Merging Catalog Services and GIS applications by component interoperability mechanisms

O. Cantán, J.A. Bañares, J. Gutierrez, J. Nogueras, F.J. Zarazaga[1]
Department of Computer Science and System Engineering
University of Zaragoza
María de Luna 3, 50015 Zaragoza (Spain)
[1]javy@posta.unizar.es

Abstract Several organisations, especially public administrations, usually manage large amounts of information with geographical references. Maybe, one of the most important problems of this kind of organisations is to share their data among their different divisions, and to enable common access to it from commercial GIS applications. Additionally, metadata describing geospatial resources is hardly available and the documentation process is usually hard and complex. This work presents a system that integrates geographic catalog services and tools with GIS applications in order to facilitate the share, discovery and access of data, and the creation of metadata.

1. Introduction

Large amounts of geospatial data have been collected for more than 35 years. Furthermore, the collection speed increases quickly with new technologies in high resolution satellites, GPS, data bases, new software technologies for processing geospatial data and the increase of people and organisations which are collecting and using this kind of data. About 80% of data bases used by public administrations have geospatial references (addresses, city distributions, cartographic coordinates, etc). From another point of view, the geographic information production market in Europe is about 10 billions of Euros per year. Information is being collected, but the market does not exploit it correctly. In many cases, geospatial data-consuming companies or people do not find the data they need and they usually pay data suppliers for custom-made products.

Most of this data is created and managed using commercial GIS tools such as ESRI products ArcView and ArcInfo. In many cases, this information is created combining new data with existing data. To do this work, these tools do not usually offer utilities for discovering data using other reference apart from file name. This problem is bigger if data across Internet have been tried to find. One possible solution should be to have access to a geographical-information distributed-catalog from the commercial application (figure 1).

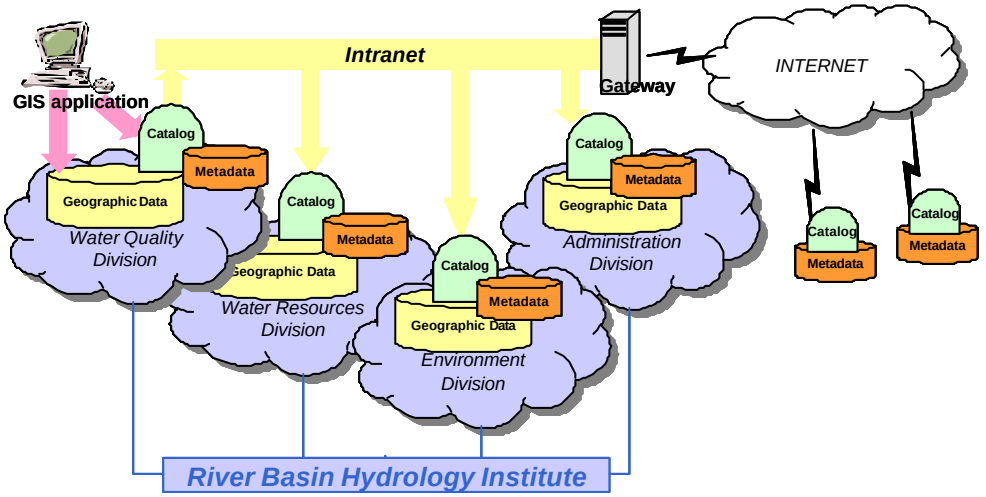


***Figure 1**: Looking for information with a commercial GIS application.*

On the other hand, most commercial GIS applications offer utilities to derive information which describes the geographic information (metadata). This functionality could be used by metadata-creation tools to facilitate the metadata-generation process. In this process, the tool should enable access to the commercial GIS application in order to derive so much metadata as possible (figure 2).

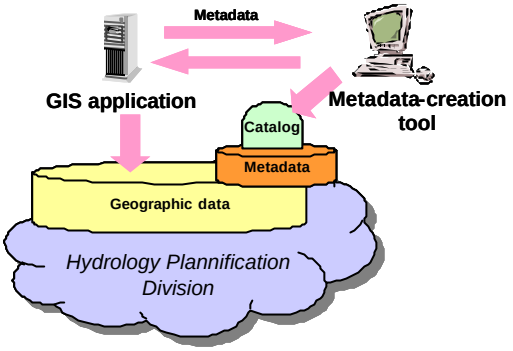


***Figure 2**: Deriving metadata using a commercial GIS application.*

This work presents a system that integrates geographic catalog services and tools with GIS applications in order to facilitate the share, discovery and access of data, and the creation of metadata. The rest of the paper has been structured as follows. Next section presents the system architecture. Section three is focused on the component developed to interoperate ArcInfo 8 and ArcView with the rest of the system. This paper ends with a conclusions and future work section.

2. System architecture

This system is composed of three main components (see the architecture presented in figure 3):

- a catalog built using the OpenGIS Catalog specification (see [OGC 1999] as a reference),
- a set of tools to catalog and discover metadata in this catalog,
- and a set of sub-components which enable the connection of some commercial GIS products (initially ArcView and ArcInfo 8 from ESRI) with the catalog.

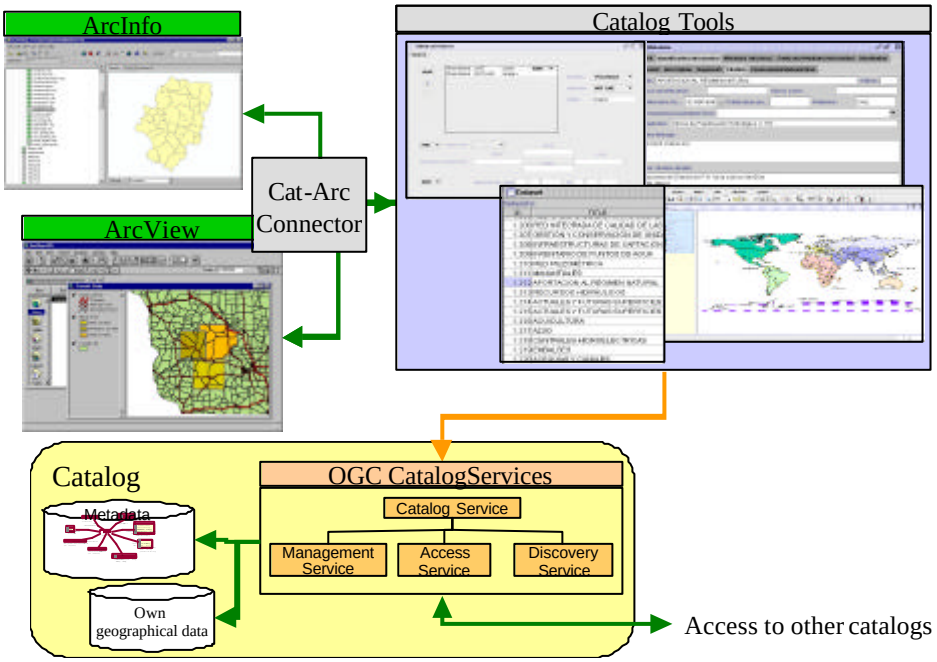


***Figure 3**: Information system architecture.*

According to the three main components of the system, three working areas have been defined to develop this project. The first one provides the utilities to create and maintain a geographic information catalog. This catalog must be able to store metadata describing the information managed by these organisations. The OpenGIS Consortium uses the term "Catalog" to describe the set of service interfaces which support organisation, discovery, and access of geospatial information. Catalog services help users or application software to find information that exists anywhere in a distributed computing environment. A Catalog can be thought of as a specialised database of information about geospatial resources available to a group or community of users.

The development of the Catalog presented by this paper has been guided by the OGC Catalog Services specification (see [Zarazaga et al]). One of the basic components in this catalog is the metadata object model. Metadata provided by the system must conform not only international standards, but also be capable of representing additional metadata items related with a particular organisation context and not supported directly by international standards. In this way, the metadata object model under development follows basically the American CSDGM of FGDC (see [FGDC 1998]), with the capacity to be extended easily. On the other hand, it must be taken in mind that there is still no commitment about what metadata standard is going to succeed 100% in the future. Therefore the focus is not to implement exactly a unique standard but to be compliant with different standards. To be compliant does not mean create a new standard consisting of an increasingly number of metadata elements covering all possibilities but to make efforts on providing translation services from/to different standards such as ISO/TC 211 (see [ISO 1998]), CEO-Recommendations on Metadata (see [CEO 1999]), or European CEN/TC 287 prENV 12657 (see [CEN 2000]).

The objective of the second working area is to offer tools for search and access to the data on an organisation Intranet, or over the Internet. These utilities will have four degrees of complexity:

1. A set of "closed-queries" that will be solved in run-time in order to include in the response the latest incorporations to the catalog. Organisation employees using a simple language could create these queries.
2. An interface enabling users to explore all the data available at the organisation. Results are shown either as a straightforward list or organised by a thematic tree customised to the organisation context.
3. Some interfaces to construct "free-text" and "map-guided" queries. The first one will offer the possibility of using natural language to ask the catalog.
4. An interface to construct OGC Common Query Language expressions and query a compliant OpenGIS Catalog integrated within a network of distributed catalogs.

Finally, the organisation catalog and the metadata-creation tool should be able to interact with the commercial products. It would be desirable for these products user to have "natural access" to these products from/to the catalog utilities. In this way, two are the jobs to be done:

1. From the point of view of data cataloguing utilities, the system must provide access to these GIS products in order to derive metadata from geographic data managed by these products.
2. From the point of view of GIS products, the system must provide access to the "search & access" utilities of the catalog in order to locate and incorporate data.

First versions of this component allow communications with ArcView and ArcInfo 8 from ESRI. They have been selected because the use of these products is very common in many public Spanish-administrations. The technology to communicate ESRI products with the catalog utilities will be different depending on the ESRI product possibilities. In this way, the connection with ArcView will be done using DDE (see [DDE 2000]), while the connection with ArcInfo will be done using COM-Java RMI technology (see [Sessions] or [Orfali et al]). The communication between ESRI products and the catalog will be done using component named *CatArcConnector.*

3. The CatArcConnector component

The widespread use of the ESRI products in public administrations makes it desirable to enable the interaction of these tools with the organisation catalog. In order to achieve this goal a new software component, called *CatArcConnector,* had to be deployed. Its purpose is to make a straightforward communication among the parties involved possible. The design of this component has been done thinking in extend the functionality to allow the connection with other GIS commercial products.

The *CatArcConnector* component is settled between the ESRI products and the organisation catalog. Figure 2 shows this configuration.

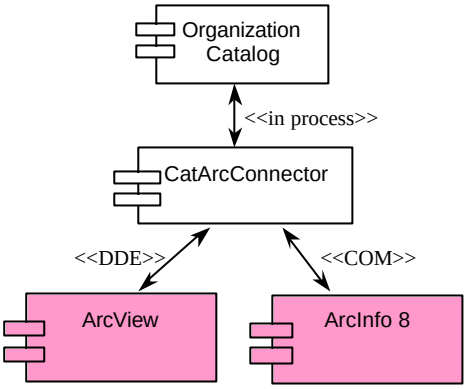


***Figure 4**: CatArcConnector configuration*

As it is shown, *CatArcConnector* has to be able to use different kinds of inter-process communication mechanisms (IPC) depending on the ESRI product to interact with. According to this, *CatArcConnector* communicates with ArcView using DDE (Dynamic Data Exchange) whereas it uses COM to communicate with ArcInfo 8. As for the interaction with the Organisation Catalog, it can be chosen the most convenient IPC mechanism.

In the following sections, we will dive into the concrete functionality and architecture of *CatArcConnector* regarding its interactions with ArcView and ArcInfo 8.

3.1 CatArcConnector interacting with ArcView

The *CatArcConnector* component enables the interaction between ArcView and the organisation catalog. One of the main objectives was that it should be easy, flexible and powerful to ask for a service from one application to the other. This implies, among others, the following capabilities:

- *CatArcConnector* should permit the reception of requests both from ArcView and from the organisation catalog tools (such as the metadata-creation tool).
- *CatArcConnector* should be able to manage the requests both synchronous and asynchronously.
- *CatArcConnector* should forward the request to the corresponding component and, after receiving the answer, redirect it to the initial solicitor.
- *CatArcConnector* should be configurable at run-time. For this component to work properly several parameters have to be specified: the IPC mechanism to be used for the communication with the organisation catalog, the set of operations that are allowed and the mode of operation for each service (synch-asynch). All this set-up is done at run time, just when the component is created, so there is no need to recompile the code to adapt to new scenarios.

In figure 3 the main parts of *CatArcConnector* for the interaction with ArcView are shown:

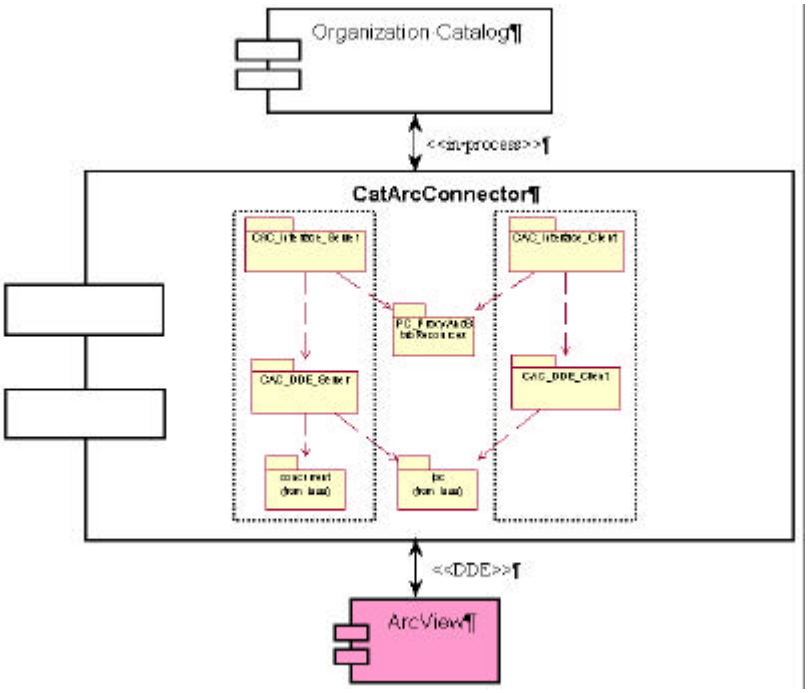


Figure 5*: CatArcConnector architecture for communications with ArcView*

The *CatArcConnector* component can be conceptually divided into two parts, which are graphically delimited by the dotted-line squares in the figure 3. This division has to do with the two roles that can be played during the interaction between ArcView and the organisation catalog tools. According to the first role, ArcView will act as a client whereas the catalog tools will be the server. In the second role, the applications will behave in the opposite way. In either case, the requests for service can be treated synchronous or asynchronously. In synchronous mode the client will wait till the server supplies the answer. Quite the opposite will occur in the asynchronous mode: the client will continue with its work and, after some unknown time, it will receive the answer. These two modes of operation make the *CatArcConnector* flexible enough to adapt gracefully to the user's needs.

Finally, it is worth mentioning that the communication with ArcView is accomplished by means of DDE (Dynamic Data Exchange). This is a sort of inter-process communication mechanisms provided by Windows operating systems. The fact that the current version of ArcView is rather old makes it impossible to use a more advanced inter-process communication mechanism, such as COM.

3.2 CatArcConnector interacting with ArcInfo 8

Another important and widespread product of the ESRI family is ArcInfo 8. Therefore, it was considered to be very convenient that the *CatArcConnector* component should permit the interaction between this product and the organisation catalog tools.

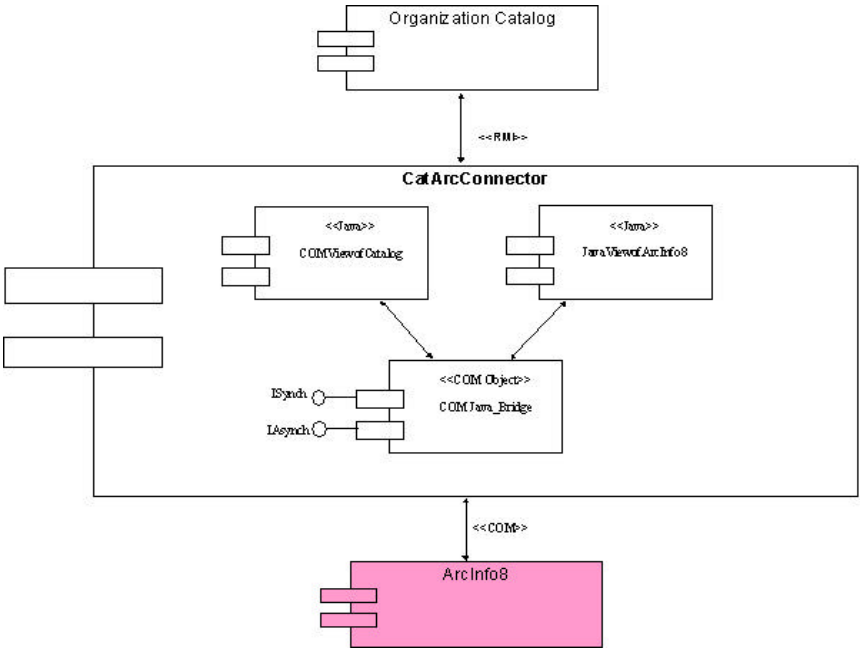


***Figure 6**: CatArcConnector architecture for communications with ArcInfo 8*

Due to the fact that ArcInfo is built using COM objects and that the catalog tools are written entirely in Java, one of the main tasks of the *CatArcConnector* component is to provide a kind of adapter between these two types of technologies. This adapter is called *Com-Java_Bridge*, as is shown in figure 4.

The *COM-Java_Bridge* adapter is a COM object written in Java and interpreted by the Microsoft Virtual Machine (MVM). Of course, this is not the sole solution for overcoming this problem. In fact, the use of a COM-CORBA bridge was also studied but it was discarded in favour of the use of the MVM.

The other two components of *CatArcConnector* are *COMViewofCatalog* and *JavaViewofArcInfo8*. The first one's function is to provide a view of the organisation catalog as if it were a COM object. That way, ArcInfo 8 can handle the organisation catalog as a COM object. The purpose of *JavaViewofArcInfo8* is quite the opposite: to provide the organisation catalog with a Java view of ArcInfo 8. With these two components, both applications, the organisation catalog and ArcInfo 8, view each other as if they were of the same class when, in fact, they are using completely different technologies.

4. Conclusions

This work has presented a system that integrates geographic catalog services and tools with GIS applications in order to facilitate the share, discovery and access of data, and the creation of metadata. This system incorporates a geographical information catalog and a metadata-creation tool developed in previous projects. Over these components, it has been necessary to build an interface which connects them to the GIS commercial applications. These connections allow GIS application users to search for geographic data using the catalog capacities. And on the other hand, metadata creators can take advantage of commercial GIS products in order to improve the metadata-generation process. In both cases, it has been possible to establish a new synergetic approach making better use of applications and data.

The solution built to interoperate GIS commercial products and the rest of the system has involved a hard work to test different solutions and to verify its technical viability. Finally, it has been necessary to develop the component using three different compilers: Visual Basic compiler to work with ArcInfo components (COM technology), 100% pure Java compiler for the development of the main part of the component, and Visual J++ (it is not 100% pure Java) to make the COM-RMI bridge.

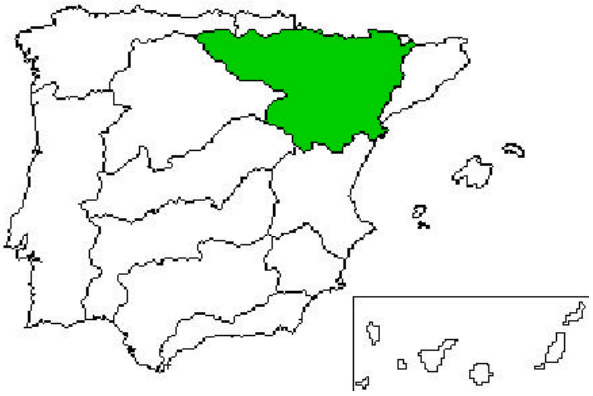


Figure 7*: The Ebro River Basin Organisation geographical domain*

Currently, there are seven people working half time in this project. The development of the system has been divided into three versions. The first version of this system will be finished by July 2001, while the final system should be operative at the end of 2001. This product will be installed as a first test platform on the Ebro River Basin Organisation (CHE: http://www.chebro.es), which is responsible for the management of water resources concerning the Ebro river, the biggest one in Spain, and its flowing ones (Figure 7).

Acknowledgements

This work was partially supported by the Comisión Interministerial de Ciencia y Tecnología (CICYT) of Spain through the project TIC2000-1568-C03-01. The work of Oscar Cantán has been partially supported by the "Gobierno de Aragón" and the "Fondo Social Europeo" through a grant.

References

OGC 1999, *OpenGIS Project Document 99-051s, OpenGIS – Catalog Interface Implementation Specification (version 1.0)*, OpenGIS Consortium, 1999

FGDC 1998, *Metadata Ad Hoc Working Group. Document FGDC-STD-001-1998 Content Standard for Digital Geospatial Metadata*, Federal Geographic Data Committee (USA), 1998

ISO 1998, *ISO/TC 211 committee draft for Geographic Information Metadata. ISO CD 15046 part 15*,. International Standardization Organization (http://www.statkart.no/isotc211/), dated 1998

CEO 1999, *A User Guide provided by the Center for Earth Observation Programme (CEO programme) Of the European Commission. Recommendations on Metadata. Describing the data, services and information you have available! (version 2.0)*, Center for Earth Observervation, February 1999.

CEN 2000, *The CEN/TC 287 Geographic Information European Prestandards, ENV 12657. Euro-norme Voluntaire for Geographic Information - Data description - Metadata,* European Committee for Standardization (http://www.cenorm.be), dated 2000

DDE 2000, *Web pages about Dynamic Data Exchange (DDE): http://www.angelfire.com/biz/rhaminisys/ddeinfo.html , http://msdn.microsoft.com/library/*

Sessions, R., *COM and DCOM: Microsoft's Vision for Distributed Objects*, Ed. John Wiley & Sons, 1997

Orfali, R., Harkey, D., *Client/Server Programming with Java and CORBA*, 2nd Edition, Ed. Wiley Computer Publishing, 1998

Zarazaga, F.J., López, R., Nogueras, J., Cantán, O., Álvarez, P., Muro-Medrano, P.R., "First Steps to Set Up Java Components for the OpenGIS Catalog Services and its Software Infrastructure", *Proceedings of 3rd AGILE Conference on Geographic Information Science*, Helsinki/Espoo (Finland) , 25-27 May 2000

European Commission
SPE 01.150 – Proceedings of the 7th EC-GI & GIS Workshop 'EGII - Managing the Mosaic'

Editors: Karen Fullerton & Alison Munro

The 7th EC-GI & GIS Workshop was organised by the GI & GIS project of the Joint Research Centre. The GI & GIS project supports the actions to create an INfrastructure for SPatial InfoRmation in Europe (INSPIRE). INSPIRE encompasses the broad policy, organisational, technical and financial arrangements necessary to support increased access to Geographic Information in Europe. It will potentially benefit many stakeholders - government and non-government organisations, education and research institutions, the commercial sector, and the general community - at the national, regional and global level. INSPIRE offers the prospect of better decision-making and thus improved economic growth, social development and environmental management.

The GI & GIS project reflects the increasing awareness of the importance of spatial aspects in policy and decision-making, and focuses on both policy and technical issues relating to spatial information and spatial information systems. The main activities of the project include:

- Assisting the GI policy making process of the Commission, helping to formalise the user requirements for INSPIRE, and contributing to the establishment of an interoperable European Geo-Statistical system;

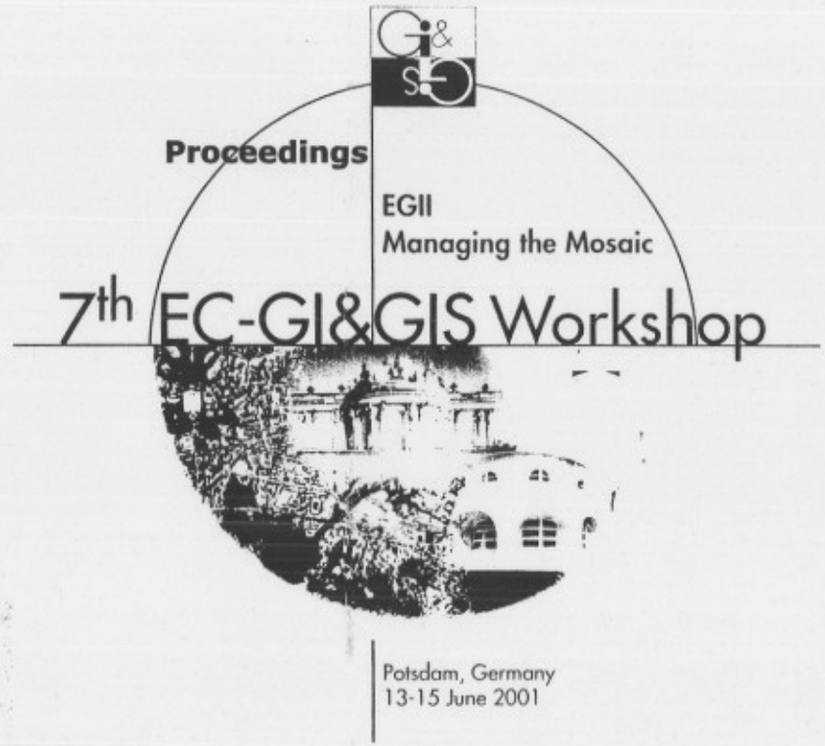
- Developing harmonised and coherent multidisciplinary Pan-European databases and analyzing spatial information across different sectoral policies and different levels of governmental organizations;

- Providing technical support to the Services of the Commission in defining base and thematic spatial data requirements at the European level, and operating [illegible] communication facilities; and

- Monitoring GI & GIS standards, interoperability, and market development through a Technology Watch activity.

For more information on GI & GIS in the European Commission please visit our website: **http://www.ec-gis.org/**

7EC-GISWS

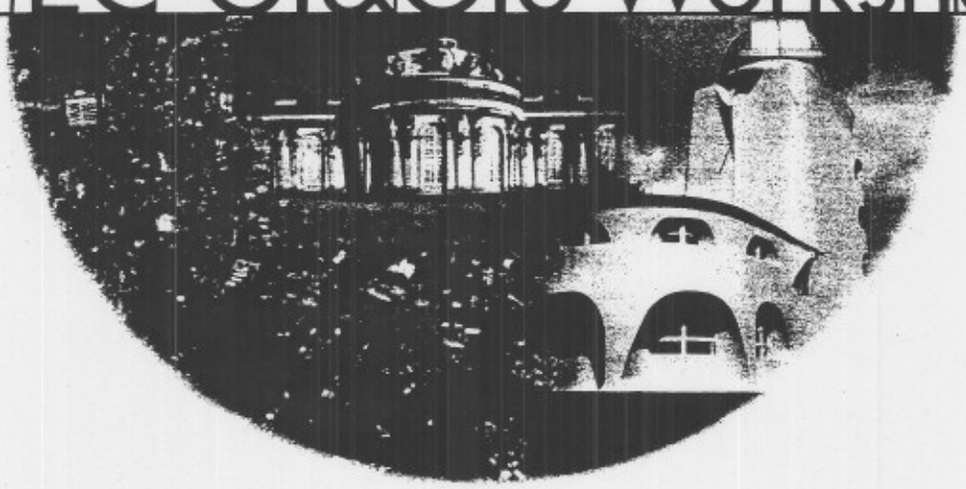


Space Applications Institute

Abstracts

EGII
Managing the Mosaic

7th EC-GI&GIS Workshop



Potsdam, Germany
13-15 June 2001

Sai Space Applications Institute

EUROPEAN COMMISSION
JOINT RESEARCH CENTRE

GFZ POTSDAM