

Processing of spatial metadata queries within federated GISs

Steve Ramroop and Richard Pascoe

Department of Information Science,
University of Otago, Dunedin, New Zealand.
Phone: +64 3 479 5608 Fax: +64 3 479 8311
Email: sramroop, rpascoe @infoscience.otago.ac.nz

Presented at SIRC 99 – The 11th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
December 13-15th 1999

ABSTRACT

A federated GIS is a collection of heterogeneous GIS that collectively act as a single source of geographical information. Spatial metadata plays an important role in accessing information from a federated GIS and furthermore, is seen as the key to achieving interoperability. Metadata usually conforms to a standard such as the Australia New Zealand Land Information Committee¹ (ANZLIC) standard; the Content Standard for Digital Geospatial Metadata² (CSDGM) from the Federal Geographic Data Committee³ (FGDC); and defacto standards such as the Facility for Interrogating the National Directory of Australian Resources (FINDAR). In this paper, we describe our research into strategies for processing metadata queries using a variety of metadata standards within a federated GIS, rather than restricting metadata to conform to a single standard. Using the two strategies, an analogy is drawn with Pascoe & Penny's (1990) work on interfacing strategies. In this analogy the two query processing strategies is compared to the interfacing strategies and conclusions are drawn regarding the appropriateness of each query processing strategy.

Keywords and phrases: Processing, metadata queries, strategies, federated GISs

1.0 INTRODUCTION

The sharing of spatial data has always been an area of interest by GIS users since sharing is a cheaper alternative to duplicate acquisition of data sets. In homogeneous GISs where data sets adhere to the same schema, the technical acquisition of data sets is not a problem. Realistically, the sharing of data sets in most instances involves users of a federation of heterogeneous GIS, each GIS having data sets conforming to schemata unique to that GIS. Such a federation is a goal of the OpenGIS whereby 'transparent access to heterogeneous geodata and geoprocessing resources in a networked environment' are provided through 'a comprehensive suite of open interface specifications that enable developers to write interoperating components that provide these capabilities', (OpenGIS, 1999).

Within a federation of GIS there are two key elements to achieving a basic level of interoperability and data sharing. First, and the focus of discussion in this paper, is providing users with a method of locating any existing data sets within the federation appropriate to their needs. Second, the transformation of these sets to conform to the schema required by the user. Research into interoperability for example, Mena et al. (1998); and Kildow (1996), indicates that metadata and more specifically metadata standards, are the key to successful interoperability and data sharing. Metadata and its role in achieving interoperability is discussed in Section 2.0

¹ <http://www.environment.gov.au/database/metadata/anzmeta/>

² <http://www.fgdc.gov/Metadata/ContStan.html>

³ <http://www.fgdc.gov/index.html>

There are many metadata standards being adopted by users for example, some of the popular standards are ANZLIC; CSDGM; Dublin Core⁴ (DC); defacto standard such as FINDAR; and so on.

The problem of having various metadata standards is compounded for users in a federated GIS because metadata is stored in different mediums, such as directories, databases, and text files. Since there are different mediums used to store metadata, there are many research approaches making use of metadata. In this paper a review in the use of metadata for describing spatial data is presented in Section 3.0 where their concepts, and spatial metadata query strategies are briefly discussed. These research approaches are compared in Section 4.0 in terms of the various metadata storage mediums, their query processing strategies, and their general operational context which can have a general or specific focus. The comparison forms the foundation upon which the earlier research work of Ramroop & Pascoe (1999), and (1998) is founded where the essential framework of their Virtual Data Centre was developed. In this paper a component of their model called the *Selector Broker* is investigated. The query processing strategies for such a broker is defined in this paper.

Lightweight Directory Access Protocol (LDAP) is being used by companies such as MessagingDirect which provides LDAP directory software products for Extranet communication; the IBM eNetwork LDAP provides an easy way to manage access to a company's range of computing platforms; and so on. According to the comparison of research making use of directory services, there has been very little research into the use of LDAP to store and query spatial metadata. In addition there is the need to address query strategies which is capable of recognizing the multiple spatial metadata standards already existing.

Based upon the comparison, two new strategies are investigated for processing spatial metadata queries within federated GISs. These strategies, discussed in Section 5, are called *unified schema* and *query translation*. In the *unified schema* strategy metadata elements from different standards are matched to one common ontology while in the *query translation* strategy the user defines a query using a familiar metadata standard. In both strategies queries are translated into the different metadata standards and results are reported back to the user. In this research the ANZLIC and CSDGM standards are used for practical experimentation.

2.0 METADATA

A well-known and simple definition of metadata is, 'data about data' (ANZLIC, 1996; Kildow, 1996; Williamson & Ezigbalike, 1998). Rossmeissl (1998) stated that metadata is critical information used to document and organize an organization's internal investment in spatial data and to provide information to process and interpret data received through a transfer from an external source. This definition provides more insight into the purpose of metadata from an organization's perspective. Simplistically, metadata is a set of elements describing each data set. Shah & Sheth's (1998) definition concurs with other definitions in that metadata represent information about the data in individual databases and data repositories. This common understanding of what is metadata can be refined to distinguish among different types of metadata as described next.

2.1 Metadata types

The types of metadata which has been developing through the years are metadata associated with software, data, and now metadata describing spatial processes. Wilson (1998) in Williamson & Ezigbalike (1998) discussed that the first, and oldest type of metadata occurs within GIS, CAD packages, and databases. In these systems there is an underlying set of rules that guides software programs of how to handle data. Similarly, DBMSs have long used metadata to describe the internal layout of the data schemas within them. Williamson & Ezigbalike (1998) explained that this description of the internal schemas is then used by DBMS to provide results to user queries. Therefore, the first type of metadata is the system's internal metadata used to assist the DBMS to locate results for queries.

The second type and most common of metadata is one where metadata sets are used to describe the characteristics of data sets which is particularly important in the management of spatial data. Williamson & Ezigbalike (1998) listed elements describing metadata characteristics such as custodian, geographic extent, data currency, data quality, contact information, and so on. These elements assist users in building queries defining data descriptions that satisfy their needs.

⁴ <http://purl.oclc.org/dc/>

The third type of metadata stems from investigation surrounding issues associated with interoperability. From the inception of the metadata paradigm, metadata was considered to be 'data about data', however, more recently there has been a change in this paradigm. This change is extended to include 'data about processes'. Therefore, the third type of metadata incorporates the characteristics of the first two types of metadata. One particular research in this area is [Marr et al. \(1998\)](#).

The relatively new concept, Meta Object Facility⁵ (MOF) is a set of CORBA interfaces that can be used to define and manipulate a set of interoperable metamodels. MOF is used as a meta-information repository within CORBA distributed systems. This facility can be used for incorporating metadata for spatial processes. However, in this research the focus is on spatial metadata stored using directory services and not on metamodels associated with CORBA interfaces. This choice was made because the use of directories is not fully explored and is a different approach to CORBA. To support data sharing and interoperability, research into metadata is needed particularly in the unification of existing multiple metadata standards.

2.2 Metadata standards

There is a number of metadata standards being used at present. For example ANZLIC; CSDGM; DC; defacto standards such as FINDAR; and so on. Each metadata standard consists of various elements describing the spatial data which is stored. In all of these standards information of one or more of the following are included in their metadata element descriptions:

- detailed information about data collection methods, integration and analysis techniques applied to various components of source data to support the preparation of scientific reports;
- information about the accuracy of source data sets, processing history, and archival procedures to effectively manage and utilise data within custodian organisations;
- information about projection specifications, scale, and data dictionary to accompany transfers to other organisations;
- adequate descriptions of the content, quality and geographic extent of data sets so potential users of existing data can assess its suitability for other purposes; and
- summary descriptions of content and quality as well as contact information for inclusion into directory systems.

As discussed in Section 2.1, to support data sharing and interoperability the second type of metadata best expresses these elements. Some research examples are: ([Abel et al., 1998](#)), ([Johnson et al., 1991](#)), ([ANZLIC, 1998](#)), and so on. The preferred approach for data sharing is by utilizing metadata describing the semantics of spatial data sets. However, metadata standards differ in their content therefore, mechanisms are needed to address these variations. For example CSDGM contains 220 elements while ANZLIC contains 32 elements. Mapping between metadata standards can be a difficult process to map similar metadata elements. In the following section is discussed particular research which utilize metadata to describe spatial data sets.

3.0 USE OF METADATA TO DESCRIBE SPATIAL DATA

The topic of *Integration* is a highly researched area. Much research is done either with a focus on the area of data integration or process integration although the two are inseparable when discussing interoperability. At the same time there is research that addresses related topics (such as data search, data standardisation, data transformation, and so on) which when put together provides the building blocks in achieving a total integration solution.

Most of the research reviewed is focussed on establishing mechanisms capable of identifying and/or transforming existing data sets. [Walker et al. \(1992\)](#) indicated that, in light of the widespread use of GIS and the consequent rapid increase in the number of data sets being created, the problem of integration is unlikely to be solved by using existing database technology to hold information about available data sets but instead requires a different approach. This approach is through the use of metadata.

In the following sections is presented a sample of recent metadata research briefly discussing their use of metadata and general characteristics. The discussion is under two categories: Ontologies; and Directory systems.

⁵ <http://www.dstc.edu.au/Meta-Object-Facility/>

3.1 Ontologies

Wiederhold (1994) defined ontology as a vocabulary of terms and a specification of their relationships. Mena et al. (1998) defined an ontology as a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. Ontologies and the interontology relationships between them are created by experts in the corresponding domain. To do this ontologies are expressed using Descriptive Logics (DL). The role in using ontology as Mena et al. (1998) discussed is one which supports semantic information brokering for providing access to numerous data repositories accessible on the Global Information Infrastructure (GII) enabled by the Web.

An ontology can be viewed as the description of concepts and relationships based upon a common vocabulary. Ontologies differ and are dependent upon their domains within which they are created. Systems making use of ontologies are general in their applications. Even though ontology is a relatively new concept there has been a few examples that are worth mentioning.

3.1.1 Ontology examples

There is a number of systems being researched that are using the ontology concept. One common example is OBSERVER⁶ (Ontology Based System Enhanced with Relationships for Vocabulary hEterogeneity Resolution). Mena et al. (1998) explained that information requests to OBSERVER are specified as a Description Logic (DL) expression based on terms chosen from some domain ontology.

Mena et al. (1998) indicated that OBSERVER supports multiple pre-existing ontologies to access heterogeneous, distributed and independently developed data repositories. The content of each data repository is described by one or more ontologies expressed in a system based on DLs. Information requests in OBSERVER are specified using concepts in a domain ontology chosen by the user. OBSERVER uses ontological inferences to determine relevant data repositories and translated DL expressions to the local query languages of the relevant data repositories. The query processing allows controlled expansion of a user's query to involve other ontologies using an extended set of relationships including synonyms, homonyms and hypernyms.

Another ontology example is InfoQuilt which is based upon the Info Harness system⁷ (Shah & Sheth, 1998). Since these ontology systems are general, their metadata query processing strategies are also general and can be applied to directory services. Consequently, the Selector Broker in Ramroop & Pascoe's (1999) model adheres to the concept of ontologies by having a unified metadata vocabulary.

3.2 Directory systems

Directory systems provides the mechanism to inform users of data set(s) descriptions. Walker et al. (1992) identified four types of directory system. They are:

- Directories holding information about data sets held on one workstation, often structured for use within one GIS software package.
- Directories holding information on datasets within a single organisation.
- Directories holding information about data sets available for a particular country.
- Directories holding information about data sets that apply for different countries, continent and the world internationally.

Walker et al.'s (1992) categories of directory systems presents a definite hierarchy of directory systems which is influenced by the extent of their implementation. The simplest being a stand alone workstation while the more challenging is the international implementation. There are a number of research investigating the use of directory systems. The following presents research that makes use of directories. Their characteristics are briefly described following which a comparison is made.

⁶ <http://siul02.si.ehu.es/jirgbd/OBSERVER/>

⁷ <http://lstdis.cs.uga.edu/proj/iq/iq.html>

Data sharing using the X.500 directory

Pascoe (1996) described the notion of an agency for sharing geographical data. The agency is based on the use of the ISODE implementation of the *X.500 Directory Service* and a collection of software agents which collaborate with each other to perform the various task associated with sharing data.

From the ISODE (1994) user guide, the X.500 Directory Service is a facility for storing and retrieving widely distributed information. Pascoe (1996) stated that the directory is hierarchically structured in the form of a Directory Information Tree (DIT) with branches of the tree stored on different computer systems and independently administered.

Pascoe (1996) discussed an object class associated with geographical data sets. He claims that within this class there is a hierarchy where there will be instances of various application object classes corresponding to different types of applications such as agents, interfaces, and other data processing software packages. Pascoe (1996) suggested a paradigm switch such that an interface which is software traditionally used for converting data from one format to another, becomes one or more specialised software agents. He indicated that in the X.500 directory there are two types of agents: a Directory User Agent (DUA) which formulates queries, wrap them in the required protocol, pass them on to directory systems agents, and display the results obtained; and a Directory System Agent (DSA), which responds to DUA directory searches. He stated that all of these agents would be described in the geographical data directory and would collectively form an agency for sharing geographical data.

ASDD

ANZLIC developed a distributed data directory prototype containing 32 data items grouped into 10 categories. The directory is know as the *Australian Data Directory (ASDD)*⁸. ANZLIC (1997) stated that the prototype uses the public domain *Isite Z39.50* software to do a live search of metadata held at the three distributed sites. Currently, the three distributed sites are accessed through a single web interface⁹. The software developed makes use of Standard Generalized Markup Language (SGML) structured using the Document Type Definition (DTD) to incorporate element guidelines. By using eXtensible Markup Language (XML), accessibility through the WWW is made possible.

FINDAR

Another directory system is the *Facility for Interrogating the National Directory of Australian Resources, (FINDAR)*. Johnson et al. (1991) stated that the FINDAR directory meta-model is a database structure that allows the attributes of any metadata description to be defined dynamically by the user so that they can be immediately displayed and searched without rebuilding the database. The attribute table contains information regarding the attribute identifiers, type, and value. Metadata elements are owner identification, data items, spatial identification, spatial coverage, data set information, currency, lineage, quality, ordering, and organization information. The software consists of three components: tables of attribute, keyword and spatial data for the directory entries; a gazetteer of geographic entities for spatial indexing and searching; and a thesaurus of standard terms for subject indexing and searching. The software is executed in several places with different versions of the attribute tables.

To perform spatial search, entities are limited to a spatial descriptor which searches for a pre-defined spatial database with defined place names. Searching is done using the *SIRO-DBMS* software developed by the *Division of Information Technology in Australia's Commonwealth Scientific Industrial Research Organization*, together with a gazetteer of pre-defined and user-defined spatial definitions. Johnson et al. (1991) stated that all spatial data, comprising of directory spatial coverages, the gazetteer and search polygons are held internally in ORACLE tables in SIRO-DBMS format.

Isaac

Project Isaac is the Internet's Scout Project¹⁰ funded through the National Science Foundation. For this project there was a need to assist with the development of resource discovery on the Internet through the use of *Scout Reports*. A Scout Report is a weekly publication offering a selection of new and newly discovered Internet resources of interest to researchers and educators. The *Scout Report Signpost* contains only the best Internet

⁸ <http://www.environment.gov.au/net/asdd/>

⁹ http://www.environment.gov.au/servlet/asdd_basic

¹⁰ <http://scout.cs.wisc.edu/index.html>

resources, as chosen by the editorial staff of the Scout Report, which have been catalogued and organized for efficient browsing and searching. Roszkowski & Lukas (1998) indicated that Signpost consists of more than 2000 catalogued Internet sites using established standards such as Library of Congress subject headings and abbreviated call letters, and emerging standards such as the Dublin Core (DC).

Roszkowski & Lukas (1998) indicated that Project Isaac's infrastructure utilizes standard Internet protocols, such as the Lightweight Directory Access Protocol (LDAP) and the Common Indexing Protocol (CIP)¹¹ to distribute queries, return results, and exchange indexes. Allen & Mealling (1998) indicated that CIP is used to pass indexing information from server to server in order to facilitate query routing. They stated that query routing is the process of redirecting and replicating queries through a distributed database system towards servers holding the desired results.

The metadata servers use LDAP for query processing and the storage and retrieval of index information and CIP Version 3 to generate and transport indexes between servers. There are two types of Isaac installations: a 'full node' which supports all of the services (metadata repository, index and search) and a 'collection node' that supports only the metadata repository service.

Allen & Mealling (1998) explained that Isaac consists of three services, a metadata repository service, an index service, and a search service. They added that the metadata repository service uses a stand-alone LDAP server to store the metadata for a collection. The service supports LDAP queries against its metadata collection. The metadata repository service can also generate indexes of the metadata and transfer the index using the CIP protocol. The metadata repository can also extract index information from the collection and transfer the index using the CIP protocol. The index service gathers index information from other Isaac nodes and stores the index information in LDAP database.

Allen & Mealling (1998) stated that the search service is Web-based utilizing HTTP and HTML. Using a browser, the user fills out a form with query information and a CGI script handles the form. The CGI script uses an LDAP client to submit the query to the local LDAP server on a 'full node,' and the LDAP client processes any referrals (based on the index information) before results are returned to the user.

Virtual Data Centre

In Ramroop & Pascoe's (1998) work, their main focus is representing data integration in a GIS environment. The problem of GIS data integration is briefly discussed and a Conceptual Model, comprising of a *Data Centre* and *Data Agencies* is presented. The problem of data integration is divided into manageable and modular components and the notation proposed by Pascoe & Penny (1995) is extended to document these components. The notation attempts to present the retrieval and processing of the data starting from the user query to that of the final data destination.

In Ramroop & Pascoe's (1999) work, their main focus was presenting the implementation architecture of their conceptual model for data integration in a GIS environment. The model was briefly discussed and extensions to the model were made. These extensions is also applied to the notation Ramroop & Pascoe (1998) used to denote the entire data integration process. Details of the actual processing steps followed within the Data Centre were discussed and the general architecture design was described. The implementation of the concept is modular, starting with the first service called the Selector Broker which is responsible for searching for spatial metadata.

After presenting a brief overview of systems making use of metadata describing metadata, a comparison between them is presented in the next section.

4.0 COMPARISON

Although the three systems reviewed in Section 3.2, which were ASDD, FINDAR, and Isaac, all make use of directory services, their approaches are different. In Table 1 a comparison is presented of some features associated with these systems. Properties chosen for comparison is associated with the search for data sets using metadata. The tools used to execute queries are identified while under the metadata property further classification is presented to include metadata concepts described; metadata standards used; management system for metadata; and the database environment.

¹¹ <http://www.ietf.org/internet-drafts/draft-ietf-find-cip-arch-02.txt>

All three systems are specific in their application and make use of different metadata concepts to achieve their objectives. For example ASDD used ANZLIC metadata; FINDAR used attribute values; and Isaac used DC. Similarly the metadata storage and tools used are also different, however, their overall intent is the search for metadata. All three systems address spatial metadata. However, Isaac's approach is new since no other research has investigated the possibility of using LDAP.

In all of the systems there is no common metadata standard followed and their search strategies are all different and do not take into consideration the existence of multiple metadata standards. Therefore the findings from this comparison is that there is a need to focus on multiple spatial metadata standards and spatial metadata query processing. This research is based upon Pascoe's (1996) work which has lead to the on-going research work of Ramroop & Pascoe (1999) and their concept of a Virtual Data Centre. Their work is substantiated by the Isaac Project which utilizes similar concepts but with a different perspective than that of the unification of metadata standards and the influence on query processing strategies.

PROPERTIES			SYSTEMS / RESEARCH			
			ASDD (ANZLIC 1998)	FINDAR (Johnson et al. 1991)	Isaac (Roszkowski & Lukas 1998)	Virtual Data Centre (Ramroop & Pascoe 1999a)
Query	Z39.50		X			X
	Other			X	X	
Metadata	Concepts described	Data	X	X	X	X
		Process				
		Both				
	Standard	Accepted Standard	X		X	
		Multiple Standard				X
		De-facto Standard		X	X	
	Management	Relational DBMS		X		
		LDAP Directory			X	X
		Other				
	Distributed Environment	CORBA				
		Agent			X	X
Other		X	X			

Table 1: Comparison of systems for metadata search

There findings there is little indication of systems/research making use of LDAP to address the various metadata standards already being used to support user queries. A recognition of the multiple standards is needed such that users can utilize whatever standard they are familiar with and still get responses for all existing standards when a query is executed. With this in mind two strategies are described in the next section.

5.0 METADATA QUERY PROCESSING STRATEGIES

The authors are researching two strategies for metadata query processing. The *Unified Schema* strategy described in Section 5.1 and the *Query Translation* strategy described in Section 5.2. These strategies are aimed at allowing user queries to be resolved within environments in which many different metadata standards are used for describing data sets. The general search principle followed in the two strategies is shown on Figure 1. The user submits a query to the Selector Broker which in turn searches all of the directories within the Federated GIS. The user query is translated into the various metadata standards followed at each Data Agency. The Selector Broker executes search strategies.

5.1 Unified schema

The *Unified Schema* method is a query processing strategy in which elements of a defined ontology are mapped onto similar elements of other standards. In this strategy, users need only to understand the unified schema and the query syntax used to express queries. The concept of the strategy within the Selector Broker is shown on Figure 2.

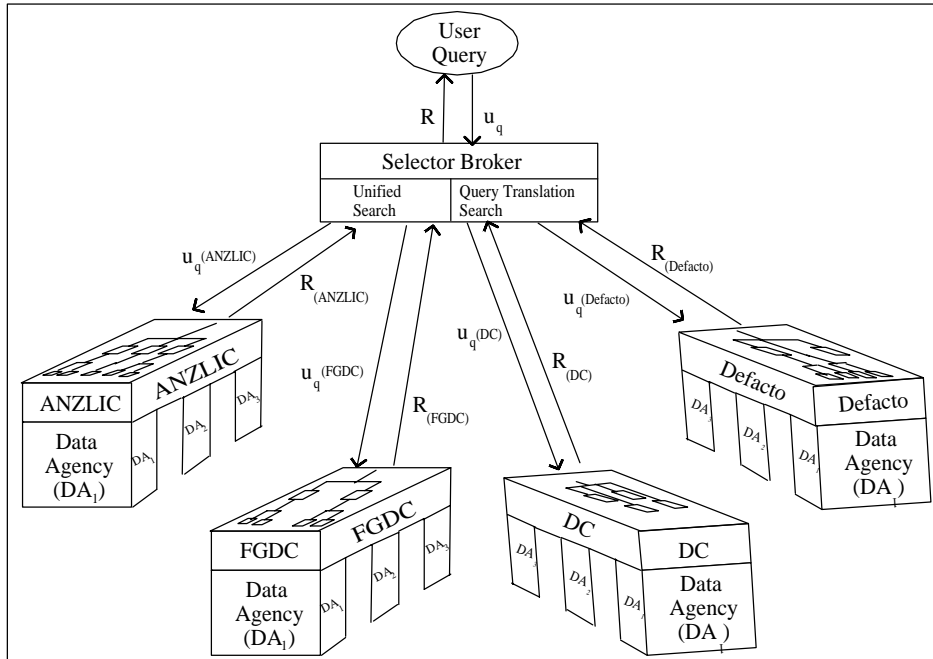


Figure 1: General search principle

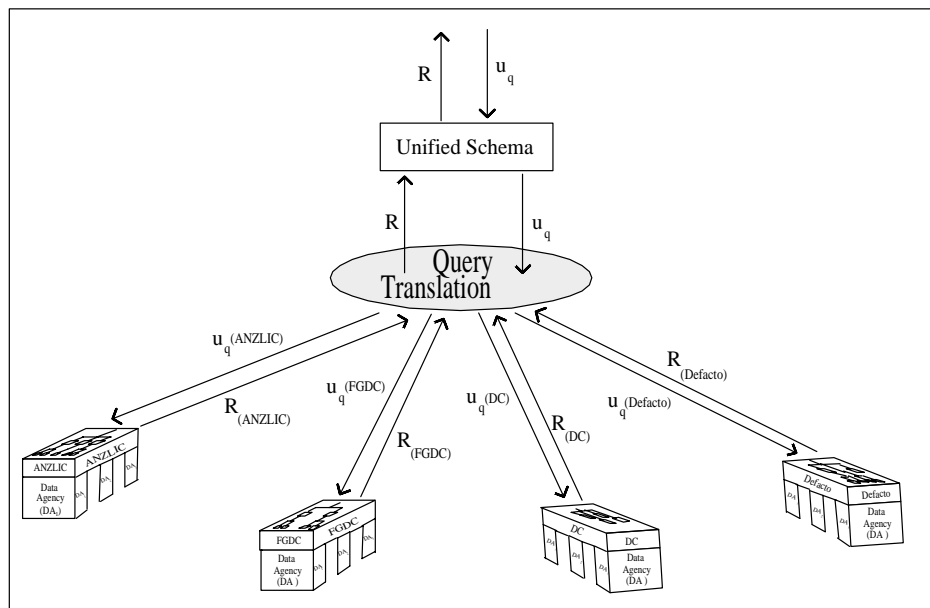


Figure 2: Unified Schema strategy

From Figure 2, a user will define a query using the unified schema definition of metadata elements. These elements are then mapped to the other metadata elements of the various metadata standards. This is similar to the five-level schema architecture of a Federated Database System (FDBS) as discussed by Sheth & Larson (1990). Initially, the user query will be processed using a parser to retrieve the unified schema metadata elements. These elements are mapped onto the elements of other standards such as ANZLIC, CSDGM, and so on. Using these mappings the Selector Broker will search each LDAP server for possible matches notwithstanding the metadata standard followed. The results are then reported back to the user using the unified schema.

Take for example the entire process followed when a user submits a query of *address = New Zealand*. In this case 'address' is one of the GEO Profile metadata elements defined by FGDC and is used as the unified schema

in this research. Therefore, 'address' will be mapped onto the 'Mail Address' element from ANZLIC, the 'Address' element from CSDGM, the 'Source' element from DC, and so on. These matches will be merged into a single response and sent back to the user.

5.2 Query translation

The Query Translation method is a query process strategy in which the user submits a query using a chosen metadata standard. In this strategy as shown on Figure 3 the user need only to understand the syntax for one metadata standard. Therefore, in this strategy the mappings of the various metadata elements from the different standards are needed. Once these mappings are established then the Selector Broker will automatically report back to the user matches from all Data Agencies not withstanding the choice of the metadata standard they are using.

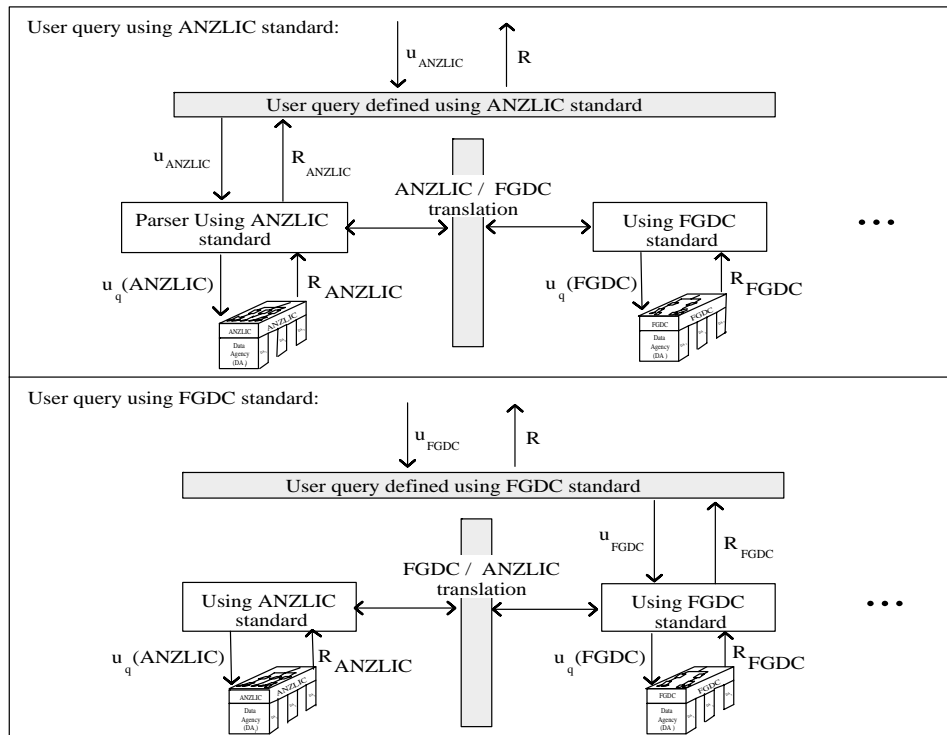


Figure 3: Query Translation

The concept depicted in this strategy is shown on where a user can submit a query using ANZLIC or CSDGM standard. If the user query is submitted using the ANZLIC standard results are reported back to the user using the ANZLIC standard. Additionally, the original query is automatically translated into another similar query for example CSDGM after which the relevant Data Agencies containing CSDGM metadata elements are then searched. Similar translations are repeated for all other Data Agencies using different metadata standards.

For a practical example assume that a user submits a query 'Mail Address = New Zealand' using ANZLIC standard. This query will be translated into the query 'Address = New Zealand' using the CSDGM standard and the relevant Data Agencies containing CSDGM metadata will be searched. This translation of queries are repeated for other Data Agencies using different metadata standards. In the next section, using the two strategies an analogy is drawn with Pascoe & Penny's (1990) work on interfacing strategies.

5.3 Analogy with interfacing strategies

Pascoe & Penny (1990) discussed three interfacing strategies. Of the three interfacing strategies two are relevant here. First, the 'Individual Interfacing Strategy' which makes use of individual interfaces for each source and target pair and second, the 'Interchange Interfacing Strategy' which provides a unified structure through which translation from any format can be made to any other format.

The Unified Schema Strategy for query processing described in Section 5.1 is analogous to the 'Interchange Interfacing Strategy' where a common metadata standard such as the CSDGM GEO Profile is utilized. As a measure of performance, for N metadata standards, will require 2N query translations. The Query Translation Strategy for query processing described in Section 5.2 involves the use of query translations between all existing metadata standards which is analogous to the 'Individual Interfacing Strategy'. As a measure of performance for N metadata standards requires N(N-1) query translations.

In the Unified Schema Strategy, one query translation strategy is needed to map the unified schema elements to the other metadata standards. Even though the GEO Profile defines the unified schema, the quality of metadata searches is dependant upon the availability of the GEO Profile attribute set to be able to define all possible metadata standards, and the accuracy of user queries during query translation. In the Query Translation Strategy no single metadata standard is followed. Pascoe & Penny (1990) showed that the 'Individual Interfacing Strategy' is best in terms of quality and performance.

6.0 CONCLUSION

In this paper a brief review of research making use of metadata was presented. Based upon this review a comparison was done which identified areas of interest regarding the storage and processing of spatial metadata queries. Two strategies were presented and analogously compared to Pascoe & Penny's (1990) work. The conclusion drawn is that the Unified Schema Strategy is not a preferred strategy because there is no single metadata standard which incorporates all metadata elements for all existing standard. Therefore, for the purposes of this research, the analogy indicates that the preferred strategy for query processing is the Query Translation Strategy which is independent upon the metadata standard followed.

Future research includes the need to establish the mappings between the various metadata standards; and translating query results back into the standard familiar to the user. Additional work is required to investigate the possibility of including the mappings in the directories located at the Data Agencies or hard coding the mappings into the mapping tables defined in the Selector Broker.

7.0 REFERENCES

- Abel, D. J., Ooi, B., K.Tan & Tan, S. (1998), 'Towards integrated geographical information processing', *International Journal for Geographic Information Systems* Vol: 12(No: 4), pp: 353-371.
- Allen, J. & Mealling, M. (1998), 'The architecture of the common indexing protocol (CIP)'. <http://www.ietf.org/internet-drafts/draft-ietf-find-cip-arch-02.txt> Accessed on 12/08/99.
- ANZLIC, ed. (1996), *Core Metadata Elements for Land and Geographic Data Directories in Australia and New Zealand*, Working Group on Metadata Australia New Zealand Land Information Council.
- ANZLIC, W.G. (1997), 'Prototype distributed data directory'. <http://www.environment.gov.au/net/directory.html>.
- ANZLIC, W. G. (1998), 'The Australian New Zealand Land Information Council (ANZLIC)'. <http://www.environment.gov.au/database/metadata/anzmeta/>.
- ISODE (1994), 'User guide directory service', *ISODE Technical report 8*. The Dome, Square Richmond TW91DTUK.
- Johnson, B., Shelly, E., Taylor, M. & Callahan, S. (1991), 'The FINDAR directory system: A meta-model for metadata', *Metadata in the Geosciences* pp. 123-137. Editors: D. Medyckyj-Scott, I. Newman, C. Ruggles and D. Walker.
- Kildow, M. (1996), 'The value of metadata (a NSDI report)', *US Fisheries and Wildlife Services*. <http://www.rl.fws.gov/metadata/meta.html> Accessed 4/9/97.
- Marr, A., Pascoe, R. T., Benwell, G. L. & Mann, S. (1998), 'Development of a generic system for modelling spatial processes', *Computer, Environment, and Urban Systems* Vol: 22(No: 1). Published by: Elsevier Science Ltd., Great Britain.
- Mena, E., Kashyap, V., Illarramendi, A. & Sheth, A. (1998), 'Domain specific ontologies for semantic information brokering on the global information infrastructure', *International Conference on Formal Ontologies*

in *Information Systems (FOIS'98), Trento (Italy)*. http://siul02.si.edu.es/jirgbdatt-bin/show_pub?type=Any&date=1985&author=s=E.+Mena&SUBMIT=Show+Them.

OpenGIS (1999), 'What is OpenGIS?', *Open GIS Consortium, Inc.* <http://www.opengis.org/opengis.htm#opengis> Accessed on 23/07/99.

Pascoe, R. (1996), 'Data sharing using x.500 directory', *First International Conference on Geoprocessing* pp. 689-698. Spatial Information Research Center (SIRC), University of Otago, New Zealand.

Pascoe, R. & Penny, J. (1990), 'Construction of interfaces for the exchange of geographic data', *International Journal of Geographical Information Systems* 4(2), pp: 147-156.

Pascoe, R. & Penny, J. (1995), 'Constructing interfaces between (and within) geographical information systems.', *International Journal of Geographic Information Systems* 9(3), pp: 275-291.

Ramroop, S. & Pascoe, R. (1998), 'Notation for national integration of geographic data', *10th Annual Colloquium of the: Spatial Information Research Centre* pp. 279-288. Ed: Peter Firms. Spatial Information Research Center (SIRC), University of Otago, New Zealand.

Ramroop, S. & Pascoe, R. (1999), 'Implementation architecture for a national data center', *Interoperating Geographic Information Systems. Second International Conference, INTEROP'99, Zurich, Switzerland*. pp. 65-74. Lecture notes in Computer Science (1580), edited by A. Vckovski, K.E. Brassel and Hans-Jorg Schek. Published by Springer-Verlag Berlin Heidelberg.

Rossmeissl, H. J. (1998), 'Information access; where do we go from here?', *The 26th Annual Conference of AURISA (Australian Urban and Regional Information Systems Association, INC.), Perth, Western Australia*. . Editor: AURISA.

Roszkowski, M. & Lukas, C. (1998), 'A distributed architecture for resource discovery using metadata', *D-Lib Magazine. ISSN 1082-9873*. Internet Scout Project, Computer Sciences Department University of Wisconsin-Madison <http://www.dlib.org/dlib/june98/scout/06roszkowski.html> Accessed on 12/08/99.

Shah, K. & Sheth, A. (1998), 'Logical information modeling of web-accessible heterogeneous digital assets', *Proceedings of the IEEE Forum on Research and Technology Advances in Digital Libraries, IEEE ADL '98, April 22-24, 1998, Santa Barbara, California, USA. IEEE Computer Society* pp. 266-275. ISBN 0-8186-8464-X.

Sheth, A. & Larson, J. (1990), 'Federated database systems for mapping distributed heterogeneous and automated database', *ACM Computing Surveys* 22(3).

Walker, D., Ruggles, C., Newman, A. & Medyckyj-Scott, D. (1992), 'A system for identifying datasets for GIS users', *International Journal for Geographical Information Systems* 6(6), pp: 511-527.

Wiederhold, G. (1994), 'Interoperation, mediation, and ontologies', *Proceedings International Symposium on Fifth Generation Computer Systems (FGCS94)*. Workshop on Heterogeneous Cooperative Knowledge-Bases, Vol. W3, pp: 33-48.

Williamson, A. P. & Ezigbalike, C. (1998), 'The importance of metadata engines in spatial data infrastructures.', *The 26th Annual Conference of AURISA (Australian Urban and Regional Information Systems Association, INC.), Perth, Western Australia*. Editor: AURISA.

Wilson, E. (1998), 'Manuals go click', *The Age*. Melbourne, 30/6/98.