



INFORME	Identificación: LA CLEF
	Fecha: 7 de noviembre de 2001
TÍTULO La Clef. An Operational Model for Unlocking Public Sector Geographic Information through e-Commerce	
PROYECTO	
RESUMEN	
<p>El objetivo del proyecto La Clef ha sido el de explorar métodos de difusión de información geográfica a través del comercio electrónico.</p> <p>Para ello se ha desarrollado una metabase de información geo-espacial y el suministro de datos geoespaciales a través de Internet con el siguiente programa de trabajo:</p> <ul style="list-style-type: none">- Análisis organizacional de servicios de información geográfica por Internet- Determinación de las necesidades de usuarios de información geográfica a nivel de metadatos- Aspectos técnicos y legales en el suministro de IG- Diseño de un modelo de metadatos compatible con los internacionales- Adquisición de los metadatos disponibles a nivel europeo- Desarrollo de una base de datos geodésica- Desarrollo e implantación de un prototipo de comercio electrónico <p>Los resultados conseguidos más significativos han sido el desarrollo de un modelo y prototipo que permite a los usuarios:</p> <ul style="list-style-type: none">- descubrir datos de información geográfica, con herramientas textuales y gráficas de recuperación de la información- determinar la adecuación de los datos a sus necesidades específicas- adquirir los datos a través de comercio electrónico <p>Además se han implantado servicios de transformación de coordenadas y de información geodésica.</p> <p>El modelo de comercio electrónico desarrollado será implantado por los Institutos Geográficos Nacionales europeos miembros del proyecto.</p>	
Autores: Luis Delgado (IGME), B. Farrell (GSO)	Responsable: Luis Delgado

LaClef

An Operational Model for Unlocking Public Sector GI
through E-Commerce

LaClef Final Report
PUB1102-LACLEF 25035/0

Public Report
Reference: WP1/MEGR/ 030 / 11

Contents

1 EXECUTIVE SUMMARY	6
1.1 PURPOSE OF LACLEF.....	6
1.2 PROCESS.....	6
1.3 ACHIEVEMENTS.....	6
1.4 SIGNIFICANT ISSUES.....	7
1.5 FUTURE PLANS	7
2 INTRODUCTION.....	8
2.1 PURPOSE OF REPORT.....	8
2.2 OBJECTIVES OF LACLEF.....	8
2.3 BACKGROUND.....	8
2.3.1 <i>INFO2000 programme</i>	8
2.3.2 <i>Metadata</i>	9
2.3.3 <i>GDDD</i>	10
2.3.4 <i>Consortium</i>	10
2.3.4.1 MEGRIN	10
2.3.4.2 EUROGI.....	10
2.3.4.3 IGN.....	11
2.3.4.4 ESRI-France	11
2.3.4.5 GIM	11
2.3.4.6 BRGM and ITGE.....	12
3 WHAT WE WANTED TO ACHIEVE	12
3.1 "LA CLEF" VISION.....	13
3.1.1 <i>Overview</i>	13
3.1.2 <i>The central metadata database</i>	14
3.1.3 <i>Exchange protocols</i>	14
3.1.3.1 Exchange format.....	14
3.1.3.2 Exchange protocols and tools	15
3.1.4 <i>Services</i>	15
3.1.5 <i>E-commerce</i>	15
3.2 PROJECT PLAN.....	15
3.3 ISSUES.....	16
3.3.1 <i>Data Availability</i>	17
3.3.2 <i>User requirements</i>	17
3.3.2.1 An attractive site.....	17
3.3.2.2 Ease of use.....	18
3.3.2.3 Confidence in the utility of the site.....	18
3.3.2.4 Confidence in security - a trustworthy site	18
3.3.3 <i>E-commerce issues</i>	19
3.3.3.1 Pricing	19
3.3.3.2 Licensing	20
3.3.3.3 IPR and data security.....	20
3.3.3.4 e-commerce payment mechanisms	21
3.3.3.5 Service Providers.....	22
3.3.3.6 Legal Issues	23
3.3.4 <i>Availability of data for sale</i>	25
4 OUTCOMES.....	25
4.1 OVERVIEW	25
4.2 ACHIEVEMENTS.....	25
4.3 DATA & STRUCTURE	26
4.3.1 <i>LaClef data model</i>	26
4.3.1.1 Workshops.....	26
4.3.1.2 ESMI (European Spatial Metadata Infrastructure) Project	28
4.3.1.3 Standards	28
4.3.1.4 Final data structure	28
4.3.2 <i>Data</i>	29
4.3.2.1 GDDD.....	29
4.3.2.2 Corine Land Cover (CLC).....	29

6 CONCLUSIONS..... 57

ANNEXE A - LACLEF METADATA ELEMENTS..... 58

 6.1.1 UML Notation..... 58

 6.1.2 Metadata Reference..... 61

 6.1.3 Dataset Quality..... 61

 6.1.4 Dataset Identification..... 62

 6.1.5 Dataset Sample..... 62

 6.1.6 Dataset Extent..... 63

 Dataset Distribution..... 64

 6.1.8 Responsible Party..... 65

 6.1.9 Element Descriptions..... 66

 6.1.10 Entities in Summary..... 67

ANNEXE B SUMMARY OF USER TESTING RESULTS..... 71

1 Executive Summary

1.1 Purpose of LaClef

The aim of LaClef was to explore the issues surrounding the unlocking of publicly held GI data, to provide a tangible demonstration of how technology can provide users with the data they want and to create a model for other organisations to utilise. LaClef has built on existing European initiatives, including MEGRIN's existing GDDD, to identify how an effective GI meta data and online data supply service can be developed.

The project focused on the following subjects:

- the improvement of the GDDD (Geographical Data Description Directory) meta-information system created by MEGRIN: more and detailed information in an advanced database management system;
- the identification and online provision of additional services to the user related to geodetic and conversion issues;
- the investigation of general issues related to GI e-commerce, and more specifically of commercial and legal issues;
- the creation of a web-interface through which the user can search for geographic information through the meta-information system and geographic interface, allowing the selection and ordering online of the required data sets, with download and payment online.

1.2 Process

The LaClef project has taken the following approach:

- Examination of organisational issues inhibiting the development of an online GI service.
- Conducting of user –testing to determine user-requirements for metadata and ancillary services.
- Identification of legal and technical issues involved in the provision of GI data online.
- Design of a metadata model consistent with international standards.
- Gathering of available pan-European metadata for inclusion in the LaClef metadata service.
- Development of a European geodetic database in conjunction with EUREF.
- Development and implementation of a prototype model to enable future progress in e-commerce and online data supply.
- Determination of an appropriate course of action given current environmental constraints.

1.3 Achievements

The project has created an internet site which demonstrates how GI data may be made available “online”. This is shown at <http://sixtus.gim.be/laclef>. The model provides a mechanism for users to:

- “discover” GI data that is of interest to them, using textual and graphical search and retrieval tools.
- determine it's “appropriateness for use” by interpretation of extensive metadata, and

- “acquire” the data using the integrated e-commerce model.

In addition, there are two additional services provided within the site:

- Co-ordinate transformation service – a user can enter a series of co-ordinates in their chosen system and convert it to another of their choice.
- Geodetic information service – a source of information about geodetic matters to help people understand the data they have – or would like to acquire.

We believe the e-commerce model developed and the lessons learnt will be of great benefit to EuroGeographics members in developing their own national e-commerce implementations.

1.4 Significant Issues

The project has found that the significant issues for GI e-commerce are organisational, financial and legal.

Because of the diverse data sources, the permissions to make data available for online data supply are difficult to obtain. This creates a difficult administrative burden on the hosting organisation, both with regard to permissions and also data acquisition.

The cost of maintaining the service is also a significant issue. The ongoing financial commitment necessary to maintain the site, develop the online metadata maintenance and complete the e-commerce and geodetic services is prohibitive.

Legal issues – specifically in relation to copyright, contract and taxation – have to be addressed. As an organisation operating within the EU, there is a need to observe the regulations within this jurisdiction. We have been able to suggest a range of possible approaches. However, there remain concerns from data providers (and users) over the supply of data “online” because of the complications and unresolved issues in this area.

1.5 Future plans

LaClef was intended to consider the range of complex issues associated with supplying GI data “online”. This information is now publicly available and will be used by the NMAs as they determine their approaches and strategies for e-commerce over the next few years.

The future of the site itself is not yet confirmed. It has been built in three modules and future development will take a modular approach:

- EMF – the discovery section. This core service provides all available information on datasets contained within the service. It allows textual and graphical search and retrieval. Some development work is still required, particularly on the online metadata maintenance.
- E-Commerce. Although some technical work is required, the primary development is in the selection of the service provider and acquisition of permissions to sell datasets.
- Geodetic information service. It is envisaged that EUREF will complete data collection and validation of pan-European geodetic information during 2001. We propose to formally adopt EUREF validation and for the process of geodetic updates to be maintained by EUREF.

These three modules can be used together or independently. It must be emphasised that development of any or all of these modules demands a financial resource that is currently unavailable. An important and ongoing part of these plans is the identification of possible future funding

2 Introduction

2.1 Purpose of report

This is the final project report for LaClef. It is intended to set out what happened within the project and what will happen in the future. This report represents deliverable ID31 of INFO2000 project PUB 1102.

2.2 Objectives of LaClef

The aim of LaClef was to explore the issues surrounding the unlocking of publicly held GI data, to provide a tangible demonstration of how technology can provide users with the data they want and to create a model for other organisations to utilise. LaClef built on existing European initiatives, including MEGRIN's existing GDDD, to identify how an effective GI meta data and online data supply service can be developed.

The project focused on the following subjects:

- the improvement of the GDDD (Geographical Data Description Directory) meta-information system created by MEGRIN: more and detailed information in an advanced database management system;
- the identification and online provision of additional services to the user related to geodetic and conversion issues;
- investigation of general issues related to GI e-commerce, and more specifically of commercial and legal issues;
- the creation of a web-interface through which the user can search for geographic information through the meta-information system and geographic interface, select and order online the required data sets, download and pay online.

2.3 Background

2.3.1 INFO2000 programme

The INFO2000 Programme in general "aims at stimulating the emerging multimedia content industry to recognise and exploit new business opportunities". The following description of INFO2000 is extracted from the introduction to the 'INFO2000 4-Year Work Programme 1996-1999' available on the INFO2000 web site, <http://www.echo.lu/info2000/> INFO2000 aims at stimulating the emerging multimedia content industry to recognise and exploit new business opportunities. The central theme of INFO2000 is the development of a European information content industry capable of competing on a global scale and able to satisfy the needs of Europe's enterprises and citizens for information content leading, on the one hand, to economic growth, competitiveness and employment and on the other hand to individual professional, social and cultural development.

The INFO2000 Programme has three main actions:

- Stimulating demand and raising awareness
 - AL1-1 Creating new markets by raising awareness at the European level with specific user groups.
 - AL1-2 Encouraging clusters of pan-European user groups.
- Exploiting Europe's public-sector information

- AL2-1 Developing policies to access and exploit European public sector information.
- AL2-2 Linking directories to public sector information.
- AL2-3 Making use of content resources in the public sector
- Triggering European multimedia potential
 - AL3-1 Catalysing high quality multimedia content
 - AL3-2 Trading multimedia intellectual property rights
- Support actions
 - AL4-1 Observing and analysing the multimedia content market
 - AL4-2 Spreading the use of multimedia content standards
 - AL4-3 Encouraging skills development at European level
 - AL4-4 Developing and exchanging best practice

2.3.2 Metadata

The following discussion of metadata is taken from the GSDI Cookbook¹

Metadata is the term used to describe the summary information or characteristics of a set of data. The term has become widely used over the past 15 years, and has become particularly common with the popularity of the World Wide Web. In the area of geospatial information or information with a geographic component this normally means the:

- What - title and description of the data set.
- Why - abstract detailing reasons for the data collection and its uses.
- When - when the data set was created and the update cycles if any.
- Who - originator, data supplier, and possibly intended audience.
- Where - the geographical extent based on latitude / longitude, co-ordinates, geographical names or administrative areas.
- How - how it was built and how to access the data.

There are different levels that metadata may be used for:

- Discovery metadata - What data sets hold the sort of data I am interested in? This enables organisations to know and publicise what data holdings they have.
- Exploration metadata - Do the identified data sets contain sufficient information to enable a sensible analysis to be made for my purposes? This is documentation to be provided with the data to ensure that others use the data correctly and wisely.
- Exploitation metadata - What is the process of obtaining and using the data that are required? This helps end users and provider organisations to effectively store, reuse, maintain and archive their data holdings.

Each of these purposes, while complementary, requires different levels of information. As such organisations should look at their overall needs and requirements before developing their metadata systems. The important aspect is for agencies to establish their business requirements first, the content specifications second and the technology and implementation methods third.

¹ Taylor, M., "Chapter 3 – Metadata" in Nerbert, D (ed), *GSDI Cookbook* (version 1.0) July 2000, <http://www.gsdi.org/pubs/cookbook/index.html>

2.3.3 GDDD

GDDD is the acronym for the Geographical Data Description Directory. It was created in 1994 to establish a descriptive listing of all the principal geographical databases available from the official National Mapping Agencies (NMAs) of Europe. The GDDD was also a pilot implementation of the European metadata standard CEN ENV 12657 of [CEN/TC287](#).

Since 1996 the main part of the information contained in the GDDD has been freely accessible on the Internet. Data about new products are regularly added and more NMAs are joining those already contributing to the service. Today more than three hundred products are described, from 36 European Mapping Agencies.

However, there are significant areas for development within the GDDD. Only high-level information is available, describing general datasets rather than specific maps, images or data clusters. The service is in a single language, English. Maintenance of the metadata is provided through a CD-based Access tool, with considerable administrative overhead and questionable data currency.

2.3.4 Consortium

The project consortium is a broad-based mixture representing the public and private sector:

- MEGRIN (Europe/France) Project Co-ordinator;
- EUROGI (Europe/The Netherlands)
- IGN-France (France)
- GIM (Belgium)
- ESRI France (France)
- BRGM (France)
- ITGE(Spain)

2.3.4.1 MEGRIN

MEGRIN was created in 1993 on the initiative of [CERCO](#) with the aim of helping the National Mapping Agencies (NMAs) of Europe to meet the increasing demand for cross-border products and services. Since November 1995 MEGRIN has had the legal statute of a GIE (Groupement d'Intérêt Economique, i.e. Economic Grouping of Interest) according to French law. MEGRIN's members, which are also CERCO members, have signed the GIE agreement and pay an annual membership fee to MEGRIN. There are today 20 MEGRIN members and other CERCO members also take part in the life of MEGRIN as observers.

MEGRIN is an acronym of "Multipurpose European Ground Related Information Network" it is a European network of geographical referenced information for use in many diverse applications. MEGRIN's budget is derived primarily from the financial contributions of its members, and from the incomes of its first commercial product [SABE](#) (Seamless Administrative Boundaries of Europe). MEGRIN also takes part in several projects partly funded by the European Commission.

A new organisation – EuroGeographics – has been established in order to strengthen the co-operation between the National Mapping Agencies of Europe. The new EuroGeographics organisation can be considered as a merger between CERCO and MEGRIN and it was driven by a desire to make as easy as possible contact with the users of Geographic Information, and their understanding of how their needs can be satisfied. It was also felt that a single administration of the two previous partners would be more efficient.

This important organisational change was made in Malmö, in southern Sweden, where CERCO and MEGRIN had their General Assemblies during 9-14 September 2000.

EuroGeographics will be a not-for-profit Association under French law and will be based in Marne la Vallée on the outskirts of Paris. 33 countries have registered as Founder Members of EuroGeographics and it is hoped that the remaining 4 CERCO members will join shortly.

2.3.4.2 EUROGI

EUROGI, the European Umbrella Organisation for Geographic Information, was set up in November 1993, as a result of a study commissioned by [DG XIII-E](#) of the European

Commission to develop a unified European approach to the use of geographic technologies. It has been established as a European Foundation under Dutch law with its headquarters in Amersfoort, The Netherlands. The Secretariat is also based in Netherlands. EUROGI now has 18 members, i.e. 17 national associations and one pan-European sectorial organisation. EUROGI objectives include:

- Define a European GI policy and facilitate a European Geographic Information Infrastructure (EGII);
- Improve communication between members of affiliated associations through conference co-ordination and directories;
- Ease data exchange at the national and European level through the development of standards and limit the impact of the constraints of the various legal issues affecting GI.

2.3.4.3 IGN

IGN is a French government agency supervised by the Ministry of Equipment, Transports and Housing. Its budget is worth 700 million French Francs per year, part of it is assured by government endowment for public-service missions, and part by commercial revenues. The main production site is located in Saint-Mandé (south-east of Paris) where 300 maps are published yearly and 4 major geographic databases are maintained. There are 6 inter-regional production centres and 16 regional agencies. A satellite image production and processing unit is also located in Toulouse (IGN-Espace). An airport, with 4 survey aircraft and 1 research aircraft, is located in Creil (north of Paris). An integrated printing workshop delivers 10 000 000 maps yearly, while the geodesy department maintains 100 000 geodetic points and 400 000 bench-marks. IGN has a research department involving 4 laboratories and employing 80 persons, including 30 survey engineers.

IGN employs 1900 highly qualified members of staff. Retail distribution is via 16 distributors and there are 10 000 sales outlets spread over France. There are 4 700 000 maps sold annually to the public. Like all mapping agencies, IGN's role has changed in recent years from providing traditional geodesy, surveying and cartography to one of spatial data managers and producers. IGN has been making pioneering steps with their use of the Internet for online sales of aerial photography. Involvement in this project will allow them to further explore the commercial opportunities opening up with e-commerce and to contribute their expertise in areas of Internet site management, commercial and legal issues.

2.3.4.4 ESRI-France

ESRI-FRANCE is an independent company, being the exclusive distributor of ESRI's GIS software such as Arc/Info and ArcView in France. In addition of this activity, ESRI-FRANCE offers services in training, expertise, development and data processing. ESRI-FRANCE has 50 employees, distributed among two offices (Meudon and Montpellier), and has a turnover of 49 million francs (1997). ESRI-FRANCE works in many sectors, for example agriculture, environment, industry, defence, local or governmental administrations, transport and marketing. ESRI software is the most used GIS technology in the world with up to 400 000 users at the end of 1997. The development of distributed Geographic Information Systems and the electronic commerce of geographic data constitute a major part in the company strategy.

ESRI-FRANCE already is experienced in the management of large amounts of geographic data and the distribution of geographic information via the Internet. La Clef will give ESRI-FRANCE the opportunity to demonstrate this know-how and to participate to a major and ambitious European project in the field of electronic commerce.

2.3.4.5 GIM

GIM NV/SA is a private SME establishing its place in the IT service market and defining itself as a proven provider of information management and consultancy in the selected areas of Information Technologies (IT), Applications Development, Remote Sensing (RS) & Geographic Information Systems (GIS). The company currently employs 20 staff with competence in GIS and RS. The team is multi-cultural and interdisciplinary in nature, holding

university degrees in geography, cartography, agricultural engineering, informatics and urban and regional planning.

GIM NV/SA has already gathered an extensive list of references and project experience at European, National and regional level in the areas of GIS, Remote Sensing, Information Technologies and related activities. A flagship example at European level is the Commission's GISCO project - Geographic Information System of the Commission of the European Communities. At EUROSTAT, GIM NV/SA holds contracted responsibilities in the areas of project management, database design, implementation and maintenance for information technology applications in the area of GIS and policy advice. GIM is committed to this project as it wishes to develop its already substantial skills even further within the area of commercial geographic information services, it is therefore seeking new markets and its staff are particularly well qualified to understand government held data sources.

2.3.4.6 BRGM and ITGE

These two organisations are members of EuroGeoSurveys (EGS) which is a new European non-profit association constituted by the Directors of the National Geological Surveys of all fifteen member states of the European Union plus Norway. The association's main aim is to provide the entire range of European institutions with expert, balanced and practical pan-European advice and information as an aid to problem-solving and policy formulation in areas such as the use of natural resources (minerals, water, energy and soils) and the sustainable management of environmental issues and natural hazards. The aims of the Association of the Geological Surveys of the European Union are:

- to bring together the Geological Surveys enabling them to jointly address European issues of common interest;
- to provide a permanent network between the Geological Surveys and a common, but not unique, gateway to each of the Surveys and their national networks;
- to assist the European Union to obtain joint technical advice from the Geological Surveys of the Member States;
- to promote, wherever appropriate, the contribution of geoscience to European Union affairs and action programmes;
- to initiate, develop, and promote geoscience inputs to co-ordinated bilateral and multilateral programmes within European and other countries.

EGS members have a desire to share experience of commercial WWW development and management. They believe that the geological surveys must take part in and influence the dissemination, exploitation and organisation of spatial geographic information.

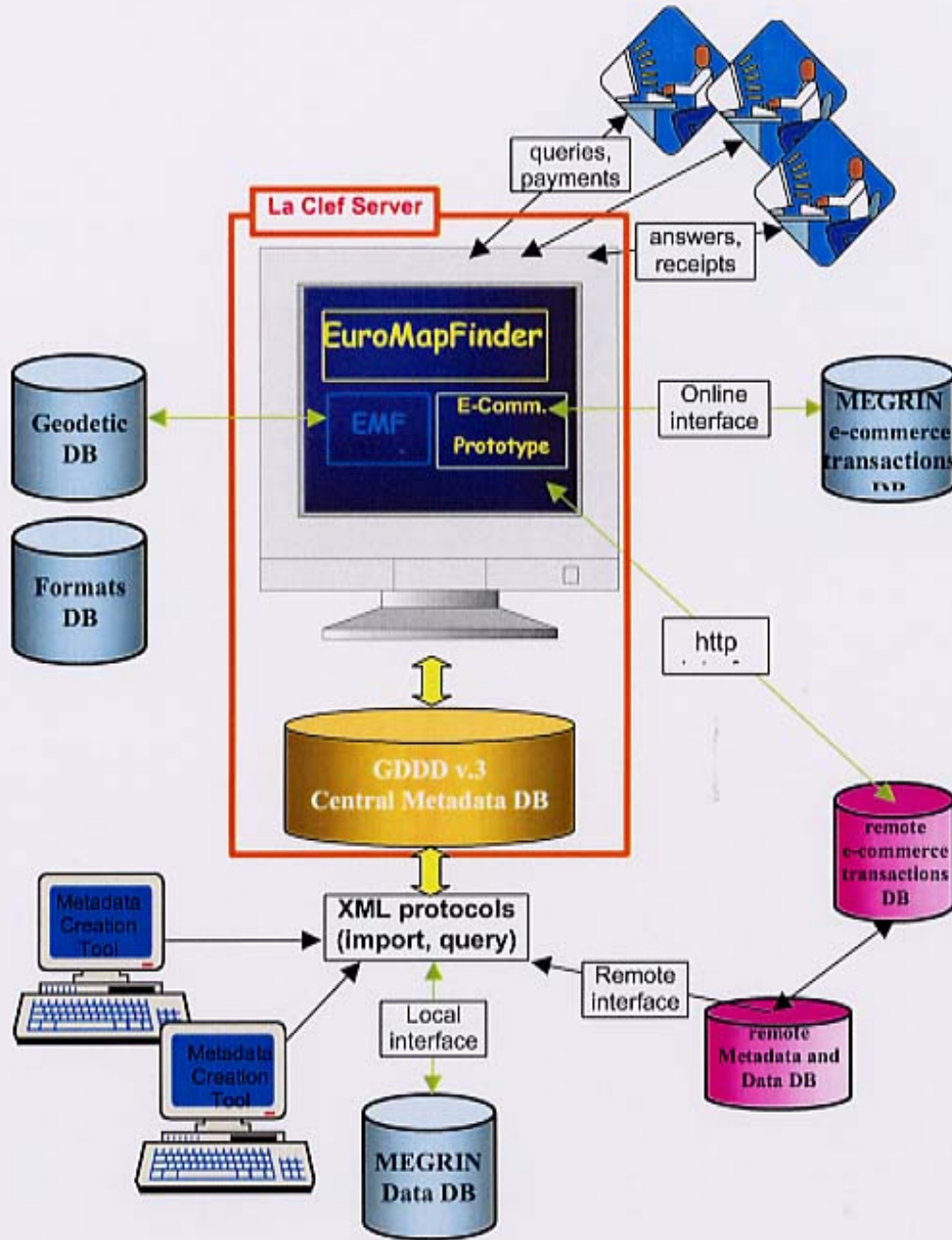
3 What we wanted to achieve

3.1 “La Clef” vision

This section sets out the vision for LaClef (as defined in the early parts of the project).

3.1.1 Overview

The diagram below shows the original concept of LaClef.



3.1.2 The central metadata database

What is described hereunder is not a central metadata database but a central metadata service or system of which the database is a part.

The “core” of the system has two main components:

- the user interface, with three sections :
- EuroMapFinder, the main metadata navigator. this navigator will be based on the structure of the Metadata DB to provide a generic browsing ;
- the e-commerce service. it will be either centralised from partners using MEGRIN as being the Vendor, or distributed by directly calling the e-commerce service for partners already having such a service ;
- EMF services, additional services such as transformation of co-ordinates. The first service La Clef is going to provide is the Import/Query Tool (XML² based) by delivering both freeware and functions libraries. Although this service does not sound as being one in the strict sense of the word, it has to be present as such in order to help clients to populate the GDDD v3 through the import/conversion facility.
- the central metadata database, GDDD v3 ;
- related databases : they cover all databases needed for the system to run (metadata, commercial data, geodetic data). They are either centralised or distributed

The system will be based on a mixed architecture made of centralised and distributed databases and services.

The EMF central database will be the “heir” of the current GDDD, and will be a unique structure that will contain:

- the full implementation of the CEN TC287 experimental metadata standard ;
- with eventual modifications/additions proposed by the MEGRIN “ Core metadata ” ;
- the recommendations on Metadata (a user guide provided by the Centre for Earth Observation Programme Of the European Commission) are of great help in the design of the metadata structure. These should be taken into account in early stages;
- with additional elements made necessary for implementation of the e-commerce, such as unique identifiers of items and/or VAT rules. It is clear that datasets as meant in the CEN standard might be or available for sale or may contain items to be sold. The client's shopping basket may also contain several datasets. Localisation of e-commerce information (items sold) might be centralised, or distributed, or both ;
- geodetic information is seen as a logical link between metadata and the geodetic database. The localisation of the latter might be centralised or remote.

It would be internally divided in three levels :

- “level 1” : fully centralised (all content to be loaded and updated within the GDDD) for elements that are needed for the primary searches ;
- “level 2” : will contain all the other elements needed by all the EMF service(s) functionalities ;
- “level 3” : will contain all the other elements that provide useful information for “professional” queries, but whose absence does not hinder the EMF functionalities.

Levels 2 and 3 elements will either contain the element value (centralised) or a link to the distributed server where the value is to be found.

3.1.3 Exchange protocols

3.1.3.1 Exchange format

The EMF will use an XML based profile as transfer format (see 4.4.1.1 for a description of XML). It will be used for all transfers of data/information between the EMF system and its associated data/metadata providers. The experience of KMS (Denmark) will be used and

² XML – Extensible Markup Language

adapted to the EMF needs. It is worth noting that ISO/TC 211 "transfer service" is also based on XML, which is an emerging technology.

3.1.3.2 Exchange protocols and tools

Different protocols will be necessary, based on the pre-defined XML format, according to the type of operations and the type/choice of the EMF data/metadata providers. In principle, LaClef would develop and provide the necessary tools, some of them might be derived from the tools developed at KMS. Main mechanisms to consider will be:

- online centralised server : provision and upgrading of the current version of the GDDD to the v.3;
- online distributed server : online queries (L2 and L3) ;
- off-line provision and updating of the GDDD from an off-line metadata databases (harvesting procedure) ;
- metadata editing tool (XML based).

3.1.4 Services

It was envisaged that a range of complementary services could be combined with the metadata service. The primary service will be a geodetic transformation and information service providing the facility to transform user-defined co-ordinates from one reference system to another.

3.1.5 E-commerce

The architecture of the EMF e-commerce prototype will be defined while implementing the previous steps. Decisions will be made on:

- the accessibility of the service (public, internal, external for selected clients);
- the payment mechanisms ;
- the data to be sold (products availability) ;
- the data delivery procedures (online, CDs, etc);
- direct links to providers servers (such as IGN-F).

3.2 Project plan

The project proposed to deliver:

- 1 **Multilingual interface**; with the use of profiles and keywords to facilitate a distributed network; high resolution metadata, index maps and data samples;
- 2 **Metadata Internet Server** which will allow users to select a geographical area and to order all related data;
- 3 **Data Internet Server** which will process secured transactions for electronic payment; cross-sector linkages to other metadata sites.
- 4 **Series of Reports** covering:
NMA metadata availability and integrity; Catalogue of complementary datasets; User requirements in terms of associated services; User requirements: e-commerce content and service specification; Services/e-commerce payment mechanisms.

The specific deliverables from each workpackage were:

- WP1
 - Deliverables:
 1. Project quality plan

2. Consortium agreement
 3. Project progress reports
 4. www news, press-releases, journal and conference papers
 5. Project final report
 6. Project final public report
- WP2
 - Deliverables:
 1. Report on NMA (meta)data availability
 2. Catalogue of complimentary publicly held datasets
 3. Inventory of selected metadata for EMF version 1
 4. Inventory of proposed metadata for EMF version 2
 - WP3
 - Deliverables:
 1. Service Specification
 2. Geodetic and other data
 3. Calculation algorithms
 - WP4
 - Deliverables:
 1. EMF metadata core service specification
 2. EMF beta version
 3. EMF V.1
 4. EMF V.2
 - WP5
 - Deliverables
 1. Report on user requirements : added services
 2. Report on user requirements : electronic commerce
 3. Report on user feedback on the first version of EMF
 - WP6
 - Deliverables:
 1. Report on services/e-payment mechanisms
 2. Report on IPR/data security
 3. Report on licensing and pricing options
 - WP7
 - Deliverables:
 1. EMF e-commerce data content specification
 2. EMF e-commerce service specification
 3. EMF e-commerce beta version
 4. EMF e-commerce V.1
 5. EMF e-commerce V.2

The initial project Gantt is shown in Annexe C.

3.3 Issues

There were many environmental issues to consider in selecting the different options. The general architecture chosen was highly dependent on available metadata from NMAs (see 4.3.1.1) and other pan-European metadata services.

The metadata model was developed at the same time as the ISO metadata standard was reaching draft stage and European data providers were shifting support from CEN to ISO. This model had to cater for identified user requirements.

Our commitment to XML as the appropriate method of online updating meant overcoming limited understanding within a rapidly changing technology. Various models for e-commerce are available, each with advantages and disadvantages. LaClef's modular approach lends itself to future integration with any of the models but a final decision remains to be made.

3.3.1 Data Availability

The primary data source at this point in time has been NMAs. As has been discussed, they have provided information about their datasets. The model defined above will require that additional information to be provided, in extent and quality. The collection (and on-going maintenance) of this information must be addressed.

3.3.2 User requirements

The subject of user requirements was investigated within the LaClef project by EUROGI. Its two reports cover the subject of user requirements from the point of view of services³, and e-commerce⁴. The EUROGI services report segmented the market into three user types:

- Traditional principal users of GI; public institutions, local administration, utility companies;
- Recent users of digital GI; for example the transport and geo-marketing sectors;
- All others; consumer use by citizens, looking for route planning or nearest facilities for example.

The EUROGI report also highlights metadata as a key requirement to help all user types find the data they are looking for. After that a number of types of service are identified, starting with data delivery, going on to customised GI products and services, and including GI related services such as format translation and geodetic services.

When looking at the generality of what users want from an Internet site, one which is attractive, easy to use, and in which they can have confidence, the following features arise:

3.3.2.1 An attractive site

The site must be attractive because of its:

- content, which users find useful for their particular requirements;
- currency – daily updates of new and relevant content;
- novelty, making the visit an amusing/stimulating experience;
- aesthetics – a “nice to look at” site with plenty of graphics, animation, sounds etc;
- “value” – many first time or casual visitors are looking for data / software for free; a data site must be able to provide free samples, and viewing software to enable the user to see the data.

The EUROGI report into User Requirements for e-commerce refers to the following “user friendly” attributes that would be necessary for the type of service that LaClef is attempting to develop:

- Obvious access to a centrally located store entrance;
- Meaningful product categories;
- Clear and informative product descriptions;

³ Report on User Requirements: Services, PUB1102-LACLEF 25035/0, Reference WP5/EURO/02 (Restricted)

⁴ Report on User Requirements: E-Commerce, PUB1102-LACLEF 25035/0, Reference WP5/EURO/03 (Restricted)

- Pictures of products and services;
- Search facilities that accept appropriate terms (multiple words and synonyms);
- Search results in a sequence that is logical to the user;
- Clear indication of which items are available from the online store;
- Feedback to the user as items are added to the shopping basket;
- An option to review the status of the order at anytime and easily remove things from the shopping basket;
- Explanations of technical terms and jargon;
- A clear indication of the security measures taken by the site;
- A value added shopping experience, e.g. product reviews.

3.3.2.2 Ease of use

Internet users want a site that is intuitive to use, they do not want to spend much time reading instructions. Guidance through the site map from page to page should be visual/graphic, with text kept to a minimum. The number of pages (clicks) between the opening page and the objective should be kept to a minimum. The aesthetics of a site is important; an attractive site will attract more users, although making an attractive site may involve more graphics - which can make the site slower to navigate. Information on sales terms and conditions must be available on the site, clear to understand and easy to reach when required, but not a distraction to the flow of discovery and access.

3.3.2.3 Confidence in the utility of the site

Users will not return to a site unless they are confident that the service will provide benefits. In the case of an information service, which is LaClef's primary aim, these benefits are provided by maintaining the quality and currency of the metadata.

- **Currency** - in general a site must be updated very often, daily if possible and in a way that makes it obvious to users that changes are being made frequently. This will give the user confidence that the site is well maintained and that it is worth returning to for new information at a later date.
- **Content** - In terms of a metadata service like GDDD, there also needs to be a critical mass of data that gives the user the confidence to know that if he/she needs to find some data there is a reasonable chance that it will be included in the site. The growth of distributed databases via a common portal will of course help to develop user confidence.
- **Follow-up** – a web site cannot be left to run itself. As well as the need for continual update there should be some means of interaction with staff.

3.3.2.4 Confidence in security - a trustworthy site

Users demand security. Without the perception of security, whether giving personal details or conducting financial transactions, users will refuse to use the site. It is vital that the site provides a high level of security and is perceived as being highly secure. It is this perception that will encourage use.

- **Financial security** - another aspect of user confidence is their confidence to make payments to a site – how secure are the financial transactions. Many users/customers are still reluctant to purchase online, so the more reliable and trustworthy the site is, the better.
- **Data Privacy & Consumer Protection** – Users must have the confidence that any information they give about themselves – their names and addresses, bank account

details, will not be misused. The key European Community consumer protection provisions relevant to electronic commerce are well documented⁵ and encompass the following areas:

- Commercial communications / advertising and Promotion
- Information to be provided before the conclusion of a contract
 - Identity of the seller
 - Information on the goods or services on offer
 - Information on the contract and contractual obligations
- Conclusion of the contract and contractual obligations
 - Right of withdrawal
 - Inertia selling
- Payment
- Guarantees
- Settlement of disputes
 - Out of court settlement of disputes
 - Court actions
- Privacy and data protection
- Applicable law
- Jurisdiction

3.3.3 E-commerce issues

3.3.3.1 Pricing

LaClef is constrained by current NMA and GSO pricing policies in the online supply of datasets.

- The cost of data supplied by NMAs and GSOs is higher than that for other geographically-related public data due to the high quality of input necessary to produce the data, and this will be reflected in the costs to the customer.
- Users want free access to more information about datasets, products, services and change.
- Metadata can be provided free of charge but the related raw and elaborated data must be disseminated at reasonable cost to cover the costs of the GI provider.
- It is important to avoid unfair competition with the private sector.

Who will pay for the products and services, or will they be made free is driven by the aims and objectives of the organisation concerned, and will follow policy driven by the mission statement and business marketing plan. The payment options were considered by the Imprimatur project⁶, which listed the options as:

1. Payment by the user
 - Direct payment for ordered goods or services;
 - Payment by paying for other goods (i.e. access to material supported by merchandising).
2. Payment by third parties

⁵ US Perspectives on Consumer Protection in the Global Marketplace – Comments by the European Commission, 21.04.99, http://europa.eu.int/comm/dg24/policy/developments/e_comm/e_comm01_en.html – an excellent review of the EC position.

⁶ <http://www.imprimatur.alcs.co.uk/>

- Advertising. Not considered by Imprimatur as the predominant revenue stream across the network.
 - Sponsorship. Imprimatur include company image advertising, charity based and public advertising in this category, and suggest that probable recipients of such revenues are sites with a cultural character or those aiming at youth, the disabled, or ethnic minorities.
3. Payment by the seller – i.e. content or physical goods offered for free for promotional purposes, PR, and/or public services.

Given the constraints of the NMAs as the primary providers of data to the service, LaClef was required to follow the first of these models, ie payment by the user.

3.3.3.2 Licensing

Currently, many digital datasets are licensed for use rather than sold outright. Many of these licences are dependent on application, eg private or commercial use; number of users etc, and are necessary irrespective of cost.

User testing found:

- Licenses may not be required for low cost products.
- Rather than seeking to categorise users LaClef should consider organising access to the web-site into categories e.g. 'free', 'leisure', 'land and property' etc. the subsequent steps of the user's visit to the site can then be tailored.
- For each category of user LaClef may need to tailor license agreements.
- As more flexible access for the customer to the web site is allowed so it is more important to control that access. The user testing report suggested that LaClef should consider the controls below:

Data integrity: data integrity control is aimed at making sure that the data has not been accidentally or fraudulently altered.

Authentication: this can be of two types, authentication of partners and authentication of the origin of information. This service principally makes it possible to make sure that the correspondent connected corresponds to the person he or she say they are, or to check the identity of the person who signs the act of purchase. The process is sometimes referred to as using "keys".

Non renouncement: this makes it possible to obtain proof that information has been sent or received. The sender or receiver cannot therefore deny that the information has been sent or received.

- Certification authority: its role is to produce and manage public key certificates used for digital signatures. The objective of a certificate is to guarantee to a person who uses a public key to check a signature that the public key really does belong to the person it is supposed to belong to. In order to do so, a certificate guarantees the link between the public key and the holder of the corresponding secret

The nature of the datasets being considered for sale through the LaClef service required the use of a single standard license for each product, e.g., one for SABE data, one for Corine Land Cover (CLC). The ultimate aim for the service will be for the service provider to manage the complexity of different licence arrangements with the data providers thus providing the user with a common license that is independent of the different products. This issue was not tackled within the project and will require further work.

3.3.3.3 IPR and data security

Intellectual Property Rights (IPR) and data security are key considerations for both data providers and end-users. For the data providers, it is necessary to ensure full protection of their assets. For the end-user, all terms and conditions of supply must be clear and transparent.

LaClef research found the following:

1. **Current EU Initiatives**. Although there are currently many differences in the way e-commerce law is applied across Europe the EU is working with urgency to resolve the differences and facilitate e-commerce among member states. On 29 September 2000 an EU spokesman said "All pending measures are to be adopted before the end of this year. The aim is to have a minimum level of common rules within the EU, to facilitate cross-

border e-commerce within the Internal Market, and to give legal guarantees to consumers and businesses.”

2. **Jurisdiction and Regulatory Issues** – LaClef should endeavour to have all transactions covered by French/EU law
3. **Copyright** – conditions of sale should include a declaration by the data provider to say that no third-party rights are involved and that data providers are to be considered as sole owners and will indemnify LaClef against third party interests.
4. **Technological Change** – the LaClef project has chosen to use XML because of its widespread support and its versatility.
5. **Trademarks and Domain Registration** – Since the Office for Harmonisation in the Internal Market (OHIM) was set up a single trade mark application to the OHIM can ensure a trademark covering the whole of the EU.
6. **Contract Terms**. If possible, the complete contract with all terms and conditions should be displayed on the Web site. Or if this is not practical, the user/buyer should be given clear notice at the time of purchase that the transaction is subject to additional terms and conditions.
7. **Technology issues - security**. Content should be secured on the server level (server-side security, firewall), on transaction level (encryption, digital signature) and eventually on data identification level (watermarking).

LaClef, as a service being created with the EU, is required to comply with regulation within that jurisdiction.

3.3.3.4 e-commerce payment mechanisms

LaClef research showed there are a number of options for who will be paying for the e-commerce products and/or services (see 3.3.3.1). The following section assumes that the end user will be paying and discusses the various payment mechanisms available.

Payment can be made off-line, or by a number of online methods:

- **Payment off-line** – For example payment by cheque, payment on delivery, sending a credit card number by post or fax, or direct debiting. Many organisations are likely to continue to use a simple system of invoice and payment. Going “off-line” to make the payment detracts from the overall convenience of an Internet-based transaction, but is important to many users in terms of confidence.
- **Payment online by credit card** - Today, the credit card is the means of payment most widely used for Internet retail purchases. Debit cards require card readers at the customer end of the transaction. For credit card purchases the card number is transmitted over the network using SSL encryption (Secure Socket Layer). This method does have some disadvantages:
 - It is not totally safe for the cardholder (it does not solve the problem of the identity of the seller, nor that of the confidentiality of the bank information provided to the seller);
 - It exposes the retailer to the risk of an unsettled bill;
 - It is inappropriate for transactions involving small amounts – micro-payments. (and in France, for example, for amounts of over 5,000 FF where “written evidence” is required).

The SET protocol (Secure Electronic Transaction) is far more robust than the SSL protocol and was accepted by European Committee of Banking Standards as the de facto standard for credit card payments over the Internet. SET 1 details how payment card transactions on the Internet and other open Networks will be secured using encryption technology and digital identification.
- **Payment online with electronic or “digital” money** – Electronic or digital money is a medium for electronic commerce, whether it takes place on the Internet or not. It is simply the electronic form of the physical money.

Although micro-payments may not become the dominant payment mode for consumer commerce, they appear useful for low cost items in particular segments such as games or pay per view. Electronic cash seems the viable option for these micro-payments, either in the form of an electronic purse/wallet or virtual cash.

3.3.3.6 Legal Issues

Many of the issues that deter organisations from embarking upon e-commerce concern the uncertainty and complexity of the legal issues involved. The situation is made especially complex by the international nature of the Internet. The law is different from nation to nation, but the Internet is global.

Contract law

In English law a contract is created after four conditions have been met. There must be an offer, an acceptance, consideration, and an intention to create legal relations. These same principles apply to many other, but by no means all, other countries. Even when dealing only with English law there are many issues to consider when interpreting the law of contract within the Internet context. Subtleties such as the difference between the making of an offer, and the "invitation to treat, for example" will be important when determining the intention behind a web page listing products and prices. Some experts assume that it is the website that is making an offer to treat, the user who is making the offer, and therefore the seller who is providing the acceptance¹⁰. There is then the question of timing – at which point is the offer accepted, the jurisdiction of which country? the validity or not of click wrap acceptance, and digital signatures.

In the United States attempts to craft a "digital signatures" law have been a high priority for the current Congress and technology companies. The industry says it needs the Electronic Signatures in Global and National Commerce Act to allow companies to do business more efficiently online.

In Europe, aspects of contract law, affecting consumer protection, have been included in the European Directive 97/7/EC of the European Parliament and of the Council of 20 May 1997 on the Protection of Consumers in respect of Distance Contracts, Official Journal L 144 of 04/06/1997. This directive is in the process of becoming national law in member states. The EC has also released a Proposal for a European Parliament and Council Directive on certain legal aspects of electronic commerce in the internal market (COM(1998) 586final). One of the main topics in the Directive is 'electronic contracts' (Chapter 3). The Directive is intended to put an obligation on the Member states to harmonise their legal systems in order to facilitate use of electronic contracts.

Another key question is the question of the identity of the parties. Another EC proposal covers the issue of digital, or electronic, signatures - Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures.

At a global level the UN (General Assembly) has approved a Model Law on Electronic Commerce. The model law has been made by the UNCITRAL (UN Commission on International Trade Law) which exists to try to harmonise law on international trade.

A study on validity of electronic contracts is made within the EC IMPRIMATEUR project: Formation and Validity of Online Contracts. The report discussed, among other things, the enforceability of online contracts in general, the UNCITRAL Model Law and the EC Directive on Distance Contracts.

In conclusion, the situation is far from clear. The global nature of the Internet means that laws should be created within a global context, guided by UNCITRAL. Meanwhile, the law in individual countries is desperately trying to keep up with the pace of Internet development, and is by definition heterogeneous. The links provided in the text will hopefully enable readers to keep up with the main initiatives as they occur globally, within the US, and in Europe.

Tax.

Once again the global nature of the Internet is causing problems. EU policy is stated in the E-Commerce and Indirect Taxation: Communication by the Commission to the Council of Ministers, the European Parliament and to the Economic and Social Committee:

¹⁰ Electronic Commerce, Law and Practice, Michael Chiswick, Alistair Kelman, Sweet & Maxwell, 1999

COM(98)374final; 17/6/98. The communication seeks clarification of many issues, but has been interpreted by some¹¹ to suggest that products sold in digital form would be taxed based on the VAT rate in the customer's country. The consequence is that companies operating in the United States would have to collect the taxes on all sales to European customers.

In the USA Congress passed the Internet Tax Freedom Act (ITFA) in 1998, that imposed a three-year moratorium on new Internet taxation. As part of the Act, Congress established the Advisory Commission on Electronic Commerce to address the issues related to Internet taxation. The Advisory Commission conducted a thorough study of Federal, State and local, and international taxation and tariff treatment of transactions using the Internet and Internet access, and other comparable intrastate, interstate, or international sales activities. The main recommendation from the Commission's report appears to be that States should be given a five year period in which to simplify and unify their tax regimes.

Security and confidentiality of exchanges

The development of electronic trade is subordinate to the use of encrypting techniques. Cryptology is a group of techniques that make it possible to protect information using a secret code. It allows the implementation of security measure that protect electronic data or transactions, including:

- Data integrity: data integrity control is aimed at making sure that the data has not been accidentally or fraudulently altered.
- Authentication: this can be of two types, authentication of partners and authentication of the origin of information. This service principally makes it possible to make sure that the correspondent connected corresponds to the person he or she say they are, or to check the identity of the person who signs the act of purchase. The process is sometimes referred to as using "keys".
- Non renouncement: this makes it possible to obtain proof that information has been sent or received. The sender or receiver cannot therefore deny that the information has been sent or received.
- Digital signature: this is a technique that makes it possible to implement both data integrity, authentication and non renouncement.
- Confidentiality: this makes it possible to make information intelligible to non-authorized third parties when it is saved, or when it is transferred. Encryption of information is the most widely used technique.
- Trusted third party: this is a organisation that is trusted by the user and which carries out certain transactions for the latter, linked to the management of confidentiality keys and/or digital signature. It is advisable to distinguish between the third party deposit functions (keys used for confidentiality) and certification authorisation functions (AC) for public keys that are only used in applications linked to the signature. This trust relies on the competencies of the organisation, the eventual approval that it receives, the contents of the contract linking the organisation to the user and the measures implemented by the organisation to guarantee protection of the user's data and keys.
- Third party deposit: In France, this is an organisation approved by the Prime Minister after its dossier requesting approval had been investigated by the SCSSI (Service Central de la Sécurité des Systèmes d'Information – Central Department for the Security of Information Systems). Its function is to keep users' secret keys that are utilised for the purpose of confidentiality in order to give them to the same users if they so request and to the legal or security authorities.
- Certification authority: its role is to produce and manage public key certificates used for digital signatures. The objective of a certificate is to guarantee to a person who uses a public key to check a signature that the public key really does belong to the person it is supposed to belong to. In order to do so, a certificate guarantees the link between the public key and the holder of the corresponding secret key to make up the digital signature.

In France, legislation introduced in March 1999 makes it considerably easier to use cryptology resources for an electronic signature and for confidentiality using an algorithm of less than 40

¹¹ <http://www.nytimes.com/library/tech/00/03/biztech/articles/02tax.html>

bits, free usage exempt from any prior formality; for confidentiality with third party deposit, free usage subject to authorisation. Moreover, a Government Bill on the legal value of an electronic signature has been filed by the government and recognises in the act of the electronic signature the same legal value as that given to the act on paper. The measures will be voted before summer 2000. In the UK the Electronic Communications Bill was due to receive Royal Assent in May 2000, meaning that it will become the Electronic Communications Act 2000 and the Secretary of State can fix the date on which it will come into force.

As with the issues relating to IPR and data security, LaClef, as a service being created with the EU, is required to comply with regulation within that jurisdiction.

3.3.4 Availability of data for sale

One of the prime requirements for an e-commerce service is having the agreement of the data owners to supply through this medium. This requires a consideration of the issues outlined above. Agreement was obtained in principle for the supply of both SABE and CLC data through the LaClef online service.

4 Outcomes

4.1 Overview

LaClef has performed a thorough investigation into the issues surrounding the supply of GI data "online" and created a prototype model to demonstrate the nature of information and data required – and the process that is involved.

A geodetic transformation service has been implemented as well as a geodetic information service.

Multilinguality has been implemented, allowing users to use the site in a number of European languages.

A metadata model has been designed that allows detailed metadata to be included within the service. Both the design of the database structure and the actual metadata model consumed considerably more resources than anticipated.

Much research and development has been conducted into the use of XML for metadata maintenance.

The project has found that the significant issues for GI e-commerce are organisational, financial and legal.

Because of the diverse data sources, the permissions to make data available for online data supply are difficult to obtain. This creates a difficult administrative burden on the hosting organisation, both with regard to permissions and also data acquisition.

The cost of maintaining the service is also a significant issue. The ongoing financial commitment necessary to maintain the site, develop the online metadata maintenance and complete the e-commerce and geodetic services is prohibitive.

Legal issues – specifically in relation to copyright, contract and taxation – have to be addressed.

4.2 Achievements

LaClef has performed a thorough investigation into the issues surrounding the supply of GI data "online" and created a prototype model to demonstrate the nature of information and data required – and the process that is involved.

The project has created an internet site which demonstrates how GI data may be made available "online". This is shown at <http://sixtus.gim.be/laclef>. The model provides a mechanism for users to:

- "discover" GI data that is of interest to them, using textual and graphical search and retrieval tools.
- determine it's "appropriateness for use" by interpretation of extensive metadata, and
- "acquire" the data using the integrated e-commerce model.

In addition, there are two additional services provided within the site:

- Co-ordinate transformation service – a user can enter a series of co-ordinates in their chosen system and convert it to another of their choice.
- Geodetic information service – a source of information about geodetic matters to help people understand the data they have – or would like to acquire.

The table below sets out the reports that are publicly available

ID	Report
6	Catalogue of complementary datasets
7	Report on NMA meta data availability
12	Project Progress reports
13	EMF Service Specification (specification for all elements of the web site – EMF (interface))
14	Report on Services/e-commerce payment mechanisms
16	EMF e-commerce data content specification
17	EMF e-commerce service specification (all aspects of e-commerce)
19	Final report on EMF services – for all services (European reference systems, geodesy information, map projection information, co-ordinate transformation service, multi-linguality)
26	Final report on EMF services development
28	Report on IPR / data security
29	Report on licensing and pricing options
30	Edited public report
31	Final project report

These documents and other information about the project are available within the EuroGeographics website: <http://www.eurogeographics.org>
The demonstration can be seen at the website: <http://sixtus.gim.be/laclef>

4.3 Data & structure

4.3.1 LaClef data model

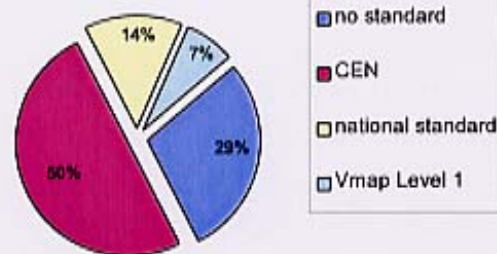
4.3.1.1 Workshops

An important first step in determining the data model was the identification of the required core-metadata elements. A necessary first step in such identification was to determine NMA's metadata implementation status and perspectives regarding core-metadata, thesaurus and metadata maintenance.

Two workshops were held to consider these issues (October 1998 and December 1999). A significant issue was the availability and status of metadata within NMAs. General points highlighted in the findings were:

- 66% of respondents have metadata in digital format but 17% of organisations claim to have no metadata available;
- 50% have implemented CEN metadata standards but 43% have no internationally recognised standard implemented;
- 47% of organisations can supply their data in their national language only (non-English);
- 20% of organisations are currently capable of supplying map-unit level metadata;
- 76% have no controlled lists defined.

Standards used to evaluate metadata



The following issues were identified:

Standard

The importance of CEN as a European standard was clear with 50% of organisations' metadata implementation to CEN standard. The fact that 43% as yet either only conform to a national standard or have no standard at all emphasises the role MEGRIN must play in developing agreed implementation standards.

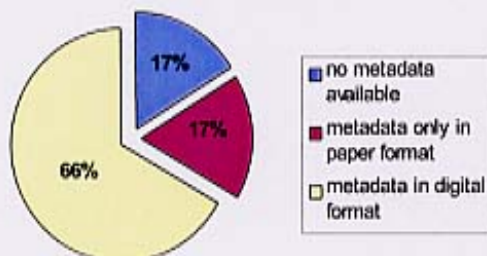
Language

While 47% of non-English speaking countries can only supply their metadata in their national language, English language implementation is still recognised as important with 33% of non-English speaking countries supporting the language. The results emphasised the need for MEGRIN to ensure its metadata service supports national languages.

Metadata Language



Metadata availability



Availability

Metadata in digital format is available from 66% of respondents organisations. A serious issue remains the 34% with no metadata or metadata in paper format only. Future development of a pan-European service will need to recognise the various stages of development within European organisations and the resource requirements required to develop a metadata service to the appropriate level.

Controlled Lists



Controlled Lists

Only 24% of organisations have developed or utilised controlled lists of keywords to facilitate searching and data entry. In developing a core-metadata, MEGRIN will need to consider providing organisations with

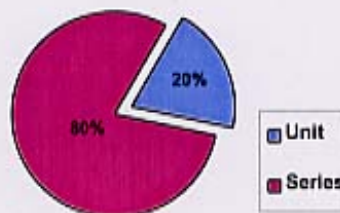
appropriate controlled lists.

Metadata level

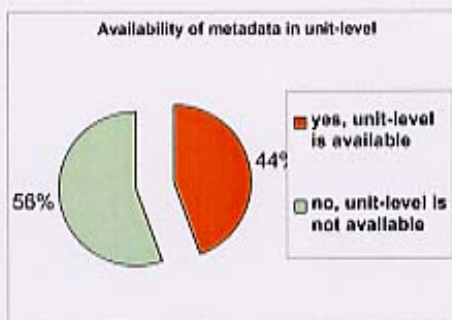
In response to the question of metadata level, 80% of organisations answered that metadata was available at series level, with 20% at unit level. It has been recognised that MEGRIN may need to clarify what is meant by 'unit'. There are implications for the MEGRIN La Clef project as this project is committed to implementing metadata at unit level. Judging by the questionnaire results, much co-ordination is required by MEGRIN to enable this implementation.

At the second workshop the stress was laid on transfer issues (e.g. by using XML for (meta)data transfer) and unit level availability rather than the core metadata structure itself.

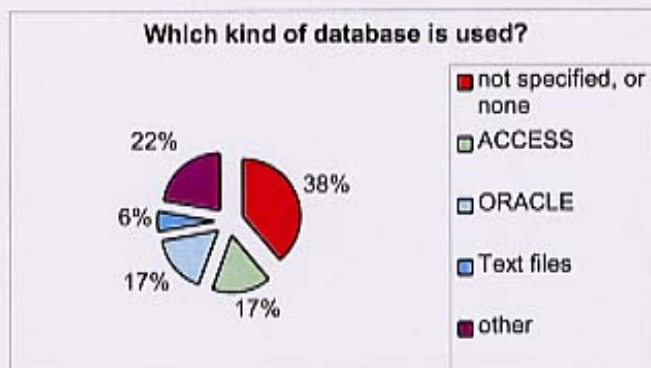
Metadata Level



The availability of metadata down to unit level has significantly increased: Whereas end of 1998 only 20% of the responding organisations had unit level metadata available, now (end of 1999, 44% of the responding organisations had unit level metadata available, and more, were willing to provide these unit level metadata to LaClef. These organisations are the NMAs of the following countries: Cyprus, France, Germany (BKG Leipzig), Hungary, Ireland, Portugal, Spain, Switzerland.



The databases used to maintain the metadata on the NMA's side range from ACCESS over ORACLE to text files. In addition to the questionnaire evaluation, the workshop highlighted a significant interest of the NMAs in using XML as an exchange tool, in order to facilitate online exchange of metadata from the producer's database (NMA) to the LaClef service (MEGRIN).



4.3.1.2 ESMI (European Spatial Metadata Infrastructure) Project

The European Spatial Metadata Infrastructure project, of which MEGRIN was a partner, finished in December 1999. This initiative was set up by several European public and private organisations to establish a framework for the distribution of geographic information by creating a universal metadata service.

MEGRIN was responsible for producing the metadata model for ESMI, largely based on their existing GDDD service. The ESMI project differed from LaClef insofar as there was no e-commerce component and consequently no need for e-commerce metadata. Another significant difference was the context of the CEN pre-standard. Whereas at the outset of ESMI it appeared the CEN pre-standard would move to an accepted European standard, it became obvious during LaClef that the ISO standard was being accepted by many European agencies. The consequence of this was a serious re-evaluation of the LaClef data model and transition from a model based on CEN to an ISO compatible model.

The implication of this transition was a serious resource burden on LaClef in redesigning both logical and physical models.

4.3.1.3 Standards

There has been a significant amount of standardisation work involved in this area. The following are the major initiatives that have been considered in this area of work:

- CEN ENV 12657 is the pre-standard that formed the basis for GDDD. The initial LaClef design was a modified version of this but essentially based on this CEN pre-standard.
- ISO. There is on-going work in ISO to create consensus across the world. In particular, ongoing activity in ISO 19115 (formerly 15046-15) as well as ISO CD 15046-12, -13 was considered.
- Dublin Core activity within CEN/MMIS.

Each of these has provided contributions to the ideas of metadata. Equally, the profusion of standards has caused some confusion within the project about which standard is more relevant to the work of the project.

The final status is that the data model is largely compatible with ISO 19115 (expected to reach final status in Summer 2001).

4.3.1.4 Final data structure

The final data structure (version 2.1) is shown in Annexe A

4.3.2 Data

In addition to the GDDD, there are a number of datasets that were considered for incorporation into LaClef. The basic criteria for considering these data sets is that they are products that do not belong to the NMAs – rather to other national organisations. Among those to be considered were:

- Geological Electronic Information Exchange System (GEIXS) project,
- Corine Land Cover (CLC) database

4.3.2.1 GDDD

As has been outlined in section 2.3.3 above, there is a need for an upgrade of the GDDD to handle:

- Language - providing multilingual capability
- Content – increasing the quality and depth of content to map-sheet level or equivalent
- Update – implementing an online maintenance mechanism to remove the burden of updating for the service provider and data providers. Primary in this is the exploitation of XML.

4.3.2.2 Corine Land Cover (CLC)

The CLC database is collected by a common methodology and according to a set of specifications in the member states of the European Community and Eastern Europe (PHARE countries). The CORINE Land Cover specifications have been defined at the beginning of the CORINE programme. Their objective is to provide the greatest possible number of users with information, which is homogeneous, fully comparable for all countries concerned and which is updated periodically. CORINE Land Cover is a European wide data set and one of the few data sets of this coverage at such a large scale of 1:100 000. It was hoped that it fits into the framework and ambitions of the new LaClef database to have a new sophisticated metadatabase.

For these reasons a catalogue of national CLC data owners was built. For all 24 European countries information on the commercial use of CLC data has been investigated. The national data dissemination strategies were identified and categorised into three groups.

The first group are those countries, selling national CLC data to all kinds of user including commercial.

Secondly those who sell their data to commercial users only with restrictions.

The third group lists countries who allow the commercialisation of value added and derived products. This is important since most countries allow this and do not allow commercial users to resell national CLC data.

Three current national pricing policy categories were identified. One policy is not to sell the data but to disseminate it for free to everybody. The second is to sell the data for fixed prices

The third pricing policy category is to make different prices for non-commercial and commercial users where the prices for the latter are much higher.

Five out of eight countries that returned the questionnaire agreed to the integration of their national CLC data in the creation of the prototype. They explicitly agreed on including their national data into the prototype and not only commercialising in general their national CLC.

4.3.2.3 GEIXS

The aim of the Geological Electronic Information Exchange System (GEIXS) project was to bring together the existing dispersed data holdings, gathered over more than 100 years, presently held by the national geological surveys, and to turn these information assets into an European resource. GEIXS is a free service of EuroGeoSurveys (the association of all European Geological Surveys). In addition to information about the datasets, commercial contact information is provided to allow users to purchase information from the relevant national Geological Surveys.

In discussions with EuroGeoSurveys, there are no legal impediments to incorporating the GEIXS data into LaClef. Technically, the GEIXS data supports the CEN 287 standard and can be made compatible with MEGRIN's GDDD and largely to LaClef. Additionally, there is interest from EuroGeoSurveys in examining e-commerce options for the supply of data.

However, during the course of LaClef - because of the status of the GEIXS project - it became apparent that we would be unable to acquire the GEIXS information and populate the database within the LaClef project lifetime. The issue of maintaining the GEIXS information also needs addressing and is an issue to be addressed in the future development of the service.

4.4 Collection and Maintenance of Data

One of the challenges the project had to face was collecting metadata for MEGRIN's central access database GDDD. At the beginning of the project, the data flow was as follows: NMAs sent either paper metadata or digital metadata and MEGRIN had to import them into the centralised access database. The texts were mainly in English. The initial aim was to provide some support to the NMAs without resources for maintaining such a database.

4.4.1 Mechanisms for collecting

The goal of the project being to connect remote NMAs to a central database using the internet, and considering the emerging technologies proposed by the W3C, it has been decided to use XML as the transfer format for both the collection and the distribution of metadata.

There were others methods considered:

1. have a distributed architecture based on Z39.50 ;
2. have a centralised architecture based on file transfer or client/server architecture.

The first approach has been rejected for the following reasons:

- the project would have had to acquire the Z39.50 technology which was not planned for the project.
- IGN-F, a LaClef partner - based on their experience within the CEO's INFEO project - was apprehensive about the Z39.50 architecture, demanding as it does that all NMAs would have to run a Z39.30 client. This approach was seen as inappropriate with the MEGRIN's mission on behalf on NMAs.
- this technology was an old one, not really in the grove of emerging ones.

The second approach (centralised architecture) has then been chosen because there were no other alternatives. Technologies like java were not accepted due to the use of ESRI software (proprietary) and seemed cumbersome for the project team. Using client/server architecture was also rejected due to the current service MEGRIN was providing to NMAs. Thus, the centralised architecture, which was at that time, used by GEIXS (also a project partner) has been analysed.

The choices the project had were to use either EXPRESS/SDAIF files (CEN/TC 287) or another Information Technology standard. A demonstration from KMS (Danish NMA) showing that Danish public agencies were using XML for transferring metadata information amongst the administration levels, was a deciding factor in the choice of XML as the relevant tool.

4.4.1.1 XML

Extensible Markup Language (XML) is a simple, very flexible text format derived from SGML (ISO 8879). Originally designed to meet the challenges of large-scale electronic publishing, XML is also playing an increasingly important role in the exchange of a wide variety of data on the internet.

XML is a set of rules for designing text formats for structured data such as metadata database fields, in a way that produces files that are easy to generate and read (by a computer), are unambiguous and platform independent.

Further information on XML is available from <http://www.w3.org/XML/>

XML was chosen as the exchange tool for LaClef metadata and a significant amount of resource was necessary to develop a DTD for exchange purposes. As a 'new' technology, there was a considerable learning curve for involved partners. Also, as a consequence of delays in developing the metadata model on which the DTD would be designed, it was not possible to implement an online exchange.

Future exploitation of LaClef will implement this XML technology and will consider the transition from DTD to Schemas.

4.5 EMF

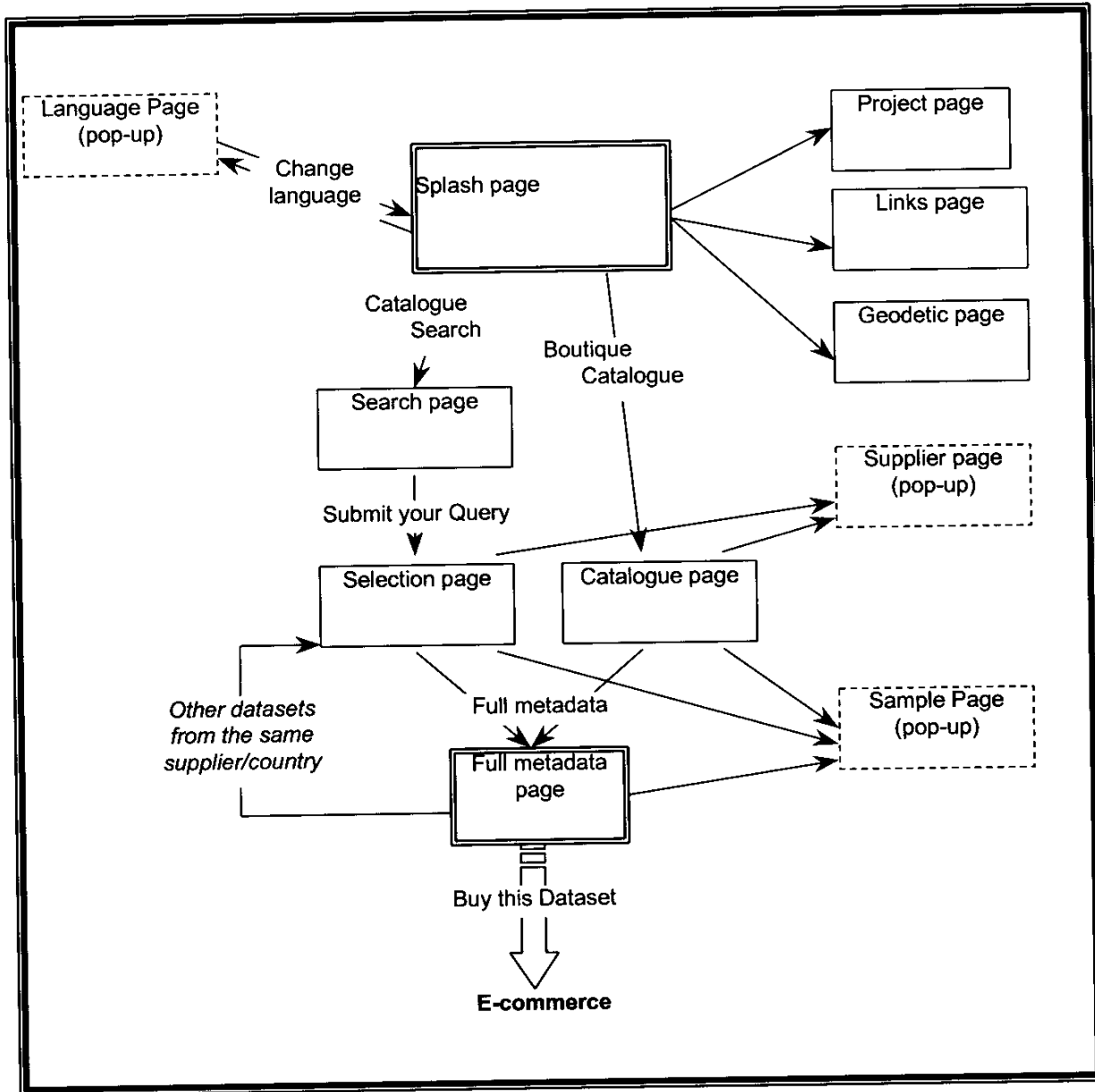


Figure 1: Summary of a session (links between pages)
 The above schema shows the links between all the pages of the EMF interface but the most effective way of explaining this interface is to present a typical sample of a scenario. Therefore, we are going to introduce its various phases: access to the site, selection of data sets, selection of metadata, e-commerce.

4.5.2 Implementation

The basic process is as follows:

Search by location using:

- Placename
- Co-ordinates
- Select from a map

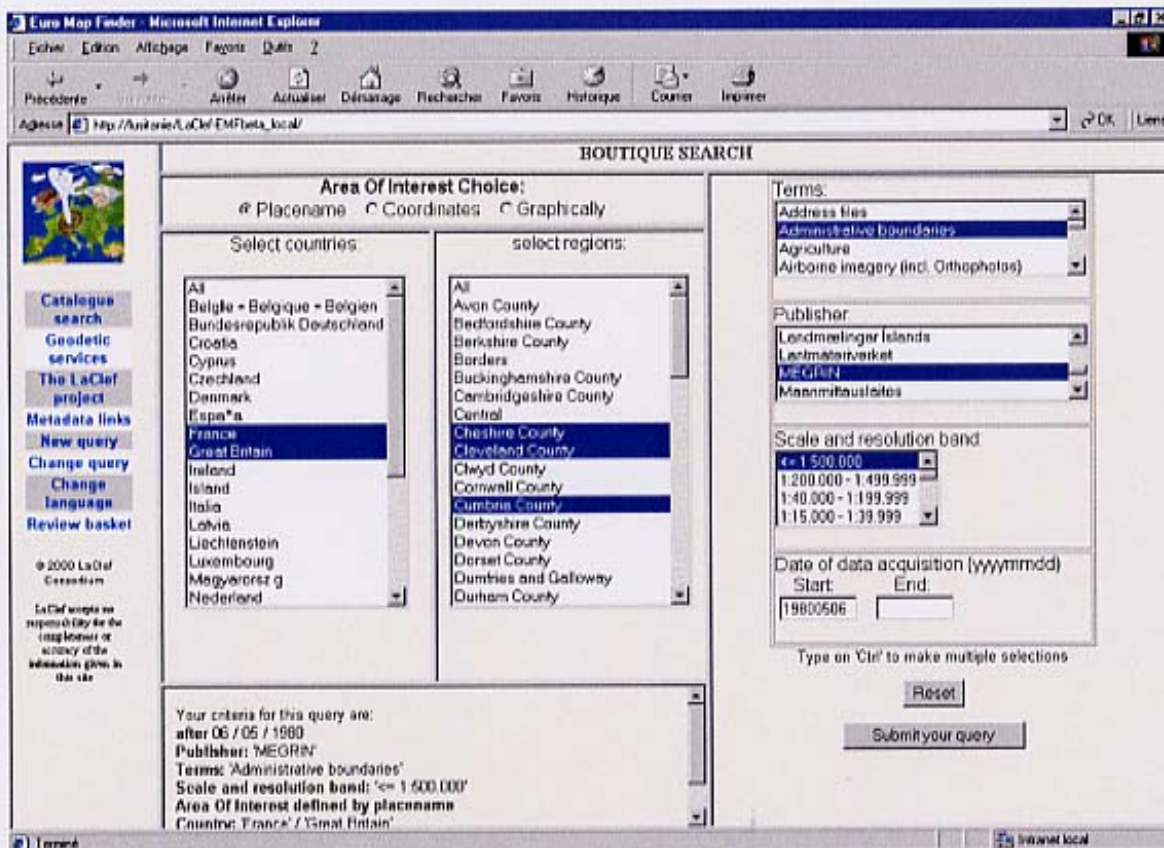


Figure 2: The Search page

The various criteria introduced in the page are summarised in a box (at the bottom left of the page).

Three methods can be used to define the area of interest.

Using the **country and region** in question (place names). The place names are taken from the SABE attribute table. The list of regions is updated when the user selects a country. Regional entities correspond with a zone level of 3. (see part II - Data). The datasets are selected when their footprints intersect the shape of the selected country or region.

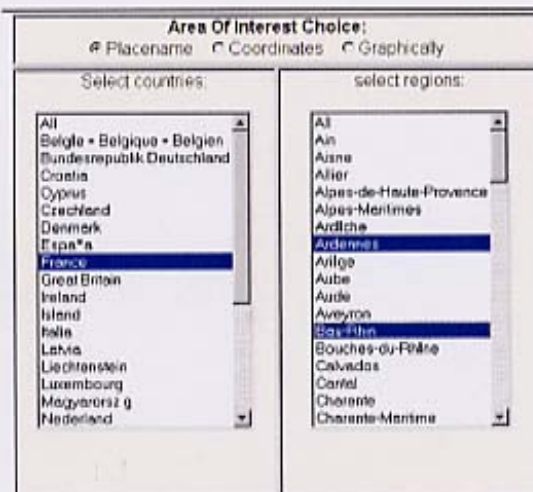


Figure 3: AOI defined by place names

The screenshot shows a web browser window titled 'Euro Map Finder - Microsoft Internet Explorer'. The address bar shows 'http://www.inra.fr/LaClef/EMF/beta_local/'. The main content area is divided into three sections:

- Left sidebar:** Contains a 'Catalogue search' button and a list of links: 'Geodetic services', 'The LaClef project', 'Metadata links', 'New query', 'Change query', 'Change language', and 'Review basket'. At the bottom, it includes copyright information: '© 2000 LaClef' and 'LaClef respects the responsibility of the accuracy of the information given in this site'.
- Map:** A map of Europe with France highlighted in yellow. Labels for 'Great Britain', 'Belgique', 'Belgien', 'Paris', 'Bundestrepublik', 'VADUZ', 'Suisse', and 'España' are visible. A compass rose is located above the map.
- Right panel (SABE France):**
 - Alternative title:** SABE France
 - Abbreviated title:** SABE FR
 - Summary:** SABE contains all low level administrative units (NUTS 5 for the European Union) for 26 European countries. SABE features: Seamless geometry, Consistent attributes, Unique identifiers to enable links to national census statistics, Details of each country's hierarchical structure, including names of the different levels. SABE is available in two levels of generalisation (30m and 200m filtering). The dataset is updated periodically. Versions exist for 1991, 1995, and 1997.
 - Terms:** # See parent #
 - Scheme type:** # See parent #
 - Character set:** ISO-10646-Unicode-Latin1
 - Scale and resolution band:** # See parent #
 - Lineage:** Created from 1:50,000 scale NMA administrative boundary source data of France. The data has been given a

Below the map, there are links for 'Other datasets from France', 'Other datasets from MEGRIN', 'Download metadata', 'No sample available', and 'Parent metadata'. A 'BUY THIS DATASET' button is located at the bottom of this section.

Figure 7: Full metadata page for the selected unit

The box devoted to 'actions' gives the following options:

to access directly data sets relating tot the same country or the same editor as the current data (return to the selection list with the criterion (editor = ... or country = ...)).

to see an extract if it is available (sample page)

to access the parent immediately above in the hierarchy ('parent metadata')

Finally, if the data is for sale on another site, this is indicated. But if it is sold on the same site, the user will be able to reach the e-commerce part developed by GIM by clicking on 'buy this dataset'.

When the selected dataset can be purchased online, the user can click on the 'Buy this dataset' button.

4.6 Services

LaClef has implemented two geodetic services:

- Co-ordinate transformation
- Additional information

Both these services use as their basis a common database.

4.6.1 Geodetic Database

The purpose of the LaClef database of European Reference Systems is to support the Geodetic services that are part of the set of EMF services. The geodetic database contains the description and definition of a wide range of European coordinate reference systems. The database has been compiled based on information supplied by EUREF, a study carried out by

the Free university of Brussels and GIM's own collection of information. Although it was agreed that EUREF would control the contents database, this has not been done because of resource constraints within EUREF.

The database consists of nine elements:

- “Map Series” - contains the map series names and their parameters.
- “Sounding Datum”,
- “Country”,
- “Vertical Reference System”
- “Coordinate System”.
- “Projection” for each co-ordinate system
- “Geodetic Datum” for each co-ordinate system
- “Ellipsoid” To each “Geodetic Datum”

The database contains 99 European map series and their related 47 reference systems from 31 European countries. The information is completed for 77 map series. The source of the parameters is EUREF’s “Datum Database Project Report”. The datum names in this report do not always correspond to LaClef’s datum names. This made it difficult to understand which datums in LaClef and EUREF are equivalent. Seven equivalent datums were identified. The eleven other datums entered into this table could not be matched with 100% confidence.

4.6.2 Coordinate transformation service

The most important of the Geodetic services is the Co-ordinate transformation service. It is also the most complicated service, from development and architecture point of view.

The user can give a set of x,y co-ordinate pairs in a co-ordinate system of his choice (to be taken from a list) and convert the set into any other co-ordinate system (to be taken from a list). The input can be typed in manually in the web page or copied from a text file on the users hard disk.

Originally it was planned to provide the service as well for GIS-formatted files (shp, mif, ...). This option has not been developed due to time constraints. However, the overall concept and architecture used to transform sets of co-ordinate pairs can be used as well for the transformation of GIS formatted files.

To implement the co-ordinate transformation service, a solution has been chosen which uses the webserver in combination with CGI/perl-programming and FME (Feature Manipulation Engine) - produced by Safe software - as the transformation engine. An overview of the architecture is given in Figure 8.

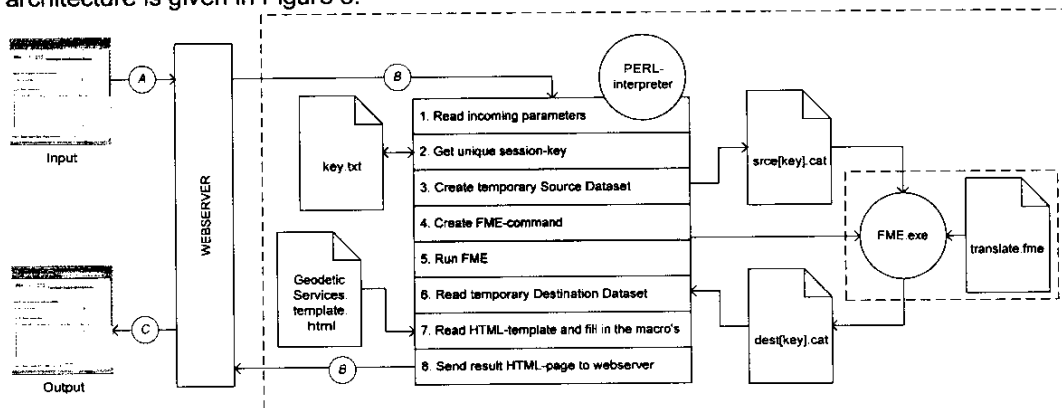


Figure 8: Quick Overview

The scheme shows the four main components of the coordinate transformation service:

1. To the left, there are two browser-windows, one for the input and one for the output. These are displayed in the user's browser and their content is sent to the webserver.
2. The user's browser sends information to the webserver that acts like a gateway for the perl-script that's co-ordinating the transformation.
3. The perl-script, in the middle of the scheme receives the information from the webserver and basically transforms this into media useful for FME to use. Finally, it assembles the resulting HTML-page which is sent back to the webserver and forwarded to the user's browser.
4. The perl-script itself does not do the transformation. The perl-interpreter uses the FME-transformation engine by sending a command to the FME console program. The results of this command are again saved in temporary files and read by the perl-script to be incorporated in the resulting HTML-page.

4.6.3 Geodetic Information service

The Geodetic Service "Map Projection Information On European Map series" gives the user the opportunity to look for information on the European Map series. The information that can be found comprises Map series, Coordinate Systems, Geodetic Datums, Map-projections, Projection-parameters, Datum-transformation-parameters, etc. All the information is taken from the "LaClef database of European Reference Systems" described above.

This database contains all the information that describes the European Reference Systems (see above). In order to "web-enable" this database extra information has been added: the labels that are displayed on the web pages in several languages. The structure of the database is as such that more languages can be added easily.

The information service also incorporates a geodetic link page to various geodesy related sites and also a Frequently Asked Questions section.

4.7 E-commerce

The project has found that the significant issues for GI e-commerce are organisational, financial and legal. Because of the diverse data sources, the permissions to make data available for online data supply are difficult to obtain. This creates a difficult administrative burden on the hosting organisation, both with regard to permissions and also data acquisition. The cost of maintaining the service is also a significant issue. The ongoing financial commitment necessary to maintain the site, develop the online metadata maintenance and complete the e-commerce and geodetic services is prohibitive.

Legal issues – specifically in relation to copyright, contract and taxation – have to be addressed. As an organisation operating within the EU, there is a need to observe the regulations within this jurisdiction. We have been able to suggest a range of possible approaches. However, there remain concerns from data providers (and users) over the supply of data "online" because of the complications and unresolved issues in this area.

4.7.1 Options

Section 3.3.3.5 detailed the options available with regard to service providers and a full e-commerce implementation.

4.7.2 The LaClef e-commerce model

For the LaClef e-commerce services no decision has been taken whether to use the first or second model described in section 3.3.3.5. Payment services are not part of the LaClef-prototype. However, the whole model has been developed to be extended with real e-commerce in the future. The different components are developed in such a general way that extending them will not be too much of a problem.

In section 4.7.2.1 the LaClef e-commerce model, or better framework, will be explained. This is still a conceptual view, though, since the way the different components are linked together will be explained. Some aspects of the implementation will be discussed in detail in the following sections.

4.7.2.1 Connecting the e-commerce model to the EMF website

Figure 9 explains the different steps the model will go through as soon as the customer selects the buy- or purchase-button of the dataset he wants to buy in the full metadata-screen at the EMF-site. The different steps are discussed below.

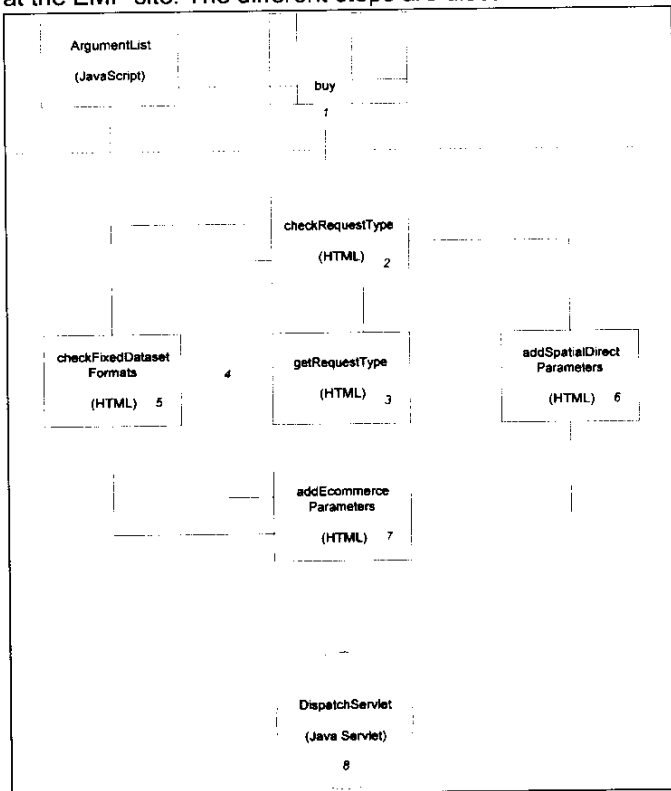


Figure 9

Step 1: Deciding to buy a dataset

When the customer wants to buy a dataset, he clicks the buy- or purchase-button in the full-metadata screen of the EMF-site. As soon as this happens, a request-URL is sent to the webpage "checkRequestType.html" (step 2). A request-URL is like an ordinary URL (Universal Resource Locator) for addressing webpages on the Internet, but also incorporates parameters. These parameters are commonly used for requests to the webserver (regarding the page that has been asked), but can also be used for transmitting information between webpages. In order to do this, a generic approach has been chosen. For the LaClef-project, a JavaScript-object has been developed to manage the incoming and outgoing parameters in the models webpages. This object can be found in the file "ArgumentList.js" and is further explained in chapter 3 of The EMF e-commerce Services Specification ref WP3/GIM/01/10. The parameters that are sent when the customer clicks the purchase-button are these:

- fixed: true/false
- services: true/false
- datasetid: a unique number identifying the dataset
- title: a title describing the dataset
- format: a list of fixed formats the dataset can be downloaded in
- minX: the "westing" border of the chosen extent
- maxX: the "easting" border of the chosen extent
- minY: the "southing" border of the chosen extent
- maxY: the "northing" border of the chosen extent
- language: an id describing the client's language

Step 2: Checking the type of request

Checking the type of request is done in the webpage "checkRequestType.html". This webpage does not really have an interface with the customer: it merely contains JavaScript-code to check what kind of dataset the customer requires. There are three possibilities:

- 1. the dataset can be delivered fixed or with services (fixed = true and services = true);
- 2. the dataset can only be delivered with services (fixed = false and services = true);
- 3. the dataset can only be delivered fixed (fixed = true and services = false).

In the first case, the customer will be shown a webpage where he can select the kind of dataset he wants, i.e. fixed or with services (step 2). In the second case, the customer will be shown a webpage where he can select the services like the co-ordinate system and GIS-format (step 6). In the third case, the customer will be shown a page to select one of the available fixed formats (step 5) or he will be guided immediately to the webpage for entering personal information, if there is only one data-format available (step 4).

Step 3: Choosing between a dataset with services or a fixed dataset

This webpage is really simple. All the customer is asked for is how he wants to purchase his data, i.e. fixed or with services (figure 10). As soon as a choice has been made, the webpage "checkRequestType.html" is loaded again (return to step 2). This time, the first possibility will not be chosen because one type of dataset has been chosen, eliminating the other one.

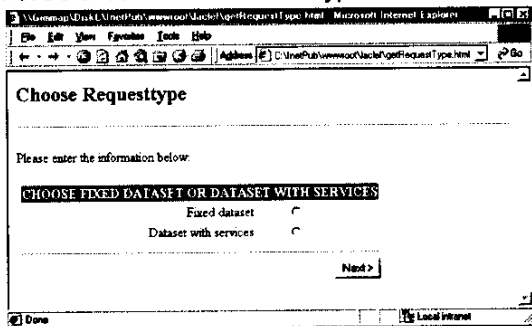


Figure 10

Step 4: A fixed dataset with one format has been chosen

When the customer wants to buy a dataset that can only be supplied fixed and in one single format, all the information for the e-commerce to proceed is known. The webpage containing a form for entering personal information is loaded (step 7).

Step 5: A fixed dataset with multiple formats has been chosen

When the customer selects a fixed dataset which can be supplied in different formats, the next page asks him to choose the dataset format (figure 11). The list of formats shown in this page is taken from the parameter "format" in the incoming URL-request. The parameter's content is parsed and the different fields are mapped to format-descriptions shown to the customer. All this is done with a considerable amount of JavaScript-code that is executed when the page gets loaded (thus, before the page is shown!).

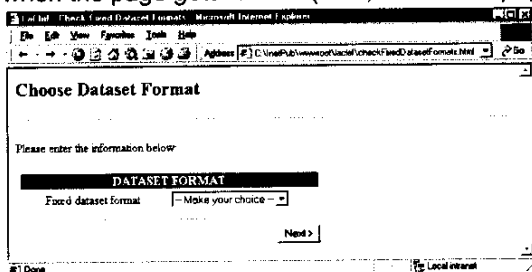


Figure 11

When the dataset format is selected and the customer presses the Next-button, he will proceed with step 7.

Step 6: A dataset with services has been chosen

When the customer selects a dataset with services, only the extent of the dataset is already known (see parameters “minX”, “maxX”, “minY” and “maxY” in the URL-request, step 1). Now, it's up to the customer to select the desired output GIS-format and co-ordinate system. All this can be done in the webpage shown in figure 12.

Figure 12

When the services are selected and the customer presses the Next-button, he will proceed with step 7.

Step 7: Adding personal information for e-commerce purposes

In step 7, the customer is asked to enter some personal information (figure 13). The required fields are his last name, first name and e-mail-address. Real address information like street or house number are less important because delivery of digital information is online. The information entered in this page is used by the Dispatch Servlet when responding to the customer (see step 8 or section 4.7.2.2 and 4.7.2.3).

Figure 13

When the required fields are filled in and the Next-button is pressed, a URL-request containing all the information will be sent to the Dispatch Servlet (step 8). The URL-request sent to the Dispatch Servlet contains these fields.

Step 8: Starting the Dispatch Servlet

The Dispatch Servlet is called like any other webpage with a URL (in this case a URL-request). The webserver is configured in such a way that instead of delivering an HTML-page to the browser, the servlet-runner starts the Java-servlet “DispatchServlet.class”. The parameters are sent to the servlet in the same way as between the webpages, i.e. with a URL-request. The working of the Dispatch Servlet will be explained in more detail in the next two paragraphs.

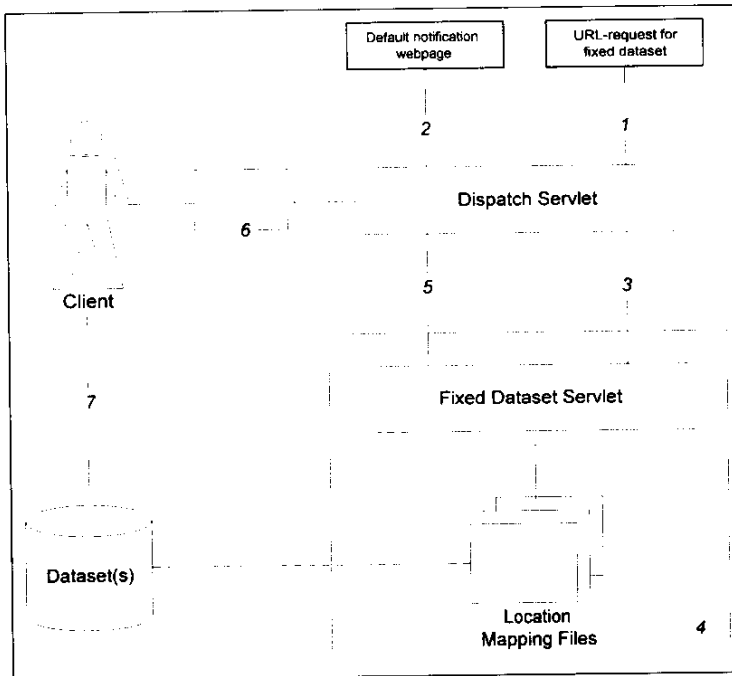


Figure 16

Step 1: Getting the request

The Dispatch Servlet is called by the client's browser when the customer commits the form in the webpage to add e-commerce parameters (this is personal information). The browser sends a request-URL to the Dispatch Servlet, containing all the information to dispatch the requested dataset. When the customer wants a fixed dataset, the request-URL contains the parameters the Fixed Dataset Servlet needs to handle the request.

Step 2: Sending a default HTML-page

The Dispatch Servlet sends a default-webpage back to the client's browser to notify the customer his request will be taken care of and the result will be send by means of an e-mail (figure 15). The customer's name and e-mail-address are filled in dynamically (these are passed as e-commerce-parameters in the request-URL to the Dispatch Servlet).

Step 3: Sending the information to the Fixed Dataset Servlet

The Dispatch Servlet parses the different fields in the incoming request and sends another requests to Fixed Dataset Servlet, containing all the necessary fields.

Step 4: Getting information from Location Mapping Files

The Fixed Dataset Servlet receives a request-URL containing an ID of the dataset to be delivered. Now, the Fixed Dataset Servlet looks up this ID in a collection of Location Mapping Files to find the URL the dataset can be downloaded from. This location typically will be the server where the servlets are running, but could be any location on the Internet world-wide.

Step 5: Sending the result to the Dispatch Servlet

The Fixed Dataset Servlet responds to the Dispatch Servlet by sending it the URL where the dataset can be downloaded.

Step 6: Sending an e-mail to the client

The Dispatch Servlet responds to the client by sending him an e-mail containing all the information he needs to download the requested fixed dataset. The e-mail will contain the URL pointing to the location where the dataset can be found. If the lookup of the ID in the Location Mapping Files by the Fixed Dataset Servlet was unsuccessful, an appropriate e-mail will be send to the client as well.

Step 7: Downloading the dataset(s)

Finally, the customer can download the requested dataset(s) by clicking the link in the e-mail-message.

4.8 Language

Euogeographics has sent all text information used within the service to its international members for translation. Quite a number of these translations have been implemented in the service and this activity is ongoing to complete all European languages.

4.8.1 Implementing multi-linguality

A main issue in the LaClef web site is multi-linguality. In the overall architecture ASP is used to provide multi-linguality. With this principle labels and contents that need to be translated is stored in "translation" database and through the use of ASP-tags the labels corresponding to a certain language variable are displayed in the user's browser.

The major drawback of this system is that the information to be translated in a database (in the LaClef case Oracle) and is not easily accessible by those who do the translation. The use of XML to provide a better system for multi-lingual web pages has been investigated. The emerging XML technology is mainly used for the transfer of data. However, combined with other techniques XML can be used to replace HTML. Microsoft's Internet Explorer browser is already ready to do so.

The concept which has been elaborated is the use of XML in combination with XSL (new technology) and CSS (old technology). The principle is that content is separated from structure, layout and style. The content is stored in XML which is 'readable' by human beings; only content is translated, layout and style which are stored in respectively XSL and CSS files should not be touched.

A small application that facilitates the translation of XML content has been created: an intuitive "XML/XSL Language Converter Tool". Because this technology is not yet supported by all web browsers (only Internet Explorer 5 and Netscape Navigator 6) the concept has only been implemented as a prototype for some LaClef web pages. However, the prototype is ready for use as soon as more users have turned to newer web browsers.

4.8.1.1 Option 1 - templates

Multi-linguality is obtained in the LaClef web pages through the use of a 'translation' database and ASP (the XXML/XSL solution is only implemented as a prototype because it is not supported by older browsers)..

The main problem with incorporating multi-linguality for the coordinate transformation service is the working together of ASP and CGI-programming with perl. These server-side technologies are substitutes instead of being complementary. Microsoft's ASP is used for letting a webserver assemble the resulting webpage dynamically. Likewise, CGI-programming let's a perl-script assemble the webpage before it gets sent to the client's browser. Thus, ASP and perl-scripting really belong to the same family of technologies (amongst others like Allaire's ColdFusion, Java servlets, JSP and PHP to name just a few). Thus, from a client's browser point of view, it's either multi-linguality with ASP or the coordinate transformation service by using a perl-script, but not both!

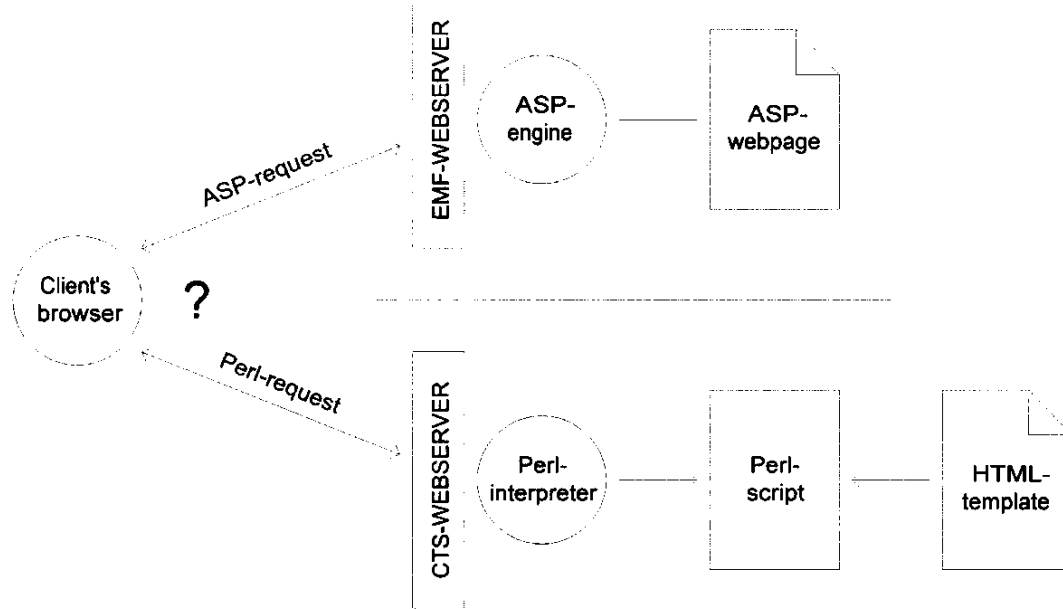


Figure 17

In the next two paragraphs, two solutions will be proposed. Each model has its advantages and disadvantages.

Model 1: multi-lingual templates

The first solution is by far the simplest one, though not the best to implement. Here, a language-id is sent to the webserver where the perl-script is running. Instead of always using the same webpage-template, the perl-script gets the template that matches the language-id (Figure 18). Thus, when the user's language is English, the perl-script reads the template with the English interface and sends it to the user after the template-fields have been replaced. If the user calls for an interface in another language, then the other language template is read and sent to the user after the template-fields have been filled in.

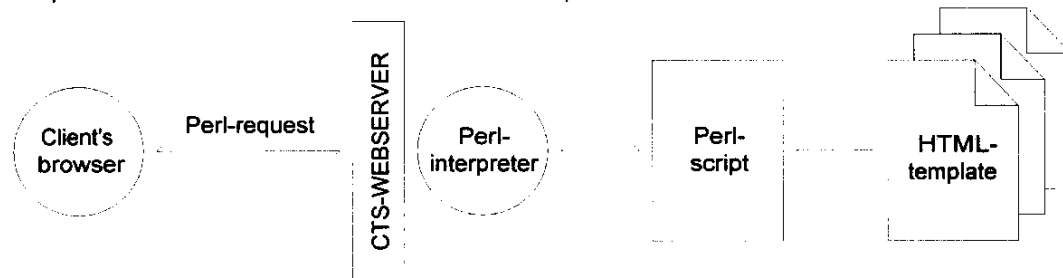


Figure 18

Implementing this solution is really simple. All that has to be done is develop the templates in all supported languages and extend the perl-script to figure out which language the user is preferring based on an extra language-argument.

On the contrary, this model is rather hard to maintain. For the EMF-site, all the language specific information is stored in one central database, which is read by the webserver solving the ASP-tags. The language-templates break with this approach. Even more, if the webserver hosting the EMF-site and the one running the perl-script are different (which is the case at the time of this writing), then the maintenance of the language-specific information needs to be done on two physically different servers.

Model 2: using CGI-programming and ASP together

This solution is the most ingenious one, but is rather hard to implement and may be difficult to maintain from a programmer's point of view. When discussing the problem with multi-

linguality and CGI-programming, it was said that ASP and perl-scripts cannot work together because they are substitutes. It's either one or the other, but not both. Well, this is true when the client's browser communicates with one of the two webservers in Figure 17. Using ASP and CGI-programming with perl is possible when one webserver, behind the scenes, communicates with the other one, as shown in Figure 19.

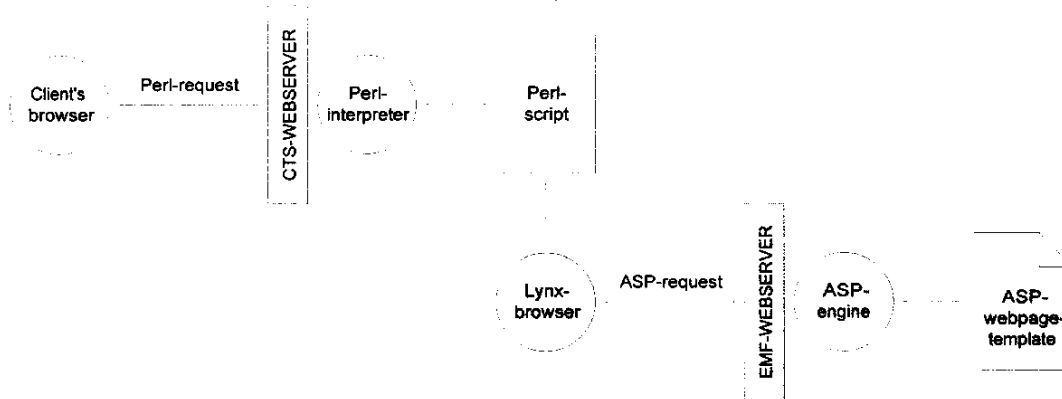


Figure 19

The client's browser still asks the webserver for a perl-script to run. Nothing changes there, but instead of reading a fixed HTML-template to send back to the client's browser for output, this time, the perl-script asks the EMF-webserver, running ASP, for this template, more or less in the same way a client's browser does in all the other pages from the EMF-site. Of course, the perl-script merely needs the content of this template: it does not care about displaying it like a browser does. That's why it uses a console browser like Lynx to retrieve the content of the HTML-template. Lynx does exactly the same thing a full-blown webbrowser like Microsoft's Internet Explorer or Sun's Netscape Navigator does, i.e. asking a webserver to send it a webpage. Now, the Lynx-browser communicates with the EMF-server sending an ASP-request and retrieving an HTML-template, which is language specific. This page still is a template, because it contains some fields the perl-script is going to replace with values just before sending the resulting webpage to the client's browser.

In this architecture, the fixed HTML-template is replaced by a sophisticated implementation for using the EMF-server's ASP-engine to incorporate multi-linguality. This is, of course, a nice work-around but is complicated nonetheless.

Implementation

For the coordinate transformation service in the EMF-prototype, no multi-linguality has been implemented yet.

4.8.1.2 Option 2 - XML/XLS

Concept of XML – XSL – CSS

HTML (HyperText Mark-up Language) is derived from SGML (Standard Generalised Mark-up Language), an international standard for defining both the structure and the content of electronic documents.

- XML (Extensible Mark-up Language) is a syntax to describe and deliver structured data from any application in a standard and consistent fashion. XML is a generic language, and lends itself well to the creation of specialised versions, designed for a particular purpose.
- XSL (Extensible Stylesheet Language) is a set of presentation rules, defining how tags should be rendered. The XSL processor is a generic engine that applies formatting patterns to elements, expands elements to text and associates CSS styles with elements.
- CSS (Cascading Style Sheets) is a means of assigning display values to page elements.

At this moment, Microsoft Internet Explorer 5.0 or higher is the only browser that directly supports XML files. Opening an XML file will automatically parse it with the related XSL and CSS file. For other browsers, server-side merging is possible, resulting in an HTML page browsed by the web client.

Use for multilingual pages

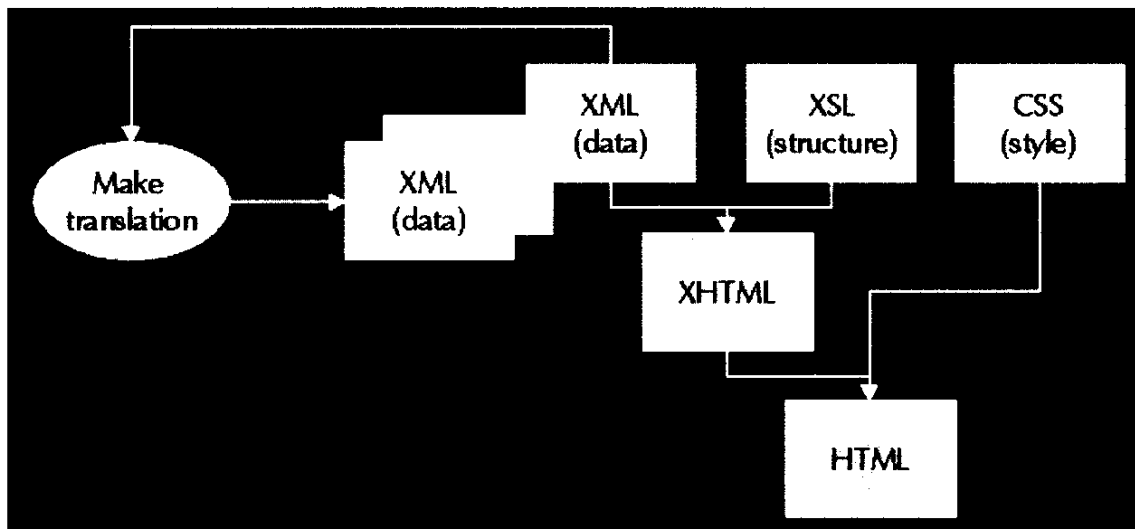


Figure 20 Use of XML/XSL/CSS to support multi-linguality

HTML has evolved into a complex and somewhat messy combination of structuring and formatting tags, which makes maintaining and translating large and complex websites very difficult.

CSS allow you to isolate formatting codes applied to many pages, so that the general stylistic changes can be made by editing a single file.

XML allow you to isolate the content of a page, so that translations of a webpage can be made by translating a single file with simple structure.

XSL put the XML data in the desired structure and adds the CSS style definitions.

Once a set of XML + XSL + CSS file has been created, it is sufficient to translate only the XML file. This can result in several XML files (one for each language), using all the same XSL and CSS file.

In order to facilitate the translation process, an intuitive "XML/XSL Language Converter Tool" was developed.

4.9 User Feedback

4.9.1 Research

Research was conducted into two distinct areas, namely user requirements for additional services and e-commerce.

The research showed that the complete range of services that could be offered via the LaClef interface could be classified as:

1. Facilities to search the metadata and display the results.
2. Provision of some standard or customised GI products using e-commerce.
3. Value-added services to the standard or customised GI products.
4. Services for data suppliers to support the provision of LaClef compatible metadata.
5. Additional GI related services independent of the metadata and the available GI products.

Some of these, notably (1), (2) and (4) were identified as fundamental to LaClef to ensure the main objective is met of offering an effective GI metadata and on-line data supply service; at least as a working prototype. It was also highlighted that careful consideration of other value-added services would be required with the appropriate allocation of resources.

4.9.1.1 User Segmentation

A simple segmentation of the user population was proposed, in decreasing order of expertise in searching for GI and identifying suitable data:

- Traditional principal users of GI: public institutions, local administrations, and utility companies.
- Recent users of digital GI: transport companies for route optimisation, geo-marketing companies.
- All others - use of GI by citizens, e.g. route planning, identification of nearest facilities.

4.9.1.2 Searching for GI using the Metadata

Via the multilingual interface users will need to be able to:

- Search the metadata databases to identify sources of GI matching their interests via:
 - a geographical area of interest: area names, administrative units, town names...
 - area selection on a graphical map background
 - a hierarchy of themes
 - keywords
 - a scale range.
- Assess the quality of the data identified (optional).
- View sample data (optional).
- Query features in the sample data for attributes (optional).

4.9.1.3 GI Product Delivery

Research indicated that LaClef would need to offer the user the opportunity to select several GI products that are then priced, paid for and delivered. These products may be from a set list of standard products or may, to a greater or lesser extent, be customised according to the needs of the user prior to delivery.

4.9.2 Customised GI products

Users may be offered additional services to the delivery of products by allowing the definition of customised GI products. Customisation of the standard GI products will clearly depend on which standard products are available via LaClef and on agreements between the LaClef project and the data providers. Possibilities for customisation include:

- Change from the default exchange format to another exchange format.
- Change from the default projection/coordinate system to another.
- Customisation of the geographical extent of a standard product: cookie cutter selection of an area or point identification.
- Combination/intersection of data from different sources (e.g. generating thematic maps that combine base maps with environmental analysis).
- Creation of the customised GI product defined by the user (mandatory).
- Ordering of and payment for a customised product (mandatory).
- On-line delivery of a customised product (optional).
- Off-line delivery of a customised product (mandatory if on-line delivery is not available).

4.9.2.1 Services Independent of the Metadata and Available GI Products

GI related services recommended for LaClef were aimed at:

- Attracting more people to the LaClef website.
- Providing a resource to help people to improve their GI knowledge.

Options include an exchange format translation service, projection/co-ordinate system translation service and a geodetic information service.

Exchange format translation services

- On-line service for exchange format translation on data provided by the user (e.g., vendor format, national formats).
- Off-line service for exchange format translation on data provided by the user.

Projection/co-ordinate system translation services

- On-line service for projection/co-ordinate system transformation on data provided by the user.
- Off-line service for projection/co-ordinate system transformation on data provided by the user.

Co-ordinate and geodetic services

- Information service on geodetic reference systems and co-ordinates used at the national or international level.
- Reference guide on geodetic semantics and transformation parameters.
- Algorithms and tools for co-ordinate conversion and geodetic transformations with accuracy statements.
- On-line service for co-ordinate conversion and geodetic transformations on data provided by the user.
- Off-line service for co-ordinate conversion and geodetic transformations on data provided by the user.
- Off-line service to convert vector data to raster data on data provided by the user.
- On-line service to convert vector data to raster data on data provided by the user.

During the LaClef project it was only be possible to implement a limited number of value-added services as examples of what can currently be done with Internet based tools.

4.9.2.2 E-commerce

Results from LaClef research into effective e-commerce showed that a customer-friendly LaClef should have (<http://www.usability.serco.com>):

- Obvious access to a centrally located store entrance.
- Meaningful product categories.
- Clear and informative product descriptions.
- Pictures of products and categories.
- Search facilities that accept appropriate terms (multiple words and synonyms).
- Search results in a sequence that is logical to the user (e.g. product name, price, category).

- Clear indication of which items are available from the on-line store.
- Feedback to the user as items are added to the shopping basket (a cumulative list of items in the basket, not simply the most recently added item).
- An option to review the status of the order at anytime and easily remove things from the shopping basket.
- Explanations of technical terms and jargon.
- A clear indications of the security measures taken by the site.
- A 'value-added' shopping experience, e.g. using product reviews.

Most problems with e-commerce sites are due to misunderstandings on the part of the site creators about how users understand the structure and elements of typical commerce transactions. Users have formed schemas to understand commerce based on their experience of off-line shopping.

Although LaClef has not developed the e-commerce functionality to the point where many of these issues come into play, the modular nature of the development allows these issues to be catered for in the future.

Research also showed that customers must see the LaClef service as trustworthy and efficient. The research found that:

- The web site will need strong branding, building on the reputation of the data suppliers and will need to be well designed to ensure that the customer does not get lost somewhere in the transaction process and fail to complete the e-commerce order.
- Customers must be convinced that the quality of service via LaClef as an intermediary exceeds that of going directly to the data supplier via the standard distribution channels.

This additional quality may come from a greater choice, easier ordering, better customer support or a whole host of other parameters. If it is more difficult for a customer to buy via LaClef than directly from the data supplier then the customer has no reason to use the e-commerce part of the LaClef service.

- Successful web sites are based on rich tools, personalisation, and channel integration.
- Customers must know at all times whether they are accessing the metadata service, the e-commerce service or additional services of LaClef. The combination of several, inter-related services in LaClef need careful management.

4.9.3 User Testing

User testing of our development has shown that the basic LaClef model is an effective tool in researching geographical datasets. The general interface is seen to be simple but intuitive. The search page options were generally accepted as appropriate. Specific areas for future development were identified both within the consortium and by external users.

Particular issues highlighted include:

- The quality of the metadata content,
- The quality of the graphic interface,
- Lack of help and clarity of language in interface,
- Need to clearly define tool functions,
- Requirement to use English in conjunction with national language,
- Lack of any price information,
- Definition of areas covered and displayed,
- Layout and design requires improvement.

Many of the issues identified were recognised during the course of the project. However, due to resource limitations within the technical development area, it was not possible to implement these changes.

All of the user comments will be considered within the context of the future development and exploitation of the site and overall service.

5 Lessons

5.1 Dynamic environment

There are two aspects of the environment which have and will continue to effect LaClef.

- Standards - As has been described in section 4.3 above, the development of the LaClef data structure was overtaken by the development of the ISO 19115 (formerly ISO 15046-15). There are clear indications that it will be well supported – and this will have implications for future development of LaClef. It should be remembered that the actual LaClef model is compatible with the ISO standard.
- Technology – The software tools that were used for the EMF, ie ESRI's IMS and SDE, have been superseded by new products. This casts doubt on the future technical support for the service. The speed of XML development meant that during the course of the project lifetime much of the work done became almost obsolete, eg transition from DTDs to Schemas.

The speed of both developments, standards being slow and technology being fast, impacted on the progress of the project. With regard to the technology, a greater awareness within the project of external events may have facilitated future development of the service.

5.2 Project planning

The nature of the work was such that there was a great deal of inter-relationships between different work packages – especially between WP 2 and the other development WPs (3, 4 and 7). When delays occurred in WP2, there were accompanying delays in the other WPs. The result of this was a great deal of confusion and misunderstanding about different aspects associated with the development of the project.

The process of creating Work Packages was laudable – attempting to identify independent areas of work that can be given to different parties. However, greater care must be taken to ensure this independence of work – especially in areas of development. Where a very significant dependence occurs, this must be addressed as a priority. (It must be noted that the division of work between the other WP – especially 3, 4 and 7 worked very)

5.3 Commitment to the project

One of the major problems with the project was that it suffered from a huge “drop - out” rate, e.g., there are only 2 consortium representatives who remained involved with the project throughout its life. Much of this interchange of staff was due to the dynamic business environment in which each of the partner's organisations compete and was completely unavoidable. However, it had a severely destabilising effect on the consortium and disrupted production.

The lesson for the future is to stabilise involvement as far as possible and to ensure commitment of all project partners.

5.4 Factors affecting progress

Other factors that impacted on progress were language and distance. Language was a factor, especially within the technical development, with clarification sometimes needed after development resources had misinterpreted direction.

With a number of different partners involved in the technical development, distance between these partners sometimes also impacted on progress.

The lesson learnt is to fully clarify all agreed decisions before technical changes are implemented, especially within a multi-cultural consortium. With regard to distance, it is

necessary to ensure work is assigned appropriately, eg through a modular approach, and there are formal lines of communication established.

5.5 Future exploitation

The aim of the LaClef project was to explore the issues concerning the unlocking of publicly held geographic data; to provide a tangible demonstration on how technology can provide users with the data they want to obtain; and to create a model for other organisations to utilise. This has been done by a combination of reports and demonstrations.

Important issues at stake were:

- multi-linguality of content and user interface;
- the use of standards;
- distributed database system and access;
- availability of other GI related services;
- online access / provision of data;
- implementation of a graphical interface for search and retrieval of information;

The LaClef project has indeed resulted in a model and well-advanced prototype/demonstrator application that facilitates the unlocking of geographic information available in the public sector domain through the use of e-commerce. The model is built around three distinct modules or services: the metadata service, the geodetic service and the e-commerce service. These modules co-operate together to provide an all-embracing service. Because of the modular approach, each service can also be used separately and independently. The project work was further elaborated by development in metadata exchange and multi-linguality.

At the end of the project, the project team has gained a much better insight into the issues at stake and it is felt that the achieved results offer a very good base for further work in this field. There are in fact several paths to follow to exploit the achievements of the project. The two most important directions we can go are:

1. To make the already well-advanced prototype operational as a service offered by Eurogeographics, possibly in a joint venture with other partners.
2. To exploit the overall concept or the separate modules targeted at Eurogeographics members (members states) or at other players in the field, both public and private sector.

5.6 SWOT Analysis

A SWOT analysis of the LaClef results has been conducted to identify the possible future direction.

- The strengths indicate the current functionality of LaClef
- The weaknesses indicate where improvements or further developments are required to turn the current LaClef prototype into an operational application.
- The opportunities indicate where and for what type of partners the LaClef model, as a whole or as single modules, can be exploited and by whom (the consortium as a whole or the individual partners)
- The threats indicate the preconditions for the successful implementation of the LaClef model.

5.7 Analysis

5.7.1 Overall LaClef model

5.7.1.1 Strengths

- All-embracing model that covers the processes from metadata collection over catalogue search to online provision of geographic data. Complementary geodetic services are added to make the offered information more complete.
- Modular: each module stands on its own and can be used separately or plugged into existing systems.
- Multi-linguality is consistently built in, both at content level and at user - interface level.
- Technology and data standards are used wherever possible.
- New technologies such as XML are investigated and used where appropriate and possible.

5.7.1.2 Weaknesses

- Although well advanced, the LaClef application is still a prototype. In order to make it operational some final development work is required.
- Some issues are, though investigated and studied, not developed in the prototype: mainly issues of data exchange, data maintenance and e-commerce payment transactions (payment clearinghouse).

5.7.1.3 Opportunities

The project team has gained a very good insight into the relevant issues. There is a very good basis established which can be used to make the all-embracing application operational. The LaClef application in its current state proves that the model is robust and a fully functional application is feasible.

Exploitation of the LaClef results by Eurogeographics

Content-wise, the LaClef project is built around geographic information available at the NMA members of Eurogeographics. The first step in the exploitation of the LaClef results is logically the use of it by Eurogeographics itself.

Eurogeographics can offer a single European platform for the distribution of European-wide geographic data of high quality. This is both a service to the data providers and to the GI user community.

From the providers viewpoint:

- They are associated with European partners of high visibility and standards through a common platform.
- Information about their data and even the data itself is made available through one common platform.
- Other public and even private data providers might be associated.
- Partners are not required to have an expensive and complex architecture in place. They can make use of the infrastructure offered by Eurogeographics.

From the customers viewpoint:

- The platform offers one - stop for searching and obtaining European wide geo data.
- The data are of a "known" provider of high and secured quality.

- The user interface is multi-lingual which makes the search engine more accessible in a European-wide context.
- The user can search and find data for areas of interest that are crossing national boundaries.
- The user can order data online and possibly even retrieve data online.
- Use LaClef building blocks as a toolkit for NMAs and other public or private organisations

More and more organisations become conscious that publicly held geographic information must be unlocked. Some organisations are already well advanced in this process, but many organisations only have just started or still need to start. The LaClef experience should be used to assist the latter organisations in the process of making their geographic data available. This can be done by providing the LaClef building blocks as toolkits or by offering the expertise of the LaClef consortium or its individual partners.

- Mapping agencies and other public or private data providers that have already an operational metadatabase in place now want to take the next step to make these metadata and the data itself publicly available online (through e-commerce). They will be assisted, either to set up their own web site or to set up links and associations with the Eurogeographics platform. The LaClef project team as a whole or the individual partners of the consortium can help the organisation in re-using and integrating the required building blocks of the LaClef project in their own systems.
- For data providers that are still in the start-up phase the LaClef application can be offered as a toolkit for the creation of metadata storage systems and e-stores for geographic information.

5.7.1.4 Threats

- The architecture and technical environment is complex. While appropriate for this particular development, it means that high-level technical expertise is required for maintenance. No application - how good it might be developed - will run by itself, especially not highly interactive internet applications. To keep such an application running, a dedicated technician is required during the life time of the application.
- The current architecture is not cheap: the LaClef building blocks are based on commercially available software, mainly Microsoft, Oracle, ESRI and Safe for which software licenses are required. Internet licenses are particularly expensive.
- A precisely defined and closely monitored process (organisational and technical) is required to keep the metadata and the data to be provided online up to date. If data are not kept up to date, the service will become useless very quickly.
- An internet service needs to be fast. Because a graphical user interface (maps) is required, sufficient bandwidth at the service provider is necessary.

5.8 Proposed exploitation

The Eurogeographics Management Board is presently considering the following options. It is envisaged that whatever option is chosen, it will be developed in partnership with some of the LaClef consortium members, members of Eurogeographics or private partners.

Various development options can be considered separately or in combinations:

- The exploitation of the metadata module only as a European cross-sector metadata service.
- The exploitation of the e-commerce module built on top of the metadata module or exploited separately.

- The exploitation of the geodetic service , included in the overall service or exploited separately.

5.8.1 European cross-sector metadata service

5.8.1.1 Finalise current development

The development of the current LaClef application is well under way but needs to be finalised and ported to an application that fulfils the Eurogeographics requirements. In this option, only the metadata service is ported.

5.8.1.2 Actions to be considered include:

- Adaptations to user- interface: graphic design, adaptation to Eurogeographics house style, removal of pages and fields which are not required (for example if e-commerce service is not exploited).
- Final and thorough functionality and performance testing
- Update contents of database and other pages
- Debugging and fine-tuning

5.8.1.3 Hosting the site and maintaining the site

An Internet site must be fast and stable. Adapted hardware adequate bandwidth is required. There are a number of options available:

- Hosting within a LaClef consortium member, eg GIM.
- Hosting externally by a dedicated Internet Service Provider.
- Hosting by Eurogeographics centrally at Head Office.

Irrespective of which option is chosen, remaining actions include:

- Evaluation of hardware requirements and purchase of equipment if necessary
- Initialisation and configuration of the server, required software and data
- Provision of network infrastructure, security, servers, etc
- Ongoing maintenance (see 'Creation of a metadata exchange tool' below).
- There are significant resources required to complete the development and maintain the required level of service.

5.8.1.4 Creation of a metadata exchange tool

The metadata exchange tool consists of three parts:

- the XML/DTD: the definition of the structure of the XML metadata exchange file;
- the tool to import/update from a supplied XML exchange file to the Oracle database;
- An XML based extraction tool for each provider to extract data from an existing metadata system.

Requirements

- Finalise XML/DTD work - validate XML (discussions with some data providers that have experience with XML and their metadata is required, especially to look at the option "conceptual" or "physical" model. It is also useful to look at tools that are becoming available for the new ISO standard.

- Create tool to import XML to update the Oracle database. This is not a simple importation tool. Validity of the given files must be checked, checking for existing data records etc. must be done and the import must be done in an intelligent way.
- Create tool to export to an XML file from the Oracle database.

5.8.1.5 Input tool for metadata

If a data provider has not yet an advanced metadata system in place, he will not be able to extract XML metadata exchange files from his database. For those providers it might be useful to develop a tool that helps them creating the XML exchange file manually. The complexity of such a tool will also depend on the defined XML/DTD. It is also useful to look at tools that are coming available for the new ISO standard.

Necessary actions:

- Evaluate if an input tool for metadata is a requirement and if yes what the functionality of such a tool should be.
- Look at tools that come available for the ISO metadata standard.

5.9 Geodetic services (added to European cross-sector metadata service)

The proposal is based on the assumption that these services are added to the metadata service and utilise the existing hardware and software.

5.9.1 FAQ and Useful hyperlinks list

To finalise development the graphics must be adapted to the house style of Eurogeographics. An input tool (based on web technology = 'extranet') is required to maintain the content of the database.

5.9.1.1 Identified actions

- Adapt interface to house style of Eurogeographics
- Create input tool for maintenance of content.

5.9.2 Information about European co-ordinate systems

To finalise development the graphics must be adapted to the house style of Eurogeographics. Information remains to be validated by EUREF and quality accreditation agreed.

5.9.2.1 Identified actions

- Adapt interface to house style of Eurogeographics
- (Minor) changes following EUREF quality accreditation.

5.9.3 Co-ordinate transformation tool

To finalise development the graphics must be adapted to the house style of Eurogeographics. Information remains to be validated by EUREF and quality accreditation agreed.

5.9.3.1 Actions:

- Adapt interface to house style of Eurogeographics

5.10 E-commerce service (added to European cross sector metadata service)

The proposal is based on the assumption that these services are added to the metadata service and the exploitation of existing hardware and software.

The e-commerce service uses the metadata service as catalogue.

The most important issue that needs still to be elaborated is the handling of payments. A clearinghouse system must be chosen and agreements with a provider must be made.

This proposal is based on the assumption that the e-commerce system is built on the metadata service and that "customised" data and "standard" products are provided online.

Administrative issues are not covered in the implementation phase, eg follow-up of client purchases, account management etc.

5.10.1 Necessary actions:

- Final evaluation and decision on payment handling and investigation of administrative issues such as client purchase follow-up, account management etc.
- Add payment handling to current e-commerce "dispatcher";
- Finalise and fine-tune the implementation
- Prepare data for "customised" online delivery
- Adapt the metadata e-commerce elements according to products sold

5.11 Other actions

5.11.1 Licenses

The nature of the datasets being considered for sale through the LaClef service required the use of a single standard license for each product, e.g., one for SABE data, one for CLC. The ultimate aim for the service will be for the service provider to manage the complexity of different license arrangements with the providers and provide the user with a common license that is independent of the different products. This issue was not tackled within the project and will require further work.

6 Conclusions

LaClef has provided an investigation into the issues and process for the supply of GI data "online". The outcomes of the project are:

- a series of detailed reports setting out the issues and possible approaches to the problems associated with the online supply of GI data, and
- a prototype that shows the processes that are required for the supply of data – from the initial "discovery" of data, the determination of it's "appropriateness for use" and finally its "acquisition".

While the project has encountered technical and organisational difficulties, it nevertheless has largely succeeded in completing its task and has built a viable model for the future.

The expected finalisation of ISO 19115 Metadata standard in 2001 will provide some stability to the service but may necessitate modification of the model to ensure conformity. The advantage of such conformity is wider software support.

The improvement of licensing arrangements and making datasets available for online sale will require some further work but will be addressed as part of the follow-up.

Further development is required in the area of online data exchange using XML and in the selection of an e-commerce service provider.

Significant issues arising from the project are the costs associated with the development of the service and the ongoing maintenance. Provisional estimates to complete the development suggest Euro200k with ongoing costs of Euro100K. This remains the most serious obstacle to final completion of the development and future exploitation.

Annexe A - LaClef Metadata Elements

6.1.1 UML Notation

The Unified Modeling Language (UML) is a language for specifying, visualizing, constructing, and documenting the artifacts of software systems, as well as for business modeling and other

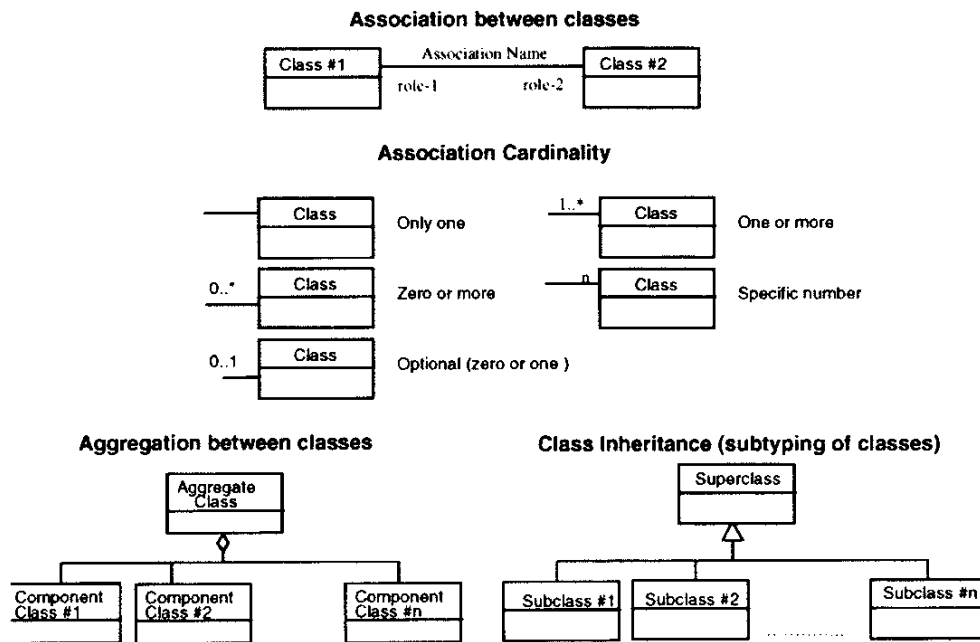
non-software systems, as reflected by the LaClef application in modeling the metadata. The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems. Good models are essential for communication among project teams, eg between the database designer and metadata manager and to assure architectural soundness. As the complexity of systems increase, so does the importance of good modeling techniques.

The modeling language includes:

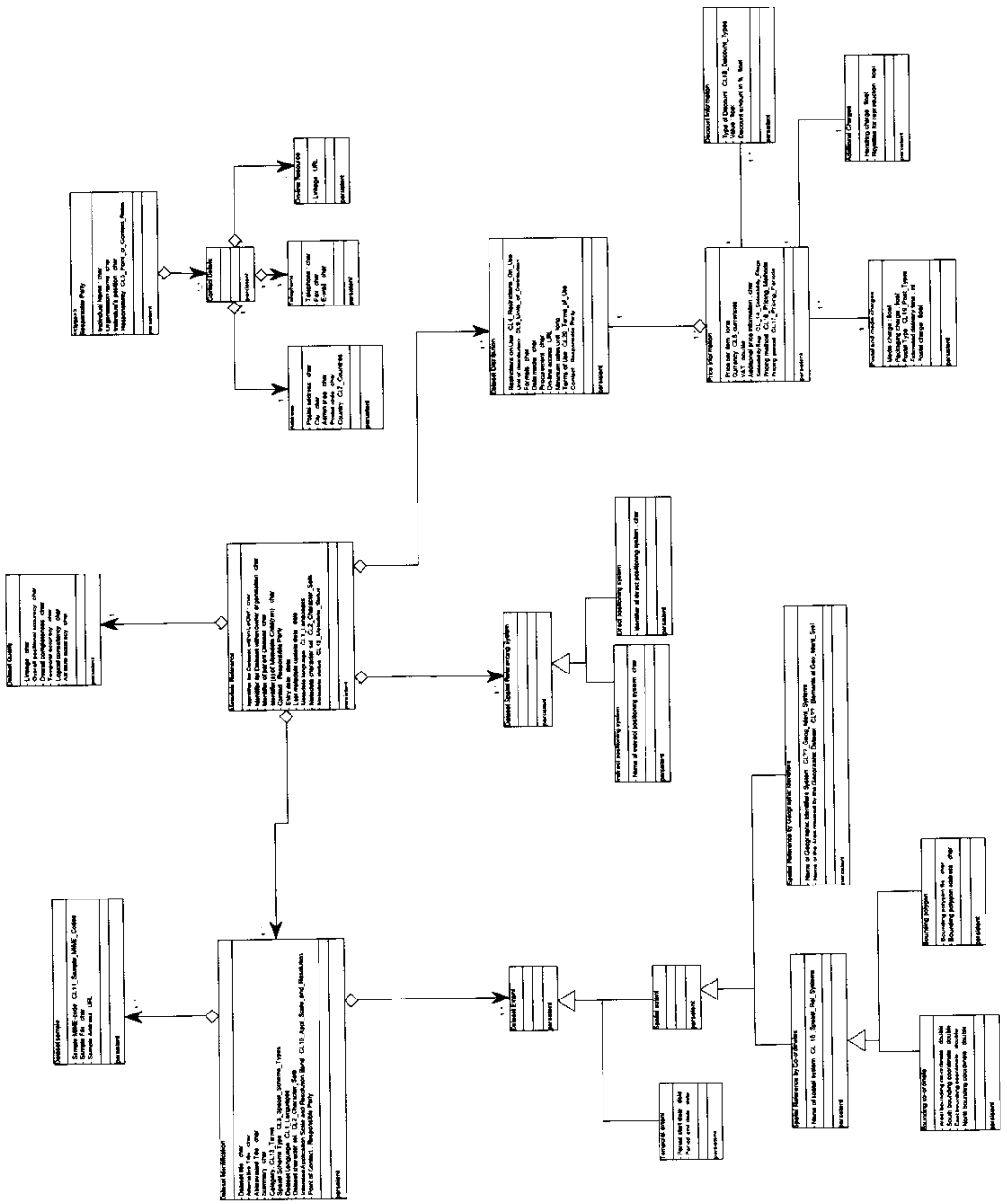
- • Model elements — fundamental modeling concepts and semantics
- • Notation — visual rendering of model elements
- • Guidelines — idioms of usage within the trade

The UML is a well-defined and widely accepted response to the need for visualising and modeling complex systems. It is the visual modeling language of choice for building object-oriented and component-based systems.

The diagrams shown are presented using the Unified Modelling Language (UML) class diagram. The notations used are shown below¹²:



¹² Taken from ISO/TC 211 N831 (30-11-1999)



6.1.2 Metadata Reference

Metadata Reference
- Identifier for Dataset within LaClef : char - Identifier for Dataset within owner organisation : char - Identifier of parent Dataset : char - Identifier(s) of Metadata Child(ren) : char - Contact : Responsible Party - Entry date : date - Last metadata update date : date - Metadata language : CL1_Languages - Metadata character set : CL2_Character_Sets - Metadata status : CL12_Metadata_Status
persistent

6.1.3 Dataset Quality

Dataset Quality
- Lineage : char - Overall positional accuracy : char - Overall completeness : char - Temporal accuracy : char - Logical consistency : char - Attribute accuracy : char
persistent

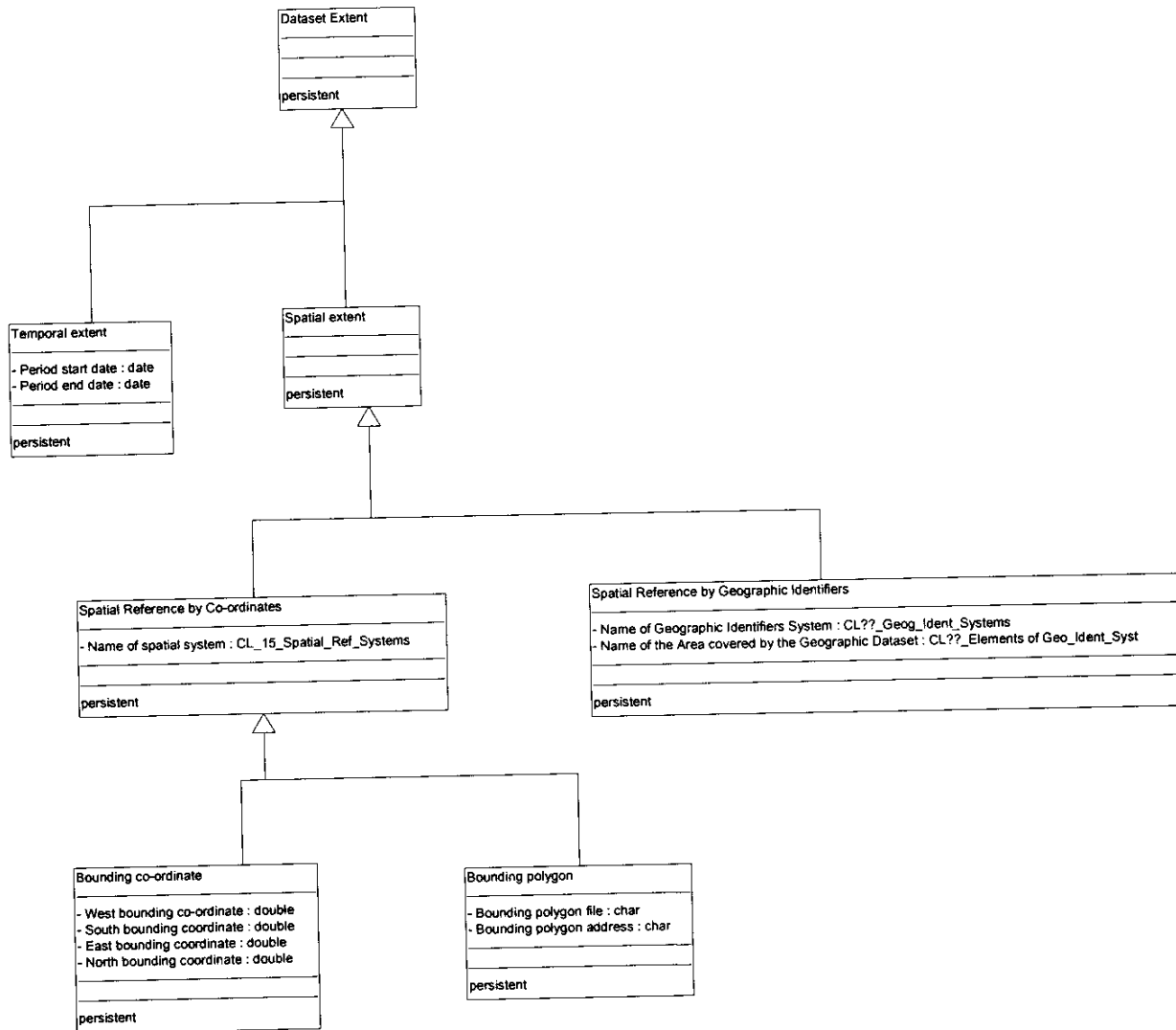
6.1.4 Dataset Identification

Dataset Identification
<ul style="list-style-type: none">- Dataset title : char- Alternative Title : char- Abbreviated Title : char- Summary : char- Category : CL13_Terms- Spatial Schema Type : CL3_Spatial_Schema_Types- Dataset Language : CL1_Languages- Dataset character set : CL2_Character_Sets- Intended Application Scale and Resolution Band : CL10_Appl_Scale_and_Resolution- Point of Contact : Responsible Party
persistent

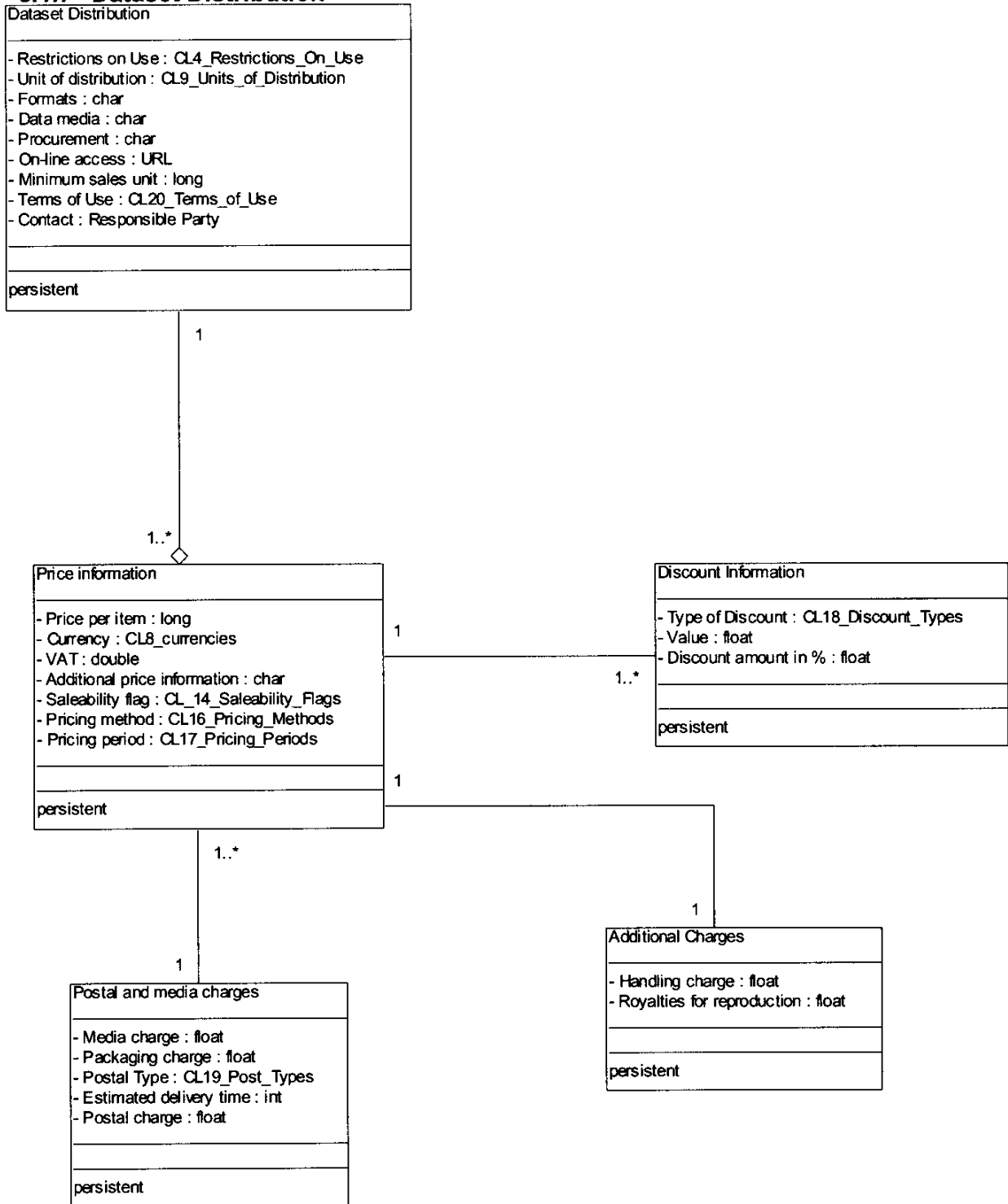
6.1.5 Dataset Sample

Dataset sample
<ul style="list-style-type: none">- Sample MIME-code : CL11_Sample_MIME_Codes- Sample File : char- Sample Address : URL
persistent

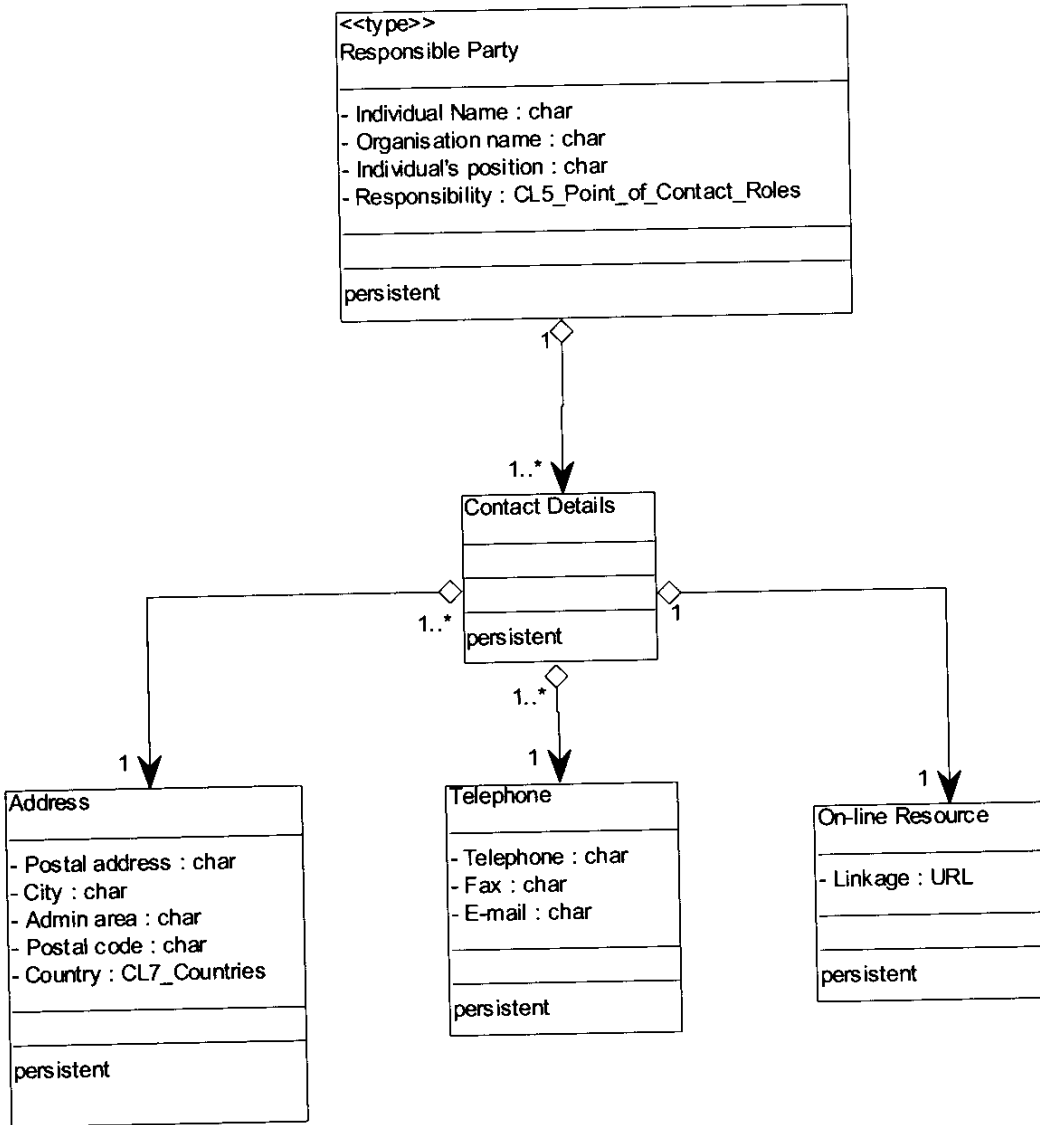
6.1.6 Dataset Extent



6.1.7 Dataset Distribution



6.1.8 Responsible Party



6.1.9 Element Descriptions

The following table provides an overview of the LaClef fundamental metadata elements. The columns contain the following information:

ID	Identifier of each fundamental metadata element
ELEMENT NAME	Name of the fundamental metadata element
OBLIGATION	Obligation – dictates whether an element may or may not be included by the metadata provider.
M	Mandatory. This category or element must be completed for all datasets by all metadata providers. These are considered important to allow potential users to determine the suitability and accessibility of particular datasets.
C	Conditional. Elements must be completed if certain constraints are satisfied. Regarded as important but only applicable in certain circumstances, eg for a particular type of dataset description. See the table of conditions below.
O	Optional. Metadata providers may choose to include or exclude this element. These elements are either inappropriate or irrelevant for certain descriptions.
CARDINALITY	The cardinality reflects the number of instances of an element that may exist or be supplied by the metadata provider. If Cardinality = 1, only one value may be given for that element. If Cardinality = N, any number of values may be given.
DATA TYPE	Data type of the metadata element. May be text, real number, date, integer, image.
DOMAIN	The domain or range of values an element may be allowed to contain. Usually free text or controlled list.
TRANSLATION	Indicates whether the metadata provider may provide the metadata element in more than one language. All providers are encouraged to provide these elements in English and their native language(s). None = no multi-lingual aspect Auto = information taken from a Controlled List. As LaClef will translate all controlled lists to European languages, these do not need to be provided by the metadata provider. Man = metadata provider should include information in additional languages

6.1.10 Entities in Summary

<i>Id</i>	<i>Int</i>	<i>Element Name</i>	<i>Obl</i>	<i>Card (Rel)</i>	<i>Card (Ling)</i>	<i>Data Type</i>	<i>Domain</i>	<i>Trans-lation</i>
1		Metadata Reference	M	1				
1.1	X	Unique Identifier within LaClef	M	1		Text	Free Text	None
1.2		Identifier for Dataset within Owner Organisation	M	1		Text	Free Text	None
1.3		Identifier of Parent dataset	O	1		Text	Free Text	None
1.4	X	Identifier of Metadata Child(ren)	O	N		Text	Free Text	None
1.5		Contact	M	N		Class	Responsible Party	
1.6		Entry Date	M	1		Date	CCYYMMDD	None
1.7		Last Metadata Update Date	M	1		Date	CCYYMMDD	None
1.8		Metadata Language	M	N	N	List	CL1_Language	Auto
1.9		Metadata Character Set	M	N	??	List	CL2_Character_Sets	None
1.10		Metadata Status	M	N	N	List	CL12_Metadata_Statu s	Auto
2		Dataset Identification	M	1				
2.1		Dataset Title	M	1		Text	Free Text	None
2.2		Alternative Title	O	N	??	Text	Free Text	Man
2.3		Abbreviated Title	O	N		Text	Free Text	Man
2.4		Summary	M	1	N	Text	Free Text	Man
2.5		Category	M	N	N	List	CL13_Categories	Auto
2.6		Spatial Schema Type	C	N	N	List	CL3_Spatial_Schema _Types	Auto
2.7		Dataset Language	M	N	N	List	CL1_Languages	Auto
2.8		Dataset Character Set	M	N	??	List	CL2_Character_Sets	None
2.9		Intended Application Scale and Resolution Band	C	N		List	CL10_Appl_Scale_an d_Resolution	None
2.10		Contact	M	N		Class	Responsible Party	
3		Sample	O	N				
3.1		Sample MIME-code Type	C	1		List	CL11_Sample_MIME_ codes	None
3.2		Sample File	C	1		Image or text file	Files in Tiff, Gif, BMP, JPEG or TXT format	None
3.3		Sample Address	C	1		Text	URL	None
4		Extent	M	N				
4.1		Temporal Extent	M	1				
4.1.1		Period Start Date	M	1		Integer	CCYYMMDD	None
4.1.2		Period End Date	M	1		Integer	CCYYMMDD	None
4.2		Spatial Extent	M	1				
4.2.1		Spatial Reference by Coordinates	C2	1				
4.2.1.1		Name of Spatial System	M	1	??	List	CL15_Spatial_Ref_Sy stems	Auto

<i>Id</i>	<i>Int</i>	<i>Element Name</i>	<i>Obl</i>	<i>Card (Rel)</i>	<i>Card (Ling)</i>	<i>Data Type</i>	<i>Domain</i>	<i>Translation</i>
4.2.1.2		Bounding Coordinates	C3	N				
4.2.1.2.1		West Bounding Coordinate	M	1		Real number	Longitude or Grid value	None
4.2.1.2.2		South Bounding Coordinate	M	1		Real number	Latitude or grid value	None
4.2.1.2.3		East Bounding Coordinate	M	1		Real number	Longitude or Grid value	None
4.2.1.2.4		North Bounding Coordinate	M	1		Real number	Latitude or grid value	None
4.2.1.3		Bounding Polygon	C3	1				
4.2.1.3.1		Bounding Polygon File	C4	1		Text	Free Text	None
4.2.1.3.2		Bounding Polygon Address	C4	1		Text	URL	None
4.2.2	X	Spatial Reference by Geographic Identifiers	C2	N				
4.2.2.1	X	Name of the Geographic Identifiers System	M	1		List	CL??_Geog_Ident_Systems	Auto
4.2.2.2	X	Name of the Area Covered by the Geographic Dataset	M	N		List	CL??_Elements_of_Geog_Ident_Sys	Auto
5		Dataset Quality	M	N				
5.1		Lineage	M	1		Text	Free Text	Man
5.2		Overall Positional Accuracy	O	1		Text	Free Text	Man
5.3		Overall Completeness	O	1		Text	Free Text	Man
5.4		Temporal Accuracy	O	1		Text	Free Text	Man
5.5		Logical Consistency	O	1		Text	Free Text	Man
5.6		Attribute Accuracy	O	1		Text	Free Text	Man
6		Dataset Spatial Referencing System	M	1				
6.1		Indirect Positioning System	C5	N				
6.1.1		Name of Indirect Positioning System	M	1	??	Text	Free Text	Man
6.2		Direct Positioning System	C5	N				
6.2.1		Identifier of direct positioning system	M	1	??	Text	Free Text	Man
7		Distribution	O	N				
7.1		Restrictions on Use	M	N	N	List	CI4_Restrictions_On_Use	Auto
7.2		Unit of Distribution	O	1	N	List	CL9_Units_of_Distribution	Auto
7.3		Formats	O	N		Text	Free Text	None
7.4		Data Media	O	N		Text	Free Text	None
7.5		Procurement	O	1	N	Text	Free Text	Man
7.6		Online Access	M	1		Text	URL	None

* This entity with the two elements 6.2.2.1 and 6.2.2.2 will be created automatically for the first version. For the final version of EMF, the data suppliers will be encouraged to provide data for these elements as well.

<i>Id</i>	<i>Int</i>	<i>Element Name</i>	<i>Obl</i>	<i>Card (Rel)</i>	<i>Card (Ling)</i>	<i>Data Type</i>	<i>Domain</i>	<i>Translation</i>
7.7		Minimum sales unit	M	1		Real Number	Real Number	None
7.9		Contact	M/C1	N		Class	Responsible Party	
8		Price Information	O	1				
8.1		Price per item	M	1		Real Number	Real Number	None
8.2		Currency	M	1	N	List	CL8_Currencies	Auto
8.3		VAT	M	1		Real Number	Real number between 0 and 100	None
8.4		Pricing method	M	??	N	List	CL16_Pricing_Methods	Auto
8.5		Pricing period	M	??	N	List	CL17_Pricing_Periods	Auto
8.6		Additional Price Information	O	1	N	Text	Free Text	Man
8.7	X	Saleability flag	M	1	N	List	CL14_Saleability_Flag	Auto
9		Discount information	O	N				
9.1		Type of Discount	O	1	N	List	CL18_Discount_Type	Auto
9.2		Value (e.g.: "5" when min. number of items has been selected)	O	1		Real number	Real number	None
9.2		Discount amount in %	O	1		Real number	Real number	None
10		Additional charges	O	1				
10.1		Handling charge	O	1		Real number	Real Number	None
10.2		Royalties for reproduction	O	1		Real number	Real number	None
11		Postal and Media Charges	O	N				
11.1		Media charge	O	1		Real number	Real number	None
11.2		Packaging charge	O	1		Real number	Real number	None
11.3		Postal type	O	1	N	List	CL19_Post_Type	Auto
11.4		Estimated delivery time	O	1	??	Real number	Real number	None
11.5		Postal charge	O	1		Real number	Real number	None
12		Responsible Party						
12.1		Individual Name	O	1		Text	Free Text	None
12.2		Organisation Name	M	1	N	Text	Free Text	Man
12.3		Individual's Position	O	1		Text	Free Text	Man
12.4		Responsibility	M	N	N	List	CL5_Point_of_Contact_Roles	Auto
13		Contact Details	M	N				
14		Address	O	1				
14.1		Postal Address	M	1		Text	Free Text	None
14.2		City	M	1		Text	Free Text	None

<i>Id</i>	<i>Int</i>	<i>Element Name</i>	<i>Obl</i>	<i>Card (Rel)</i>	<i>Card (Ling)</i>	<i>Data Type</i>	<i>Domain</i>	<i>Translation</i>
14.3		Administrative Area	O	1		Text	Free Text	None
14.4		Postal Code	O	1		Text	Free Text	None
14.5		Country	M	1	N	List	CL7_Countries	Auto
15		Telephone	O	1				
15.1		Telephone	O	N		Text	Free Text	None
15.2		Fax	O	N		Text	Free Text	None
15.3		e-mail	O	N		Text	Free Text	None
16		Online Resource	O	1				
16.1		URL	M	1		Text	URL	None

Annexe B Summary of User Testing Results

Tester: Kirsi Mäkinen, GI specialist
Organisation: National Land Survey of Finland

General comments

The quality of the metadata makes the testing sometimes a bit confusing. For example when trying to find out which datasets are available from Finland I get selection of dataset from all Nordic countries. Also for most datasets there are defined any terms which makes it difficult to test that kind of search. If the metadata is correct selecting the datasets would be quite easy and straightforward business.

I had some technical problems when testing the service from my home computer via ISDN. I didn't get the map no matter how long I waited. Also all other response from the server was a bit slow. From office everything worked smoothly. It seems the service can be used only with straight Internet connection.

The user interface

The graphical out look

The visual look of the graphical user interface is very ugly. The thick boundary lines between every sub-window should be removed or changed to hairlines. The Terms and Publisher selection windows could be higher so that more list items could be seen on the selection list.

The quality and visual look of the maps are very poor. The map projection used makes Europe look very odd. In addition to administrative boundaries and road network the map must definitely have more information, coastal line is a must in countries like Sweden and Finland. Our country is almost unrecognisable. Also the colour for areas not included the service should be something else than blue. Colour Blue is associated to water and my first thought when I saw the map was "how the Gulf of Bothnia can be so narrow and why on earth Tallinn is in the water!"

And when you zoom in you won't get any other details than road network. More place names and other details are needed in order to orientate the map. The present maps are almost useless.

The icon for resetting map is not intuitive. I thought it meant start searching.

Animated globe and compass are just hype not necessarily needed in the user interface. Either one should be enough to show something is happening.

Making queries

Despite the visual look wasn't the one I like making queries is made rather easy. You could easily make queries and there were not any major problems with that. The path to information was easily understood. It is a good thing that the main window shows all aspects of the data query.

In the country list a English name in addition to the name in country's official language is a must. You can't expect that people recognise i.e. which country is Suomi or Magyarország.

More space should be reserved in the database for the publisher name. The are many names that are cut and the user may have to guess. Even though it

is usually easy to guess the complete name it raises the image of the service if all the names are complete.

It's good that the summary of the query is shown on the query window but it would be more user friendly if the summary is placed below the submit query - button.

I could not find any difference between New query and Change query. I suppose that when selecting Change query I would have got a query window with the old selection in which I could easily have changed some aspects of the data query and then try again. Anyway change query resets the query window and no old input is left.

If the graphical search has been selected, Reset -button should reset the map window and menus on the right sub-window, not return to the search by name function as it does now.

A possibility to browse dataset names would be nice.

It also might be convenient to be able to select the data by data type: raster/vector/image.

Ordering data

I could test only the ordering of SABE data. The corine data order didn't work at all. Ordering was easy to do.

In order form a place for organisation name was missing. I think most buyers are institutional not private persons!

The price of the data should be clearly seen before ordering. Now it can't be found anywhere. In practice it would also be important to be able to order also a part of the dataset for example define a coordinates of rectangle from which the data is needed. For example our customers seldom order data from whole Finland.

Data content

The output

The brief data description has a nice clear layout. The same can't be said of the more detailed data description. A bigger window should be reserved for the textual description and I prefer to have on the left side of the application window. The data should be grouped with some subtitles i.e. Contact information, Accuracy etc. to make it easy to glance through the data description.

I did not understand what theme was displayed on the map window of the detailed dataset description. It is maybe reserved for showing the areal coverage the dataset but there isn't any information about that.

When I look at the description of SABE it has link for other datasets from France. Though the data covers many more countries.

Testers: **Mikko Hynninen**
M.Sc. (Eng.) GIS-Specialist
GIS and Remote Sensing Unit
Computer Service Division
Finnish Environment Institute

Yki Laine
M.Sc. (Eng.) GIS-Specialist
GIS and Remote Sensing Unit
Computer Service Division
Finnish Environment Institute

Is it simple?

- *interface is easy to use*

Is it logical?

- *All country names should be in english if language selection is english. Region names can be left in their original language.*
- *There are some inconsistencies in country names. E.g. Finland is called Suomi in catalogue search -country list and Finland in [geodetic services\map projection information] -country list.*
- *Functions New query and Change query should be grouped under the link Catalogue search because they belong to it.*
- *When selecting the search area with coordinate boundaries user has to manually convert coordinates to lat, long. This procedure should be included in the Apply extent-procedure.*
- *Catalogue search did not return any data when searching data from Finland with coordinate boundaries (coordinate system used: ED50 lat, long, boundaries: 21-31 eastern, 62-70 northern)*

Is it user friendly?

- *Correct phrases should be used in the user interface (e.g. Type on 'Ctrl' to make multiple selections should be hold down 'Ctrl' to make ...)*
- *There are no links to the previous window or view in the full metadata -window.*
- *The prototype does not contain any help-document. There is no description about the metadata-fields.*
- *Those options, which are not available because of earlier choices, should be invisible. For example if a user has selected Suomi (Finland) in Catalogue search, only those publishers who have data from Finland should be in the Publisher list*

Do you understand the path to information?

- *Yes, it is quite straightforward.*

Do you understand the information received?

- *Yes.*

Where the information is confusing, what development would you recommend?

- *development: help-document should be added, some coordinate systems are missing (e.g. Finland Gauss-Krüger Grid)*
- *When the prototype is searching data, transfer interrupt error is shown in the frame. Error does not seem to stop the execution of the search procedure.*

What graphical improvements would you recommend?

- *Symbology and appearance on the map where area of interest is defined should be clearer. Now text and line features overlap and are difficult to distinguish.*
- *Map window should include scale bar.*
- *North arrow should be still.*
- *User interface might be clearer with invisible frame borders.*

What information is unnecessary?

- -

What information is missing and should be included?

- *Help-documents*

General comments and areas for development?

- *Without help it is not clear that map windows pan function works through the north arrows points.*
- *Zoom out function should act so that the map is zoomed out directly after a user has clicked on the map window. The user shouldn't need to draw a box.*
- *In Catalogue search an error message is received after a query every time when a region has been selected for Finland or Sweden.*

Tester: Antti Vertanen,
Organisation: Ministry of agriculture and forestry Finland, senior advisor
in GI policy

The main requirement from the user testing relates to the interface:

Is it simple?

Yes, in a positive manner. Still the question remains to whom the service is mainly addressed to? For schools etc. the terminology might still be too professional.

Is it logical?

Yes, however I propose to add coverage information to other general dataset information in the first query result sheet

Is it user friendly?

Generally yes.
The small windows for the Terms, Publisher and Scale are causing some trouble to see and compare the most relevant categories.
The surfing between the new and old windows/selections needs some fine tuning.

Do you understand the path to information?

Yes, I still wonder if the publisher selection is necessary in the first window?

Do you understand the information received?

Mostly, however I work within the GI business. Clearly there must be an online help to translate all the terms like 'Thematic GI data' or 'Overall completeness' etc.

Where the information is confusing, what development would you recommend?

The information of the coverage (full/partial) of the datasets is important in the first query result window (or graphical view of the real coverages later).
The button 'show higher level information' did not open to me completely.

What graphical improvements would you recommend?

Information in table form (e.g. excel) of optional/selected datasets for comparison and calculations might be a necessary tool for preparatory tasks, meetings etc.

Some kind of active area division (country, province, commune) with place names might be useful when navigating and making selections in graphical mode (now only box).

What information is unnecessary?

The publisher selection in the first window.

What information is missing and should be included?

There is a need for a dedicated category for online data services (XML ...)?

General comments and areas for development?

I did not get a clear picture what I bought with the online service - more feedback/documents needed !

Annexe C Project Plan

ID	Task Name	Duration	Start	1st Quarter			2nd Quarter			3rd Quarter				
				Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
1	WP1 Project Management	521 days	Fri 01/01/99	[Gantt bar spanning from Jan to Nov]										
2	1.1 co-ordination	521 days	Fri 01/01/99	[Gantt bar spanning from Jan to Nov]										
3	1.2 quality plan	44 days	Fri 30/04/99	[Gantt bar from Apr to May]										
4	1.3 consortium agreement	129 days	Fri 01/01/99	[Gantt bar from Jan to Mar]										
5	1.4 results dissemination	390 days	Thu 25/03/99	[Gantt bar from Mar to Nov]										
6	WP2 Methodology acquisition	293 days	Fri 01/01/99	[Gantt bar from Jan to Oct]										
7	2.2 (meta)date availability	144 days	Thu 01/07/99	[Gantt bar from Jul to Oct]										
11	2.3 selection and specification	253 days	Fri 01/01/99	[Gantt bar from Jan to Oct]										
16	2.4 acquisition and formatting	144 days	Thu 01/07/99	[Gantt bar from Jul to Oct]										
19	2.5 metadata update and compl.	40 days	Wed 22/12/99	[Gantt bar from Dec to Jan]										
22	WP3 EMF services	158 days	Thu 25/03/99	[Gantt bar from Mar to Oct]										
23	3.1 specifications	90 days	Thu 25/03/99	[Gantt bar from Mar to May]										
24	3.2 data acquisition	28 days	Thu 25/07/99	[Gantt bar from Jul to Aug]										
27	3.3 services development	40 days	Thu 25/08/99	[Gantt bar from Aug to Sep]										
28	3.3.1 meta coord transfo	40 days	Thu 25/08/99	[Gantt bar from Aug to Sep]										
29	3.3.2 other services	40 days	Thu 25/08/99	[Gantt bar from Aug to Sep]										
30	WP4 EuroMagFINDER development	435 days	Mon 01/03/99	[Gantt bar from Mar to Nov]										
31	4.1 specification	132 days	Mon 01/03/99	[Gantt bar from Mar to Jun]										
32	web interface spec	1 day	Wed 01/03/99	[Gantt bar on Mar 1st]										
33	4.2 development beta ver.s.	65 days	Thu 02/09/99	[Gantt bar from Sep to Oct]										
34	4.3 identifier, version 1	37 days	Thu 02/12/99	[Gantt bar from Dec to Jan]										
35	4.4 version 2	25 days	Mon 25/09/00	[Gantt bar from Sep to Oct]										
36	WP5 Workshops	350 days	Mon 01/03/99	[Gantt bar from Mar to Nov]										
37	5.1 meta & services user requirements	65 days	Wed 31/03/99	[Gantt bar from Mar to Apr]										
38	5.2 e-comm user requirements	88 days	Mon 01/03/99	[Gantt bar from Mar to May]										
39	5.3 user feedback	41 days	Fri 05/05/99	[Gantt bar from May to Jun]										
42	WP6 Commercial & legal issues	400 days	Tue 26/01/99	[Gantt bar from Jan to Oct]										
43	6.1 payment mechanisms	60 days	Fri 09/07/99	[Gantt bar from Jul to Aug]										
44	6.2 IRI investigation	480 days	Tue 26/01/99	[Gantt bar from Jan to Oct]										
45	6.3 data security	480 days	Tue 26/01/99	[Gantt bar from Jan to Oct]										
46	6.4 licensing and pricing policy	480 days	Tue 26/01/99	[Gantt bar from Jan to Oct]										
47	WP7 On-line order & delivery	487 days	Mon 01/02/99	[Gantt bar from Jan to Oct]										
48	7.1 data content	188 days	Thu 01/07/99	[Gantt bar from Jul to Oct]										
51	7.2 e-comm specifications	125 days	Mon 01/02/99	[Gantt bar from Jan to Apr]										
52	e-commerce	1 day	Fri 30/07/99	[Gantt bar on Jul 30th]										
53	7.3 Beta version	44 days	Fri 01/10/99	[Gantt bar from Oct to Nov]										
57	7.4 Version 1	112 days	Wed 01/12/99	[Gantt bar from Dec to Feb]										
62	7.5 Version 2	77 days	Mon 31/07/99	[Gantt bar from Jul to Sep]										