GIS Interoperability: An Overview

Kokou Yétongnon, Djamal Benslimane, *Robert Laurini
Eric Leclercq, Fabrice Jouanot, Nadine Cullot

LE2I - Équipe Ingénierie Informatique - Université de Bourgogne
9, rue Alain Savary - BP 400
21011 DIJON Cedex - FRANCE
e-mail : {FirsName.LastName}@u-bourgogne.fr

LISI* - 502 - INSA de Lyon- Université Lyon I
F - 69621 Villeurbanne Cedex
e-mail : laurini@if.insa-lyon.fr

Abstract

The rapid evolution of communication and networking technologies has lead to the development of approaches to allow data exchange, information sharing and interoperation of spatial information. In this paper we present a survey of spatial interoperability approaches and show how they have evolved with the evolution of concepts such as data model translators, integration methodology, semantic representation (metadata, context, ontology) and data conflicts resolution methods that are used to support interoperability. We present a taxonomy of the different approaches in three interoperability levels: platform, syntactic and application.

1 Introduction

In the emerging open and distributed environments, interoperability is essential for many systems including GIS-based applications. Interoperability involves making multiple information sources access, manipulate and share data across their boundaries. In the last ten years or so, interest in this research area has rapidly grown with (1) the continuous evolution of networking and communication technologies and (2) the explosive growth of data (both traditional and spatial) sources that can be accessed via interconnected computers. A large amount of this research effort has been directed towards the development of methodologies, architectures, models and tools to allow sharing, exchange and control of data.

Despite the diversity of types and formats that exist in data sources, a great deal of early research on interoperability has been centered on traditional information systems. This research has identified approaches for integrating or merging heterogeneous data sources: distributed databases, federated databases and interoperable information systems. Recently, there has been an increasing number of attempts at extending the above approaches to non-traditional information systems including spatial information systems and GIS-based applications. GIS have evolved from monolithic large software systems designed for main frame computers to smaller component-oriented information systems that integrate one or more functional layers [?]. GIS-oriented tools such as ArcView and Mapinfo are increasingly available on desktop computers and widely used in various applications. Furthermore, GIS-based applications have evolved from isolated single-system-application to complex decision support systems that are based on the cooperation of heterogeneous information systems. Typical examples of GIS-based applications include urban planning, traffic control and natural resource management. They often require tools and architectures to (1) combine or merge multiple data sources to provide support for decision making,(2) share relevant geo-referenced information and services, and (3) acquire spatial data in a cost efficient manner. For example, consider an urban planning application to schedule and evaluate the cost of constructing or repairing roads in a city. It typically involves the coordination and cooperation among the information systems used to manage the city's infrastructures such as road, traffic networks, or utilities (water, gas, electric and telephone).

Interoperability of spatial information systems is motivated by several factors includin: (1) sharing the existing large collections of data that are created and maintained by different institutions and government agencies for public use [20, 27, 62], and (2) data reusability to reduce data acquisition costs and avoid duplicating the effort of creating spatial data. In this paper we present an overview of the main issues and approaches underlying GIS interoperability. First, we discuss several important criteria for defining the major features of interoperable GIS. Then we present a taxonomy of interoperability approaches organized in three levels: platform, syntactic and application. For each level, we present the main objective, tools and concepts used to resolve heterogeneity conflicts. Several examples of solutions are given for each level.

The remainder of the paper is organized as follows. Section 2 presents different types of data conflicts. The next three sections are devoted to platform, syntactic and application interoperability respectively. Section 6 present a comparative study of the different approaches and concludes the paper.

2 Problematic of Interoberability

This section presents some issues of interoperability and a taxonomy of approaches in three levels: platform, syntactic, and application.

2.1 Properties

Several properties including autonomy, extensibility, transparency and composability have been used in the literature to describe and evaluate interoperable information systems [52, 28, 38].

Site autonomy. This property refers to the ability of a system to control its contents and operations, determine what data it can share with other participants and what non local queries it can accept and process. A component cannot be forced to comply with global processing or cooperation rules. Sheth in [56] distinguishes three different types of site autonomy: design, communication and execution autonomy.

Extensibility. It is the ability to add or remove systems to the interoperable environment. A component can freely join or leave an interoperable architecture. It can modify both schema and data contents to take into account changes in local processing requirements.

Transparency. It allows transparent data localization and access: a component can access remote data without a priory knowledge of the location or the representation (model and physical structures) of the data.

Composability. This issue pertains to query-content based interoperation. It represents the ability of a system to dynamically search and combine dynamically information that are useful for computing the queries. This property is very essential in dynamic and evolving networking environments such as the internet.

2.2 Data Conflicts

Data conflicts arise when there are conceptual, structural and semantic differences when the same entities are modelled by different information systems. Data conflicts of both traditional and spatial information systems have been discussed extensively in the litterature [3, 14, 55, 46, 24, 67, 33, 15].

Three main types of data conflicts can be distinguished. **Data model conflicts** are due to the diversity of data models used in spatial information systems. Generally, two different approaches can be used to represent spatial information. The field oriented approach is appropriate for modeling continuous or spatio-temporal data. It is based on a subdivision of space into regular (or non regular) cells associated with some characteristics of data. The vector oriented (object oriented) approach views space a set of objects and topological relations among them. **Schematic conflicts** are due to differences in model concepts and schema structures used to represent data. For example, an entity can be represented by one or more tables (relations) in the relational model whereas it is viewed as a complex structured object in the object oriented model. Schematic data conflict can also arise from differences in abstraction (generalization or specialization) links between objects. For example, a site can use only one concept to identify all types of roads whereas another site can differentiate between several types of roads, thus for example, using two concepts to distinguish *highway* from *Routes* that are both specialization of a more generalized concept *road.* concept. Finally, **Semantic conflicts** arise from differences in the meanings or interpretations associated with real world objects [14, 19, 55]. Different types of semantic conflicts can be distinguished. Name conflicts appear when two sites use different names or labels for the same entity. Naming conflicts are closely related to the taxonomic properties (synonymy, homonymy, polysemy, hypernymy, etc.) used by different systems. For example, the terms *bushy* and *forest* can be used to designate a wooded area and introduce synonymy conflicts. Value conflicts represent the second semantic conflicts category, they are composed of value naming conflicts, unit conflicts and value range conflicts (continuous or discontinuous interval). Spatial semantic conflicts result from differences in spatial localization and the geographical modeling of objects. Localization conflicts result from the positioning (direct or relative), the projection plan (Gauss, UTM, Lambert, etc.), the measurement units (kilometer vs. mile) and the precision scale (meter or centimeter).

2.3 Level of Interoperability

Different levels of interoperability have been proposed in the litterature. For example, Bishr [6] classifies interoperability approaches in six levels, ranging from network protocols to a semantic interoperability. Ouksel and Sheth [54] classify solutions in two levels: (1) a system level dedicated to resolving discrepancies among hardware and network functionalities; and (2) an information level devoted to the resolution of heterogeneity (model, schema and semantic). In this paper, we propose a taxonomy of spatial interoperability approaches based on the tools and concepts used to carry out the localization,discovery, understanding and management of distributed spatial data. We distinguish three levels of interoperability: platform, syntactic and application. Table 1 presents the issues addressed by each level. The different levels are described in detail in the next three sections.

Table 1. Taxonomy of spatial interoperability solutions

Interoperability Levels	**Related issues**
Plateform	System interconnection Data exchange formats
Syntactic	Data model conflicts Language conflicts
Application	Schematic conflicts Semantic conflits

3 Plateform Level Interoperability

The objective is to allow a collection of information systems to interoperate by transferring data from one system to another. Plateform level interoperability is based on 1) networking techniques and protocols such as TCP/IP, FTP, NFS, HTTP/CGI,CORBA, JAVA/RMI for exchanging data among different systems; and 2) translation tools for conversion between pairs of data formats. To cope with the large number of existing data models and reduce the number of required translators, data model translation is generally carried out in two steps. First, step a source data format is mapped to an intermediate common standard. Then the intermediate format is mapped to a target data format.

Different solutions have been proposed at this level. They can be classified in two main groups: ad-hoc and gateways approaches. The ad hoc approach consists of case tools that allow users to exchange data sets and to carry out multi-format translation (figure 1)of retrieved data. The gateway approach consists of software components to connect two or more information systems and allow one system to access data from other systems via a common protocol. They extend the ad-hoc approach by providing users with support to search, transfer and translate relevant data sets. Users are increasingly provided with web type interfaces to query one or more GIS and to view data via a web navigator (figure 2).

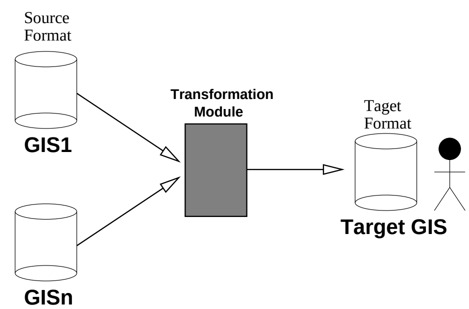

Fig. 1. Ad hoc approach to spatial interoperability

Meta-data can be used to facilitate searching and localization of relevant data sets by allowing explicit descriptions of the content(type, format, semantics) of data sources. Kashyap et al. [29] characterize meta-data by their role in abstracting and capturing the semantics of data irrespective of representation details. Several types of meta-data have been proposed, including: the CSDGM (Content Standard for Digital Geospatial Metadata)model defined by the FGDC (Federal Geographic Data Committee) [22],the CEN TC 287 [9] proposed by the European Commitee for Standardisation and ISO TC 211 model defined by the International Standard Organization[26].

Feature Manipulation Engine (FME) [23, 12] is an ad-hoc approach consisting of a multi-format translation tool for constructing interoperable architectures. It allows the conversion of data set from a format X to another format Y using an intermediate representation format. It defines a semi-automatic data conversion process in which translation scripts are created and and carried using various translation operations. These helper operations include 1) topological operations that verify the tological consistency of spatial data, 2) geometric operations for cartographical generalization (deletion of elements, geometric filtering according to critical treshold, conversion between geometric types, etc), and 3) non spatial operations for processing classical attributes (import of textual data associated with spatial entities).

COSIMA [40, 66] is an adhoc approach. It consists of a set of components that can be combined to achieve interoperable architecture. It defines two main types of services. Data access services provide a data manager, spatial storage manager and record storage manager. Processing services that define a set of operations on spatial data, consisting of geometric algorithms and statistical packages. The services are supportd by a spatial data model that contains a minimum set of data types (floating point, XYpoint, XYZpoint, string, label) and complex types constructors (Sequence, SetBag, Array, Record). COSIMA can be used in three different modes: the connected mode allows data exchange in batch processing mode, session mode implements transaction processing, and the query mode performs a single action on the server. The GEOserver prototype [39] is based on COSIMA. It implements simple 2D geometric types (point, line and polygon).

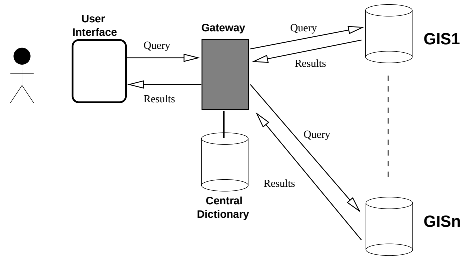

Fig. 2. Platform interoperability: gateway approach

GIS-WWW gateway [7] is a web type gateway approach. It allows a user to query different GIS and view the results of the query as maps expressed in JPEG or GIS image picture format. The gateway, called a switch, contains a data dictionary. It consists of simplified schema of each GIS and mappings between the gateway schema and the individual schema of the participants. Participants GIS are organized in themes (or thematic layers). Querying GIS-WWW consists in choosing thematic layers, selecting attributes corresponding to the layers, and displaying properties or requirements (colors,fonts, symbols ...). A user query is a set of pairs `(variable name, value)` corresponding to the elements of a schema stored in the switch. Depending on the chosen thematic layer, the user query is dispatched to the appropriate GIS for processing. Gateway queries are converted in the local systems by a dedicated translation tool that creates a sequence of (API) procedure calls to the local GIS. The results are pictures sent to the user display terminal. GIS-WWW is simple to implement, requiring the combination of few components. It provides a form based interface for accessing multiple GIS. However, only one GIS can be accessed at a time, and the user can not reuse the result of a query in subsequent queries.

GeoWEB [49] is gateway approach developped by NSDI (*National Spatial Data Infrastructure*). Its goal is to create a central repository (*Spatial Data Clearing-*

house) for localizing spatial data. Two types of data can be stored in the spatial clearinghouse: cartographical or statistical data (in USGS DLG, SDTS [21] or Arc/Info formats). GeoWEB contains describes data source in the form of meta-data represented in the standard *Content Standard for Digital Geographic Meta-data* defined by FGDC. User queries are expressed in terms of the meta-data to allow a search by keyword or by spatial selection. The interface also allows a user to define textual criteria (on thematic layers) and localization criteria (a point on a map). A query is processed by choosing a list of sites (addressed by URL) that contain relevant information. The implementation of GeoWEB is based on the WAIS (Wide Area Information Servers) [13] indexing scheme accessible through WEB browser.

4 Syntaxic Level Interoperability

The objective of syntactic level interoperability is to allow transparent access to collections of spatial data sources via schema expressed in a common data model. A typical example of syntactic level interoperability is the multidabase language interoperabibility approach proposed by Litwin et al [37] in which heterogenous information systems can be queried and accessed thrum a uniform interface. In the GIS field, syntactic level interoperability aims at resolving data model conflicts and requires: 1) common data models that include spatial primitives (OQL/ODMG [8], CEN/TC 287 [9], ISO/TC 211 [26], Open Geodata Model [45]) and 2) spatial transformation algorithms to resolve conflicts implied by the two representation modes of space (objects oriented vs field oriented). For example, Puppo [50] and Ramirez [51] describe a common data model that include both raster and vector capabilities. Piwowar et al. [48] describe algorithms to transform a raster representation to an object representation and conversely. Topological conflicts can be solved either by adding calculated relations to data (weak data model) [61] or using a common topological representation [5] [50].

Different solutions have been proposed and use technologies developed for distributed systems such as CORBA (Common Object Request Broker Architecture [43]) or object oriented language with network capabilities such as Java. Solutions of the syntactic level can be classified in two categories: (1) programming language based solutions which provide library of functions and homogeneous data structures, (2) high level languages based solutions which define a visual or declarative language to query schema provided by each source. The data model resolve the syntactic conflicts and the language provide operation to resolve manually schematic conflicts. Anyway, users have to localize data, understand the semantics of the spatial data and resolve schematic conflicts by using operators of the language. Sometimes metadata can help users in these tasks.

In the following studied solutions we try to highlight features used to resolve syntactic conflicts: common language, common data model and prototype implementation.

The Open Geospatial DataStore Interface (OGDI) solution [10] follows the way of Open Database Connectivity (ODBC) [44], an API which provides an

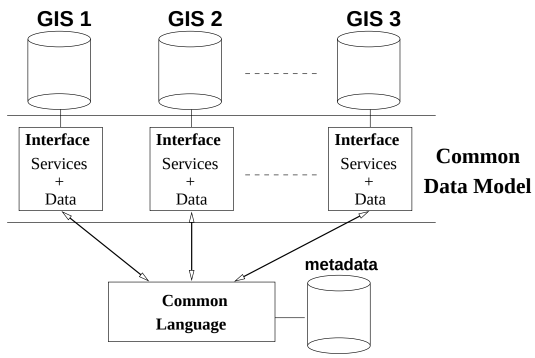

Fig. 3. Tools used by syntactic level interoperability

SQL access to database. OGDI uses drivers to access each spatial data source. It is based on the TCP/IP protocol and use Internet to distribute spatial data. OGDI provides a set of C functions including: 1) conversion of various format into a common data model called Transient Data Model, 2) transparent adjustment of coordinate system and cartographic projections, 3) transparent transformation of platform dependent representation and 4) retrieval of both spatial and attribute data. The syntactic conflicts are solved by the Transient Data Model which support both geometric and attribute data as a subset of DIGEST [17] Geometric data can be vector (point, line, area) or raster data. Metadata can be added to express spatial coverage or cartographic projection. OGDI implementation is free, it provides different drivers including Digital Terrain Elevation Data (DTED), ARC Digitized Raster Graphic (ADRG), GRASS, Arc/Info or GeoTIFF.

The Virtual Data Sets solution [63, 64] is close to OpenGIS approach. Virtual Data Sets (VDS) address problems of reliability and re-usability in field representations. VDS is based on an extension of a data set with methods that implement a model of the fields under consideration. A VDS contains methods to generate new representation or describes directly a new representation. Each VDS is able to connect to the source dedicated to relevant data. The VDS structure is implemented using a distributed object-oriented technology. Therefore each VDS is a distributed object which encapsulated services which can be queried by other applications to obtain data. A prototype based on Java is used to realize a distributed geoprocessing environment.

The Geographical Object Oriented heterogeneous Data Information Server (GEO2DIS) solution [58] implements a client-server systems for documenting, accessing, extracting and transforming spatial data. The client software lets users to navigate interactively in a catalog of geographical metadata. The client can formulate alphanumeric queries helped by metadata whose contain the following main characteristics: type (vector or raster), scale, last update, theme, structure, name of the producer and distributor. The graphic interface allows users to define spatial queries using a subset of spatial operations (intersect, fully within, zooming). The requests are sent to the relevant server which manages data. The server parses the generic request expressed in GeoQL, a spatial extension of the ODMG OQL, and translate it to the specific GIS language. Both O2 DBMS and ArcInfo translators have been developed. GEO2DIS also provides administration tools such as a metadata editor which allows to interactively create edit and delete description of data producers.

The Geographical Ante-server solution [60] aims at providing a completely open server architecture to connect and query heterogeneous information sources via a visual language interface. This language is hybrid: the user draws the configuration of the objects he is interested in and defines spatial constraints between these objects. Then he specifies the object type (town, roads, rail-roads etc.) and the thematic (textual) properties of the objects. Spatial objects are classified in point, line and zone. Different types of query analysers have been studied: 1) a simple topological and directional constraint analyser, 2) a meaning oriented analyser for metric or thematic relations and constraints, and 3) a deeper topological and directional constraints analyser. The first experimental application of the ante-server is a web server for thematic maps. It's architecture uses an hpptd daemon and a CGI interface to connect the httpd server and the ante-server. The ante-server is connected to different GIS using CORBA, proprietary API or Unix pipe. The visual language is implemented as a Java applet.

The OpenMap solution [**11**] is a distributed mapping system that allows displaying geographic data obtained from different data sources. OpenMap is a product suite developed by BBN technology in a DARPA sponsored project. It's goal is to demonstrate the capabilities of CORBA-based mapping environment. OpenMap specificies an interface between GIS and users. It includes a user interface client, a client/server interface implemented with CORBA and a set of specialists that implements the server interface, making available a particular kind of data source. The central component of OpenMap is the OpenMap Browser written in Java. It includes a map viewing area and navigation controls. A map is a collection of spatial objects which comes from data servers. The browser supports queries which perform a selection of objects in a rectangle. This objects are described by line, segments, circles, rectangles polylines, polygones or raster images. The coordinates can be specified either in latitude/longitude coordinates or in screen coordinates. The prototype implements a CORBA specialist to access SAND GIS.

5 Application Level Interoperability

The objective of application level interoperability is to provide a transparent interface which allow users to see information according to their domain or context of application. This level of interoperability takes into account localization, understanding and composability issues. Different solutions have been proposed which aim to resolve schematic and semantic conflicts. Solutions can be summarized to the federated approach which is based on schema integration and the mediation approach which is based on a more flexible and dynamic integration methodology.

5.1 The Federation Approach

The main characteristic is the definition of integrated schema which federate all information sources. Exported schemas are grouped into one or more schemas called federated schemas which allow a global view of the cooperation. The user can query federated schemas to access to information in a homogeneous way. Structural and semantic conflicts are resolved during the integration process. Sheth et Larson [56] propose a five levels federated architecture: local schemas, local schemas expressed in a common model, exported schemas, federated schemas and user views. Some recent works on GIS interoperation extend integration methodologies to spatial domain [31, 32, 47, 16]. Other works select methodologies especially adapted to spatial data, these ones are presented below.

Koschel et al. [30] propose a federated architecture based on distributed CORBA components. This is a web oriented architecture with a set of services and tools to help to construct federation of GIS. Horizontal services are dedicated to data access and to the definition of HTML user interface. Vertical services handle semantic aspects of data information and help users to find relevant data using a federated schema and metadata expressed in UDK format [36]. The query service is limited to the selection of one target region and some thematic levels.

Abel et al. [1] defines a virtual schema to integrate GIS sources. The integrated schema is represented in a common object oriented model. A research tool provides help to query virtual schema. The query process is composed of several phases : first the translation of the initial query expressed on the federated schema to sub-queries expressed on local schemas, then the query execution and optimization of these local queries on local GIS and last, the fusion of results from the different sites. A prototype implements the virtual GIS as a database containing the federated schema and mapping links to the components of the federation. The virtual GIS is a frontal component on which users can be connected to query the cooperation.

5.2 The Mediation approach

The mediation approach is an enhancement of the federated approach. Mediation simplificates the access and the understanding of multi-sources data in a cooperative environment [65]. Two software components are in charge to resolve the whole of data conflicts on users behalf: the **mediator** which resolves semantic conflicts and ensure the query process, the **wrapper** which resolves syntactic conflicts. A mediator can be used to convert units, to change data structures, to change or to translate name of entity, to classify objects and to group information according to semantic properties. A wrapper provides only an homogeneous access interface to information sources. It ensures query and data translation. Two types of mediation are used according to the management of semantic conflicts: (1) the schema mediation approach which resolves this kind of conflicts in a static way, (2) the context mediation approach which resolves them dynamically during the query process.

The Schema Mediation is based on a set of pre-established knowledge in the mediator which forms a mediation schema. This schema localizes useful information sources in a static way, it contains sufficient information to resolve schematic and semantic conflicts and mediator functionalities allows to combine and to restructure information from data sources into the mediation schema. The schema mediation has been extended to GIS interoperability by adding spatial concepts into the common mediation model, generally based on ODMG.

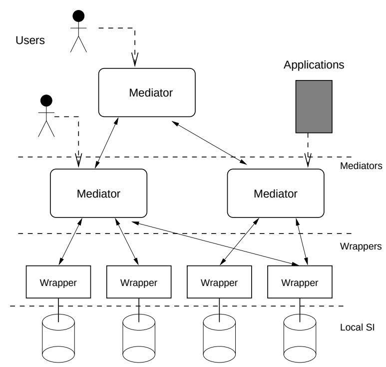

Fig. 4. Schema mediation architecture

Amann [**2**] uses a distributed object-based architectures on top of CORBA to connect spatial servers. This project has been initiated from a cooperation between the French Ministry of Culture and the CNAM to propose a mediation architecture which allows an interconnect access to spatial and multimedia documents. Costumer applications access to data sources using object mediator interface which provides an integrated view of systems (O_2, Postgres and mSQL). Interfaces are expressed in ODL language (Object Definition Language) which is a specification from ODMG-93 [8]. The geometry of spatial objects is described in a spaghetti model with primitives *point*, *line*, *polygon* and the topology of objects uses spatial operators. OQL requests allow to query the system through object interfaces. They hold spatial, alphanumeric and presentation criteria. Wrappers, mediators and interfaces communicates using ORBLine - a CORBA implementation.

The OASIS project [**41**] facilitates scientific data analysis, knowledge discovery, visualization and collaboration. Shared data are presented in a spatial object model which provides spatial types from OGM (OpenGIS [45]). OASIS proposes some many relevant services for scientific data: a catalog service to explore and localize available information, an integration and datamining service called Conquest which ensures parallel execution of user queries [53], a virtual warehouse service called geoPOM [42] which stores spatial data in a common data model. The communication architecture of OASIS uses CORBA to manage distributed objects.

The Geochange project [18] starts in 1996 to facilitate management of the growing urban development of Hong kong metropolis. Data from information sources are represented in a common model called CGDF (Common GeoChange Data Format). GeoChange provides a catalog of metadata which helps to transform and to localize data. User query are based on sessions which consist of: (1) browsing in a metada hierarchy, (2) building a user profile, (3) extract and fuse data which match this profile. GeoChange combine distributed object environment with federated approach in a loosely coupled and dynamic architecture.

The Context Mediation approach has been proposed to adapt the schema mediation to open and dynamic environments. The data integration process and the building of federated schema from many sources, which can evolve, appear and disappear, becomes quickly an impracticable design. Context mediation uses the Semantic of information to understand object meaning, to localize similar objects and to combine heterogeneous data representation in an homogeneous format which hold with user or application context. The semantic representation uses context and ontology features. **Context describes** explicitly the Semantic of information structure, information value, information relations and information functionalities. A context can be modelized using formula or logic rules, conceptual graphs and metadata or terms from ontology. **Ontology** is the conceptualization of a real world entity dedicated to an application domain [25]. An ontology consists of a set of classified terms which are linked together by semantic relations (homonym, synonymy, role). An ontology is used to homogenize the description language between systems. Ontology defines a common vocabulary to compare the semantic of shared entity.

In a context mediation approach, no integrated information is built a priori to define a cooperation of systems. These links are established dynamically during the query process using a context reconciliation mechanism. This process focus on the semantic hold by a query which allows to reconciliate contexts and to identify, to transform and to combine relevant data (figure 5). The mediator uses different conceptual tools (context, metadata, ontology) to provide semantic reconciliation and dynamic query resolution.

The SEMWEB project [6] uses context to explicit the semantic of information in the form of rules and constraints associated to object definitions. Local information are linked to a specific common context, called proxy context, to define the semantic of local data according to a common interpretation. A proxy context plays the role of the mediator by reconciling several contexts. It contains shared objects in a common context and mapping rules to transform local representation into this context. The object mediation model, called SFDS (*Se-*

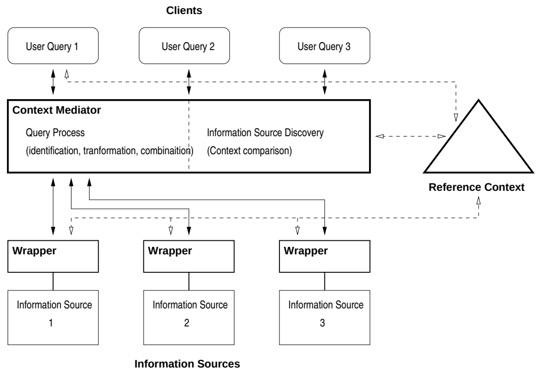

Fig. 5. Context mediation architecture

mantic Formal Data Structure), is composed of tree layers dedicated respectively to syntactic, schematic and semantic heterogeneity. The syntactic layer defines sufficient spatial primitives to represent simple geometrical objects and topologic relationships. The schematic layer uses a federated schema to resolve structural conflicts. The semantic layer adds context descriptions based on logic formula. An ontology helps to write context and to discover information.

The ISIS project [34, 35] is based on the reconciliation between one reference context and cooperation contexts. Each site interprets the reference context according to its own application domain. A generic ontology provides basic concept of language to build a context. The context mediation uses a three level architecture (ontological, cooperative and wrapper) and an extended spatial object oriented model called AMUN. A multi-level representation and interpretation allows ISIS to identify relevant sites, to focus on relevant information and to apply context transformations on data. An agent based architecture is defined which respects the three functional levels of the mediation [4]: an ontology agent defines common contexts, a broker agent provides a list of sites which match semantic hold by query and wrapper agents connect to data sources.

The SI^{3}CO project[57] (Spatial Information Infrastructure Interoperability Consortium) is intended to professional urban management (such as earthquake disaster planning system) and customer applications (embedded systems). This project results from a collaboration between two university, eight enterprises and NSDIPA (National Spatial Data Infrastructure Promoting Association in Japan). SI^{3}CO proposes a prototype based on a distributed object architecture (CORBA - VisiBroker). A three-tier mediation approach provides web client interfaces (applets), legacy database wrappers and Geo Spatial Mediator (GSM). GSM implements OpenGIS functionalities such as data localization, data transformation and data combination. GSM implements a negotiation protocol helped by metadata to find relevant information: at first, a control service of distributed retrieval receives requirements and determines relevant objects, then a spatial object trading service locates systems which contain these relevant objects, and finally a spatial object compensation service resolves schematic and semantic data conflicts.

6 Discussion and Conclusion

Figure 6 summarizes how the different solutions respect interoperability properties according to three dimensions: level of interoperability, properties or comparison criteria and solutions.

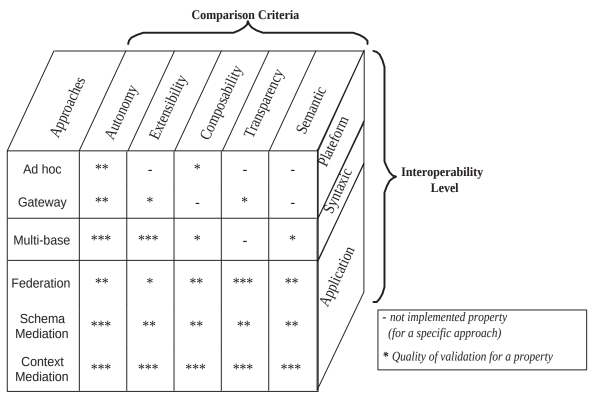

Fig. 6. Solutions and Basic Interoperability Properties

Autonomy is validated by most interoperability approaches. Generally, the participation of a system in a cooperation does not violate or modify its local functionalities. The validation of this property is achieved through a component (driver, wrapper) which is an interface between a data source and the other systems via a functional library (API) or a specific protocol (JDBC, ODBC).

Extensibility is easily achieved in most approaches. However, the schema integration based approach pay a high cost for allowing extensibility. In fact, the addition of a new system to the cooperative environment requires the definition of new integrated schema.

Composability is realized transparently in the federated approach. In the platform interoperability, the discovery and access to the relevant systems is a burden for the users.

Transparency is not validated to a satisfactory level by most approaches, particularly in the the platform and syntactic approaches. The mediation based approaches are most optimum in achieving this property. However, the transparency in the mediation depends on the component that carry out semantic similarities among the systems.

Figure 7 and 8 present an overviews of the approaches. Four comparisons criteria are used: modelization, query and access methods, implementation and semantics. The major points concerning these criteria are as follows. First, several categories of data model and query are found in the different approaches, ranging from traditional programming languages, to classical database models and legacy representation models.

Second, the implementation of most approaches is often based on low level structures and protocols. Distributed processing environment primitives such as Corba, Dcom and Java provide the foundations for most approaches. Finally, semantic information is used in different ways by the different approaches. They are used to assist information discovery in the cooperation. A user query contains semantic information that are compared with the semantics of local system to determine if a system is relevant to processing the query. They are also used in the integration step of the construction of federated schemas. The semantics are used to establish interschema links. And lastly they are used in dynamic query execution. This is mainly required for dynamic resolution of data conflict during the execution of a query. Mediation based solutions are the only approaches that uses semantic information dynamically.

In this paper, we have presented an overview of issues and approaches to achieve interoperable GIS. The focus of the discussion has been on the architectures, methodologies and the related tools to allow data sharing and exchange among heterogenous systems. Several important issues in the development of emerging cooperative information systems have been discussed. First, we have pointed the evolution of communication techniques and the development of WEB processing and the new requirements such systems are introducing in the development of information systems. Next, we have defined the major criteria and properties that are desirable in interoperable systems. We propose three levels of interoperability, including plateform, syntactic and application. Finally, we describe for each of the interoperability level, a generic approach that identify the objcetives, the tools and the architecure used to achieve interoperation. The advantages and shortfalls of each approach is discused. We conclude the paper with a comparaison study of the various approaches.

References

1. D. Abel, B. Ooi, K. Tan, and S. Huat Tan. Towards integrated geographical information processing. *International Journal of Geographical Information Science (IJGIS)*, 12(4):353–371, June 1998.
2. B. Amann. Integration gis components with mediators and corba. Technical report, Centre d'Étude et de Recherche en Informatique (CEDRIC), CNAM, Paris, 1997. ftp://sikkim.cnam.fr/pub/Reports/GISMed.ps.gz.
3. C. Batini, M. Lenzerini, and S. Navathe. A comparative analysis of methodologies for database schema integration. *ACM Computing Surveys*, 18(4):323–364, December 1986.

4. D. Benslimane. Systèmes d'informations coopératifs : une approche à base de médiation et d'agents. Habilitation à Diriger des Recherches, Université de Bourgogne, 1999.
5. M. Bertolotto, L. Floriani (De), and E. Puppo. Multiresolution topological maps. In M. Molenaar and S. Hoop (De), editors, *Proceedings of the Advanced Geographic Data Modelling Conference (AGDM'94), Spatial Data Modelling and Query Languages for 2D and 3D Applications*, number 40, pages 179–190. Delft, The Netherlands, September 1994.
6. Y. Bishr. *Semantic Aspects of Interoperable GIS.* PhD thesis, ITC Publication Series, Enschede Netherlands, ISBN 90-6164-1411, 1997.
7. C. Boehner, P. Haastrup, and A. Reggiori. The GIS-WWW gateway. In *Proceedings of the 3rd Joint European Conference and Exhibition on Geographical Information*, pages 125–134. Vienna, Austria, 1997.
8. R. Cattel., D. Barry, and D. Bartels, editors. *The Object Database Standard: ODMG 2.0.* Morgan Kaufmann, 1993.
9. CEN/TC287. *Geographic Information Reference Model, European Pre-standard*, 1996. http://old.ob.dk/ds/it/d2lsystem/128700.htm.
10. G. Clement, C. Larouche, D. Goiun, P. Morin, and H. Kucera. OGDI: toward interoperability among geospatial databases. *ACM SIGMOD Record*, 26(3):18–23, September 1997.
11. C. Cranston, F. Brabec, G. Hjaltason, D. Nebert, and H. Samet. Adding an interoperable server interface to a spatial database: Implementation experiences with openmap. In A. Včkovski, K. Brassel, and H-J. Schek, editors, *Interoperating Geographic Information Systems, 2nd International Conference, (INTEROP'99)*, pages 115–128. Zurich, Switzerland, Lecture Notes in Computer Science, Vol. 1580, Springer, ISBN 3-540-65725-8, March 1999.
12. R. Culpepper. Fme 2.2. *GeoWorld*, 1999.
13. F. Davis, B. Kahle, H. Morris, J. Salem, T. Shen, R. Wang, J. Sui, and M. Grinbaum. Wais interface protocol prototype functional specification, April 1990.
14. U. Dayal and H. Hwang. View definition and generalization for database integration in a multidatabse system. *IEEE Transaction on Software Engineering*, 10(6):628–645, 1984.
15. T. Devogele. *Processus d'intégration et d'appariement de bases de données géographiques - Application à une base de données routières multi-échelles.* PhD thesis, Institut Géographique National (IGN), Laboratoire COGIT, 1997.
16. T. Devogele, C. Parent, and S. Spaccapietra. On spatial database integration. *International Journal of Geographical Information Science (IJGIS)*, 12(4):335–352, June 1998.
17. DIGEST. The digital geographic information exchange standard (DIGEST 2.0). Technical report, Directorate of Geomatics Department of National Defence, Canada, June 1997.
18. P. Drew and J. Ying. GeoChange: an experiment in wide-area database services for geographic information exchange. In *Proceedings of the 3rd Forum on Research and Technology Advances in Digital Library, (ADL'96)*, pages 14–23. Washington DC, USA, IEEE Computer Society Press, ISBN 0-8186-7402-4, May 1996.
19. Y. Dupont. Problématique et résolution contextuelle des conflits de fragmentation dans l'intégration de schéma. *Ingénierie des systèmes d'information, Numéro spécial Bases de Données Avancées*, 3(1):29–58, 1995.
20. ETOPO5. 5-minute gridded elevations/bathymetry for the world. Data Announcement 93-MGG-01, 1993.

21. FGCS. Draft spatial data transfer standard (sdts), point profile. Technical report, Federal Geodetic Control Subcommittee, 1996. http://www.ngs.noaa.goc/FGCS/sdtsp.html.
22. FGDC. Content standard for digital geospatial metadata version 2, June 1998. FGDC-STD-001-1998.
23. FME. *Feature Manipulation Engine Overview.* Safe Software Inc., 1997. http://www.safe.com.
24. C. Goh, S. Madnick, and M. Siegel. Context interchange: Overcoming the challenges of large-scale interoperable database systems in a dynamic environment. In *Proceedings of the 3rd International Conference on Information and Knowledge Management (CIKM'94)*, pages 337–346. Gaithersburg, Maryland, USA, ACM Press, ISBN 0-89791-674-3, December 1994.
25. T. Gruber. A translation approach to portable ontology specifications. *International Journal of Knowledge Acquisition for Knowledge-based Systems*, 5(2), June 1993.
26. ISO/TC211/WG1. *Conceptual Schema Language (Working Draft 1.0)*, 1996.
27. M.E. James and S.N.V Kalluri. The pathfinder avhrr land data set: An improved coarse resolution data set for terrestrial monitoring. *International Journal of Remote Sensing*, 1994.
28. V. Kashyap and A. Sheth. Semantics-based information brokering. In *Proceedings of the 3rd International Conference on Information and Knowledge Management (CIKM'94)*, pages 363–370. Gaithersburg, Maryland, USA, ACM Press, ISBN 0-89791-674-3, December 1994.
29. V. Kashyap and A. Sheth. Schematic and semantic similarities between database objects: A context-based approach. *VLDB Journal*, 5(4):276–304, 1996.
30. A. Koschel, R. Kramer, R. Nikolai, W. Hagg, J. Wiesel, and H. Jacobs. A federation architecture for an environmental information system incorporating gis, the world wide web and corba. In *3rd International Conference Workshop on Integrating GIS and Environmental Modeling*. National Center For Geographic Information and Analysis (NCGIA), Santa Fe, New Mexico, USA, January 1996. http://ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/program.html.
31. R. Laurini. Raccordement géométrique de bases de données géographiques fédérées. *Ingénierie des Systèmes d'Information*, 4(3):361–388, 1996.
32. R. Laurini. Spatial multidabase topological continuity and indexing: a step towards seamless gis data interoperability. *International Journal of Geographical Information Science (IJGIS)*, 12(4):373–402, June 1998.
33. R. Laurini and F. Milleret-Raffort. Distributed geographical databases: some specific problems and solutions. In *Proceedings of the 7th Parallel and Distributed Computing Systems (PDCS'94)*, volume 2, pages 276–283. Las Vegas, International Society for Computers and their Applications, October 1994.
34. E. Leclercq, D. Benslimane, and K. Yétongnon. Amun: An object-oriented model for cooperative spatial information systems. In *IEEE Knowledge and Data Engineering Exchange Workshop, Newport Beach, USA*, pages 73–80. IEEE Computer Society, ISBN 0-8186-8230-2, November 1997.
35. E. Leclercq, D. Benslimane, and K. Yétongnon. Semantic mediation between cooperative spatial information systems:the Amun data model. In *6th Advances in Digital Libraries (ADL'99)*, pages 16–27. Baltimore, USA, IEEE Computer Society, ISBN 0-7695-0219-9, May 1999.
36. H. Lessing, W. Swoboda, and O. Günther. Udk : A european environmental data catalogue. In *3rd International Conference Workshop on Integrating GIS and En-*

vironmental Modeling. National Center for Geographic Information and Analysis (NCGIA), Santa Fe, New Mexico, USA, January 1996.
37. W. Litwin, L. Mark, and N. Roussopoulos. Interoperability of multiple autonomous databases. *ACM Computing Surveys*, 22(2):265–293, June 1990.
38. L. Liu and C. Pu. The diom approach to large-scale interoperable database systems. Technical Report TR95-16, Department of Computing Science, University of Alberta, 1995.
39. M. (De) Lorenzi. The xyz geoserver for geometric computation. In J. Nievergelt, T. Roos, H-J. Schek, and P. Widmayer, editors, *Geographic Information Systems, International Workshop on Advanced Information Systems (IGIS'94)*, pages 202–213. Lecture Notes in Computer Science, Vol. 884, Springer 1994, ISBN 3-540-58795-0, February-March 1994.
40. M. (De) Lorenzi and A. Wolf. A protocol for cooperative spatial information managers. In *Proceedings of DBTA Workshop on Interoperability of Database Systems and Database Applications*, 1993.
41. E. Mesrobian, R. Muntz, E. Shek, S. Nittel, C. Larouche, and M. Kriguer. Oasis: An open architecture scientific information system. In *6th International Workshop on Research Issues in Data Engineering (RIDE'96) - Interoperability of Nontraditional Database Systems*, pages 107–116. New Orleans, Lousiana, USA, IEEE Computer Society Press, ISBN 0-8186-7289-7, February 1996.
42. S. Nittel, R. Muntz, and E. Mesrobian. geoPOM: a heterogeneous geoscientific persistent object system. In *9th International Conference on Scientific and Statistical Database Management (SSDBM'97)*. Olympia, Washington, USA, IEEE Computer Society Press, August 1997.
43. Object Management Group (OMG). *CORBA 2.0 specifications*, 1996.
44. ODBC. *Microsoft Open Database Connectivity Software Development Kit: Programmer's Reference*. Microsoft, Redmond, Washington, 1992.
45. OGC. The opengis abstract specification: An object model for interoperable geoprocessing. Technical Report 96-015R1, OpenGIS Consortium and the OGIS Technical Committee, 1996.
46. A. Ouksel and C. Naiman. Towards the design of semantic cooperation protocol in heterogeneous database systems. In H. Schek, A. Sheth, and B. Czejdo, editors, *3th International Workshop on Research Issues in Data Engineering (RIDE'93) - Interoperability in Multidatabase Systems*, pages 184–187. Vienna, Austria, IEEE Computer Society Press, ISBN 0-8186-3710-2, April 1993.
47. C. Parent, S. Spaccapietra, and T. Devogele. Conflicts in spatial database integration. In K. Yétongnon and S. Hariri, editors, *Proceedings of the 9th Parallel and Distributed Computing Systems (PDCS'96)*, volume 2, pages 772–778. Dijon, France, International Society for Computers and their Applications, ISBN 1-880-843-17-x, September 1996.
48. J. Piwowar, E. Ledrew, and D. Dudycha. Integration of spatial data in vector and raster formats in a geographic information system environment. *International Journal on Geographical Information Systems (IJGIS)*, 4(4):429–444, 1990.
49. B. Plewe. The GeoWeb project: Using WAIS and the World Wide Web to aid location of distributed data sets. In *2nd International World-Wide Web Conference Mosaic and the Web*. Chicago, USA, National Center for Supercomputer Applications, University of Illinois at Urbana-Champaign, October 1994. http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/Autools/plewe/plewe.html.
50. E. Puppo and G. Dettori. Towards a formal model for multi-resolution spatial maps. In M. Egenhofer and J. Herring, editors, *Advances in Spatial Databases,*

4th International Symposium, (SSD'95), pages 152–169. Portland, Maine, USA, Lecture Notes in Computer Science, Vol. 951 Springer, ISBN 3-540-60159-7, August 1995.

51. R. Ramirez. Development of a common framework to express raster and vector datasets. In *Auto Carto*, pages 155–163, 1997.
52. E. Sciore, M. Siegel, and A. Rosenthal. Using semantic values to facilitate interoperability among heterogeneous information systems. *ACM Transactions on Database Systems (TODS)*, 19(2):254–290, 1994.
53. E. Shek, E. Mesrobian, and R. Muntz. On heterogeneous distributed scientific query processing. In *6th International Workshop on Research Issues in Data Engineering (RIDE'96) - Interoperability of Nontraditional Database Systems*, pages 98–106. New Orleans, Lousiana, USA, IEEE Computer Society Press, ISBN 0-8186-7289-7, February 1996.
54. A. Sheth. *Interoperating Geographic Information Systems*, chapter Changing Focus on Interoperability in Information Systems: from System, Syntax, Structure to Semantics. Kluwer Academic Publishers, ISBN 0-7923-8443-69, 1999.
55. A. Sheth and V. Kashyap. So far (schematically) yet so near (semantically). In D. Hsiao, E. Neuhold, and R. Sacks-Davis, editors, *Proceedings of the IFIP WG 2.6 Database Semantics Conference on Interoperable Database Systems (DS-5)*, pages 283–312. Lorne, Victoria, Australia, North-Holland, 1993, ISBN 0-444-89879-4, November 1992.
56. A. Sheth and J. Larson. Federated database systems for managing distributed, heterogeneous and autonomous databases. *ACM Computing Surveys*, 22(3):183–236, September 1990.
57. S. Shimada and H. Fukui. Geospatial mediator functions and container-based fast transfer interface in si3co test-bed. In A. Včkovski, K. Brassel, and H-J. Schek, editors, *Interoperating Geographic Information Systems, 2nd International Conference, (INTEROP'99)*, pages 265–276. Zurich, Switzerland, Lecture Notes in Computer Science, Vol. 1580, Springer, ISBN 3-540-65725-8, March 1999.
58. Intecs Sistemi. *GEO2DIS: a Client-Server Architecture on Internet to Document and Access Geodata Stored on Heterogeneous GISs*, 1997. `http://www.pisa.intecs.it/project/GEO2DIS`.
59. L. Spery and T. Libourel. Vers une structuration des métadonnées. *Revue Internationale de Géomatique - Actes du congrès CASINI*, 8(1-2):59–74, Novembre 1998.
60. M. Szmurlo, M. Gaio, and J. Madelaine. The geographical anteserver: A client/server architecture for gis. In J. Strobl and C. Best, editors, *Proceedings of the Earth Observation and Geo-Spatial Web and Internet Workshop*, 1998.
61. T. Ubeda and M. Egenhofer. Topological error correcting in gis. In M. Scholl and A. Voisard, editors, *5th International Symposium on Advances in Spatial Databases, (SSD'97)*, pages 283–297. Berlin, Germany, Lecture Notes in Computer Science, Vol. 1262, Springer, ISBN 3-540-63238-7, July 1997.
62. USGS. Global land information system, 1995. http://edcwww.cr.usgs.gov/glis/glis.html.
63. A. Včkovski. Java as a software system for distributed and interoperable geoprocessing. In K. Yétongnon and S. Hariri, editors, *Proceedings of Parallel and Distributed Computing Systems (PDCS'96)*, volume 2, pages 779–783. Dijon, France, International Society for Computers and their Applications, ISBN 1-880843-17-x, September 1996.
64. A. Včkovski. *Interoperable and Distributed Processing in GIS*. PhD thesis, University of Zurich, Switzerland, 1998.

65. G. Wiederhold. Mediators in the architecture of future information systems. *IEEE Computer Magazine*, 25(3):38–49, March 1992.
66. A. Wolf, M. Lorenzi (De), T. Ohler, and V.H. Nguyen. Cosima, a network based architecture for gis. In Jürg Nievergelt, Thomas Roos, Hans-Jörg Schek, and Peter Widmayer, editors, *International Workshop on Geographic Information Systems*, pages 192–201. Lecture Notes in Computer Science, Vol. 884, Springer 1994, ISBN 3-540-58795-0, February 1994.
67. M. Worboys and S. Deen. Semantic heterogeneity in distributed databases. *ACM SIGMOD Record*, 20(4):30–34, 1991.

Criteria \ Solutions	Platform Level				SyntaxicLevel				
	FME	COSIMA	GIS-WWW	GeoWeb	OGDI	VDS	GEO2DIS	OpenMAP	Anté serveur
Modelization									
Spatial Model	legacy	vector	bitmap	vector raster	DIGEST	raster	vector raster	vector raster	vector raster
Thematic Model	relational	minimal (num. value)	*NA*	FGDC	C	VDS (OO)	ODMG	*NA*	OO
Domain	-	geometric calculus	-	-	GRASS, VRF	geoscience	urban data	cartography	documentary
Query									
Language	scripts	C	form	form	C	Java	visual GeoQL	Java	visual spatial constraints
Interface	C/C++	*NA*	Web	Web	API	API	graphic	graphic	graphic
Implementation									
Architecture	*NA*	client / server	3-tiers switch	3-tiers switch	client / server	3-tiers	client / server	3-tiers	client / server
Technology	*NA*	RPC / XDR	HTTP / CGI	HTTP / CGI	TCP / IP socket	Java RMI	Corba / XML	Corba	HTTP / CGI / Corba
Prototype	business	research	operational	operational	business	research	operational	research	research
Semantic									
Objectives	-	-	-	research	-	-	*NA*	-	research
Method	-	-	metadata (dictionary)	metadata (dictionary)	-	-	metadata (CEN TC287)	-	-

Fig. 7. Comparison of Interoperability Solutions

Application Level								Solutions / Criteria
Federation		Schema Mediation			Context Mediation			
Koschel	Abel	Amann	OASIS	Geochange	SEMWEB	ISIS	SI3CO	
								Modelization
vector raster	vecteur	spaghetti	OpenGIS	CGDF	vector	vector raster	OpenGIS	Spatial Mode l
relational	OODM	ODMG	ODMG	ODMG	OO	Amun (OO)	ODMG	Thematic Model
GRASS, GNUPLOT	-	-	geoscience (HDF, HIRS/MSU, AVHRR)	urban data	-	-	-	Domain
								Query
MSQL	*NA*	OQL / C++	ODL / SQL / C++	SQL+	SQL	OQL / Java / KQML	*NA*	Language
-	-	graphic	Web	*NA*	graphic	no	*NA*	Interface
								Implementation
federation	federation	mediator / wrapper	mediator / wrapper	mediator / wrapper	mediator / wrapper	mediator / wrapper / agent	3 levels	Architecture
Corba / HTTP	-	Corba	Corba / HTTP / CGI	*NA*	*NA*	Corba / Java	Corba	Technology
research	research	operational	operational	research	research	research	prototype	Prototype
								Semantic
conflicts	conflicts	conflicts	research	research	research / conflicts	research / conflicts	conflicts	Objectives
integration + metadata s UDK	integration	integration	catalogue	metadata / ontology	RDM	ontology / contexts	metadata	Method

Fig. 8. (Continued) Comparison of Interoperability Solutions