INSPIRE

Infrastructure for Spatial Information in Europe

Reference Data and Metadata Position Paper

Title	Reference Data and Metadata Position Paper
Creator	RDM Working Group – Editors: Daniel Rase, Anette Björnsson, Mark Probert, Marie-France Haupt
Date	2002-10-02
Subject	INSPIRE RDM Position Paper
Status	Final
Publisher	EUROSTAT
Type	Text
Description	Inspire RDM Position Paper
Contributors	See annex
Format	MS Word 95/2000 (doc)
Source	Not applicable
Rights	INSPIRE working group members.
Identifier	INSPIRE RDM PP v4-3 en
Language	En
Relation	Not applicable
Coverage	Project duration

These are Dublin Core metadata elements. See for more details and examples http://www .dublincore.org/.

Version history:

Version number	**Date**	**Modified by**	**Comments**
1.0	2002-05-22	Daniel Rase	First draft
2.0	2002-06-20	D. Rase, A. Björnsson, M.-F. Haupt	Second draft
3.0	2002-06-26	D. Rase, A. Björnsson, M. Probert, M.-F. Haupt	Third draft
4.0	2002-09-27	D. Rase, A. Björnsson, M.-F. Haupt	Fourth draft
4.1	2002-10-02	A. Annoni	Inclusion of introduction chapter
4.2	2002-10-02	M. Probert	Language changes

Table of contents

1. Purpose of the document

This is the Position Paper of the Reference Data and Metadata working group. It is the major deliverable of the group to be presented at the Expert Group Meeting in Athens, October 30-31, 2002.The present document is also intended for dissemination on the Web. It contributes to the creation of the **INSPIRE** legislative framework and to its future implementation.

The content of this document is based on the working group action plan and on the physical meetings of February 8, 2002 in Luxembourg, March 19, 2002 in Paris and May 27-28, 2002 in Bonn.

The position paper has been developed in a short period of time and with voluntary non-funded work done by the Member states and candidate countries. The position paper could therefore not cover all issues in depth. Consequently, a lot of work is still to be done for the future. For this work to be carried out and to ensure high quality results, appropriate funding will need to be considered.

As co-ordinator of the RDM working group, EUROSTAT would like to thank the group members for the quality of their contributions and their continuous efforts to meet the short deadlines.

2. Introduction

(Chapter developed by the INSPIRE secretariate/ WG leaders)

2.1 What is the INSPIRE initiative and why is it needed?

Good policy relies on quality information. The increasing complexity and interconnectedness of issues that affect the quality of life today is recognized by the policy-makers and influences the way new policies are being prepared today. The Sixth Environmental Action Programme[1] for instance emphasises the need to base environmental policy-making on sound knowledge and participation, principles that will influence the Union environmental policy-making for the next decade.

INSPIRE is an initiative currently being prepared by the Commission to support the availability of spatial information for the formulation, implementation and evaluation of Union policies. It intends to set the legal framework for the gradual creation of a spatial information infrastructure. INSPIRE will initially focus on environmental policy needs but, being a cross-sectoral initiative, will gradually be extended to other sectors (e.g. agriculture, transport, …) as other interested Commission services participate.

What is a spatial information infrastructure?

The INSPIRE initiative intends to trigger the creation of a European spatial information infrastructure that delivers to the users integrated <u>spatial information services</u>. These services should allow the users to identify and access spatial or geographical information from a wide range of sources, from the local level to the global level, in an inter-operable way for a variety of uses. The target users of INSPIRE include policy-makers, planners and managers at European, national and local level and the citizens and their organisations. Possible services are the visualisation of information layers, overlay of information from different sources, spatial and temporal analysis, etc.

The spatial information infrastructure addresses both technical and non-technical issues, ranging from technical standards and protocols, organisational issues, data policy issues including data access policy and the creation and maintenance of geographical information for a wide range of themes, starting with the environmental sector.

The INSPIRE initiative recognises the fact that most of the quality spatial information is available at local and regional level, but that this information is difficult to exploit in a broader context for a variety of reasons. The situation on spatial information in Europe is one of fragmentation, gaps in availability of geographical information[2], duplication of information collection and problems of identifying, accessing or using data that is available. As a result of these problems, effective Union policy actions suffer because of lack of monitoring and assessment capabilities that take into account the spatial dimension[3].

Fortunately, awareness is growing at national and at EU level about the need of quality geo-referenced information for understanding the complexity and consequently for containing the negative impacts of the ever-increasing human activity on the EU territory. Many regional and national

[1] http://europa.eu.int/comm/environment/newprg/index.htm
[2] For example, only a few pan-European geographical information layers exist, often designed for specific purposes that limit the possibilities of their wider use e.g. CORINE Land Cover and the SABE dataset (Seamless Administrative Boundaries of Europe) from EuroGeographics.
[3] For example: insufficient monitoring capabilities are key obstacles to the further development of a range of priority themes of the 6th Environmental action programme, such as soil, bio-diversity, health and environment and marine policy.

initiatives are being taken[4] and numerous stakeholders both in the Member States and candidate countries collaborate with the Commission services for the preparation of the INSPIRE initiative.

Successful implementation of the INSPIRE initiative would contribute to reach the objectives set out in the Commission's White Paper on European Governance[5]. It would help the Commission to establish more coherence in its policies by better integrating the common territorial dimension. This will also help to improve policy co-ordination, an issue that is identified by the Community Sustainable Development Strategy[6] as part of a new approach to policy-making. It will allow better participation by presenting information in a clear, understandable way at national and local level. Finally, it will help to make European governance more effective by supporting the evaluation of future impact and past experience for EU policies.

2.2 Context and vision

Recent global advances in moving from paper to digital data and information has created hitherto undreamed of opportunities to revolutionise access to data, communication of information and for informed decision-making at all levels of society. This move from back room to open door access to information presents new challenges for those acquiring, handling, and providing access to electronic data and information.

The data are often of unsatisfactory or undefined quality, based on proprietary geographic information systems and not accessible to the public or other users at local, regional, national and international level. Therefore, projects that combine data coming from various sources to provide policy-relevant information and tools are often time consuming and costly. Policies need to be put in place to reduce the duplication in collection, harmonisation efforts and to facilitate and promote wide dissemination of the data. These policies should free funds to be invested in improving the availability and quality of spatial information. The increased availability of data will in turn stimulate innovation among data and information providers in the commercial sector.

INSPIRE Principles

The INSPIRE initiative intends to improve the current situation by triggering the creation of a European Spatial Data Infrastructure for the access and use of spatial information built on the basis of the following principles:

- √ Data should be collected once and maintained at the level where this can be done most effectively
- √ It must be possible to combine seamlessly spatial information from different sources across Europe and share it between many users and applications
- √ It must be possible for information collected at one level to be shared between all the different levels, e.g. detailed for detailed investigations, general for strategic purposes
- √ Geographic information needed for good governance at all levels should be abundant and widely available under conditions that do not restrain its extensive use
- √ It must be easy to discover which geographic information is available, fits the needs for a particular use and under what conditions it can be acquired and used
- √ Geographic data must become easy to understand and interpret because it can be visualised within the appropriate context and selected in a user-

[4] See Examples of regional and national initiatives to create a spatial information infrastructure in GE, UK, PT on the Internet http://www.ec-gis.org/inspire/

[5] COM(2001)428 – European Governance - a White Paper . The White Paper refers to five principles of good governance: openness, participation, accountability, effectiveness and coherence

[6] Presidency Conclusions – Göteborg European Council, 15 and 16 June 2001

friendly way.

The INSPIRE policy vision is to make harmonised and high quality geographic information readily available for formulating, implementing, monitoring and evaluating Community policy and for the citizen to access spatial information, whether local, regional, national or international[7]. This vision is illustrated in the diagram at Figure 1.

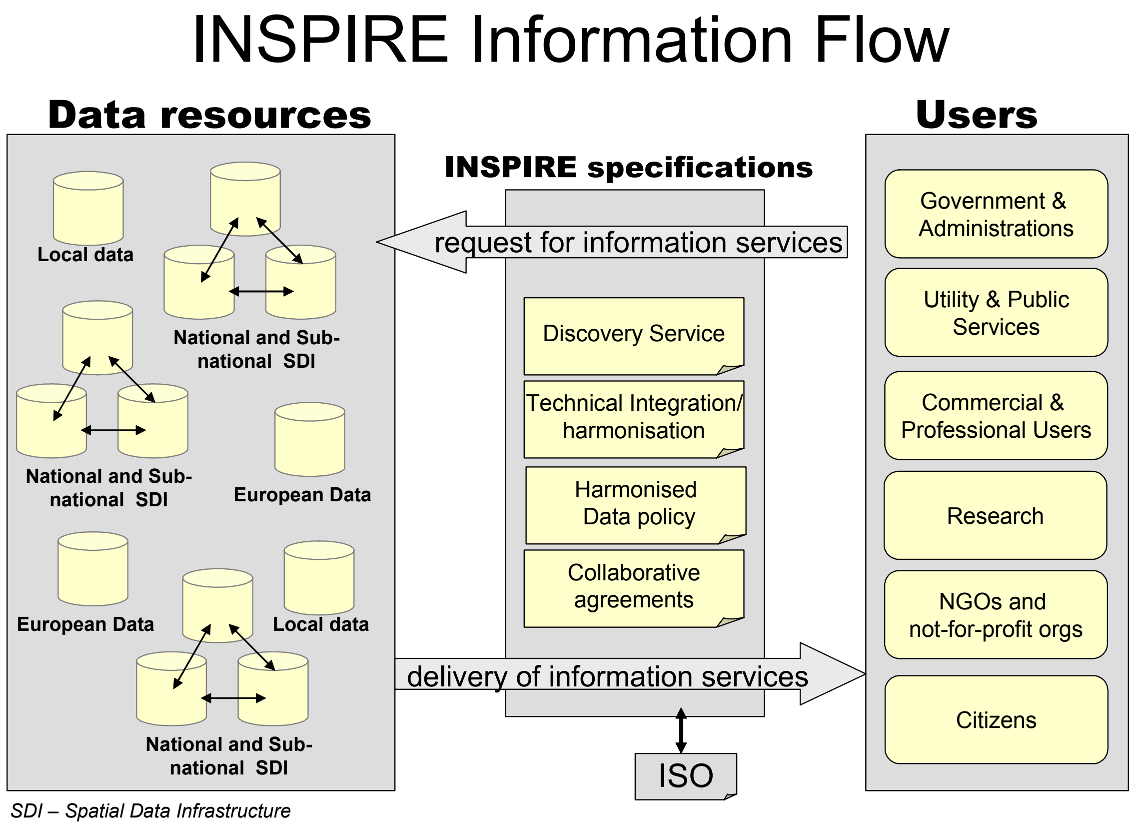

Figure 1: Diagrammatic View of the INSPIRE Vision

2.3 Stepwise approach

The INSPIRE implementation will follow a step-wise approach, starting with unlocking the potential of existing spatial data and spatial data infrastructures and then gradually harmonising data and information services allowing eventually the seamless integration of systems and datasets at different levels into a coherent European spatial data infrastructure. Achieving this objective will require the establishment of appropriate coordination mechanisms and common rules for data policies. Where relevant, synergies with the GMES initiative will be sought in order to ensure coherence between INSPIRE and GMES [8].

The first step will focus on harmonisation of documenting existing datasets (metadata) and on the necessary tools to make this documentation accessible.

The second step will primarily aim at providing common ways to access the spatial data sets themselves allowing uncomplicated analysis of data on different themes coming from different

[7] The INSPIRE initiative will link with relevant initiatives at the global level such as the work to develop the Global Spatial Data Infrastructure (GSDI).
[8] Sec(2001) 993 of 16/06/2001 Commission Staff Working Paper – Joint document from Commission services and European Space Agency

sources. An example of such analysis is visual inspection of spatial relations between phenomena by overlay of datasets.

The third step will target the establishment of common models of the objects in the environment for which spatial data is collected, such as transport networks, forests, … This will allow to map existing datasets to a common set of models, the start of the creation of a really harmonised spatial data infrastructure that will facilitate the combination of information of various sources and more advanced analysis work.

The fourth and last step will build upon the previous steps and concentrate on completing the common models and on providing the services to fully integrated data from various sources and various levels, from the local to the European level into coherent seamless datasets supporting the same standards and protocols. This step will allow real time access to up-to-date data across the whole of Europe.

These steps will partly be carried out in parallel, depending on user needs and degree of availability and harmonisation of existing information. All these steps involve actions of standardisation, of harmonisation and integration of data and services.

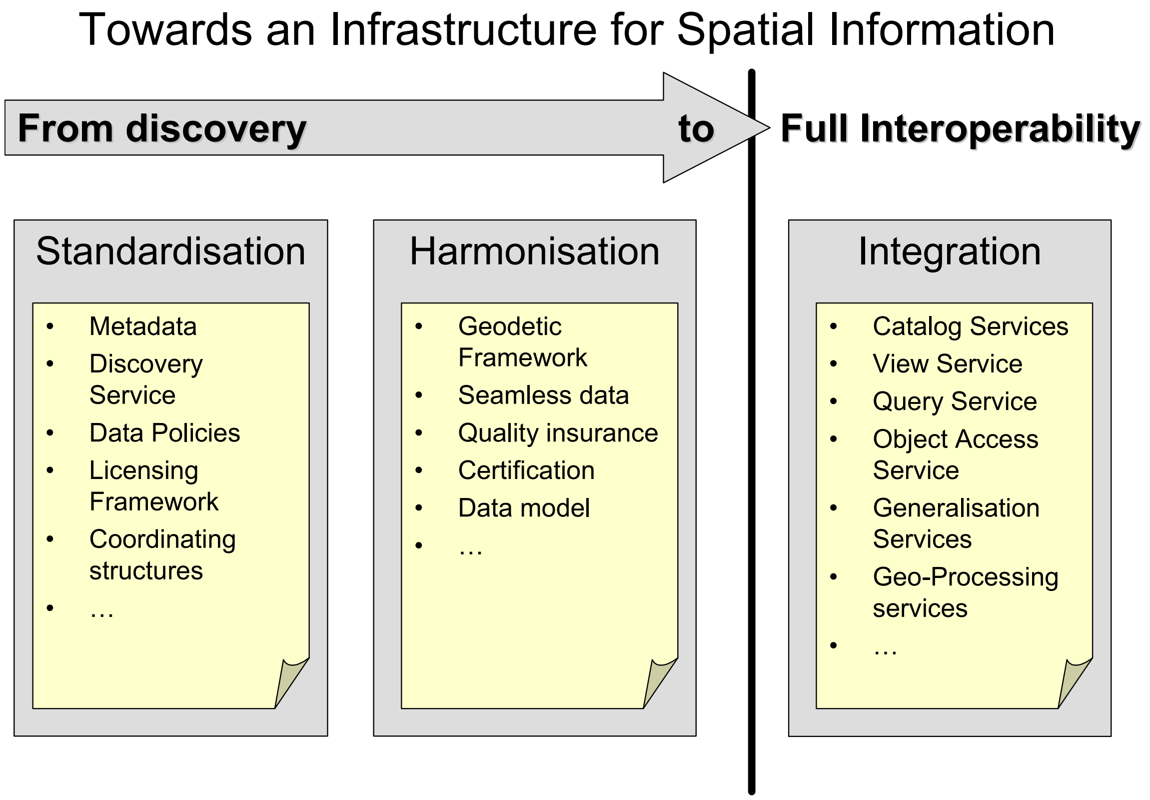

Figure 2: Towards an Infrastructure for Spatial Information

INSPIRE is conceived as a cross-sectoral initiative covering the main Community sectors with a spatial impact such as transport, energy, agriculture, … but will target initially information needed to support environmental policy. Indeed the 6th Environmental Action Programme highlights the need for better knowledge and sound science in environmental policy-making and geographical information will therefore be increasingly required to achieve this. Therefore, a horizontal framework is needed in order to ensure a coherent approach to information collection and distribution. Moreover, the requirement by the Treaty for all policy sectors to integrate environment concerns will provide a first link from environmental policy to other policy sectors that can be further extended at later stages.

2.4 Users, Producers and other Stakeholders

2.4.1 Users

Environmental users are many and various, and include users who need spatial data for planning, management, assessment, monitoring and reporting. Hence the user community is very broad and diverse and includes:

- Governments & Administrations
  - EU
  - National
  - Regional
  - Local
- Utility and Public Services, including
  - Transport
  - Health
  - Emergency services
  - Utilities (e.g. water, telecommunications, gas, electricity).
- Research and development
  - Universities
  - Public and Private Institutes
  - Application Developers for IT Systems
- Commercial & Professional End Users
  - Tourism
  - Value Added Resellers
  - Surveyors
  - Property Developers
  - Insurance
- Non Governmental Organisations (NGOs) and not-for-profit organizations
- Citizens

Different user categories must be considered because their requirements in terms of data access can vary significantly.

2.4.2 Producers

The producers of spatial information within the public sector include national environmental protection agencies, mapping agencies, national geological surveys, national maritime administrations, cadastral, land registration and other land administration organisations, local authorities and utilities.

It should also be noted that, under certain circumstances, private data producers may offer production capacity to public bodies, or possibly sell data directly onto the market themselves. In some Member States there is a thriving private sector geographic information industry supplying data and services directly to the commercial market.

Most spatially organised data and information are either used internally by public bodies, or are supplied to other public sector organisations under various types of agreement. A relatively small but growing number of government departments or agencies conduct commercial business with the private sector or with the general public. It is in the area of data use that it is important to recognise the difference between sharing data and trading data.

The simplified diagram at Figure 3 clearly shows this distinction in the context of three transaction streams which can be combined in varying proportions by any public sector body developing an overall information sharing and trading strategy, subject to common rules defined under INSPIRE.

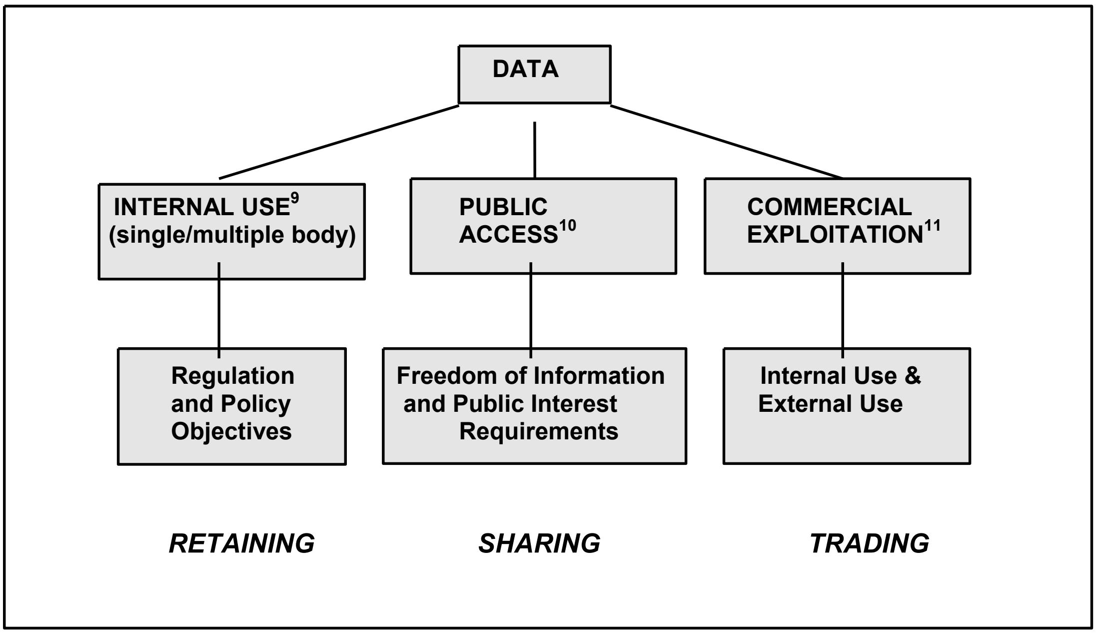

Figure 3: Simplified Diagram Illustrating Public Sector Data Uses

2.4.3 Other Stakeholders

The delivery of INSPIRE, like initiatives such as *e*Europe and *e*Government, is dependent on information technology. It will have a profound impact on a variety of disciplines and professions, affecting many individuals and organizations that cannot be categorised as users or producers. Conversely, this group of other stakeholders will also have an important role in the process of shaping the infrastructure. Examples of other stakeholders are:

- The Information and Communication Technology (ITC) sector, and in particular product providers who offer software, hardware, and related systems, and service providers who offer system development, database development operations support, and consulting services;
- Standardisation bodies like ISO, CEN, and national standardization organizations;
- Co-ordinators and regulators, including European and national associations.

[9] **Internal Use** means spatial information used exclusively within the originating public body, or shared among any public body at local, regional, national or international level.

[10] **Public access** means spatial information provided by public bodies free of charge or marginal cost of supply free of charge or marginal cost of supply for viewing or use by citizens of the European Union (including NGOs, academia, and research institutes).

[11] **Commercial exploitation** means the utilisation of public sector spatial information in commercial information products.

3. Executive Summary

The EteMII "White paper" reminds us that the origin of the concept **"reference data"** is based on two main ideas:

- It is a series of dataset that **everyone involved with geographic information** uses to reference his/her own data as part of their work.
- It provides a common link between applications and thereby provides a mechanism for the sharing of knowledge and information amongst people.

In addition, the RDM group wants to stress that it is used as a common base to which thematic data may be referenced.

The main functional requirements that geographical reference data must fulfil are:

- Provide an unambiguous location for a user's information
- Enable the merging of data from various sources
- Provide a context to allow others to better understand the information that is being presented

The INSPIRE working group on Reference Data and Metadata (RDM) has described the content of necessary geographical reference data and its metadata on the basis of these 3 principles including the principles of the INSPIRE initiative.

The RDM-group has agreed that the following components will form the geographical reference data:

- Geodetic reference data
- Units of administration
- Units of property rights (parcels, buildings)
- Addresses
- Selected topographic themes (hydrography, transport, height)
- Orthoimagery
- Geographical names

The RDM-group has identified a list of common aspects which have been described and where recommendations are given. The components are described in a separate chapter with respect to a definition, comments on the current and longer-term availability and specific conditions.

The list of the common aspects is the following:

1. Geodetic reference system
2. Quality
3. Maintenance
4. Interoperability
5. Resolution/scale and implementation priorities
6. Language and Culture
7. Metadata

The **geodetic reference system** is considered to be a component as well as a common aspect in which the projections are treated. The **recommendation** of the group concerning this issue is **to adopt the recommendations from the AST position paper.**

Vector data should always be based on ETRS89 whereas projection of raster data will be the choice of the data provider.

Data quality is an important issue. Data should be of an acceptable quality. The quality of both reference and metadata should be addressed. Quality definition and quality control is primarily linked to the spatial content of a data and its attribute. The quality of metadata is also important. ISO standards define the quality principles and quality evaluation procedures.

The recommendations from the RDM group are:

- *The quality of the reference data should be known.*
- *To adopt ISO19113 quality principles and ISO19114 quality evaluation procedures.*
- *To document the results of the quality measurements in dedicated ISO 19115 fields.*
- ***The settings of data quality parameter levels will require further study.***

Reference data must be **maintained** regularly. In most of the current GIS systems two important related questions are not properly tackled; the management of time and of changes.

The RDM group recommends

– *That reference data providers adopt methods and technologies that will permit users to access “change only” updates, but in the short term, they shall retain the traditional “snapshot” approach, according to minimum INSPIRE recommended update intervals. Tracking of incremental changes requires feature-based metadata that is not required in all occasions.* ***The complex issue of "change only updates" with time stamp will require further analysis.***

One of the basic principles of INSPIRE “is to keep the data where it is and to provide access to it”. Therefore, the issue of **interoperability** becomes critical. As different groups of users have different models of the Earth, it is not possible to use data from another database. Therefore, a common conceptual model is required on the basis of this.

The group recommends:

– *To adopt the related standards proposed by the AST group.*
– *To define a conceptual model for the reference data components.* ***This again will require a specific study.***

Data from different countries will have different languages. **Language and cultural** aspects are important factors to be taken into account in several areas, e.g. metadata, specifications, and the spatial data flowing in the infrastructure.

Metadata should be made available in any necessary European language. Data set specifications, feature catalogues and nomenclatures should be written in the official European languages[12] and have translations also in other languages whenever necessary. Free text information will be in the original language. Translated versions can only be expected when data providers see this as a valuable investment.

The RDM group recommends:

– *That reference data specifications are created and described in a way that is commonly understood and which takes into account cultural differences.*
– *To use international standards for the storing of alphanumeric character sets. RDM working group recommends the usage of the UNICODE standard.*
– *To agree of common definitions for objects and their attributes belonging to the components of the reference data.*

Scale (or resolution) of the reference data components is a complex issue with huge impact on the costs and timeframe of the INSPIRE implementation. It could not be adequately covered within the scope of the RDM WG nor within this position paper. Nevertheless, some initial ideas are proposed below that the group feels can contribute as preliminary input for future work (daughter legislations).

[12] The eleven languages of the European Union

The group suggest the following equivalences for scale/resolution:

Geographical level	Resolution range	Scale level	Scale range
European	> 100 m	Small scale	< 1:250.000
National	~ 25 m	Medium scale	1:100.000 ~ 1:250.000
Regional	~ 10 m	Medium scale	1:25.000 ~ 1:50.000
Local	< 2.5 m	Large scale	> 1:25.000

Please note that these ranges are only provided as an indication; further discussion will refine these parameters during the INSPIRE implementation phase.

Following the first principle of INSPIRE (data should be collected once and maintained at the level where this can be done most effectively) the data should be maintained at the local level where there is the highest level of accuracy. A further study is necessary to identify needs and the reference data specifications needed to accommodate for the necessary flexibility or variable resolutions.

It is important to note that the short term development timetable (small-medium scales) and the long term development timetable (large scale) will have to start simultaneously and run in parallel. Transition mechanisms between the two tracks will need to be provided in due course. The sharing of the data at one level between all different levels is a very complex task which needs further work. Therefore RDM suggests the following indicative timeframe for the development:

- 3 years : small scale and metadata
- 6 years : medium scale
- 10 years : large scale

The general recommendations from the RDM group are:

– *That the primary reference data components should be collected and maintained at the largest possible scale, generally at the local level.*
– *To carry out a complementary study to identify the resolution (horizontal,vertical, time) requirements for reference data that may vary according to feature/component and geographic areas.*
– *To define and implement mechanisms that allow the update information to flow from the local to the European level of the reference data.*
– *To focus the INSPIRE framework implementation (comprising architecture, standards, specifications, processes) on large-scale reference data components, but approach the content implementation (reference data sets) in a pragmatic and step-by-step way, based on the current state of play.*
– *To adopt a two-track approach to be started simultaneously and run in parallel.*

Metadata are the information and documentation, which makes data understandable and shareable for users over time.

The RDM group recommends:

– *All the reference data should be documented by metadata. Metadata should be kept up-to-date by a component of authority - to be identified.*
– *The three aspects of metadata must be considered: discovery, access and use.*
– *A metadata profile compatible with ISO 19115 must be developed. It will become mandatory inside the INSPIRE infrastructure.* ***<u>It is therefore recommended to carry out a specific study on this issue with the involvement of all the stakeholders.</u>***
– *Metadata shall be produced for all data that will be made accessible within the future legislation.*
– *Metadata shall be kept up-to-date. Whenever data changes occur that might affect current metadata content, metadata has to be updated as well.*
– *The Member states shall identify a competent authority for coordinating the national producers of data and for managing the metadata information systems.*
– *That priority should be given to create a one stop Internet "European GI" portal for discovering and accessing GI data – similar to the ESMI / La Clef concepts, and backed by appropriate funding and legislation. This would require further study to establish resource requirements.*

4. Methodological approach

4.1 Aim of the RDM working group

For a complete description of the INSPIRE vision, the reader is invited to refer to the common vision of INSPIRE [1].

The basic principles that give direction to the realisation of INSPIRE's aim are:

- Data should be **collected once** and **maintained at the level where this can be done most effectively**;
- **it should be possible to seamlessly combine spatial information from different sources** across Europe and share it between many users and applications;
- **it should be possible for information collected at one level to be shared between all the different levels**, detailed for detailed investigations, general for strategic purposes;
- geographic information needed for good governance at all levels should be abundant **under conditions that do not refrain its extensive use**;
- it should be **easy to discover** which geographic information is available, fits the needs for a particular use and under which conditions it can be acquired and used;
- **geographic data should become easy to understand and interpret** because it can be visualised within the appropriate context selected in a user-friendly way.

In this context, the mandate of the RDM working group was to take a clear position on the contents of the reference data components of the INSPIRE initiative and to produce a position paper by September 27, 2002.

The remit of the RDM group was to:

- Agree on the reference data elements
- Define the necessary resolution and quality according to user requirements
- Provide, if possible, information on data availability, quality, copyright rules, price and maintenance
- Provide, if possible, input to measure the impact of various choices of data in terms of accessibility and costs
- Provide a time scale for the availability of data.

The group needs to know more about the state of play in the European countries in order to provide information on data availability and to measure the impact of the choice of data at country level. EUROSTAT, together with the Directorate General of Environment, recently launched a call for tenders related to this question. Unfortunately, the first results of this survey will only be available by the end of 2002. Considering the tight deadline of September 27 for the drafting of the position paper, the RDM group members could not make much progress in this area.

On the basis of this remit, the group decided to start from the contents of two key documents: The "SDI cookbook" from GSDI [2] and the ETeMII "White paper" [3].

4.2 Definition of reference data and related metadata

The EteMII "White paper" reminds us that the origin of the concept **"reference data"** is based on two main ideas:

- It is a series of dataset that **<u>everyone involved with geographic information</u>** uses to reference his/her own data as part of their work.
- It provides a common link between applications and thereby provides a mechanism for the sharing of knowledge and information amongst people.

In addition, the RDM group wants to stress that it is used as a common base to which thematic data may be referenced.

Reference data must fulfil three functional requirements:

- Provide an **unambiguous** location for a user's information
- Enable the **merging of data from various sources**
- **Provide a context** to allow others to better understand the information that is being presented

The issue of reference data should be addressed at all the territorial levels: European, national, regional and local.

The metadata is a part of the reference data. It is information about and documentation of the reference data, which will make data understandable and sharable for users over time. Metadata shall be used to facilitate access and use of the related geographical reference data.

4.3 Agreed components of reference data

The group decided to take advantage of the important work already carried out in the framework of the ETeMII project and to follow the recommendations of the "White paper".

Finally, the group agreed on the following components:

1. **Geodetic reference data**
2. **Units of administration**
3. **Units of property rights (parcels, buildings)**
4. **Addresses**
5. **Selected topographic themes (hydrography, transport, height)**
6. **Orthoimagery**
7. **Geographical names**

The EteMII "White paper" only contains a short description of the components mentioned above. Therefore, the RDM action plan was built around the idea of providing initial proposals for these descriptions.

5. Common aspects of reference data

5.1 Introduction

The components of the reference data have common aspects. The following section briefly describes these complex issues. Each of them could be the subject of a separate project. It is clear that further work will be needed to provide clear guidance for the implementation of the future infrastructure. For each topic, the RDM group provides recommendations.

5.2 Geodetic Reference system and projections

The RDM group proposes to adopt the recommendations on the Geodetic Reference system and projections contained in the AST position paper (please refer to this document for a complete description).

It has been recognised that ETRS89 is the most appropriate geodetic datum to use within Europe. There remains a need to accept the use of a European height reference for vertical measurements (it is proposed to use EVRF2000).

Vector data should be based on ETRS89. Depending on the type of application users will apply projections corresponding to their needs.

Raster data is always projected and therefore the data provider must make a choice.

The AST position paper proposes 3 different projections to be used in Europe (the countries are allowed to use other ETRS89 based projections).

It is important that the parameters required to convert from the national reference system to ETRS89 are provided.

AST Position Paper recommends the following:

- Use ETRS89 as geodetic datum and to express and store positions, as far as possible, in ellipsoidal coordinates, with the underlying GRS80 ellipsoid [ETRS89]. To further use EVRF2000 for expressing orthometric heights.
- Use ETRS89 Lambert Azimuthal Equal Area coordinate reference system of 2001 [ETRS - LAEA], for statistical analysis and display.
- Use ETRS89 Lambert Conic Conformal coordinate reference system of 2001 [ETRS –LCC] for conformal pan-European mapping at scales smaller or equal to 1:500,000.
- Use ETRS89 Transverse Mercator coordinate reference systems [ETRS-TMzn], for conformal pan-European mapping at scales larger than 1:500,000.
- Each country will provide algorithms to convert coordinates from their national coordinate reference system to ETRS89.

Each country will, as well, provide the necessary conversion algorithms to allow transformation into one of the 3 projections proposed in the AST document.

Recommendations of the RDM group: *It is proposed to adopt the recommendations contained in the AST position paper*

5.3 Data quality

5.3.1 Data quality

Data should be of an acceptable quality. The general concept is that quality should fulfil the requirements of the users. Concerning reference data, these are multi-purpose data to be used by different users, in different sectors and at different levels. Reference data will be used as a basis for linking other data.

It is expected that the producers have the necessary contact with the users and that product specification and development reflects their needs. This has not always been the case in the production of reference data. Therefore, actions may be required to stimulate this contact.

Quality definition and quality control is primarily linked to the spatial content of a data and its attributes. The quality of metadata is also very important.

5.3.2 Technical quality and quality control

To measure data quality, one needs to have elements of comparison. A geographical data set can **be compared to certain common specifications** that are themselves a representation of the real world. In this case, certification or **conformance testing** involves measuring the quality of data samples and comparing them to common specifications. If such a technical quality test is to be carried out certain elements have to be in place:

- **Data set specification:** The specification defines the standard content of the data set. The data set specification is a product specification. There are ISO standards for how product specifications are to be made, and for spatial data there are other ISO standards for data set specifications. There is a need to make an INSPIRE profile defining the minimum content of such a data set specification and templates for implementation. The data set specifications should, whenever available, feature catalogues (data dictionaries) containing shared definitions for features, attributes and codes. The data set specifications should define the accepted range of values for each of the quality elements.

- **Quality elements:** A quality test should focus on specific elements. ISO standards define these elements. Some quality elements are more important than others, and INSPIRE should define which are the most needed (or mandatory) in quality control. These will include:
  - Accuracy
    - Spatial accuracy
    - Temporal accuracy
    - Thematic accuracy
  - Precision and resolution
    - Spatial resolution
    - Temporal resolution
    - Thematic resolution
  - Logical consistency
  - Completeness
    - Data completeness
    - Model completeness
    - Attribute completeness
    - Value completeness

5.3.3 Quality testing

Different ways of quality control may be possible:

- Producers could carry out quality control themselves.
- An independent body could carry out quality control.

Using either method data quality could be measured by comparing the datasets to common specifications (defined as conformance testing in the document ISO 19113**).**

Tools for testing datasets could be made available that would read the specification and give a report on the deviations and values found.

5.3.4 Quality flagging

Geographical data is complex and its quality difficult to measure. In most of the current applications and spatial analysis this dimension is not sufficiently taken into consideration. People continue to use geographical data as if it was perfect. However, research in the field of sensitivity and uncertainty analysis in spatial modelling has shown that error propagation in modelling has an enormous impact on the final results. **In the case of decision systems, this can lead to erroneous political decisions**.

In order to minimise these problems several measures should be taken:

- in the metadata there should be information on certain aspects of quality, especially accuracy, completeness;
- some information on quality, date, accuracy etc, should also be available at feature level, as attribute to the spatial elements in the database;
- a test documentation should be provided to the user, presenting the results where the data set is compared with the data set specification.

The RDM group recommends:

- *The quality of the reference data should be known.*
- *To adopt ISO19113 quality principles and ISO19114 quality evaluation procedures.*
- *To document the results of the quality measurements in dedicated ISO 19115 fields.*
- *The settings of data quality parameter levels will require further study*[13]*.*

5.4 Maintenance

Reference data must be regularly updated. It is clear, however, that some data sets require more frequent updates than others (e.g. roads vs. height).

In most of the current GIS systems, keeping data up-to-date often refers to replace one data layer by a complete new one. One can keep the old version of the layer for historical reasons but very often nobody knows exactly what has changed in the new one (this could be only a few segments or a few polygons).

This means that two important related questions are not properly tackled in most current systems: **the management of time** and **the management of changes**.

The management of time is important for making the comparison between two periods of observation or to follow the evolution of spatially related phenomena. Thanks to new technological developments, geometry can now be stored in relational databases. These database management systems allow for time stamping. The problem here is **to receive appropriate change data from the producers**. Today, most of the data producers, for technical or organisational reasons, cannot provide change data with a time stamp.

The RDM group recommends:

- *That reference data providers adopt methods and technologies that will permit users to access “change only” updates, but in the short term, they shall retain the traditional “snapshot” approach, according to minimum INSPIRE recommended update intervals.* ***The complex issue of "change only updates" with time stamp will require further analysis.***

5.5 Interoperability

Interoperability is a critical issue in the context of INSPIRE because one of its basic principle is to keep the data where it is and to provide access to it.

Most of the current GIS systems are proprietary systems. This means that the internal file format is rarely documented and accessible. To exchange geographic data, people need data converters provided by the software vendors. Exporting data to an intermediate format is another possibility, but unfortunately, this can result in a loss of information.

The trend is to keep the data where it is, and to allow access via a software interface. The work of OpenGIS has centred on this approach.

Accessing the data with a technical infrastructure is not sufficient however; people must also be able to use it.

[13] The cadastral administrations of all German states are currently developing the Official Cadastral Information System "ALKIS" which will integrate all cadastral data and will guarantee a redundant-free data set. The data model of ALKIS is identical to the updated Authoritative Topographic and Cartographic Information System "ATKIS". In addition a systematic semantic harmonisation was applied to ALKIS and ATKIS to allow the "vertical integration" of data as a first step to the general approach that data should only be collected once and should be used for different scales.

Different groups of users have different models of the Earth. Without **a common conceptual model** the data coming from another database is useless. Defining such a model is not possible at this (position paper) stage but it is an essential element of the infrastructure required to make reference data interoperable.

The RDM group recommends:

- *To adopt the related standards proposed by the AST group.*
- *To define a conceptual model for the reference data components.* ***This again will require a specific study.***

5.6 Language and culture

The European Spatial Data Infrastructure will be a system promoting and allowing the flow of data from all European countries. Data will come from local, regional, national and European levels.

Data from different countries will have different languages. Today, for instance, EU has eleven official languages. Dealing with data provided throughout Europe the number of languages will be significantly higher.

Language and cultural aspects are important factors to be taken into account in several areas, e.g. metadata, specifications, and the spatial data flowing in the infrastructure.

5.6.1 Metadata

Metadata are data about the dataset. The metadata comprises two major parts, discovery metadata and access/exploration/user metadata.

- Metadata definitions (fields etc) and template files for explaining and filling in metadata should be made available in any European language necessary.
- The actual metadata for each data set should follow the ISO categories, and should be made available in one of the national/ official languages. The provider of information will benefit from the addition or translation of metadata in English or one of the other languages commonly used across Europe. It is expected that this will apply for metadata for data sets at national and European levels.
- As the metadata definition, feature catalogues and data set specifications will exist in all languages, it will be possible to translate metadata for fields with precoded values.

5.6.2 Feature catalogue and data set specification

Several cultural aspects are important when looking at products and specifications:

- Feature catalogues (data dictionaries), nomenclature, and concepts: The cultural aspects have an impact on the objects descriptions. For example, the concepts of 'city' or 'agglomeration' can vary considerably from one country to another; a 'mountain area' can have various definitions, etc. Harmonising a common understanding of these concepts is crucial but, due to the extent of the task, it should be done step by step. In the development of such products it is essential that different countries are involved. The feature catalogues are descriptions of spatial phenomena found in parts or the whole of Europe. Without the involvement of a spectrum of countries and administrative levels in the development of feature catalogues and nomenclature, the products may be “skewed” in its classification not taking into account the Pan-European context. This in turn, may lead to lack of involvement by some countries.
- Data set specification: The same issues explained above are also relevant in deciding dataset specifications.

- Data set specifications, feature catalogues and nomenclatures should be multi-lingual, and written in the official European Languages and have translations also in other languages whenever necessary.

5.6.3 Data content

- As feature catalogues and data set specifications will exist in all languages, it will be possible to make available explanation texts for attributes with precoded values and feature definitions.
- It is foreseen that free text information will be held in the original language. Translated versions can only be expected when data providers see this as a valuable investment

5.6.4 Geographical names

Geographical names give a name to a location or a landscape object. The geographical names on a specific landscape object can be different in the different languages. In some datasets their primary purpose is to depict geographical locations and in others they may be attributes, and of secondary importance.

- Geographical names should in both cases be provided in the official form and language of the country. For European coverage data sets one of the official EU languages should be used. Providers of data sets to the infrastructure could include secondary name sets in other widely used languages if the data is to be used across borders and at the Pan-European level.

5.6.5 Character sets

Data and their metadata from different countries may contain various character sets. This is not specific to geographical data, therefore, the RDM group suggests using existing related international standards.

The RDM group recommends:

- *That reference data specifications are created and described in a way that is commonly understood and which takes into account cultural differences.*
- *To use international standards for the storing of alphanumeric character sets.*
- *To agree of common definitions for objects and their attributes belonging to the components of the reference data.*

5.7 Resolution/scale and implementation priorities

Scale (or resolution) of the reference data components is a complex issue with huge impact on the costs and timeframe of the INSPIRE implementation. It could not be adequately covered within the scope of the RDM WG nor within this position paper. Nevertheless, some initial ideas are proposed below that the group feels can contribute as preliminary input for future work (daughter legislations).

1. "***Resolution***" (or granularity) would be a more appropriate term to use in describing reference data than "***scale***", which relates more to the representation (on screen, for example, or hard copy) of the data. However it is recognised that the word 'scale' is still of preferable usage for some users, due to its long history. The group therefore proposes the following equivalences[14]:

[14] Time and vertical resolution is not addressed here.

Geographical level	Resolution range	Scale level	Scale range
European	> 100 m	Small scale	< 1:250.000
National	~ 25 m	Medium scale	1:100.000 ~ 1:250.000
Regional	~ 10 m	Medium scale	1:25.000 ~ 1:50.000
Local	< 2.5 m	Large scale	> 1:25.000

Please note that these ranges are only provided as an indication; further discussion will refine these parameters during the INSPIRE implementation phase.

2. Following the spirit of the first INSPIRE principle (*Data should be* ***collected once*** *and* ***maintained at the level where this can be done most effectively***), the primary reference data components should be collected and maintained at the ***highest level of accuracy, i.e. at the local level*** (metric and sub-metric resolution).

   However it must be noted that the requirements may differ for different parts of the territory of Europe, e.g., dense urban areas requiring very large-scale data (decimetric resolution), while scarcely populated areas and wasteland would not require more than medium scale data.

   Moreover, some components may require different accuracy according to the topography of the area, e.g., elevation data needs higher resolution in flat and floodable zones than in mountainous areas.

   Finally, specific applications may require higher resolution for some of the components, e.g., on-board vehicle security systems would need high resolution (horizontal, vertical, timely) data on the rail and road networks.

   Further study is necessary to identify the needs, and the reference data specifications will need to accommodate for the necessary ***flexibility***, or ***variable resolutions***.

3. There are needs for harmonised and uniform reference data all over Europe, typically at small and very small scale. Today's technology does not allow for a fully automatic generalisation of complex data from large (or very large) scales.

   Therefore the four levels mentioned above must be addressed in parallel. Moreover the third INSPIRE funding principle ('**it** ***should be possible for information collected at one level to be shared between all the different levels***, detailed for detailed investigations, general for strategic purposes) requires that mechanisms must be defined and implemented to allow the ***flow of the update information collected at the highest resolution*** (typically at the local level) and to allow the ***consistent maintenance*** of the lower scale reference data sets. These are clearly complex issues requiring further work and definition.

4. The points developed above advocate that future INSPIRE activities should focus on large scale reference data. However it is crucial to consider the current state of play in Europe, and of the huge costs and long timeframe necessary to implement a pan-European large scale reference data infrastructure.

   Therefore, a pragmatic approach is to focus on an INSPIRE framework implementation for large-scale reference data, that comprises the architecture, the standards, the specifications and all the processes and mechanisms necessary for the infrastructure to operate.

   Creating the content (i.e. the data) will obviously need a long term perspective, and should realistically be approached on a step by step basis, considering the current country by country status and on the production capacity of each country.

   An indicative time frame may be proposed as follows:

   - 3 years: small scale and metadata
   - 6 years: medium scale
   - 10 years: large scale

It is important to note that the short term development timetable (small-medium scales) and the long term development timetable (large scale) will have to start simultaneously and run in parallel. Transition mechanisms between the two tracks will need to be provided in due course.

The RDM group recommends:

- *That the primary reference data components should be collected and maintained at the largest possible scale, generally at the local level.*
- *To carry out a complementary study to identify the resolution (horizontal, vertical, time) requirements for reference data that may vary according to feature/component and geographic areas.*
- *To define and implement mechanisms that allow the update information to flow from the local to the European level of the reference data.*
- *To focus the INSPIRE framework implementation (comprising architecture, standards, specifications, processes) on large-scale reference data components, but approach the content implementation (reference data sets) in a pragmatic and step-by-step way, based on the current state of play.*
- *To adopt a two-track approach to be started simultaneously and run in parallel.*

6. Description of reference data components

Component	*2^{nd} level*	*Definition/Description (1)*	*Comments*
1. Geodetic reference data		**Geodetic control points** Includes: marker id, access information, coordinates **Levelling benchmarks** Includes: marker id, access information, coordinates **Permanent satellite observation stations** Includes: marker id, access information, coordinates, Also raw or processed observations **Tide gauges** Includes: id, access information, coordinates Also raw or processed observations **System definition and transformation data** Horizontal and vertical systems standard id. plus all necessary parameters and algorithms that defines the systems Parameters and algorithms that allow transformation to and from the European Geodetic reference systems if required Geoid definition that allows transformation between physical heights and ellipsoidal heights	Ideally the coordinates should be available in the European Geodetic reference systems (i.e. ETRS89 and EUVN2000). If other system(s) are widely use nationally or locally, coordinates must also be available in these systems Precision of the coordinates and of the transformations must also be known It is essential that all coordinates are associated with an unambiguous and perfectly defined reference system, therefore using well-known standards. This also applies for the coordinates of all the other components of reference data.

2. Units of administration		**Each national territory is divided into administrative units. The administrative units are divided by administrative boundaries.** **The administrative division forms an indirect spatial reference system. The reference to an administrative unit provides a spatial dimension to data without using coordinates.**	On the national level , data sets of administrative boundaries are available in most European countries. The national data sets differ with respect to resolution, data model and geometry of international boundaries. Harmonisation can be achieved in the mid term. Consistency of administrative boundaries with other topographic reference data is a difficult issue. It can be solved only in the long term. Administrative boundaries are the key to horizontal interoperability between the products of national data custodians. Neighbours should agree on international boundaries with shared geometry at the best possible resolution. The semantic models of other units for territorial management such as the statistical division, the post code division and the areas of restricted use (e.g. national parks) equal the model of the administrative units. The concept of the reference data component might be expanded to cover the other units of territorial division.
3. Units of property rights	**Parcels**	**A parcel is a piece of land with defined boundaries, on which a property right of an individual person or a legal entity applies.**	Parcels, as the fundamental features of the cadastre (or land administration system), give reliable and complete information of the legal situation of land by providing - basic information for planning institutions, for economic development, for transparency of administration activities, - information for taxation, - a basis for planning and real estate regulations, - a proof for the scope of any kind of rights on real properties. In this context parcels (as part of the cadastre or land administration system) are only applicable at the local level.

			The data is already available in several MS. The costs for data collection are very high and time consuming. To build up a parcel cadastre takes at least 10 years. It is recommended to have this kind of data available in the long term.
3. Units of property rights	**Buildings**	**A building is a covered facility, usable for the protection of humans, animals, things or the production of economic goods.**	A building is a key element to define a property. This element is requested at the local and regional level for many property related applications. Scale(s)/resolution(s) are different with respect to the source (cadastre/national mapping). Therefore it is recommended to have this kind of data available in the medium term.
4. Addresses		**An address is the local or officially determined designation of the position of buildings and/or parcels, which consists of a defined (unique) georeferenced location. This unique location is generally realised through the postal address (house number, street and city) and is related to coordinates.**	The address is the fundamental navigation instrument to find a location. It could be used to connect information of other non geometrical data sets, e.g. owners, land value, taxation. Addresses will be very important for future Location Based Services (LBS) applications. Therefore it is recommended to have this kind of data available in the medium term.
5. Selected topographic themes	**Hydrography**	**Hydrography data include surface water features such as lakes and ponds, streams and rivers, canals, oceans and shorelines. Each of these features has the attributes of a name and feature identification code.**	The hydrological components should constitute an integrated water network.
5. Selected topographic themes	**Transport**	**The transport component should comprise an integrated transport network, and related features, that are seamless within each**	Transportation data includes topographic features related to transport by road, rail, water, and air. It is important that the features form networks where appropriate, and that links

		national border.	between different networks are established i.e. multi-modal nodes, especially at the local level, in order to satisfy the requirements for intelligent transport systems such as location based services (LBS) and telematics. The transport network should also reflect the transport flow to enable our navigation services. With GALILEO now under development, future investment in transport reference data should aspire to 1m accuracy.
5. Selected topographic themes	**Height**	**Height data should be available in two forms; contours and Digital Elevation Models (DEM s).**	Contour data – showing heights by isolines, and including within the same data set spot heights, high and low water lines, breaklines, and bathymetry. DEM data - showing spot heights in a regular grid. DEMs can be of two types DTM and DSM (see glossary). European, national, regional and local level grids should be DTMs but the pan-European could be DSM.
6. Orthoimagery		**Ortho-imagery** ▪ **is airborne or spaceborne image data of the surface of the earth** ▪ **is rectified to fit to a defined coordinate reference and cartographic projection system at a defined accuracy** ▪ **should be presented in digital format at a defined pixel resolution** ▪ **should be acquired by optical sensors with different spectral characteristics, i.e. panchromatic, true-colour, infrared** ▪ **can be used to extract reference data components** **should allow for multi-temporal analysis, implying the supply of images with different acquisition dates**	Apart from the potential use as background layer, ortho-imagery is playing an important role in change detection and updating, both of reference data components and thematic information components. Remote sensing imagery is available at different resolutions: with low and medium resolution to be used in applications at regional, national and European level and very high resolution being generally more adapted for local use. Remote sensing imagery is available with different spectral bands (including e.g. infrared), each of which may have its importance for specific environmental application fields. The remote sensing requirements on resolution, accuracy, and spectral characteristics - specifically for environmental monitoring users - are currently tackled specifically by the GMES initiative. This activity might contribute for example to the differentiation between common reference layer, common thematic layer or thematic layer.

			There is in all the Member States a strong operational use of digital aerial photography and intense activities in the production of orthophotographs. Regarding orthoimagery, there are no significant problems of data availability but it is suggested to pay particular attention to data policy and data pricing issues.
7. Geographical names		**Geographical names – definition:** **Geographical names describe features on Earth. Often the term topographical name (or toponym) is used to emphasize the spatial dependency and relation to the adjacent topographical features.** Geographical names can be associated to different kind of spatial features: **Areal features** (e.g. geographical regions, lakes, forests…). **Linear features** (e.g. rivers, railways, shipping lines, boundary lines…). point **features** (e.g. spot heights, monuments, villages, buildings…). **Gazetteer – definition:** According to the definition in ISO19112 a gazetteer provides a master record of all location instances for a particular location type or types. Therefore gazetteers are not just geographical names' indexes but may be records of any kind of feature type or types. The positional information may include a coordinate reference, but it may be purely descriptive.	The gazetteer shall as a minimum include all the names that are part of the reference data.

7. Metadata

7.1 Metadata needs and standards

Metadata are the information and documentation, which makes data understandable and shareable for users over time.

Metadata shall be used to **discover**, **access** and **use** the related data.

Metadata shall inform users on the existence of data. Users shall be able to understand the content of the data. This includes information on the spatial reference system and the spatial representation of the data. Users shall be able to assess the fitness-for-use of the data for their own purpose. Metadata shall contain information on the distribution of the data. Metadata shall contain information on legal or security constraints that affect the use of the data. Metadata should include information about planned, and frequency of, updates.

Metadata shall be produced for all data that will be made accessible within the future legislation.

The content of the metadata shall be sufficient to satisfy the needs stated in the previous paragraphs.

Electronic searching and exchange of metadata require standardisation. Metadata shall follow the ISO 19115 standard for metadata.

7.2 Metadata profiles

The member states and the European Union shall develop a common metadata profile that satisfies the needs of the objectives of the legislation. The metadata profile shall follow the guidelines in ISO 19115 for creating metadata profiles.

The metadata profile shall include a model for metadata and define common methods and formats for metadata exchange. The metadata profile shall be applicable to data sets and in addition to other appropriate levels of the data hierarchy.

The profile shall include the core elements and additional elements that are identified as necessary. The profile shall be mandatory for the participants in the infrastructure.

7.3. Metadata implementation

Metadata shall be kept up-to-date. Whenever data changes occur that might affect current metadata content, metadata has to be updated as well.

The member states shall identify a competent authority for co-ordinating the national producers of data, for collecting and for managing the metadata.

The metadata profile shall cover multilingual aspects. Codelists shall be defined in all official languages of the European Union. A thesaurus shall be generated to define the relationship between corresponding names in the different languages. Also text presentation should be possible in all European languages.

The metadata profile should be developed together with the competent authority

The agreed metadata profile shall be implemented within a geographic data service (clearinghouse) on a wide area network. The member states shall allow access to metadata via catalogues. There shall be a direct link between metadata and the described data.

The metadata profile shall be reviewed in regular time spans and if necessary adapted to new needs or developments in the GIS field.

Recommendations of the RDM group:

- *All the reference data should be documented by metadata. Metadata should be kept up-to-date by a component of authority - to be identified.*
- *The three aspects of metadata must be considered: discovery, access and use.*
- *A metadata profile compatible with ISO 19115 must be developed. It will become mandatory inside the INSPIRE infrastructure.* ***It is therefore recommended to carry out a specific study on this issue with the involvement of all the stakeholders.***
- *Metadata shall be produced for all data that will be made accessible within the future legislation.*
- *Metadata shall be kept up-to-date. Whenever data changes occur that might affect current metadata content, metadata has to be updated as well.*
- *The Member states shall identify a competent authority for coordinating the national producers of data and for managing the metadata information systems.*
- *That priority should be given to create a one stop Internet "European GI" portal for discovering and accessing GI data – similar to the ESMI / La Clef concepts, and backed by appropriate funding and legislation. This would require further study to establish resource requirements.*

8. Inter-relationships with other position papers

8.1 ETC - Environmental thematic co-ordination

A new version of the ETC position paper became available on September 15, 2002.

Due to the short deadline of September 27 for the submission of the modified position paper, it was impossible to organise a new physical meeting of the RDM group.

The impact of its content on the RDM position paper could not be estimated in due time.

An important point is that the new paper gives clear indications on the priorities for what concerns reference data:

<u>Priority 1 data:</u>

1) Geodetic Reference System
2) Official Administrative Units
3) Addresses
4) Hydrography
5) Orthophotos
6) Elevation, Bathymetry, Coastline
7) Transport networks

A European GRID system is also mentioned as a key reference data element. The majority of the RDM group members believe **that this is not data**. Nevertheless, it has already been stressed in previous versions of the RDM position paper that the use of multiple territorial breakdowns and zoning systems for data collection in the field of environment (catchment areas, NATURA 2000 areas, coastal zones, etc.) could lead to major difficulties to compare and analyse data at a later stage.

This is a valid argument to consider the definition of a hierarchy of grids at European level. However, this will require further study because it is a complex issue.

A lot of statistical data, for example, are collected on the basis of administrative units. Their transfer into a grid system could require data disaggregation. On the one hand, Eurostat made already a couple of studies in this field showing that there is no trivial solution to this problem. For each variable considered, a different method of disaggregation might be necessary.On the other hand, scandinavian countries based their systems on registers, addresses and point data. It is therefore easy for them to transfer this data into a grid system or any other territorial breakdown.

8.2 LDP - Legal issues and data policies

We endorse the basic principles of INSPIRE: "*Geographic information needed for good governance at all levels should be abundant and widely available under conditions that do not refrain its extensive use and we stress that the conditions to use the data should be simple and harmonised".*

8.3 AST - Architecture and standards

All the international standards to be used in the context of INSPIRE should be identified as well as guidelines for their use.

8.4 IAS - Impact analysis

The IAS document should contain a complete description of the methodology proposed to carry out the impact analysis.

8.5 ISF - Implementing Structure and Funding

ISF should define the mandates of the bodies to be responsible for the management of the INSPIRE infrastructure at all the levels.

9. References

[1] The common vision for INSPIRE

[2] The SDI Cookbook – GSDI (Global Spatial Data Infrastructure) Version 1.1-15 May 2001

[3] The ETEMII white paper version 6.2.2.

[4] Geographical Information Systems - Second Edition vol. 1 & 2 by Longley, P.A., Goodchild, M.F., Maguire, D.J., & Rhind, D.W.

10. Glossary

A

Accuracy
Closeness of agreement between a test result and the accepted reference value. Accuracy is the inverse of error. Errors can further be subdivided into spatial, temporal, and thematic error for a particular entity as, respectively, the discrepancies in the encoded spatial, temporal, and thematic attribute values. Accuracy is a relative measure rather than an absolute one, since it depends on the intended form and content of the database.

Altitude
Elevation above or below a reference surface.

Analysis
The process of identifying a question or issue to be addressed, modelling the issue, investigating model results, interpreting the results, and possibly making a recommendation. See also model and spatial analysis.

Area
A generic term for a bounded, continuous, two-dimensional object that may or may not include its boundary.

Attribute
A defined characteristic of an entity type (e.g., composition).

Attribute value
A specific quality or quantity assigned to an attribute (e.g., steel), for a specific entity instance.

Azimuth
A horizontal direction measured as the angular distance between true north and an object. Often used to define an oblique cylindrical projection or the angle of a geodesic between two points. A compass direction.

B

Background (layer)
Display of an orthoimage in the background of other spatial data providing an information on the spatial context.

Band
A measure of a characteristic or quality of the features observed in a raster. Some rasters have a single band; others have more than one. For example, satellite imagery commonly has multiple bands representing different wavelengths of energy along the electromagnetic spectrum.

Bathymetry
Data set or map describing depths of water bodies.

Breakline
A line that marks a change in the shape of a surface.

C

Class
Description of a set of objects that share the same attributes, operations, methods, relationships, and semantics.

Classification
The process of sorting or arranging attribute values into groups or categories.

Clearinghouse
A central agency for the collection, classification, and distribution especially of information; broadly: an informal channel for distributing information or assistance.

Completeness
'Data completeness' is a measurable error of omission and commission observed between the database and the specification. 'Model completeness' refers to the agreement between the database specification and the abstract universe that is required for a particular database application.
'Attribute completeness' is the degree to which all relevant attributes of a feature have been encoded.
'Value completeness' refers to the degree to which values are present for all attributes.

Conceptual model
Structural set of concepts established according to the relations between them, each component being determined by its position in this set.

Conformal projection
A projection on which all angles at each point are preserved. Also called an orthomorphic projection (Snyder and Voxland, 1989).

Conformance testing
Testing of a candidate product for required characteristics to determine the extent to which the product is a conforming implementation (which satisfies the conformance requirements, consistent with the capabilities stated in ICS - ISO 19105)

Consistency
Consistency refers to the absence of apparent contradictions in a database. For spatial data, the term is used primarily to specify conformance with certain topological rules.
Elimination of topological inconsistencies is usually a prerequisite for GIS processing.

Contour
A line that connects points of equal value on a terrain surface (an isoline).

Control points
Points you establish on a paper map whose coordinates represent known ground points or specific locations. Control points are used to register a paper map before you begin digitizing features on it using a digitizer.

Coordinates
Pairs of numbers (abscissa and ordinate) expressing horizontal distances along orthogonal axes; alternatively, triplets of numbers measuring horizontal and vertical distances.

D

Data
A collection of related facts usually arranged in a particular format and gathered for a particular purpose.

Database
A collection of related data organized for efficient retrieval of information. A logical collection of interrelated information managed and stored as a unit, usually on some form of mass storage system such as magnetic tape or disk. A GIS database includes data about the spatial location and shape of geographic features recorded as points, lines, areas, pixels, grid cells, or TINs, as well as their attributes.

Data element
A logically primitive item of data.

Data layer
See layer

Data model
An abstraction of the real world which incorporates only those properties thought to be relevant to the application at hand. The data model would normally define specific groups of entities, and their attributes and the relationships between these entities. A data model is independent of a computer system and its associated data structures. A map is one example of an analogue data model.

Dataset
Any set of data which has a common theme or similar attributes.

Datum
A model of the earth's shape used for Geodetic calculations.

DBMS
Database management system. A set of computer programs for organizing the information in a database. A DBMS supports the structuring of the database in a standard format and provides tools for data input, verification, and storage

DEM (Digital elevation model)
A generic term describing a digital representation of a topographic surface. The surface elevation values can be represented in various forms e.g. by contours / spot heights / breaklines, a regular grid, TIN etc. A DEM may also include surface features such as buildings, vegetation etc..

Digital image
A two-dimensional array of regularly spaced picture elements (pixels) constituting a picture, typically produced by an optical or electronic device. Common examples include remotely sensed data (for example, satellite data), scanned data, and photographs. An image is stored as a raster dataset of binary or integer values that represent the intensity of reflected light, heat, or other range of values on the electromagnetic spectrum. An image may contain one or more bands.

DSM (Digital Surface Model)
A digital surface model representing the upper surface, including buildings, woodland etc..

DTM (Digital Terrain Model(ling))
A DEM primarily defining the ground surface. This will normally exclude ground features such as buildings, woodland etc..

Documentation
Text in an item's metadata describing where the data came from, attribute definitions, and so on.

Domain

In the definition of the elements in the metadata standard, the domain identifies valid values for a data element.

E

Elevation
conforming to Federal Information Processing Standard 70-1, the term "altitude" is used in this standard, rather than the common term elevation.

Ellipsoid
When used to represent the earth, the three-dimensional shape obtained by rotating an ellipse about its minor axis. This is an oblate ellipsoid of revolution, also called a spheroid.

Entity
A real world object that cannot be further subdivided into similar objects, for example a road, or a buidling.

ETRS89
European Terrestrial Reference System 1989

EVRF2000
European Vertical Reference Frame 2000

F

Feature
A set of points, lines or polygons in a spatial database that represent a real-world entity.

Fitness-for-use
The suitability of a data set for a specific purpose or project.

G

Gazetteer
A work of geographic reference that supplies place name and location information. When a place name is known, a gazetteer can provide the co-ordinates of the place. Most atlases contain gazetteers.

Geographic coordinates
A measurement of a location on the earth's surface expressed in degrees of latitude and longitude.

Geographic data
The locations and descriptions of geographic features. The composite of spatial data and descriptive data.

Georeferencing
Process of determining the relation between the position of data in the instrument coordinate system and its geographic or map location.

GIS
Geographic information system. An organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information.
A GIS may be used for a project (also called project GIS, or single-user GIS), by a department of an organization to support a key function of that department (called departmental GIS), or by an organization to support daily activities and strategic decision making (called enterprise GIS).

Grid
A geographic representation of the world as an array of equally sized square cells arranged in rows and columns. Each grid cell is referenced by its geographic x,y location.

H

Horizontal
Tangent to the geoid or parallel to a plane that is tangent to the geoid.

I

Image
See Digital image .

Index
A special data structure used in a database to speed searching for records in tables or spatial features in geographic datasets.

INSPIRE
Infrastructure for Spatial Information in Europe.

Interoperability
The ability of two or more systems or components to exchange information and to use the information that has been exchanged [IEEE 90].

ISO
International Standards Organization. A worldwide federation of national standards bodies (for example, ANSI from the United States). ISO maintains many computing standards, including a SQL standard.

Isolines
Lines of equal value. Height isolines are often called contours.

L

Lambert Azimuthal Equal Area Projection
An azimuthal projection that sacrifices shape and distance, but preserves area. Useful for comparing features in which area is important, such as population densities. Often used for polar projections because it is originated in a centre point.

Lambert Conic Conformal
A projection of the earth's surface on a tangent cone, normally based on two standard parallels. A projection is called conformal, when the scale of a map at any point on the map is the same in any direction. Meridians (lines of longitude) and parallels (lines of latitude) intersect at right angles. Shape is preserved locally on conformal maps.

Layer
A collection of similar geographic features—such as rivers, lakes, counties, or cities—in a particular area or place referenced together for display on a map.

Line
A set of ordered co-ordinates that represent the shape of geographic features too narrow to be displayed as an area at the given scale (contours, street centrelines, or streams), or linear features with no area (county boundary lines).

Location
Identifiable part (place) of the real world.

M

Map
A graphical representation of geographic information. It includes geographic data and other elements such as a title, north arrow, legend, and scale bar.

Map projection

See projection

Map scale
See scale

Metadata
Data about the content, quality, condition, and other characteristics of data.

Model
An abstraction of reality used to represent objects, processes, or events. A set of clearly defined analytical procedures used to derive new information.
A data representation of reality (for example, vector data model, TIN data model, raster data model).

N

Network
An interconnected set of arcs or lines representing possible paths for the movement of resources from one location to another.

Node
1. The beginning and ending locations of a line. A node is topologically linked to all lines that meet at the node.
2. A zero-dimensional object that is a topological junction of two or more links, or an end point of a link in a network.

O

Object
The representation of a real-world entity. An object has properties and behaviour and relationships to other objects.

OpenGIS Consortium - OGC
OpenGIS is defined as transparent access to heterogeneous geodata and geoprocessing resources in a networked environment. The goal is to provide a comprehensive suite of open interface specifications that enable developers to write interoperational components that provide these capabilities.

Orthoimagery
Airborne or spaceborne image data of the surface of the earth having the properties of an orthographic projection. It is derived from a conventional perspective image by simple of differential rectification so that image displacements caused by camera tilt and terrain relief are removed.

P

Phenomenon
A fact, occurrence or circumstance.

Pixel
Picture element. A raster cell in images.

Point
A zero-dimensional abstraction of an object represented by a single X,Y co-ordinate. A point normally represents a geographic feature too small to be displayed as a line or area; for example, the location of a building location on a small-scale map, or the location of a service cover on a medium scale map.

Polygon
A two-dimensional feature representing an area such as a state or county. A polygon is defined by the lines that make up its boundary.

Position
Spatial reference of a point or an object.

Positional accuracy
Spatial or positional accuracy refers to the accuracy of the spatial component of a database. For points: error is usually defined as the discrepancy (Euclidean distance) between the encoded location and the location as defined in the specifications. The most common measures are horizontal errors (distance measured in x and y simultaneously) and vertical error (distance measured in z). For lines and areas. The problem is more complex since there is no simple statistical measure of error that can be adopted from statistics.

Precision
The exactness with which a value is expressed, whether the value be right or wrong.

Projection
A mathematical model that transforms the locations of features on the Earth's surface to locations on a two-dimensional surface. Because the Earth is three-dimensional, some method must be used to depict a map in two dimensions. Some projections preserve shape; others preserve accuracy of area, distance, or direction.

Q

Quality
An essential or distinguishing characteristic necessary for cartographic data to be fit for use.

R

Raster data
Represents any data source that uses a grid structure to store geographic information. See grid and image.

Reference system
A method for identifying positions on the globe. This is often constructed with a grid that either refers to earth's latitude and longitude (graticule), or a planar equivalent that divides grid lines by a fixed length from a predefined point of origin.

Remotely sensed data/Remote sensing:
Data measured and mapped using indirect methods exploiting their properties. For example mapping vegetation measuring temperature differences.

Resolution
Resolution expresses the size of the smallest object in a spatial dataset, that can be described. It refers to the amount of detail that can be discerned. It is also known as granularity. Resolution is also limited because geo-spatial databases are intentionally generalised. Resolution affects the degree to which a database is suitable for a specific application.
In a raster data model, resolution expresses the size of the raster cell in the real world.

S

Scale
The relation between the dimensions of features on a map and the geographic objects they represent on the earth, commonly expressed as a fraction or a ratio. A map scale of 1/100,000 or 1:100,000 means that one unit of measure on the map equals 100,000 of the same unit on the earth.

Sensor
A device used to map remotely sensed data.

Source
A component of an extended element. The name of the individual or organization creating an extended element.

Spatial accuracy
See positional accuracy.

Spatial analysis
The study of the locations and shapes of geographic features and the relationships between them. The process of modelling, examining, and interpreting model results. Spatial analysis is useful for evaluating suitability and capability, for estimating and predicting, and for interpreting and understanding. There are four traditional types of spatial analysis: topological overlay and contiguity analysis; surface analysis; linear analysis; and raster analysis.

Spatial data
Information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth. This information may be derived from, among other things, remote sensing, mapping, and surveying technologies.

Spatial modelling
Analytical procedures applied with a GIS. Three categories of spatial modelling functions can be applied to geographic features within a GIS: geometric models (calculating the Euclidean distance between features, generating buffers, calculating areas and perimeters); coincidence models (topological overlay); adjacency models (pathfinding, redistricting, and allocation).
All three model categories support operations on spatial data, including points, lines, polygons, TINs, and grids. Functions are organised in a sequence of steps to derive the desired information for analysis.

Spatial resolution
In the field of remote sensing, it is defined in terms of the ground dimensions of the picture elements, or pixels, making up a digital image. For vector data, the smallest feature that can be discerned is usually defined in terms of rules for minimum mapping unit size which depends on the map scale.

Spot heights
Descrete heights of a surface.

Surface
A set of continuous data such as elevation or air temperature over an area, or the boundary between two distinct materials or processes.

T

Temporal accuracy
Temporal accuracy refers to the agreement between encoded and 'actual' temporal coordinates. Time is often not dealt with explicitly in spatial databases. Temporal information is often omitted, except in databases designed for explicitly historical purposes.

Temporal resolution
Temporal resolution refers to the minimum duration of an event that is discernible. It is affected by the interaction between the duration of the recording interval and the rate of change in the event.

Thematic accuracy or attribute accuracy
For quantitative attributes, metrics are similar to those used to measure spatial accuracy for point features. For categorical data, most of the research into data quality has come from the field of classification accuracy assessment in remote sensing.

Thematic resolution
In the thematic domain, the meaning of resolution depends on the measurement scale. For quantitative data, resolution is determined by the precision of the measurement device. For categorical data, resolution is defined in terms of the fineness of category definitions.

The European level
In the context of INSPIRE is spatial phenomena that affect the whole continent.

The national level
In the context of INSPIRE is spatial phenomena that affect a nation or some neighboring nations.

The local level
In the context of INSPIRE is spatial phenomena that affect only local communities.

The regional level
In the context of INSPIRE is spatial phenomena that affect only regional areas.

TIN
Triangulated irregular network. A surface representation derived from irregularly spaced sample points and breaklines. The TIN data set includes topological relationships between points and their neighbouring triangles. Each sample point has an x,y coordinate and a surface, or z-value. These points are connected by edges to form a set of non-overlapping triangles used to represent the surface. TINs are also called irregular triangular mesh or irregular triangular surface model.

Topographical name
A name in a gazetteer that doesn't correspond to a specific spatial phenomenon . Mount Alps, Central European are examples of a Topographical Name.

Topology
Properties of geometric forms that remain invariant when the forms are deformed or transformed by bending, stretching, and shrinking. Among the topological properties of concern in GIS are connectivity, order, and neighbourhood.
Topology is useful in GIS because many spatial modelling operations don't require co-ordinates, only topological information. For example, to find an optimal path between two points requires a list of the lines or arcs that connect to each other and the cost to traverse each line in each direction. Co-ordinates are only needed for drawing the path after it is calculated.

Toponym (see also Topographical Names)

Transverse Mercator coordinate reference systems
Transverse Mercator projections result from projecting the sphere onto a cylinder tangent to a central meridian. Distortion of scale, distance, direction and area increase away from the central meridian. The Universal Transverse Mercator (UTM) projection is used to define horizontal, positions world-wide by dividing the surface of the Earth into 6 degree zones, each mapped by the Transverse Mercator projection with a central meridian in the centre of the zone. UTM zone numbers designate 6 degree longitudinal strips extending from 80 degrees South latitude to 84 degrees North latitude

V

Vector or Vector Data
A coordinate-based data structure commonly used to represent linear geographic features. Each linear feature is represented as an ordered list of vertices. Traditional vector data structures include double-digitized polygons and arc–node models.

Vertical
At right angles to the horizontal; includes altitude and depth.

W

Web
(see WWW) World Wide Web

1. Annex

Contributors:

DENMARK
Mr. HANSEN, Henning Sten
National Environmental Research Institute
Department of Policy Analysis
Frederiksborgvej 399
4000 Roskilde
e-mail: HSH@DMU.DK

GERMANY
Dr. ILLERT, Andreas
Federal Agency for Cartography and Geodesy
Bundesamt für Kartographie und Geodäsie
Richard-Strauss-Allee 11
D - 60598 Frankfurt/M.
e-mail: illert@ifag.de

Dr. RIECKEN, Jens
Landesvermessungsamt Nordrheinwestfalen
Muffendorfer Str. 12-21
53117 Bonn
e-mail: riecken@lverma.nrw.de

GREECE
Dr. DOGANIS, Thanos
Hellenic Ministry for the Environment, Physical Planning and Public Works
Terra Ltd.
42, Tsitouri Str.
GR - 15231 Chalandri
e-mail: Thanos@terra.gr

SPAIN
RODRÍGUEZ PASCUAL, Antonio
Instituto Geográfico Nacional
C/General Ibáñez de Ibero, 3
28003 Madrid
e-mail: Afrodriguez@mfom.es

FRANCE/CNIG
SALGE, François
Secrétaire Général du CNIG
CNIG - Conseil National de l'Information Géographique
136, bis rue de Grenelle
75700 Paris
e-mail: Francois.salge@cnig.fr

GEIGER, Patrice
CNIG - Conseil National de l'Information Géographique
136, bis rue de Grenelle
75700 Paris
e-mail: Patrice.geiger@cnig.fr

ITALY
SAMBUCINI, Valter
Ministry of Environment & Land Protection

AUSTRIA
Dipl.-Ing. GROHSNEGGER, Valentin
Bundesamt für Eich- u. Vermessungswesen
Schiffamtsgasse 1-3
1025 Wien
e-mail: Valentin.grohsnegger@bev.gv.at

LUX, Felix
Federal Environment Agency Austria
e-mail: Lux@ubavie.gv.at

PORTUGAL
COUCELO, Carlos
Istituto do Ambiente
Consultoria Geomática Lda
Taguspark,
Núcleo Central 242
P - 2780-920 Porto Salvo
e-mail: Carlos@cccgeo.com

SUOMI/FINLAND
SAARIKOSKI, Antti
Spatial Data Management
126, Opastinsilta
FIN - 00521 Helsinki
e-mail: Antti.saarikoski@nls.fi

SUCKSDORFF, Yrjö
Finnish Environment Institute/Geoinformatics and Land Use Division
P.O.Box 140
00251 Helsinki
e-mail: Yrjo.sucksdorff@ymparisto.fi

SWEDEN
ENGBERG, Agneta
National Land Survey - Lantmateriet
Lantmäterigatan 2
S - 801 82 Gävle
e-mail: Agneta.engberg@lm.se

HAGGBOM, Jannica
Swedish Environmental Protection Agency
e-mail: Jannica.haggbom@environ.se

UNITED KINGDOM
PROBERT, Mark
Ordnance Survey
Romsey Road,
Maybush,
Southampton
United Kingdom
e-mail: Mprobert@ordsvy.gov.uk

HUNGARY
PRAJCZER, Tamás
GeoX Limited
Elsõ Magyar Térinformatikai Adatáruház

e-mail: Prajczer@geox.hu

POLAND
DR. BIELECKA, Elzbieta
Institute of Geodesy and Cartography
Ul. Jasna 2/4
00-950 Warszawa
e-mail: Elab@igik.edu.pl

PROF. DR. LINSENBARTH, Adam
Institute of Geodesy and Cartography
Ul. Jasna 2/4
00-950 Warszawa
e-mail: Igik@igik.edu.pl

SLOVAK REPUBLIC
OFÚKANÝ, Miloslav
Geodetic and Cartographic Institute Bratislava
Dept. of Development and International Cooperation
Chlumeckétro 4
827 45 Bratislava
e-mail: ofukany@gku.sk

KLOBUSIAK, Matej
Geodetic and Cartographic Institute Bratislava
Dept. of Development and International Cooperation
Chlumeckétro 4
827 45 Bratislava
e-mail: Klobusiak@gku.sk

EUROGEOGRAPHICS
LUZET, Claude
6-8, av. Blaise Pascal
Cité Descartes
F - 77455 Champs-sur-Marne
Marne-La-Vallée Cedex 2
e-mail: claude.luzet@eurogeographics.org

LAND, Nick
e-mail: Nick.land@eurogeographics.org

EEA - European Environment Agency
LILLETHUN, Arvid
Kongens Nytorv 6
DK - 1050 Copenhagen-K
e-mail: arvid.lillethun@eea.eu.int

STEENMANS, Chris
e-mail: chris.steenmans@eea.eu.int

ESA - EUROPEAN SPACE AGENCY
LANDGRAF, **Günther**
ESA
Earth Observation Applications Department
Via Galileo Galilei
I - 00044 FRASCATI (Roma)
e-mail: Guenther.Landgraf@esa.int

COMMISSION OF EC-EUROSTAT
RASE, Daniel
Statistical Office of the European Communities - Eurostat
5, rue Alphonse Weicker
2721 Luxembourg-Kirchberg
e-mail: daniel.rase@cec.eu.int

BJÖRNSSON, Anette
Statistical Office of the European Communities - Eurostat
5, rue Alphonse Weicker
2721 Luxembourg-Kirchberg
e-mail: Anette.bjoernsson@cec.eu.int

WIRTHMANN, Albrecht
Statistical Office of the European Communities - Eurostat
5, rue Alphonse Weicker
2721 Luxembourg-Kirchberg
e-mail: albrecht.wirthmann@cec.eu.int

COMMISSION OF EC - ENV
VANDERHAEGEN, Marc
ENV/B03
BU-5 04/124
Avenue de Beaulieu 5
B - 1160 Bruxelles
e-mail: marc.vanderhaegen@cec.eu.int

JRC - JOINT RESEARCH CENTRE
DALLEMAND, Jean François
Joint Research Centre
Institute for Environment & Sustainability
Land Management Unit
TP 262
Via E. Fermi 1
I - 21020 Ispra (Varese)
e-mail: Jean-francois.dallemand@jrc.it