ENVISION, a Consensus Finding Process for the European Next Generation Interoperable Spatial Information Network

Günther Landgraf, Ivan Petiteville[1]

[1] European Space Agency (ESA)
Directorate of Earth and Environment Monitoring from Space
Earth Observation Applications Department, ESRIN

Via Galileo Galilei, I-00044 Frascati (Roma), Italy

guenther.landgraf@esa.int
ivan.petiteville@esa.int

Abstract. For many geospatial 'business' applications – in particular in the environmental sector - combined access to earth observation as well as geographic data is essential, implying an interoperable infrastructure between the GIS and the earth observation data providers. World-wide and in Europe many initiatives have been started to define a basis for an interoperable infrastructure, without that any of them has lead - or seems to lead in short-term - to a usable implementation that is generally accepted. Among many different reasons the probably most outstanding one is insufficient attention to involve real information market share-holders in a digital economy, in particular the application community. The present paper describes the current state of ESA analysis on the subject, in preparation for consensus finding process on a unique European next generation infrastructure.

1 Introduction and Problem Illustration

Both the geographic as well as the spatial data community have independently identified the need to provide better and homogeneous access to data and services of their widely distributed provider network. Each of them has identified 'resource identification' ("yellow pages"), 'metadata access' ("catalogue"), on-line data access and on-line ordering as the essential services to be provided in an interoperable way, as a basis for better data usage and to better exploit the large market growth opportunities in this sector. This implies agreement on both services as well as underlying data formats.

Also value-adding applications builders - in particular in the manyfold environmental domains - have identified the need to combine data from heterogeneous sources, including combined access to geographic as well as earth observation data, implying a strong preference for an integrated infrastructure between the GIS and the earth observation data providers. However, some application areas are more advanced than others, with an increasing risk to loose the opportunity to agree on a horizontal multi-domain infrastructure providing interoperable access to GIS and Earth Observation data and services. Some applications might in fact generate enough 'value' to completely fund their own end-to-end infrastructure, e.g. disaster management or applications providing information on important environmental issues. Without preventive measures and proactive deployment of a horizontal infrastructure there is a concrete risk, that 'rich' applications develop and operate a dedicated 'closed' infrastructure, not accessible to anyone else. This would be a considerable impediment to the growth of 'poorer' applications or application communities where it is more difficult to establish a common focus, which on the other hand are often more suitable for business development of small and medium enterprises. These application areas where deployment and operation requires a framework of an existing and reusable open infrastructure to maintain the cost below the revenue (or benefit translated into 'commercial value'), risk to not arrive to operational exploitation and give their full benefit to the European citizen. This lack of full exploitation of the opportunities that the geospatial sector offers would result in a reduction of employment opportunities in high qualification

areas and reduced number of services offered to the European citizen by the private and public sector. In particular for the space data sector, where for the building and operation of satellite infrastructure still a major public investment is required, it has become a primary necessity to maximise the use of the acquired information and demonstrate benefit of investment to the tax payer.

Many initiatives to provide common access to geospatial data and metadata have been undertaken in the past or are still ongoing. The small snapshot in chapter 3 - far from being complete - illustrates the fact that currently application developers and data providers are confronted with a 'jungle' of different infrastructures, protocols and standardisation activities in the wide sense, not allowing them to understand in which direction they should invest.

ESA has experienced the problem both in its role as the largest European data provider on remote sensing data, as well as in its application developments with industry. As a data provider it requires an continued and not negligible effort both in development as well as in operation to ensure on-line access via a multitude of protocols. In particular for remote sensing data providers, where the data are of relevance for multiple applications, it is unrealistic to assume that data providers can support in long term a variety of application-independent access protocols (IMS, CIP, OpenGIS), in addition specialised infrastructures and protocols developed by application communities (disaster management, water management, atmosphere, agriculture, fishery, ...) and even specific interfaces required by the single applications (which is the typical case today). Practical life shows that ESA did not have enough resources to maintain support for IMS when providing access via CIP, or that SpotImage has lost connectivity with INFEO when phasing in their new catalogue system.

For what regards applications, one of the most advanced applications currently using ERS satellite radar data is the RAMSES oil-slick detection application [39]. When the need to feed this application with data from additional data sources - concretely RADARSAT - has emerged, the cost to add the interface for this additional data provider was found significant, and the time for deployment about 6 months.

The above considerations clearly motivate the potential benefit of a consensus of data providers and application community on the basis for an interoperable infrastructure. For the Agency's purpose this consensus finding process has been given the working title 'ENVISION' (Environmental Interoperable Spatial Information Open Network) and will be referred to with this name from now on.

2 Identified Active Institutional Players

The following table gives an overview of the main active institutional players that the Agency has identified for the time being, with an indication of their primary focus (remote-sensing earth observation, geographical information or applications).

	Long name	EO	GI	Appl.
ISO	International Standardisation Organisation	X	X	
CEN	European Committee for Standardisation		X	
EC	European Commission	X	X	X
CEOS	Committee on Earth Observation Satellites	X		
GSDI	Global Spatial Data Infrastructure		X	
OpenGIS		x	X	
IOC	Intergovernmental Oceanographic Commission			X
EEA	European Environmental Agency			X
EuroGI	European Umbrella Organisation for Geographic Information		X	
Eumetsat		X		x
ESA	European Space Agency	X	x	x
	European National Space Agencies	X		
	European GI Bodies		X	
NASA	(US) National Aeronautics and Space Administration	X		
FGDC	(US) Federal Geographic Data Committee		X	
NIMA	(US) National Imagery and Mapping Agency		X	X
USGS	U.S. Geological Survey			X
CCRS	Canadian Centre for Remote Sensing	X	X	X

ISO: ISO working group TC211 is currently working on a new series of standards for access to geospatial data [1]. The most important one for geographical metadata was recently issued as draft international standard ISO 19115.

CEN working group TC287 has issued several standards on geographic information [2]. The most significant one for interoperability is ENV 12657:1998, the standard on metadata.

EC: Most of the European activities have been funded by EC programmes, out of the fourth or fifth framework programme. The activities that have been particularly considered in the analysis are:

- INFEO (an infrastructure implementation of the CEOS CIP protocol) [3] ,
- the SeaSearch network for the Oceanographic application community, including in particular the EDMED database for resource discovery [4]
- ESMI (European Spatial Metadata Infrastructure) [5], implementing a GI access infrastructure based initially on CEN TC286 and later also on the CORBA profile of the OpenGIS catalogue specification.

CEOS [6]: Within this body remote sensing data providers (NASA, ESA, NASDA, CNES, DLR, CCRS, BNSC, etc.) agree on issues regarding access to their data. The most significant CEOS results are:

- IDN, the International Directory Network, providing EO data resource discovery [7]
- IMS, the first interoperable protocol, still used in the NASA EOSDIS infrastructure [8]
- CIP [9], the latest endorsed interoperable protocol, implemented as Z39.50 profile and used in the INFEO[3] infrastructure; CIP is an initiative initiated by ESA (which finalised version 2.2) and which was continued by EC/JRC, leading the finalisation of version 2.4

GSDI: Within this body world-wide geographic data providers aim to harmonise access to their data, similar to the remote-sensing data providers in CEOS. Practically all geographic information providers listed above are participating in GSDI [10].

The **Open GIS Consortium** (OGC) [11] is an industrially organised body with institutional, industrial and academic members, both from the earth observation as well as from the geographic information community. OGC vision is 'A world in which everyone benefits from geographic information and services made available across any network, application, or platform'. OGC core mission is to deliver spatial interface specifications that are openly available for global use. Open interface specifications shall enable content providers, application developers and integrators to focus on delivering more capable products and services to consumers in less time, at less cost, and with more flexibility.

IOC [12] has launched the IODE (Int. Oceanographic Data and information Exchange) programme [13] in 1961 to "enhance marine research, exploration, and development by facilitating the exchange of oceanographic data and information between participating Member States". Most relevant activities are

- the existing MEDI (Marine Environmental Data Information Referral System) software
- the Marine XML specification which is currently ongoing
- the long-term Global Ocean Observing System (GOOS) [14], the end-to-end data management system (being defined), including
- the Integrated Global Ocean Services System (IGOSS) [15], which aims to implement the international system for the collection and exchange of ocean data (such as temperature and salinity) and the preparation and dissemination of oceanic products and services

EEA[16]**:** The European Environment Agency aims to support sustainable development and to help achieve significant and measurable improvement in Europe's environment through the provision of timely, targeted, relevant and reliable information to policy making agents and the public. The aim of the European Environment Agency is to establish a seamless environmental information system EIONET [17].

Currently the ETC/CDS (European Topic Centre/Catalogue of Data Sources) provides an on-line resource discovery service based on the GEMET environmental thesaurus [18].

EUMETSAT currently operates the Meteosat satellites and is the operational provider of meteorological data, relevant for many geospatial applications. EUMETSAT is developing multi-mission user services in the UMARF project, which will enable users to access to data from past and future satellites (e.g. EPS) in a uniform way.

ESA: The European Space Agency [19] is one of the largest European EO data providers and handles data from a numerous international Earth Observation Satellites (ERS, Landsat, JERS, NOAA, SeaWiFS, IRS, Nimbus, MOS and ENVISAT in near future) [20]. Several thousand users world-wide have online access via Internet to related meta-data, small size data and derived information products. To support access also via a global standard infrastructure, ESA has taken over operations of INFEO [3] from JRC as of January 2001.

The Earth Observation Applications Department at ESRIN is taking actions for promoting a market for EO applications aiming at self-sustainability. It is considering specific investments for the development of the Down Stream Industry involved in the utilisation of EO data and their combination with GIS data to create new products and derived information. Target applications are currently Fire, Flood, Coastal Pollution with specific attention to oil spills, Volcano, Landslide and Earthquakes, Ozone.

GIS is starting to play a more and more important role in these application projects and typically application products produced are geographic information. The second pillar of ESA interest in the geographic domain is in navigation with the newly approved Galileo programme [21].

European National Space Agencies have their independent remote sensing missions and co-ordinate at CEOS level. The currently most important players on the subject of interoperable infrastructures are:

- **CNES**[22], the French Space Agency, which is the largest national space agency in Europe. CNES has and is developing the SPOT satellite series, for which data are distributed by SpotImage. CNES will distribute data themselves for some future missions, e.g. POLDER.
- **DLR** [23], the German Space Agency and in particular its remote sensing centre (DFD) has an extensive network of collaboration with external partners. It is providing a large set of Earth Observation data from their international collaboration, including the NASA SRTM mission, Russian satellites, ESA ERS and the Indian IRS MOS satellite [24]. The DFD is also hosting the commercial distribution services of Euromap.
- **NLR** [25], the Dutch Space Agency has deployed a national infrastructure called NEONET [26].

EUROGI [27], the European Umbrella Organisation for Geographic Information, was set up in November 1993, as a result of a study commissioned by DGXIII-E of the European Commission to develop a unified European approach to the use of geographic technologies. EUROGI mission is to maximise the effective use of geographic information for the benefit of the citizen, good governance and commerce in Europe and to represent the views of the geographic information community. EUROGI objectives are:

- Support the definition and implementation of a European geographic information (GI) policy and facilitate the development of the European Geographic Information Infrastructure (EGII).
- Raise awareness of the value of GI and its associated technologies. Improve the sharing of knowledge between members themselves and between EUROGI and the European Commission.
- Encourage greater use of GI in Europe, through improved availability of and access to GI, removal of legal and economic constraints to use and promotion of the use of standards.
- Represent the European view in the development of the Global Spatial Data Infrastructure (GSDI) and be the European regional contact for GSDI.

- Work towards the development in all European countries of a strong national GI organisation with particular emphasis on organisations within the Central and Eastern European countries so that the enlargement of the European Union is enabled.

EUROGI members currently include a pan-European organisation EuroGeographics/CERCO and the following national bodies: AESIG (E), AFIGÉO (F), AGI (UK), CC Belgium (B), CNIG (P), GeoForum Denmark (DK), DDGI (D), GeoForum (N), GISPOL (PL), GTIM-SIG (L), HUNAGI (H), IRLOGI (IRL), NDC (Greece), ProGIS (SF), Ravi (NL), SOGI (CH), ULI (S)

European Geographic Information Bodies are mostly represented in EUROGI and are therefore not listed separately. Technically one of the most active one is **MEGRIN** [28], created by CERCO (a EUROGI member) and standing for "Multipurpose European Ground Related Information Network", which is a European network of geographical referenced information for use in many diverse applications. MEGRIN is operating the GDDD (Geographical Data Description Directory), currently upgraded within the LaClef project.

NASA has extended its historic IMS-based infrastructure, which is globally installed at various remote sensing providers, as the first version of the future distributed infrastructure (ESDIS)[8]. Currently NASA is linking this IMS infrastructure with the European INFEO infrastructure. NASA has taken the lead in the CEOS contribution to the Open GIS Catalogue Interface Implementation Specification.

FGDC has defined its own Z39.50 profile - GEO[29] - for access to geographic data, and implemented a distributed infrastructure called 'Clearinghouse' [30]. GEO and CIP profiles are harmonised to a large degree, but not fully [31]. FGDC has been extremely active in the generation of the ISO 19115 standard and has also contributed to the Open GIS Catalogue Interface Implementation Specification.

NIMA has been actively working with NASA and FGDC to contribute to the Open GIS Catalogue Interface Implementation Specification.

USGS is one of the world-wide largest holders of earth observation and mapping data. It is participating in GSDI and is a principal member of OpenGIS.

CCRS has very successfully deployed a distributed infrastructure named CEONET [32], which via 3 different protocols - 'Simple Search', IMS and GEO - making a high number of geospatial data - both from the earth observation as well as from the geographic domain - available via a comfortable interface.

3 Current State of the Art

Numerous activities have been undertaken to arrive to an infrastructure at various levels. The following table identifies the activities for which ESA has performed at least a minimum analysis so far. It has to be acknowledged that in particular in the GIS area the number of different infrastructure is so high, that a pre-selection had to be done. The main criteria for selection were:

- involvement of real operational bodies
- provision of an operational service
- consideration of state-of-the-art interoperability standards[1]

However, many projects that are not included have acquired a lot of valid experiences (e.g. GISEDI, Geoserve, Hypergeo, DISGIS, etc.), which will be important to be fed into the more detailed step of the consensus finding process.

Project	Protocol/ Standard	Org.	Domain	Service Coverage[2]			
				Dir	Inv	DA	O
IDN	DIF	CEOS	EO	X			
EDG	IMS	NASA	EO	X	X	x	x
CEONET	simple search, IMS, GEO	CCRS	EO, GI	X	X	x	X
Clearinghouse	GEO	FGDC	GI		X		
OGDI	GLTP	Global Geomatics	GI		X		X
GDDD		MEGRIN	GI	X			
LaClef		MEGRIN	GI	X	X	x	X
ESMI	Z39.50/CEN, OGC/Corba	EC	GI		X		
ETC/CDS	GELOS	EEA	Appl.	X			
EDMED		EC,BODC	Appl.	X			
GEIXS		EC	Appl.	X			
NEONET		NLR	EO	X	X	X	
INFEO	CIP	CEOS,EC	EO, GI	X	X		x
WMT	OGC	OGC	GI	X	X	X	
	Dublin Core			X			
	ISO 19115	ISO	EO, GI				
	OGC/Catalogue	OGC	EO, GI	X	X		X

[1] in particular, the Agency currently works on the - to be confirmed assumption - that a future European consensus finding process will have to tae Open GIS results into account

[2] DIR: directory/resource discovery (service returns descriptions of available data collections)
INV: inventory (service returns descriptions of available data)
DA: data access (service allows to access data directly)
O: ordering

- The **CEOS International Directory Network (IDN)** [7] is an international effort to assist researchers in locating information on available data sets. It provides free, on-line access to information on world-wide scientific data including Earth sciences (geoscience, hydrospheric, biospheric, satellite remote sensing, atmospheric sciences), space physics, solar physics, planetary science and astronomy/astrophysics. The CEOS IDN describes data held by university departments, government agencies, and other organisations. There are co-ordinating nodes of the CEOS IDN for the Asian (NASDA), American (NASA), European (ESA), and African (UNEP) continents, and another co-ordinating node for those countries participating in Antarctic research (JCADM/GCMD). In addition there are several world-wide co-operating nodes. Data exchange is based on the CEOS Directory Interchange Format (DIF[33]) which includes keyword classification on the four-level GCMD thesaurus.
- The **EDG** (Earth Observing System Data Gateway) - formerly EOSDIS - is NASA's web-based search-and-order tool, provides a way for users to search for Earth Science data from multiple participating archives [8]. It uses the IMS protocol for distributed search.
- **CEONet** [32] is a key access component of the Canadian Geospatial Data Infrastructure (CGDI) initiative.
  It was designed and built to enable geospatial resource discovery, evaluation and access. CEONet uses NASA IMS and FGDC GEO protocol to enrich its domain by access to US data holdings, while using its native 'simple search' protocol (based on simple html-encapsulated search and result syntax) to access to a high number of additional sites.
- The Geospatial Data **Clearinghouse**[30] - developed by the FGDC is a collection of over 250 (US) spatial data servers, that have digital geographic data primarily for use in Geographic Information Systems (GIS), image processing systems, and other modelling software. These data collections can be searched through a single interface based on their descriptions, or "metadata." It is based the Geospatial Metadata Application Profile (GEO Version 2) for Z39.50[29]. The current Version 2.2 is the current document in process and includes specific mention of interoperability with the Government Information Locator Service (GILS) and the Catalogue Interoperability Profile (CIP) [31].
- The **'Open Geospatial Datastore Interface'** (OGDI), based on **the Geographic Library Transfer Protocol** (GLTP) was initially developed by Global Geomatics [38] for the Canadian Ministry of Defence and was made available as open source code. Later Global Geomatics has modified this standard and is using a de facto proprietary standard with the same denomination within their own products
- **GDDD** [28] 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 [2]. Since 1996 the main part of the information contained in the GDDD has been freely accessible on the Internet. The GDDD is becoming increasingly easy to update through semi-automated procedures making the content increasingly valuable. 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. MEGRIN also co-ordinates the LaClef project which aims to develop a new version of the GDDD, including, for example, a multilingual service, metadata of a higher resolution than at present, and a prototype electronic commerce service.
- The **ESMI** (European Spatial Metadata Infrastructure) [5] project aimed to create a single access gate to linked internet services that describe geographic information, thereby giving users access to a wealth of information about geographic data. The development was based initially on the Clearinghouse software of the FGDC, but the GEO profile has been adapted to CEN ENV 12657. In the latest project phase also OGC Corba Catalogue interface [11] was used as a protocol for the distributed search.

- For what regards the scope of ENVISION, the mission (limited to the scope of ENVISION) of the **ETC/CDS** (European Topic Centre/Catalogue of Data Sources) [18] is to implement a locator system for environmental metainformation from all Network members in Europe - the Catalogue of Data Sources (CDS). ETC/CDS is based on the GELOS (Global Environmental Locator Service) profile of Z39.50, which is an interpretation of GILS (Global Information Locator Service) and using a multilingual environmental thesaurus developed for the EEA (GEMET - the GEneral Multilingual Environmental Thesaurus).
- The **European Directory of Marine Environmental Datasets** (EDMED) [4] was initiated in 1991 by BODC (UK) within the EC-MAST framework and has established itself as a de-facto European standard for indexing and searching datasets relating to the marine environment. It covers a wide range of disciplines and is a high level inventory, describing both Datasets and Data Holding Centres. At present, EDMED already describes more than 2300 Datasets from over 500 Data Holding Centres across Europe. All national directories are assembled into a single centralised system managed by the BODC and are made available via the Sea-Search website (see above). Activities are now underway by the 16 European Sea-Search partners, from 14 coastal states, to update their national EDMED entries and to develop and install an innovative infrastructure for updating the EDMED database by means of the Internet. An interlink with the Information Service for Earth Observation Data (INFEO) of the Centre of Earth Observation (CEO) is being studied. Furthermore, contacts are underway with the IOC-IODE [13] to harmonise formats in order to establish a uniform global metadata format that will be the basis for further EDMED developments.
- The **GEIXS** (European Geological Information eXchange System) project has implemented "The European Geological Data Catalogue" [34], giving a dataset description through the geographic coverage of the data, key words from lexicons and free text.
- **Neonet** [26]is an initiative of the National Aerospace Laboratory NLR [25]and was carried out as a project on behalf of the Netherlands Remote Sensing Board (BCRS) and the Space Research Organization Netherlands (SRON). It is a Dutch information network where users of remote sensing data can find and present information data related to remote sensing. Its search engine and information infrastructure is specifically intended for remote sensing information and implements a distributed network of 'topical nodes', which synchronise their information content off-line, with the capability to perform only a selective propagation between nodes. Neonet uses a native email-based protocol for this purpose.
- The **INFEO** [3] system offers access to Earth observation and geographic information & services, and data catalogues around the world. INFEO has been developed in the framework of the CEO programme within the Joint Research Centre of the European Commission. INFEO includes a middleware implementation for a distributed search infrastructure as defined by the CEOS CIP protocol [9], but includes backends to access also GEO [29] and IMS-based [8] servers. INFEO provides common access to both remote sensing as well as geographic datasets on a European and global scale. This characteristic was considered essential in support of geospatial applications development, which is a primary focus of the EO Applications Department and was one of the main reasons for the European Space Agency to take over operations of INFEO in 2001.
- The **Web Mapping Testbed** (WMT) [35] was the first of OGC's planned Interoperability Initiatives, which involve sponsors and participants. OpenGIS Specifications resulting from the Web Mapping Testbed build on OGC's OpenGIS Specifications for Grid Coverages, Simple Features, and Catalog Services. OpenGIS Grid Coverage Specifications address satellite images, aerial photos, digital elevation data, and other kinds of "gridded" data. OpenGIS Simple Features Specifications address "vector" geodata, i.e. digital map information (streets, population, land use zones, property lines, watersheds, etc.) represented in polygons and lines. OpenGIS Catalog Services Specifications provide a common architecture for online automated directories of web-based geospatial data and geoprocessing services, rather like "spatial search engines." As vendors implement these open standards, web users will easily find, view, overlay, and combine different thematic maps for a given region. Currently

a follow-on project (WMT-2) is ongoing, with the aim to finalize and publish the first three OpenGIS Web Map Server Interface standards and test and agree upon minor improvements to GetFeatureInfo, GetMap, and GetCapabilities, approved by OGC in February, 2000. Examples of WMT applications can also be found on some ESA information servers [40][41].

- The **Dublin Core** Metadata Initiative (DCMI) [36] is an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models. DCMI's activities include consensus-driven working groups, global workshops, conferences, standards liaison, and educational efforts to promote widespread acceptance of metadata standards and practices.
- **ISO/TC211** [1] provides standardisation in the field of digital geographic information aiming to establish a structured set of standards for information concerning objects or phenomena that are directly or indirectly associated with a location relative to the Earth. These standards may specify, for geographic information, methods, tools and services for data management (including definition and description), acquiring, processing, analysing, accessing, presenting and transferring such data in digital/electronic form between different users, systems and locations. The standard affecting ENVISION mostly is '19115 (15046-15) Geographic information - Metadata'. This standard defines the schema required for describing geographic information and services. Note that the standards defines only the data schema, not the underlying transport protocol. Consequently different ISO19115 implementations can exist without that they are really interoperable. However it should be relatively easy to build gateways between them.
- The **OpenGIS Catalogue Interface Implementation Specification** [11] standardises directory, inventory and ordering services and foresees 4 different profiles: The CORBA fine-grained and the CORBA course-grained, the OLEDB profile and the WWW profile based on a HTTP/XML solution. The CORBA coarse-grain and the WWW profile are de factor a mapping of the current CIP and GEO profiles performed by NASA and the FGDC with some support of EC/JRC.

The resulting impression is that the ‘state of the art’ are interoperability islands for specific communities, mostly based Web pages for human use or Z39.50 profiles. Operational services deployed by application communities typically limit themselves to directory function, for others on-line access and on-line ordering are frequently not addressed. In the rare cases that they are addressed, they typically do not consider commercial restrictions and restrictions from European or national data policy. We have therefore to conclude that it is high time for a consensus finding process, possibly on an infrastructure with enhanced functionality.

4 Preliminary Assessment and Recommendations

As explained in the introduction, successful operational deployment and commercial exploitation of geospatial applications will in many cases require the existence of a suitable interoperable infrastructure basis. Some preliminary assessment has been done and there are some impressions and findings that the Agency can feed into a consensus finding process. 'Preliminary', as these findings are based on the limited input that could be collected in the 'working area' of the Agency (ESA), which could include only very small feedback from the geographic and also the application community. It is evident that the Agency's internal thinking that is presented it in the following is ongoing and needs to be verified and refined against additional inputs from the other communities. The ideas presented as a 'question and answer' session herafter has therefore to be considered as an initial input that the Agency can feed into a European-level consensus finding process.

What kind of interfaces do geospatial applications need in an operational scenario?

In an operational scenario, it is assumed that geospatial applications require access to data and services in an automated way, typically without human intervention and based on a stable 'Application Programmer Interface' (API). This implies that ENVISION services cannot be limited to a Web interface for human users, as this is typically frequently changing and usually 'interactive' (i.e. the input of a single operation is acquired in successive steps and also the result is presented at different levels in an interactive way). On the contrary, an application interface needs to be 'frozen', and preferably is based on a self-containing request/response paradigm.

This consideration eliminates Web-only infrastructures and require that the services are based on a defined and stable protocol. However, the data model behind the just 'eliminated' projects should be considered and fed into the consensus finding process at the appropriate moment.

What kind of the existing services do geospatial applications need in an operational scenario?

It is difficult to imagine that geospatial applications select their data sources 'dynamically' via a directory search and then transparently access directly to on-line data. In a typical operational scenario the data sources are assumed to be predetermined and changes (e.g. extension to an additional data provider) require human intervention. While "**directory services**" have a primary enabling role in the research and development phase of an application, they seem to have secondary priority for the specific goals of ENVISION. This might ease the consensus finding process considerably, as one of the most difficult issues to solve would be on indexing and thesaurus usage, where there is a considerable diversity between existing infrastructures. However, for the sake of applications development it will still be desirable to also agree on a standard for directory entry description and exchange in long term.

"**Inventory services**" are instead considered the key service for automated access to the real data. Preferably applications will then directly access to the data, but in some specific cases ordering will still play a role. In particular ordering of future products will be important for EO mission planning (this does not exclude that the products are then fetched from an on-line archive, once they are available).

As many of the existing (diverging) infrastructures limit themselves to the directory function, an initial low priority for this service will probably contribute to accelerate the overall consensus finding process. It has to be noted that the only protocol-based operational infrastructure, going beyond the directory service, where a European participation remains, is INFEO [3]. Although we will see later, why INFEO and CIP are probably not a satisfactory basis for a consensus, this existing infrastructure allows us to gain valuable experience to be considered in ENVISION.

Any additional services required?

Current infrastructures typically base their service paradigm on a client-server 'pull' approach. In mature applications, e.g. meteorology, we can instead observe that in most cases data are automatically 'pushed' to the institution responsible for application elaboration. This is e.g. the case for typical near-

real-time chains served today by ESA satellites. In most cases these data are simply for free and the customers are public bodies, so definitely such an indiscriminated 'push' would not be suitable in all environments. However at least for the data acquisitions that trigger the application elaboration it has to be considered that 'push' of newly acquired data - filtered by some criteria - will be a valid choice. The 'pull' approach will still remain valid for many situations where the cost of data is critical, and maybe an application operator will have to actively confirm the purchase of some data before they are actually retrieved by the application.

To support such 'filtered push' or 'moderated pull' data access, it may be considered a valid approach to have a 'data sniffing' service. Such a need has been identified for the RAMSES oil slick detection, where a preliminary assessment of a low-resolution image ('sniffing') would give some confidence if it is an interesting candidate for the application, and only in this case the data would be purchased.

Single-stop shop or service federation?

It is evident that for the human user a 'single-stop shop' is preferable and many current infrastructures implement this model. The typical technical drawback is the creation of bottlenecks. In addition it has to be understood, that data providers may feel uncomfortable to be accessed and 'presented' through an interface that they cannot control. This could go so far to create reluctance to attach the services to an interoperable infrastructure, e.g. because the data provider is not satisfied with the way his data are presented.

In the relative discussion, where pro's and con's will have to be debated, it will also be essential to distinguish between discovery ('directory') services and the actual inventory and data access services offered by the providers. For discovery services there also exist good examples of centralised approaches (e.g. 'Yahoo!', 'Altavista') while for the actual data access services a federated model seems to be more appropriate for a European-size infrastructure possibly aiming at global scalability. However, it should be envisaged that the small providers might prefer to share a common access point, e.g. organised at national level. Some protocols support this scenario from a protocol point of view, but typically the aspect of user management is not solved satisfactorily in the context of a federated approach.

How powerful does a protocol have to be to be successfully promoted to data providers and application builders?

Federated operational infrastructures have to face the problem of convincing the single data providers to offer a suitable protocol front-end to the middleware. Clearinghouse has solved the problem by distributing their data server software (ISITE) for free, but this approach cannot work with the big providers with large legacy systems. INFEO has developed a configurable CIP/ODBC gateway, but experience with this specific implementation shows many cases where such a generic approach has lead to unacceptable performance losses. In fact, providers like ESA, DLR or Spotimage have developed their own native gateway. This has proven a costly and time-consuming exercise with a powerful protocol like CIP and therefore the effort to have a performing gateway is one of the main obstacles to connect data providers in a fully operational way. An interesting example for a different approach is CEONET. CCRS has successfully connected many providers in a performant way via their 'simple search' protocol. It provides only essential functionality, but proves sufficient for most cases and is very simple to use.

By now ESA has also acquired some experience about the needs of application builders and the result is that also for them a too powerful protocol is an even greater obstacle. It requires a too long learning curve in a non-core business area and typically only a very small subset of functionality is required.

It is therefore useful to be aware of the dangers of 'academic completeness', which many approaches are aiming for. A possible way forward would be to separate between a 'simple access' and a 'full-fletched' protocol. The ENVISION consensus finding process could focus on the 'simple access' protocol, targeting it to the application builder community, while it should give only secondary priority to a 'full-fletched' protocol, which would be more useful for closer co-operations, e.g. data exchange be-

tween large data providers. Also from a programmatic point of view it will be easier to initially agree and implement a mandatory core and amend it later with optional extensions.

What transport layer technology can be successfully promoted to data providers and application builders?

Another obstacle that is common in promoting an infrastructure to data providers or application builders is the technology chosen for the transport layer. This is heavily experienced for the Z39.50 protocols (e.g. INFEO), which were the state-of-the art when these protocols were defined in the early 90ies, but are not mainstream information technology anymore. It is evident that an application builder cannot spend his time in looking for Z39.50 skilled people on the market. Unfortunately this eliminates most of current operational protocols from the consensus finding process. Considering that the choice should be multi-platform protocol suitable for Internet access, based on off-the-shelf mainstream information technology, from the currently available choices only the Open GIS Corba and WWW profile remain on the table. Only for the WWW profile some practical experience exists, that was acquired within the World Mapping Testbed (WMT) pathfinder initiative. Although the scope was very limited, at least some proof of technology has been successfully performed.

Which domain should drive the process? EO, GI or applications?

Who takes the lead? If this question is not spoken out openly, it could harm a sane consensus finding process. EO has proven programmatic capability to implement infrastructures on a global scale, but ESA assessment would be that EO is just a (relatively small) subset of the GI domain.

The application community is the target customer and if it would define a standard it would like to use, it would be easy for the data providers to simply follow. Unfortunately it cannot be expected that there is an easy consensus between different application domains. It is more realistic that the underlying 'generic' EO/GI community define a generic standard and then the application domains specialise for their purpose in the same way as one level below standard information technology (e.g. Z39.50,XML) is being 'profiled' for the geospatial information domain.

The GI world should take the lead considering that they should be able to do the right level of abstraction and probably have the most significant 'overall weight'. On the other hand, in Europe the heterogeneity is an order of magnitude higher than in the EO world, and also than in the US. It seems therefore quite clear that no one will be able to take the lead and play a predominant role. The full consciousness that no one will be able to 'impose' his standard to the others could be essential to ensure a common positive attitude to the consensus finding process.

What standards should be considered?

To guarantee stability, a standard should be maintained by a body, and not be the property of a single institution or company. This eliminates any proprietary standards. For standards qualifying according to this criterion the answer seems rather straight-forward. De facto - and admittedly thanks to the US 'players' - data models of existing protocol-based infrastructures have been fed to both to ISO as well as to OpenGIS. OpenGIS also has a strong liaison with ISO/TC211 and therefore it can be expected that OpenGIS solutions will be ISO compliant. Also the CEN standards have been input to ISO, and therefore we assume that CIP, GEO and CEN are superseded by an ISO-aligned OpenGIS specification for the next generation infrastructure.

However, all involved parties should be aware of the current status and quality of the ISO and OpenGIS specifications. The application community did not have the occasion to participate in the standardisation process and current specifications are theoretical specification standards that have to be verified in practical implementations.

The ISO standard is at a quite high-level, leaving considerable freedom for the final implementation (e.g. choice of transport layer protocol), de facto not guaranteeing interoperability at communication level between ISO-91115 compliant infrastructures, without agreeing on the implementation details.

The current OpenGIS Catalogue Implementation specification contains de facto four different 'stan-

dards' of which one will have to be selected. Only for the WWW profile some practical experience exists, that was acquired within the World Mapping Testbed (WMT) pathfinder initiative, which had a very limited scope. The OpenGIS Catalogue implementation is also quite rich in content and it would have to be evaluated, if it is feasible to base a consensus finding process on the entire width of the specification or if it is more promising to start on an subset of the specification.

In any case it cannot be recommended that in the current context ENVISION should start a 'lighter protocol' from scratch, but instead it should stick as closely as possible to these standards. As a compromise it will probably be wise to find consensus on an initial subset of ISO 19115/OpenGIS, and enrich it with elements required by the application community.

European or Global?

ESA experience in CEOS shows that typically the technically most active participants in the global bodies come from the US. Discussion and consensus building in these bodies is sometimes heavily influenced by US agenda and the need to solve divergences between different US bodies. As an example we can find the NASA requirement to harmonize the CIP version 2.4 with the FGDC GEO profile which had a considerable impact on the delay to issue this revised specification and also on the INFEO implementation. Also in the current 'OpenGIS Catalogue Specification' NASA/FGDC and NIMA have made different proposals for CORBA profiles, which have required substantial work to get aligned, without full success (there are still two different CORBA profiles with significant overlap).

It should be clear that in Europe the bodies that need to be involved in the consensus finding process are even more than in the US and that this will require a strong European cooperation. This process will be at risk in an environment where the agenda and speed is heavily influenced by the needs of the US bodies. It seems therefore evident that a reserved European forum for the Europe-internal consensus finding process is required.

The fastest way to build a European infrastructure would therefore probably be to discuss and define a 'adapted' solution for European purposes, derived from existing standards. Considering the size of the European market, the application sector could probably go far with this. While such a scenario might be satisfactory for many GIS data providers, typically offering data on a regional scale, such a solution would be a major step back for he EO community. Satellite technology implies the opportunity of acquiring data globally and many problems in the understanding of our planet have to be solved in a global scientific effort. Applications based on EO data (but producing GI data) can benefit from the global availability of acquisitions from the same data source could also benefit from a global approach that eases world-wide 'export'. Considering the successful results in aiming at a global infrastructure, a European 'adapted' solution - not necessarily compatible with global standards anymore - would be considered a major step back.

The value of a European forum should therefore possibly be seen as an opportunity to arrive at a common position in international bodies, and there should be a recognised necessity for a feedback process between international and European activities, with European partners being more active in the global bodies.

It is a highly political and strategic issue to find the best compromise between two extreme positions, which highly influences the programmatic way forward.

Commercial or Scientific approach?

Available infrastructures and protocols typically make very liberal assumptions (typically 'no restrictions') on the right of a data provider to distribute data to any user and also do not satisfactorily address security problems of electronic commerce, in particular non-repudiation of consumed data. This drawback has many sources: in part it may be due to the fact that in the US there are much less data policy restrictions than in Europe, in part because in the US there is a different culture on public accessibility of data acquired by public institutions, but mostly there is also an inherent complexity of the issue.

We should not necessarily assume that data are for free, but that they may have to be retrieved from commercial providers in accordance with commercial restrictions on security, confidentiality and data policy. Depending on the chosen approach this may require a considerable upgrade of existing standards, requiring that appropriate effort in the ENVISION consensus finding process is dedicated to this issue. A good part of the problem might be solvable by addressing security issues in a modern public/private key (as used for digital signatures) environment instead of the traditional userid/password approach.

How many different 'standards'?

The currently most imminent issue is the harmonisation between access to EO and GI data. From the experience gained in alignment of CEOS and FGDC protocols [31] it is judged possible to arrive at a common standard wit only a few specific extensions.

A similar problem of divergence may occur for the application communities. It can be assumed that a specific application will need some application-specific information in the metadata model. Taking into consideration that the primary requirement for application infrastructure builders will be to satisfy their own information needs, this can easily lead to a completely specialised data access infrastructure, decoupled from any more generic standard.

It will be important to ensure that application communities recognize that any of their data may be useful for other external applications that may build on it. It should therefore be a primary goal of a harmonization initiative like ENVISION to ensure openness of information infrastructures to support also interdisciplinary data access. In the Agency's vision an appropriate balance between communality and diversity should be found in a technical approach that aims for a common core with the possibility for application-specific extensions.

The value-adding industry, which is supposed to build the information chain that generates application-specific information products may play an important role in this process. This sector has both the need to access lower-level 'raw data' to feed their information production engines, as well as to distribute the finally generated products in a way suitable for the application end users. In addition they have a motivation to possibly re-sell their information products also to other customers that the application end-users for which they were perceived.

5 Conclusion

This current ESA analysis indicates an opportunity to base a European consensus on OGC standards. A certain number of issues are identified (need for verification by implementation, 'heavy' against 'light' approach, lack of consideration for security and data policy constraints). Out first assessment is that the standard is not defined in sufficient detail to be implementable straight-forward, but will require some refinement and revision.

The answer to question 'European or Global' will needed to be answered with primary importance. Shall we just take a snapshot of the current standard specifications and adapted it independently? Or shall we try to do the necessary adaptations in a close liaison with the global bodies, discussing them within a global forum? The latter approach might e.g. require active work and funding of European bodies within OpenGIS to get the European problems addressed and solved quickly. In addition there would also be a need to raise the awareness of the application community and their institutional bodies (e.g. IOC, EEA) and involve them more tightly into the consensus finding process.

Europe should avoid to repeat the US experience, where the GI sector lead by FGDC and the EO sector lead by NASA have developed incompatible infrastructures that now need to be harmonised with substantial overhead. This is not a trivial task, as the only partially successful attempt of GEO and CIP harmonisation shows. At the end FGDC and NASA will be able to harmonise only in the next generation infrastructure, for which seemingly agreement has been reached on the OpenGIS protocol. With the currently extremely divergent lines of activities between the EO, GI, and application sector communities, Europe is at risk to pass the same experience, with considerable damage to the growth of an important sector of digital economy.

It is therefore essential to finally start working on a consensus on how to proceed in Europe. This need has been outlined in particular for the environmental sector, e.g. in occasion of the "1st IST concertation meeting on environment applications", held in Brussels 4-5 May 2000. Following this meeting there seemed to be an opportunity to implement the required consensus finding process via an IST accompanying measure. Unfortunately the ESA attempt to meet these opportunities by submitting a proposal for the ENVISION activity has not been successful, preventing any rapid progress on this issue. At this point the EO community will have to analyse if it still makes sense to address the topic in short term, or if it is more useful to have it incorporated it into a future framework like GMES for instance.

References

1. ISO TC211 Geographic Information/Geomatics http://www.statkart.no/isotc211/
2. http://forum.afnor.fr/afnor/WORK/AFNOR/GPN2/Z13C/PUBLIC/WEB/ENGLISH/pren.htm: CEN/TC287: The Geographic Information European Prestandards and CEN Reports
3. www.infeo.org
4. www.sea-search.net
5. www.megrin.org/PROJECTS/ESMI/Esmi.html
6. www.ceos.org: Committee for Earth Observation Satellites
7. gcmd.gsfc.nasa.gov/ceosidn
8. eos.nasa.gov/imswelcome
9. Catalogue Interoperability Protocol (CIP) Specification - Release B: CEOS/WGISS/PTT/CIP-B, Issue 2.4., June 1998, CEOS WGISS - Protocol Task Team, see www.dfd.dlr.de/ftp/pub/CIP_documents/
10. www.gsdi.org
11. www.opengis.org
12. ioc.unesco.org
13. ioc.unesco.org/iode
14. ioc.unesco.org/goos
15. www.ioc.unesco.org/igossweb/igoshome.htm
16. www.eea.eu.int
17. eea.eionet.eu.int/etc
18. www.mu.niedersachsen.de/cds/
19. www.esa.int
20. earth.esa.int
21. www.esa.int/export/esaSA/navigation.html
22. www.cnes.fr
23. www.dlr.de
24. eoweb.dlr.de
25. www.nlr.nl
26. www.neonet.nl
27. www.eurogi.org
28. www.megrin.org
29. Z39.50 Application Profile for the Content Specification for Digital Geospatial Metadata or "GEO", 22 April 1997, Federal Geographic Data Committee
30. www.fgdc.gov/clearinghouse/clearinghouse.html
31. Best, C., Hayes, L., Nebert, D., Reich, L., Perceival, G., Hilton, J., Smith, S.: Alignment of CIP and GEO Profiles. Published in: Strobl, J., Best, C. (Eds.), 1998: Proceedings of the Earth Observation & Geo-Spatial Web and Internet Workshop. ISBN: 3-85283-014-1, see www.dfd.dlr.de/ftp/pub/CIP_documents/
32. ceonet.ccrs.nrcan.gc.ca
33. gcmd.gsfc.nasa.gov/difguide/difman.html
34. geixs.brgm.fr/
35. www.opengis.org/wmt
36. dublincore.org
37. Ian Foster, Carl Kesselmann: The GRID: Blueprint for a New Computing Infrastructure, Morgan-Kaufmann Publishers, San Francisco, 1999, ISBN 1-55860-475-8
38. www.globalgeo.com
39. ramses.esrin.esa.it
40. mapserv.esrin.esa.it/ionic/atsr
41. odisseo.esrin.esa.it/eoli/eoli.html