEC funding.

In this contest INTECS is interested in exploit the Modules M1 (the Client part) and M2 (the Server part) of GEO2DIS based on ARC-INFO.

The Server is based on the adoption of the ARC-INFO GIS software v. 7.0.2 (by ESRI) for the Server part, plus ArcView2 (by ESRI) and PC-NFS (by Sun Microsystems), for the Client part, and

FMS/Arc application kit for the application. The basic software technology is completed by the TCP/IP protocol suite, Windows 3.11, and Solaris 2.3. ArcView2 merits a further, though short, description. In fact ArcView2 brings the user world together on the desktop. It's a complete system for accessing, displaying, querying, analysing and publishing the user's organisation data. ArcView2 links traditional data analysis tools, such as spreadsheets and business graphics, with maps for a completely integrated analysis system. Available, with the same functionality, on all major platforms (Unix, Mac, PC), ArcView can be used a stand-alone project system, or extended into an entire department, division or organisation.

Note that, being the Object-Oriented technology so leading nowadays, ESRI couldn’t have missed to comply with it in its latest product. In fact ArcView2 is developed on top of a new object-oriented programming language and development environment known as Avenue. Avenue is a full-function, object-oriented development environment with a scripting language designed specifically for GIS. With Avenue the user can completely customise his own ArcView2-based application, addressing his own needs.

Formally delivered on January '95, after an unexpected long "planning stage", ArcView2 has been further enhanced and consolidated by ESRI (as it is always for very innovative products). In a short time it will be substituted by ArcView3 thus representing the "innovation" edge of the balance previously mentioned.

Namely the following are the “key" items which represent the competitive technological results of GEO2DIS:

- architecture open to WAN distribution (thus IPC, Internet);

- software implemented both on Unix and PC/Windows, but open for being migrated to Mac and Windows NT (thus multi-platform availability);

- Catalogue functions rehostable on different DBMSs (thus low impact and needed resources for a Customer);

- Server Functions available on different GIS at the sole cost of developing a new translator.

While the majority of LAs have facilities to host the Client part, they do not have yet anything able to host the Server part, and only a few of them have a Geo-db (such as ARC-INFO): the openness of the GEO2DIS Server to WAN represents a good opportunity to attract also those potential users. There is a significant immaturity in this respect in the LAs, therefore the experience gained in Client/Server applications on WAN (through Internet), represents an "offer" with a potential very large demand. Note that, apart of the specific handling of geographic data, the WAN communication infrastructure is open to host many other software applications (say for transferring multimedia documents, browsing remote archives, e-mails, teleconferences, etc.).

The adoption of portable software on Unix and PC excellently fits the scenario, either of medium-large Municipalities as well as of small-medium ones. In addition the Client software on PC is relatively at low cost (ArcView2 single runtime license less than 2 Millions liras), affordable by any LA. ESRI products are fortunately prevailing on the market, thus minimising the impact on those LAs which haven't. PC/Windows are still prevailing. In any case should Windows NT or later technologies emerge, there are no impediments, since ArcView2 already runs on such platform (and

Mac too).
Lastly there is to point out the light impact that the design of the Client has on the typical customer resources. Since ArcView2 can access any popular DBMS (either on PC or Unix or Mac), there are no additional constraints for a customer. The GEO2DIS Client can be re-hosted on the Customer-preferred DBMS, at negligible cost.

Being the medium-small size Municipalities the target Customer of INTECS, the low-cost, popularity, customisability of the Client software are of the utmost importance.
From the functionality point of view, it is worth mentioning that the GEO2DIS Client/Server architecture has "built-in" the answer to that sort of organisational lack in the scenario. Uniformising the concept of Geo-Data-Set (GDS), providing its general model regardless its structure, type, format, Geo-db, etc., and allowing its retrieval upon additional user-defined clauses (restrictions), represents the kernel of an integrated Region-level Geographic Information System. Note that, in this framework, the Catalogue brings with itself an implicit form of standardisation, therefore assuring that a support is also provided for those LAs which are WAN connectionless.

9.6.3 INTECS Sistemi Exploitation Channels

The activities follow the rationale described so far. INTECS Sistemi is in the pleasant situation of having the GEO2DIS End User as his own "true" Customer (Regione Toscana). There is a considerable expectation in Regione Toscana for the GEO2DIS system.
The exploitation activities follow several directives. They can be grouped into direct and indirect ones.

9.6.3.1 Direct activities

Direct exploitation is first assured by the envisaged installation of the GEO2DIS system within Regione Toscana premises, in particular in the Urban Planning Department (i.e. those actually involved in the Project). A first, and easier marketing opportunity, is anyway given by the GEO2DIS promotion towards the other Region departments: Agriculture, Transports, Environment, etc.
The Regione Toscana represents the central authority of the region of course, and they can handle implicit promotional tools towards the other LAs (10 Provinces and 278 Municipalities), but also outside. It is worth mentioning that the Regione Toscana Associated Contract contains explicit clauses where the Regione Toscana commits itself to divulge and promote the results of the project in the Region.
Both the Province and the Municipality of Pisa have enough HW and SW equipment's to accommodate the whole Server: in particular ArcView2 on PC, LAN-connected to ARC-INFO on Unix, plus PC-NFS. Those represent two additional installation sites, having in turn an implicit divulging capability.

Another direct exploitation initiative is the dissemination of the results in favour of the other "Business Units" of the Company. The adopted basic technology can open new marketing opportunities also in nearby fields (multimedia telematic, Client/Server applications, Internet services, etc.). There are several user application fields where the adopted technology can see a significant re-use (say Tourism, National Health Service). Therefore the people of the Company, not strictly involved in the project, will be instructed of the results.

Another direct initiative aims at divulging the results in other equivalent contexts, such as other Regions, Provinces and Municipalities. As for the Regione Toscana, the Company sees opportunities not only for those Departments having the direct responsibility of the territory planning (Urban Planning Departments), but also for those which operate on the territory (Tourism, Environment, Agriculture). The principle is that GEO2DIS will represent a potential answer to all those distributed Organisations which operate on a "shared" territory, with shared geographic data. We wish to highlight that the approach above has already given successful results. In fact the Company, on the basis of the "competitive edge position" given by the GEO2DIS project participation, has been awarded, together with the Consorzio Pisa Ricerche and a pool of public & private organisations of the Prato Province area, of a contract for the realisation of a pollution monitoring system of the Arno river, in the Leather Industry Area (Comprensorio del Cuoio).The project, named POLLUTION, is funded by Ministry of Environment Protection and involves the Arno Basin Authority, the Pisa Province and the Health Local Authorities of the area. It aims at the realisation of a geo-referenced database of pollution-recorded data, thus integrating with multi source geo-data.

The same has happened for a less worth contract, awarded by the Province of Pisa and won by the Company, and concerning the actual realisation, on paper and vector forms, of the Territory Co-ordination Plan (PTC) for the same Province. Though limited in the budget, in Italy it represents one of few (maybe the first) PTC fully developed with automatic means (by means of ESRI products, such ARC-INFO and ArcCAD).

The same will hopefully happen for another big ITT (Invitation To Tender), issued by the Regione Toscana, and concerning the realisation of a Regional Environment Information System (namely SIRA, Sistema Informativo Regionale per l'Ambiente). INTECS Sistemi have arranged a pool of Companies and expects to exploit, at the best, the competitive edge position gained during the GEO2DIS project participation. This also because the required technologies for the SIRA project include ones very similar to those used in GEO2DIS.

9.6.4 Indirect activities

Indirect promotion sees several known “channels": brochures, articles on specialised magazines (but common newspapers too, especially having a large diffusion in the Region), seminars, congresses and exhibitions.

9.7 Exploitation of the GEO2dIS project by O_2 Technology

9.7.1 Technology description

Within GEO2DIS, O_2 Technology is involved in three O_2 enhancements : C++ methods activation from OQL, index clustering and index extensibility and spatial index. The benefits resulting from these enhancements are performance improvements of the O_2 DBMS for spatial applications (one or two orders of magnitude) and a more standard and effective programming environment. More specifically, O_2 will be enhanced as follows :

- **C+ methods activation from OQL** : ODMG is the leading organisation for object database standards. In ODMG, all the major database vendors are represented, including O_2 Technology. ODMG has delivered specifications concerning object model, C++ bindings and object query language (OQL). O_2 Technology is now working at delivering the first ODMG compliant system. O_2 already supports the full OQL and most of the C++ bindings. The possibility to invoke C++ methods from OQL is one of the last pieces missing. This is essential to have a smooth integration between C++ and OQL. This is also very valuable for building spatial data extraction programs.

- **Index extensibility and spatial index** : The O_2 system implements various indices to optimise the manipulation of large collections of objects. However, some applications (text retrieval, geographic information system), require specific indexing technologies. Because these indexing methods are many and specific, O_2 Technology decided that the first priority was to open the O_2 database engine so that new indices can be added by users. This is what is done in GEO2DIS : O_2 Technology opens the O_2 engine while MATRA SYSTEMES & INFORMATION implements a spatial index in it.

- **Spatial and index clustering** : In database systems, performance are critical and, most of the time, disk accesses are the bottleneck of the system. Among the various means to optimise disk accesses, clustering refers to the way objects are grouped within disk pages. Objects can be clustered according to different criteria. One of these criteria which make a lot of sense for spatial data management is to cluster objets according to (spatial) index sort. Index clustering combined with spatial indexing will deliver the performance required in GEO2DIS. Clustering strategies are defined by the database administrator. Then, the database goes through a number of stages : creation, exploitation, reorganisation, etc. The clustering quality decreases in time and it is necessary to have a tool to reorganise the physical space to fit the suite clustering. Index clustering will improve O_2 performance in many applications, especially those dealing with spatial data.

9.7.2 Business Opportunity

Among the market areas we have identified, more than many are concerned spatial data and therefore with the O_2 enhancements performed in the context of GEO2DIS. These applications also represented 20% of O_2 revenues in 1994.

The market segments concerned with spatial data is performance driven. Solution providers are looking for tools that can provide the level of performance required by applications and end-users. Cartographic systems commonly deal with hundred of Gigabytes of data up to several Terabytes as well as many concurrent users. These performance are not available with traditional tools such as relational databases.

For all these applications, the O_2 enhancements undertaken within GEO2DIS will improve performance in a 20 to 100 ratio compared to the previous O_2 release and its competitors (including relational and object databases). We expect to gain market advantages by industrialising the GEO2DIS result rapidly.

9.7.3 Positioning

The object database market generally and the database market for spatial application specifically is driven by performance and standards. The O_2 Technology strategy is to be very competitive on these two points and GEO2DIS results contribute to this strategy. Let us analyse these 2 issues :

- **Standards** : choosing a DBMS is usually a strategic choice for any organisation. The DBMS handles corporate information and should last for many years. A critical issue then is whether the DBMS supplier is a sound company that will be supportive for the a organisation as long as required. Relational suppliers are now large profitable companies that makes user confident with them. But object DBMS vendors are still very small and none of them can guarantee to be here in 5 or 10 years from now. The only way this question can be answered is standards. By providing effective portability standards, customers can move their system from one product to another with limited costs. These standards exist today with ODMG and OODBMS vendors are committed to deliver compliant product. O_2 Technology is now shipping a new release (4.6) with a fully ODMG C++ and OQL interface. Thanks to GEO2DIS, the integration of OQL and C++ is complete.

- **Performance** : when products are highly comparable in terms of functions (and this is the case as soon as standards such as SQL or ODMG) are in effect, the only way to choose one DBMS is performance. Besides, performance is a critical issue for many application. The object database market and growth is based on 2 facts : (1) though relational vendors will propose object extensions to their engine, the only way to have a full object model is to implement a full blown object engine, and (2) most applications dealing with large volume of complex objects require performance that only pure OODBMS can deliver. The performance ratio is today in the range of 10 to 100 and we expect this to grow in the coming years. Geographic systems are dealing with huge amount (100 Gigabytes or Terabytes) of complex 2D objects. Experts knows that relational systems are not adequate and, though it is a little young, the object technology is the best money can buy today. GEO2DIS enabled us to enhance the O_2 Engine to deliver the appropriate level of performance for a geographic server. O_2 Technology will continue enhancing the performance because that is the only way to bit the competition.

9.7.4 Customer profile

The major segments where O_2 Technology plans to promote the GEO2DIS results are defence and services :

- **Defence** : management of very large volume of images (satellite or air observation pictures) and maps are needed and performance are critical. Spatial index and clustering are decisive competitive advantages against other database solutions.

- **Services** : electricity distribution is concerned with network management (similar to telecommunication applications). Other applications such as train simulation, water supply, ground traffic control are also intensively manipulating spatial data and driven by database performance.

- Other customers may be interested (for instance some telecom system) but we believe defence and services are the first priority targets.

9.7.5 Industrialisation plan

- **C+ methods activation from OQL** : This extension to the current O_2 has been delivered in the 4.6 release available for customers in November 1995.

- **Index extensibility and spatial index** : The new API available to O_2 users to implement specific indices will be available confidentially in release 4.6 (limited availability) and shipped widely in the O_2 product release 4.7 (available Mid 1996). It should be outline that this does not concern the spatial index which belongs to MATRA SYSTEMES & INFORMATION.

- **Spatial and index clustering** : These extensions will be added to the O_2 product in release 4.6 with restricted availability and integrated in the standard product for release 4.7 (mid 1996).

9.7.6 Pricing

The current price for the O_2 system depends mainly on 3 parameters : the platform on which the system runs, the number of concurrent database connections allowed and whether the system is used for developing an application or just running it. The product comes as 2 pieces : server and clients. The price of the server ranges from 2000 ECU's on a single PC, for one connection in run-time till 50 000 ECU's on a UNIX box, with 24 connections for development purpose. The development tools are priced per seat between 1500 ECU's (a one seat C++ API) till 10000 ECU's for a full environment for one seat. Clients are run time free.

We also have a pricing policy for VAR (when O_2 is embedded in another software system) which is usually defined as a percentage of the total software with of course upper and lower bounds which depends on quantity.

9.7.7 Sales and support organisation

Since its creation in 1990, O_2 Technology has set up a world-wide sales and customer support organisation. O_2 Technology is currently organised as follows :

- R&D as well as quality control and software production are located in Versailles - France

- Sales, marketing and customer support (training, hotline, consultancy) is performed by subsidiaries in USA, UK, Germany and France. UK covers directly northern Europe. Germany also covers Austria. France takes care of Switzerland, southern Europe (Spain, Portugal) and the regions of the world not covered by a specific local organisation.

- Sales, marketing and customer support is done by distributors in Italy, Belgium, Japan, Korea, China and Hong Kong by distributors.

9.8 Exploitation of the GEO2DIs project by the IGN

A central data server could be a very important feature of IGN, since it could be used to support the geographic databases produced by IGN, and possibly play a role in their dissemination. However, a recent attempt to gather the main geographic databases of IGN (BDTopo, BDCarto, and GeoRoute) on a central server within a GIS failed because of poor performance. Questions are being asked in response: Do we need one server per database? Could we make a loose coupling between them to share certain common functions for demonstration purposes, or for database maintenance? Until these questions are resolved the future of GEO2DIS will remain uncertain.

However, IGN is committed to the project, and anticipates significant benefits from this participation (new technical developments such as the implementation of a spatial index in O2 are potentially valuable). Performance tests of GEO2DIS under realistic conditions (retrieval of data sets from BDTopo, for example) are likely to be made in the near future. If the results of these tests are interesting (they have not yet been communicated to us), more extensive experiments might be made over the next year..

9.9 Exploitation of the GEO2DIs project by the « Regione Toscana » (RTO)

There is no RTO Exploitation activity defined in the associated contract,and RTO has not planned any activity of promotion toward other Local Governments.

The GEO2DIS prototype has been developed according to system requirements which are a subset of the user requirements. In this state, the Geo2dis system is not completely compliant to the needs of the RTO, so RTO is not in the condition to plan a widespread use of the prototype.

In addition the following consideration can be made:

european standards concerning the transfer format of geographical informations as well as metainformation are near to be accepted and it is expected that all the european countries must conform to these standards

from a technical point of view the use of Internet and related servicesare more and more pervasive; this implies the use of technical solutionwhich are much cheaper than the one proposed by Geo2dis

It should be noted also that the activation of a GeoDataServer implies theuse of a certain set of resources by Regione Toscana and an organization effort that Regione Toscana will perform according to its own plans which in turn will be defined according to the demand of geographical information, which is not growing at the expected rate.

Nevertheless the Geo2dis project has been an exciting way for Regione Toscana to face the problem of Distributing geographical information and to create contacts with industries and scientific institutions outside of the traditional channels. As a final result Geo2dis was an opportunity for Regione Toscana for cultural grow and for increasing the sensibility toward dissemination of geodata.

9.10 Exploitation of the GEO2DIs project by the « CNUCE/CNR » (CNR)

CNUCE is involved with Geographic Data Servers since several years and it has many opportunities to exploit the experience acquired with the GEO2DIs project. A long term goal is to set up a Geo-Data Server for internal use, which then should be organised as a service for the whole CNR. In this case the experience of GEO2DIs is important because this system shall manage very heterogeneous data coming from many different disciplinary fields; a detailed on-line catalogue (together with data accessibility) is mandatory to provide a good service. A short term application for GEO2DIs is in the field of PNRA (the Italian National Program for Antarctic Research): at present CNR participates to the PNRA with several operating units and one of these is active at CNUCE with the specific task to set up a Geo-Data Server for the Antarctic scientific community. To do this we are evaluating the possibility to adopt a GEO2DIs-like philosophy. CNUCE is the owner of the Transformer for ARC/INFO that can be used as is to interface the cartographic archive (most data are ARC/INFO coverages). As the Client part of the GEO2DIs system is not available for CNR, the aim is to develop a system, based on WWW technology, which permits both to access the meta information (standard TC/287) and to issue orders to ARC/INFO via the Transformer.

Another form of exploitation is represented by economical exploitation: due to its nature of scientific organisation, CNR has neither the interest nor the resources to commercialise the Transformer for ARC/INFO, so the best way is to Patent it and to gain duties.

ANNEXES

10. Annex 1 : Discussion of the possible general operational architectures

There are many DUs, and many DPs too and as a consequence the system must offer a lot of heterogeneous data.

There are many possible architectures for the system, with various levels of complexity :

- Lev. 1 There is just one DD which collects data from many DPs and distributes this data to many DUs. It is the simplest way to activate the system and is well defined from a territorial institutional point of view (a good application for the ILLAs like a Région in France, Regione in Italy, and Lander in Germany).

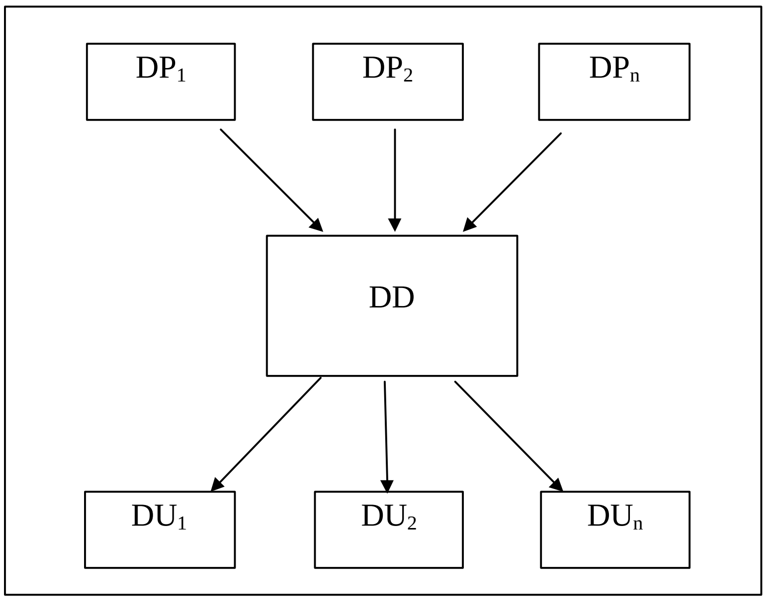


Figure 10-1 : Level 1 implementation

- Lev. 2 There are many DDs. The technological solution is complex, and the territorial dimension is higher than for a ILLA (it may be national or even international). The DPs may have a more active role, and they themselves may become DDs. This level can be subdivided in relation to how the data and the corresponding documentation (CAT) are organised .

- lev. 2a There are many DDs, and each DD can supply all the data of all the DDs, each DD is in fact a replica of the other DDs. The advantage is only in terms of access for the DUs; the drawbacks are organisational problems and duplication of resources.

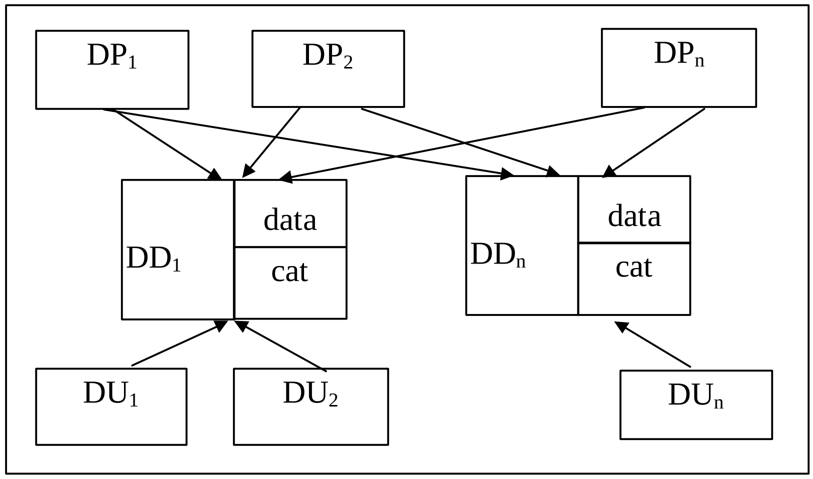


Figure 10-2 : Level 2a implementation

- lev. 2b There are many DDs, and each DD has (and therefore can supply to DUs) only their own specific data which is chosen following various criteria (e.g. by theme, territorial responsibility, institutional responsibility). However each DD also has the documentation (CAT) of the data of all the other DDs. DUs can connect with the most convenient DD and check for the data they need; afterwards, if the data required is available they can get them directly from the same DD, otherwise they have to make a link with the DD where the data is. This method has the advantage of gaining access to the DUs and doesn't lead to geodata being duplicated, but it does entail some organisation.

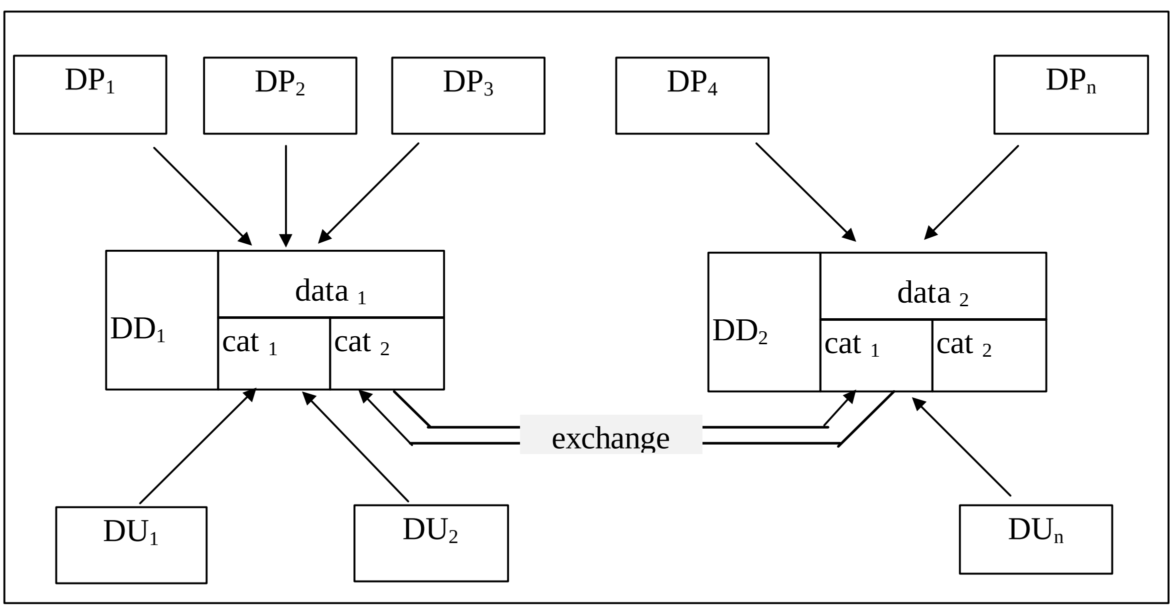


Figure 10-3 : Level 2b implementation

- lev. 2c There are many DDs, and each DD has (and therefore can supply to DUs) only their own specific data which are chosen following various criteria (e.g. by theme, territorial responsibility, institutional responsibility). In addition, each DD only has the documentation (CAT) of the data it manages. All the DDs are linked together through a telecommunication network, so that the DU perceives only one unified service and not the individual DDs. This doesn't lead to data being duplicated and requires a little less organisation; it is, however, technically more complex because each DD must provide a generic DU with all the services it offers by itself along with the services offered by the other Dds. Response times will also be longer and information sent to the DU depends on the number of DDs that are currently active.

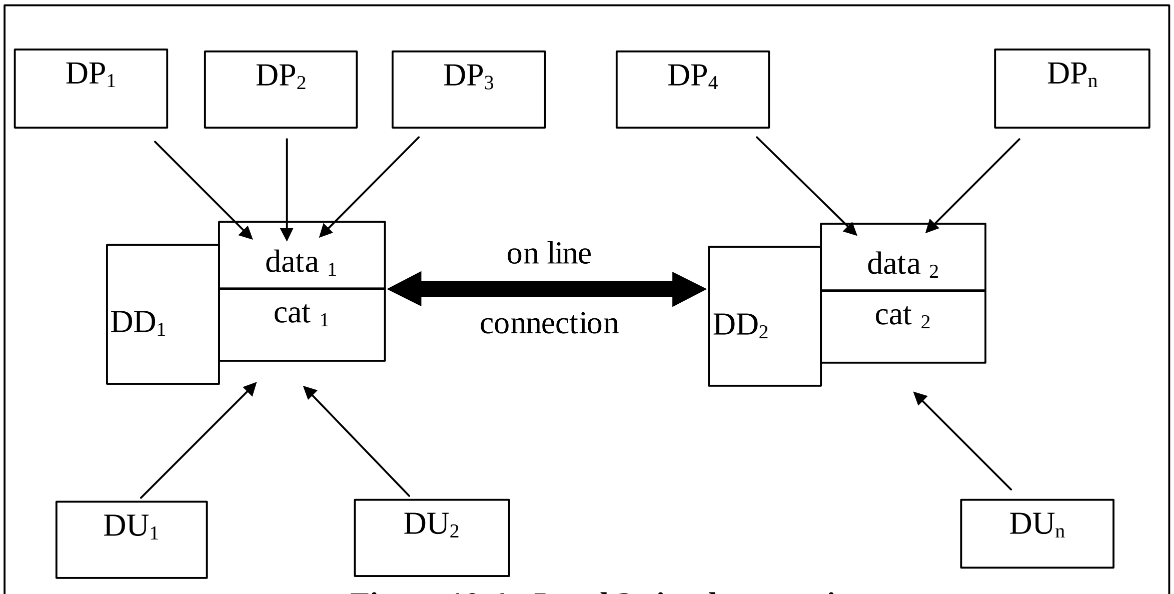


Figure 10-4 : Level 2c implementation

We believe that the solution outlined in 2b is the most suitable for the project; some of the requirements given in the main document and in annexes 2 and 3 are given taking into account this hypothesis.

11. Annex 2 : Users requirements related to the Data

11.1 GENERAL INFORMATION

The data is what a DD sells via GEO2DIS; it is therefore a critical element in the system.

Producing geodata and digitising already existing geodata is not the task of a DD. However a DD wants to carry out operations on the data he receives from a DP so as to offer a service that is more complete than that offered by the DPs. In fact a DD offers data coming from several DPs, a check on the quality of the data, a comprehensive catalogue and distribution facilities.

The data managed by GEO2DIS relates to the territory, and therefore has a geographical component. A lot of data is conceived without the geographical component being explicit (a lot of statistical, legal and normative data), but since it is applied to territorial entities it has a geographical value. If we take this idea to the extreme then all the information we manage is geo information. GEO2DIS must be able to manage the kinds of data described in the following sections, which cover most of geo data types, but will also be open to new types of data.

11.2 Data managed by GEO2DIS

GEO2DIS can manage the following data :

11.2.1 Structured vector data, topologically corrected, whose geometric primitives are points, lines and areas.

- All this data has attributes, which are generally alphanumeric and maybe numerous, this represents the real richness of the data.
- This data exists both for large scale (for example cadastral data) and for very small scale (for example at the national level).
- The quantity of data in terms of bytes is high, but not excessive.
- Availability will be in continual increase, and there will be continuous updates of both graphics and attributes.
- Several more simple types of data can be grouped together into data with higher complexity, which gives a more complete model of the geographical entity.

11.2.2 Non-structured vector data, whose geometric primitives are points, lines and areas.

- This is data from the digitising of geographical maps, whose various geographical entities have not been identified and selected. In practice, the lines appear as such, though they may be associated with a code that qualifies them and associates them with a plot technique (this means that a line is qualified as a "line that describes the edge of a road", but not as "a road").
- This data doesn't need to be topologically corrected.

- It comes from large scale maps and has no attributes.
- The quantity in terms of bytes is high, but not excessive.
- It is mainly used for plot operations.
- A plotting file produced by any software system (and therefore by a GIS system) can also be considered in this category.

11.2.3 Tabular data to be associated with vector data, as described in 11.2.1

- This data consists of alphanumerical and numerical fields.
- The quantity of data in terms of bytes is high, but not excessive.
- The availability is in continual increase over time, and there will be considerable amounts of such data in the future.
- Tabular data not directly associated with vector data, as described in 11.2.
- The availability of this data will be in continual increase over time, and there will be considerable amounts of such data in the future.
- Raster data based on a regular grid. This data is either concerned with a specific acquisition process, or is being modeled like this because there is no need or possibility to model the data as vector data.
- Since the geographical entities are not defined there are no attributes but only values associated with pixels. However, there is information associated such as geographical map references, the meaning of the pixel values (indispensable in a symbolic raster map), a histogram of values, etc.
- There is a high quantity of data, but compression algorithms can be activated that can reduce the size of the data from around a factor of 5 (typical for remote sensing images) to 100 (which may be achieved for black and white scanned maps).
- Data from remote sensing is acquired with high frequency, therefore the need to have up-to-date data entails managing an enormous quantity of data.

11.2.4 Maps on paper and photos.

- This concerns documents of the territory that are on paper and will probably never be inserted into a computer (for example aerial photos or historical maps).
- The existence of such data is documented via a description of the data, although the data itself is accessible via a non-computerised procedure outside GEO2DIS.

- This procedure can also be used for existing cartography that has not yet been computerised.

11.2.5 Texts that are directly and indirectly linked to geographical entities

This is the case of norms and laws which can be expressed on formatted fields or as free text. In free texts the document must be formed by ASCII characters or be in raster whereas if it can be expressed using formatted fields, we go back to the tabular data case.

11.3 Input formats

- The data that the DPs send to the DDs are coded in accordance with a standard that will be defined by the DPs or by an agreement between the DPs and the DDs.

- Since GEO2DIS is an open system, it doesn't have its own input standard nor does it rest on a national standard, but rather on international and market standards.

- International standards are currently being studied and so far few have actually been defined and are in use. The market standards are, however, quite consolidated and are those which we consider should be used, while still being open to future international standards.

- The use of market standards, as well as national and international ones, has also the role to stop the proliferation of proprietary standards.

11.4 Pre-processing the data

- Data from a DP to a DD has to undergo a series of controls, some general ones (quality and consistency), and others which are more specific and depend on constraints imposed by GEO2DIS (special controls).

- Whether these controls are made or not depends on the role of the DDs and their technical and organisational know-how. In any case, the system must be able to support such activities with adequate procedures. The functionalities performed by the procedures can only be defined after the data on which the procedures operate has been identified.

- The quality control procedures must allow some data characteristics (on the geometry and attributes) to be measured, and thus compare the quality of the data with the quality claimed by the DP. The DD can subsequently be responsible for certifying the good quality of the data or for declaring inferior quality.

- A geographical quality control procedure consists in defining a common reference base (on a larger scale or equal to those of the data treated) and to verify whether the data under examination, compared to the reference data, differs within the declared tolerance levels. The result is a measurement or an estimate of the average or maximum geometric displacement between the data being examined and the reference data.

- The consistency procedures must allow the numerical values to be verified, both those relating to the attributes and to the geometry, and then allow the contents of the data itself to be modified as a function of the consistency criteria in the data itself.

- The procedures for special controls must allow the numerical values to be verified, those relating to the attributes and where necessary those relating to the geometry, and then allow the contents of the data to be modified as a function of the constraints typical of GEO2DIS.

- Procedures for checks on quality and consistency, and for special controls cannot be automated procedures. They will be interactive procedures, with some help by the system for the operator in charge of controls.

- Another important operation is to organise the vector data into a logic of continuum if the DP supplies this data with different organisational methods. This would entail slight modifications to the vector data both with automatic criteria and in graphical editing, subsequently acting on the attributes.

- DDs can modify the data according to an agreement between themselves and the DPs. Any modifications to the data should be technically locked according to the local legal constraints. In any case, if the data is modified, the original version supplied by the DP must in any case be available for the DU.

11.5 Description of the data

- To be efficiently accessed via GEO2DIS, the data must be supplied with a description (Catalogue abbr. CAT) outlining the following points for the data:

- technical description of the geographical domain, the attributes present, the origin, and the quality;

- applicable areas (e.g. urban planning, waste disposal);

- cost and access administrative procedures;

- security classification;

- a sample.

- The user needs to know which DDs possess the data required.

- The exact contents of the CAT has been defined in the Task 160.

- During the creation of the catalogue data duplication needs to be verified along with any inconsistencies (in geometry and attributes) inside the DD facilities; some control may also be needed amongst the various DDs' data depending on operational constraints.

11.6 Organisation of data

- An archive of homogenous data, organised by a DD and made available for the DUs is a geo data set. For example, a geological map of Bosnia, the administrative borders of Normandy, and the rainfall in Baveria, are different geo data sets and each has its own description in the catalogue.

- Within a geo data set, the various elements have characteristics of homogeneity. For example, the geological areas follow the same legend, the administrative borders come from the same scale, and the rainfall data is measured with the same criteria.

- In many cases the elements that make up a geo data set are not homogenous, either for historical reasons or due to difficulties in updating; this may cause problems in cataloguing.

- The best strategy is to manage geo data sets with the greatest possible homogeneity. Therefore if a geo data set has internal elements that differ significantly, it is best to split the geo data set into two, each of which contains absolutely homogenous geo data within itself; for example, if the individual areas of a geological map for the north of Italy follow a legend that is different from the center and south, then the policy is to have two separate geo data sets (or informative layers). The proliferation of similar geo data sets from the thematic point of view can create problems for the DU in the analysis phase of the CAT; this should be resolved at a UI level.

- Data is organised within GEO2DIS, according to its characteristics; the most common organisations are a geographical continuum or a regular or irregular polygon-based organisation. Satellite images have their own organisation. GEO2DIS must support these types of organisations.

11.7 DATA UPDATING

The data will be modified for two reasons: to correct data where an error has been found, and to update data with more recent data. Corrected data completely replaces the previous data, whereas when data is updated, the previous data is still kept available for the DUs. If the amount of data becomes too great, historical data can be stored off-line, but will still appear in the catalogue though with different delivery times.

This section has been expanded in task 150.

12. Annex 3 : Users requirements related to the functionalities

12.1 General Information

GEO2DIS is based on two types of physical station: the DD's station and the DU's station. On the DD's station the data is prepared for distribution, the information in the catalogue is organised and managed, and communications with the DUs are kept active. The DU's station has to allow users to submit queries to the catalogue, and to define and issue a request for data and receive it, all this within one session alone. Subsequent processing by the DU is independent of GEO2DIS and must make use of suitable software, whose owner is the DU.

The DD's station has to operate on a workstation with a UNIX operating system; the DU's station must be able to operate on:

- a workstation with a UNIX operating system
- PCs with DOS 6.2 operating system with WINDOWS 3.1 or a later version.
- PCs with WINDOWS-NT
- Macintosh systems.

The DD's station must be designed so as to be easily adaptable to new types of data from the DPs, to new formats, and to new GIS software that may appear on the market.

12.2 FUNCTIONALITIES

12.2.1 DD’s station functionalities

The DD's station must have the following functionalities:

- it must be possible to access information in the catalogue both on-line (direct link between the DU and the DD) and off-line. In the latter case, the DU must be able to acquire a product composed by the contents of the catalogue and by a query software, all hypothetically on a CD-ROM. This functionality will hereafter be referred to as off-line analysis.
- it must be possible to select a subset of data, following the specific requests of the DU and to send him the data either on the telematic network or by mail.
- it must be possible to process geo data resident on the DD's station and to create a Reference Background, a support to help the DU in the identification of the most important geographical entities.

12.2.2 General functionalities

The combined use of the server and the DU's stations must enable the DU to:

- identify the data needed for a specific application
- select from available data
- see a generic sample of the selected data
- select a subset of data related to a specific geographical area, typically a polygon area
- select a subset of geographical elements from those available (e.g. select only the main roads from all the roads)
- select a subset of attributes from those available and sort them freely
- preview the data selected, that means allow subsampled vector or pixel information to be displayed on the DU's screen so that the DU can evaluate the real data he will receive if he purchases it.
- transform the geographical projection and the reference system
- convert the data into a required data transfer format
- activate data delivery from the DD to the DU.

12.2.3 Local facilities

Once the data have been received, the DU can exploit it in relation to the hardware and software that he has available and that are outside the GEO2DIS; in any case the DUs have to be able to :

draw the data as it is on a plotter

process it with GIS software and with other software (for example CAD or statistics)

transport it onto equipment with different characteristics.

12.3 Creating the catalogue

A catalogue cannot be created automatically. There needs to be a procedure that can extract some data from the geo data set (informative layer), and subsequently interactively allow new fields to be added and to edit those already present.

During this phase the control procedures will be carried out.

The information available in the catalogue may be presented to the DUs in various ways depending on the DU profile; this is a task of the User.

12.4 Data distribution

Data have to be distributed in the format chosen by the DU from a subset of possible formats. Commercial standards are preferred, whilst still being open to future international standards.

Distribution can take place :

- on-line, via local networks (LAN) or geographical networks (WAN),
- off-line via CCT, cassettes, floppy disks sent by mail (numerical data),
- off-line by paper or photos sent by mail (maps on paper).

12.5 User Interface

Building an efficient user interface is a fundamental part of the project since the UI is what DUs see of the system and what they dialogue with.

The software that manages the UI operates on the DU's station and the interaction with the DD's station takes place only by submitting requests and transmitting catalogue, preview and download data.

The DU interacts with the system using a mouse, keyboard and if possible by voice, using national languages (French, German, Italian, English) and as much as possible icons.

The DUs introduce themselves to the system with a name that the system knows. On the basis of this name (and of a description of the user associated with it) the system behaves accordingly, and presents personalised information on:

- interaction techniques
- interaction language
- degree of detail of the information in the catalogue (more expert users will receive more concise information)
- administrative information.

A fundamental element in the UI is the tool offered to the DU for selecting the desired data .

The system has to present the DU with a representation of the territory so that the DU can define the area to operate on through the identification of the main reference geographical entities (Reference Background or RB).

The RB is managed within a logic of geographical continuum and contains physical data (for example rivers, lakes or coastlines), infrastructural (for example roads, railways or urban areas), administrative (for example administrative borders) and a reduced toponymy.

This representation of the territory is obtained with a combination of raster and vector data that the DU can make visible or invisible, can zoom and pan, and can change the presentation techniques (for vector data : for example type of line or colour ; for raster data : for example contrast or brightness).

The RB is presented at the beginning of a session on the basis of the user profile.

The DU can select entities of any type on the RB, or by choosing one or more existing entities (for example a river or a certain number of contiguous administrative entities) or by defining new entities (for example a polygon, a line or a point). We will call these entities "SSUs" (Selection Support Unit). However the SSUs are chosen they are, in any case, points, lines or areas.

The SSUs are used to select geographically the data to extract from the DD on the basis of geometric operators:

if the SSU is an area, the operations are:

- the data within that area
- the data that intersect that area

if the SSU is a line, the operation is:

- the data that intersect that line.

The point and line SSUs can be transformed into areas through buffering, thus allowing a vast number of selection criteria. In this way the operation *the data near to a line less than X* becomes a two step operation: the first step is to transform the line into a polygon through a buffering with size X and the second is the operation « the data within that area. The same for a point.

It is desirable, when extracting the data, to maintain data consistency, that is geo objects should not be cut.

The DU is further supported by a user manual for the system, and on-line help tools, which, if there is a direct link between the DU and the DD, can be located on the DD.

Additional support tools for the DU:

- estimate of the cost of an extraction after the extraction criteria have been defined, but before the process has been confirmed;

- being able to cancel an extraction request if the procedure has not yet been activated.

12.6 *Additional Value Added Services (AVASs)*

The DP can also offer the DU additional services, supplied on request, with an extra cost added to the normal cost charged by a DP.

These services may be useful when the DU is not suitably equipped or doesn't have a specific skill.

Examples of AVASs include correcting geometrical distortions of remote sensed images, statistic processing, and electrostatic plotting.

These services are based on software made available to the DP, and not included in GEO2DIS. GEO2DIS must only support the requests of AVASs.

12.7 *Performance*

System response times vary considerably according to the specific request:

- off-line analysis of the catalogue: immediate response (less than 5 seconds, but tolerable up to 15 sec depending of the hardware, the catalogue extent and the type of request).

- on-line analysis of the catalogue: immediate response (less than 5 sec for LAN and 15 for WAN)

- preview of selected data (obviously only on-line access through LAN or WAN): quasi immediate response (less than 60 sec for LAN and less than 5 min. for WAN)

- data extraction: due to the different types of data (numerical or on paper), the various mailing methods (LAN, WAN, network, courier, post) and the various performances of the geodbs, the response times cannot be defined. The DU wishes to have the data as quickly as possible, but, on the other hand, the mailing criteria cannot be interfered with; as far as the extraction time is concerned, it should be as close as possible to the maximum performances of the geodb involved .

- Data can be programmed to be sent at a pre-arranged date and time.

12.8 *Organisation*

12.8.1 Data Distributor

The DD, from an organisational point of view, will be a public organisation created for this goal with respect to a mixed pool of private and public organisations, depending on the political and administrative situation in each country.

The DD will deliver data to a set of DUs within a non-profit making approach. This does not mean necessarily that the service is free, but something may be paid to compensate for the investment. In the future a profit making approach may be planned.

In any case the DD should promote Geographic Information and its use through Geographic Information Systems.

The DP is the primary element responsible for the quality of the data; whereas the DD is responsible for the global service, i.e. data documentation, access to data and any other value added service. The DU perceives the DD as the one who provides the data, thus giving the DD some kind of global responsibility: this should be borne in mind during operational activity.

The DD will provide the DU with a software running on a set of platforms (hardware and operating system) that will enable the DU to use the facilities of the DD. Optionally the DD may provide the hardware too. It is important to maintain active a technological link between the DD and the DU as, in many cases, the DU is unaware of technological developments.

12.8.2 Organisation Of The Data Distributor

The DD may be considered as an organisation whose main items are hardware, software, data and skilled people.

People		
Hardware	Software	Data

The main components of a DD organisation.

The first three items concern technology; people are concerned with organisational aspects.

As far as this special point is concerned, the DD will activate some special offices (one or more) whose main jobs are to:

- promote the services offered by the DD itself
- authorise users to use the system
- define a profile of the user
- inform users of new data or new products
- monitor user activity

- prepare bills for users.

The commercial strategy of the DD may differ according to the environment it is operating in. With a closed group of users within the same organisation a financial account may not be necessary, or it may only be useful for monitoring the DU's activity. In other cases an exact account of the DU's activity must be kept.

In each DD installation there will be a DBA (Data Base Administrator) whose job is to:

- insert the data furnished by the DP and update it
- create the catalogue
- manage the whole system from a technical point of view
- monitor the system
- offer technical assistance to the DUs.

12.8.3 Data User

DUs belong to many categories: public organisations, private companies and professional individuals.

At present only a few dozen DUs are active, equally balanced between public and private users. In a few years they will increase, in some cases up to some hundreds. In particular, the number of private users will increase. These figures come from the implementation areas planned in the project, thus being only a small part of the whole market.

The number of DU varies considerably according to country.

DU access to the system also depends on the various situations. At most each user will query the catalogue once a day, but only a few queries (from 2% to 10%) will be followed by a request for data. The maximum load of the system will be 40 queries to the catalogue with one request for data per day. This figure will rise to 100-200 and the number of requests for data will go up to five.

In some cases requests for data involve delivering a map on paper or other means, instead of a telematic transmission of information.

As can be seen, accessing catalogues is rather heavy, and if an off-line catalogue has not yet been implemented, concurrent access may exist. This is also subject to the duration of a session which may be from some minutes up to one hour depending on the DU's skill and the performance of the system.