

# Sustainable Mobility Information System (SMIS)

## A tool for decision-making.

JM ROTTIERS\*, K. MAESEN\*\*, F. MEURIS\*\*, L. ALBRECHTS\*\*\*, W. HECQ\* - 1998

\* CEESE, Université Libre de Bruxelles Belgium; \*\* GfG, Katholieke Universiteit Leuven Belgium; \*\*\* ISRO, Katholieke Universiteit Leuven Belgium.

## 1. Introduction

### 1.1 Context

In 1992, the United Nations Conference for Environment and Development was held in Rio. One of the main documents adopted during this conference was Agenda 21<sup>1</sup>. This document deals with the promotion of actions to be carried out by the different actors of society in order to ensure socio-economic growth in different countries without endangering this for future generations. The last part of the document deals with practical methods of implementation of the actions to be carried out. In particular, Chapter 40, called "Information for Decision-Making"<sup>2</sup> deals with the subject that I will develop here.

The two main objectives of research to be carried out are:

- bridging the "data gap";
- improving information availability and make the information useable.

These two actions require the development of tools like :

- standardisation;
- utilisation of electronic resources;
- all other methods allowing rapid and efficient dissemination of all relevant resources.

These two research goals are the two requested conditions necessary for the construction of a powerful information system, but it must be constructed with other priorities, such as simplicity and speed of the information's research.

Following the Rio Conference, in March 1996 the Belgian federal government implemented the scientific action plan in support of Sustainable Development policies through the Office of Scientific, Technical, and Cultural Affairs of the Prime Minister (OSTC)<sup>3 4</sup>. Our research is a part of the 'Mobility' component of this program.

### 1.2 Objectives

Based on the preceding information, the goals of the work are the following:

- research all relevant information related to sustainable mobility in Belgium;
- make this information accessible and useable;

- be a model for the construction of another, larger, information system covering sustainable growth in general;
- demonstrate the importance of the interaction between usually separated actors to allow the development of efficient policies for sustainable growth.

One of the main problems with which researchers and decision-makers are confronted is the relevant information research related to their work. This problem may constitute a waste of time; the idea is to centralise the sources of data and references.

On the other hand, some information-holders are not always able to share their information, leading to a potential loss of information. Our aim is to propose a catalogue system allowing a quick location and in parallel offer the possibility to the information-holders to put the resources in their possession at others' disposal.

Once the information is collected, we have to ensure its diffusion. The emergence of the information technologies—in particular the 'Internet'—creates opportunities that we can make use of. The data-encoding must be done with all respect of standardisation rules in order to ensure the compatibility of the system with other existing or emerging systems<sup>2</sup>.

Our work serves as a model for the larger project of the Belgian databank for sustainable growth. Considering the extent of the field to be covered, indeed, it will help determine the potential problems with which the construction of such a tool will be confronted. The interest of our work lies in the more restricted character of mobility, while offering a sufficiently broad field of investigation to make it possible to consider, study and find solutions for these potential problems.

A scientific objective of the SMIS is also to demonstrate in a practical manner that close interaction between policies, disciplines and actors will be an essential element of sustainable development. A standard information system crossing the different disciplines involved in sustainable mobility (economical, institutional, sociological, environmental) is a first step to make needed bridges between usually separate disciplines and actors (researchers, administrators, and the public).

## **2. Metadata**

### **2.1 Concept of metadata**

Presently, information research, especially using electronic services, shows a huge disparity in the structure of answers made to requests. The reasons for this disparity are diverse:

- disciplines related to the research theme (sciences, engineering, human resources, economics...);
- information sources (governmental, private, academic...);
- resources types (documents, people, web sites...);
- processes for the creation of data (used standards, languages, software...);
- geographical factors (language, character sets, sociocultural context...);
- ...

All of these factors influence the information research processes and the quality of the results shown at the end of the request. Regarding the amount of information actually available, Agenda 21 proposes in Chapter 40.24<sup>1</sup> : "the organs and organisations of the United Nations system, as well as other governmental and other non governmental organisations, should document and share information about the sources of available information in their respective organisations"<sup>2</sup>. This process of documentation of the data sources is the creation of metadata.

The term metadata can be defined as "data about data". A metadata is made of a set of elements describing the datasource without directly giving its integrity. This way of proceeding allows the reduction of the quantity of information that the search engines will have to process during a request. The basis metadata will be a list of key words which are likely to be used in a request.

A collection of metadata constitutes a metadatabase or a metadatabank. The role of these metadatabases is to accelerate the research processes.

We have to notice the fact that a metadata is not especially electronic. Indeed, a library index is no more or less than a metadatabase. Each paper record contains a description in terms of title, author, keywords, or even a description. The consultation of these files allow the reader to quickly locate the desired book and access it.

The creation of metadata, therefore, presents a real interest in terms of facility and speed of research. But the more you reach various subjects, such as description of : databases, organisations, people, samples, documents..., the more it becomes necessary to use complex description systems. And the dimension of language and cultural identity problems can add supplementary issues to this difficulty of choice of description fields<sup>5</sup>.

So, we have to proceed to a standardisation of the metadatas. Indeed, if independently of the language dimension, the different datasources agree on a set of common description elements and adapted to the description of datasources, for example we shall use the description fields : author, title, name, address, keywords, and description to describe a scientific paper. Once we have defined the standards on the structural level and the way to encode metadata, then we can establish automatic translation or multi-lingual research systems allowing the avoidance of the language problem.

But the definition of a standard must also be made keeping in mind the limitation of the quantity of proposed information. To summarise, we could say that too much information kills the information. Indeed, if we don't limit the number of description fields, the duration of the request will grow and also the risk of misinterpretation of the contents of the fields. It will naturally lead to results without relationship to the expected results of the user. It is the paradox of the information research<sup>6</sup>.

To solve these problems, several organisations have developed metadata standards, notably:

- the HTML language (HyperText Markup Language), the encoding standard for the creation of the proposed documents on the world wide web; this standard<sup>7</sup> contains several fields which

are not automatically viewed, but which are visited by the search engine: author name, creation date, description, keywords;

- the Dublin Core<sup>8</sup>, which is a description standard with 15 elements developed to facilitate the electronic resources research;
- the Catalogue of Data Sources (CDS)<sup>9</sup>, developed by the European Environment Agency (EEA), to indicate who possesses which information in Europe, in which form, and where and how to access the information ; the CDS is also multilingual.

## **2.2 Development of the metadata standard specific to our information system**

Regarding the previous information, we have developed a metadata system adapted to our research objectives and compatible with the multilingual character of Belgium (the two main national languages are French and Dutch, and also English as a language of reference for many scientific publications) which brings a supplementary dimension to the problem.<sup>10</sup>.

Our standard was developed to be, as much as possible, compatible with other information systems having environmental character. Especially the Catalogue of Data Sources<sup>9</sup> and the Dublin Core<sup>8</sup> which have been mentioned previously, the Federal Geographic Data Committee (FGDC)<sup>11</sup> standard developed for geographical information, the one of The U.S. Environmental Protection Agency (US EPA)<sup>12</sup>, several norms of the International Organisation for Standardisation (ISO)<sup>13</sup> for the fields of date, character set, etc.

In order to be able to identify in a unique way each described resource, we have first defined the 'identifier' field. Then we defined the main types of data and datasources that we will take into account. This kind of information was labelled as "object type", the possible entries may be subdivided into sub-categories to clarify the description which are :

- Organisation : public, governmental, academic, private...
- People : expert, promoter, researcher, documentalist...
- Activity : program, project, conference...
- Product : statistical data, publication, report, software...

Then, we introduced the 'Name' field, this field taking several formats depending on the described resource type with references to several fields such as address, family name, phone, fax, URL, ... allowing the location of the resource.

Several fields must allow the user to know the state of the proposed such as: Metadata\_Status, Metadata\_Date, Metadata\_Language... Other fields allow the user to take knowledge of the access constraints (price, timetable, needed equipment, confidentiality rules, ...) to the described data.

Finally, an important field<sup>14</sup> for the research is the one called "Keywords". For the moment, the system works with a keyword set taken from the General European Multilingual Environmental Thesaurus (GEMET)<sup>15</sup> developed by the EEA. The use of a predefined thesaurus allows the

users to know the words to use in order to make efficient requests inside the information system. One of the other advantages of the GEMET is its multilingual character, which is one of our predefined needs.

The different fields used in the information system are summarised at the Figure 1.

### **3. Application of the metadata concept, construction of the information system**

#### ***3.1 Requirements of the systems and selected solutions***

Based on the previous information, some requirements must be fulfilled to be able to build a performance information system. They are the following :

- speed, simplicity and accessibility for a maximum number of users ;
- to limit the amount of data, the system must not be the information-holder, but a tool to help the quick location of the researched data ; this approach also allows us to avoid the problems of copyright by the means by which the system only signals the existence of the resource, describing it and giving the access constraints ;
- low cost with consideration to the other requirements of the project (character multi platform, non restricted accessibility in time, exemption from payment of the service...);
- multilingual character with consideration of the requirements specific for Belgium (French, Dutch and English).

The solutions adopted to satisfy these requirements are:

- use of metadata makes it possible to give to the user information necessary to the localisation of the resource;
- usage of " Internet " technologies makes it possible to ensure a maximum dispersion of the information at a moderate cost, unrestricted by time or a number of users;
- work by partnership between two research institutes belonging to two different speech communities makes it possible to ensure the multilingual character French/Dutch/English of the tool.

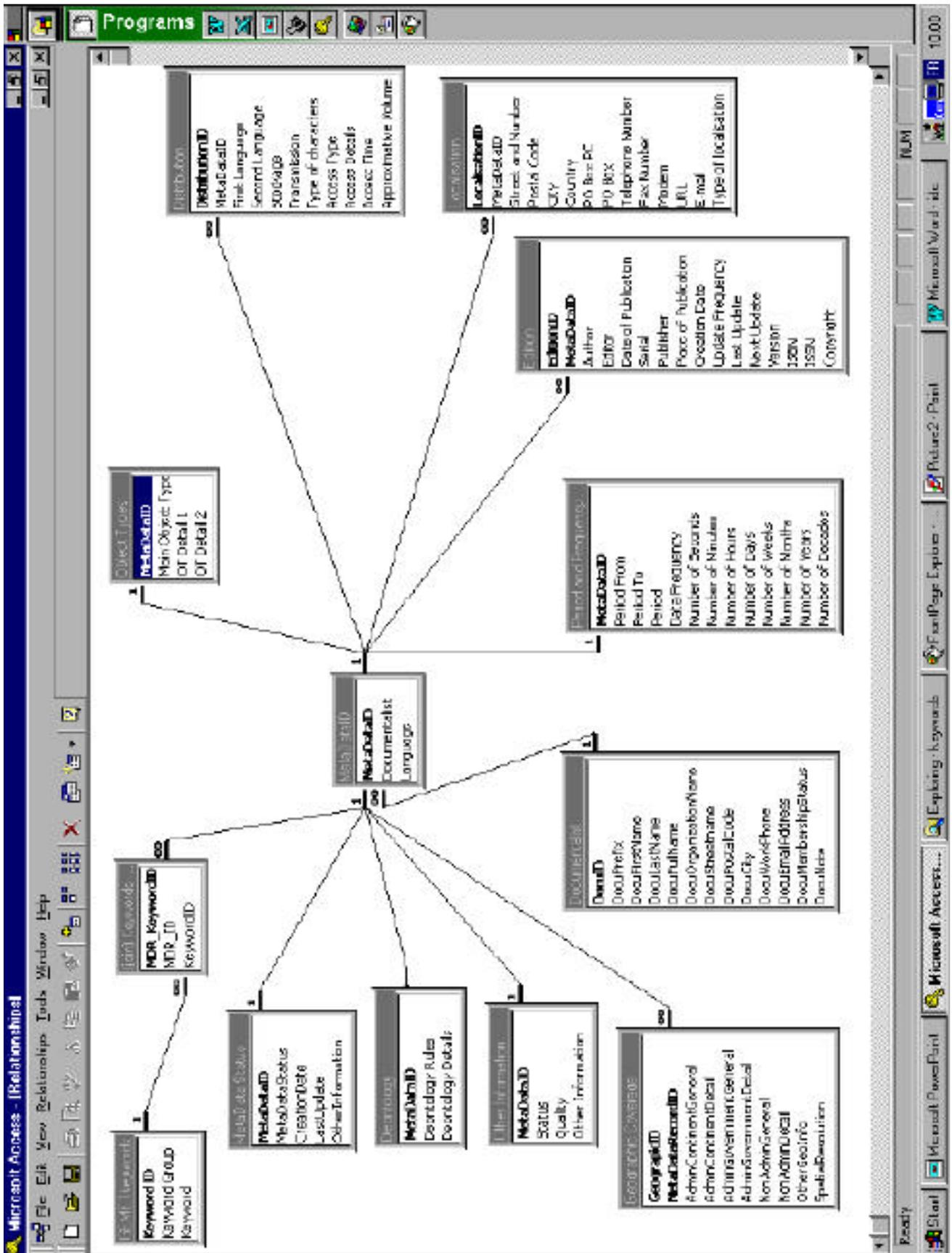
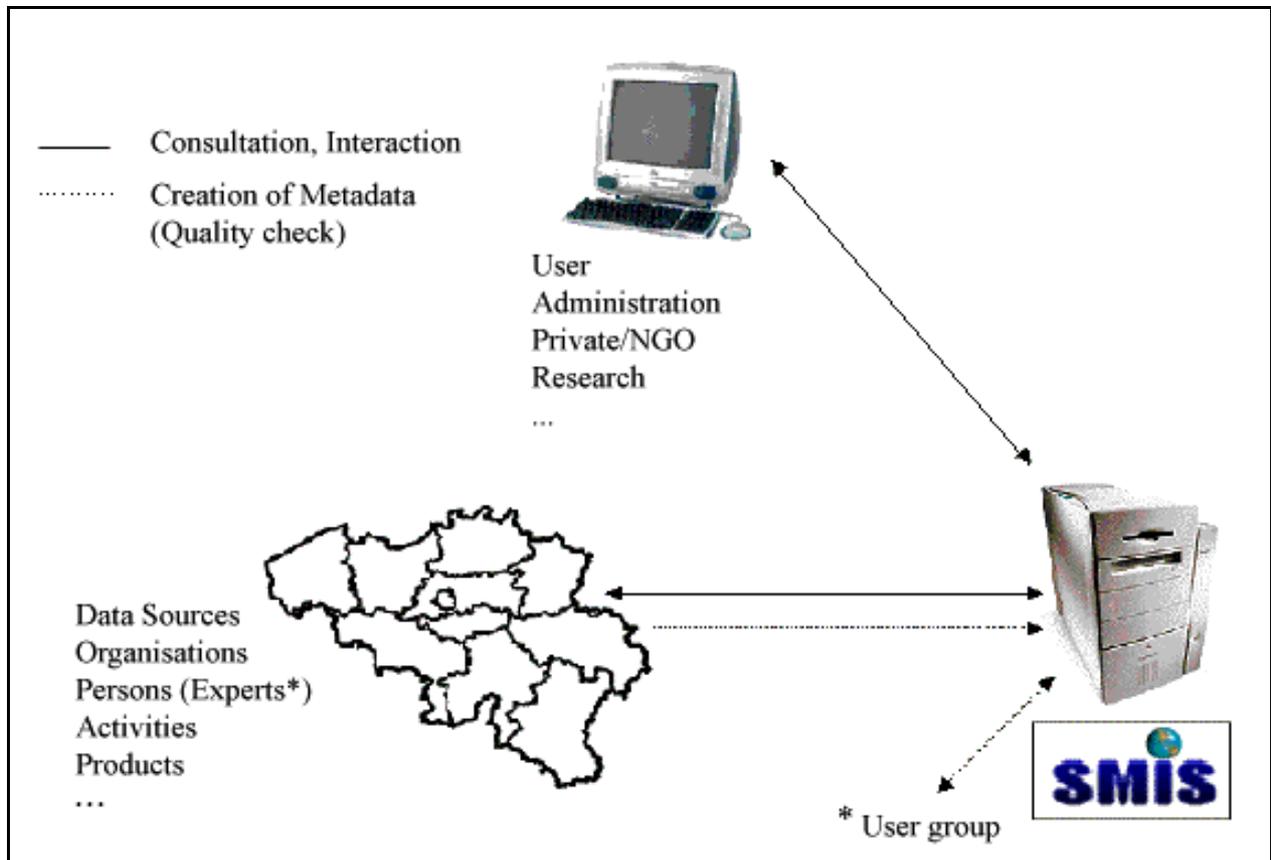


Figure 1. Different fields used in the information system

### 3.2 Development of the Sustainable Mobility Information System (SMIS)

For approaching the construction of the information system, it was necessary to define the relations between the various actors who were going to be brought together in and around the system.

Figure 2. summarises these relations.



**Figure 1. Relationships within the SMIS**

There are three types of relations:

- consultation of the metadata by the users who can be administrations, researchers, members of the civil company;
- for the creation of metadata by the datasource, the SMIS must offer at the end the possibility to the data source to create their own metadata inside the system (institution, organisation, expert...); it should be noted that certain experts have a special role to play since they belong to the "user group" made up in order to control the quality of metadata proposed in the information system;
- the possibility of interaction by the intermediary of the electronic forum of discussion developed within the system.

The partnership of which we made mention in our objectives appears here clearly. The actors involved in sustainable mobility in Belgium and usually separated, find themselves joined together in a common system of data sources where they can interact and come into contact with each other, particularly by the means of an electronic forum of discussion devoted to the research

topic. In addition the fact that the development of the system is ensured by the members of the two principal speech communities of the country is another factor of success.

### 3.2.1 Preliminary phase, setting-up of a basic tool<sup>16</sup>

Our work plan suggested that, in a preliminary phase, the SMIS was proposed in the form of a set of HTML pages connected by hypertext links. The search for metadata was initially made on the realisation of the inventory of the great data sources and metadata already existing, in particular on the level of official and academic authorities. The access to metadata was made through alphabetical thematic indexes.

This way of proceeding presented a great facility of creation of documents and a great flexibility of adaptation in the measurement where the standard of metadata was still prone to small modifications. But as the number of records was growing, the management of these indexes became increasingly burdensome and time intensive. The limit of management of such a system is estimated to be around a thousand documents. It was thus necessary to pass to management by automatic tools of the type of relational database. This type of tool has the advantage of offering the possibility of more advanced search possibilities, in particular by keyword or textual search within the information, which is, from the point of view of the user, an enormous advantage compared to the heaviness of indexes.

### 3.2.2 Phase of development, creation of the evolved site with search engine<sup>17</sup>

The quantity of currently collected metadata is about 4000 recordings which were consultable according to methods of the old version. These recordings were imported in a relational Access database, which is consultable via the Internet through a web interface especially developed for our system.

The advantages of such a system are numerous insofar as the management of the tool is not limited by the number of records and the update is quasi-automatic. Its drawback is the long time needed to develop the tool. It is for this reason that we developed the preliminary site.

Table 1. summarises the principal characteristics of the two approaches.

|                            | Preliminary Site   | Final Site  |
|----------------------------|--|---|
| Kind of support            | HTML Pages   | Relational database   |
| Maximum number of records  | 1000   | unlimited (taking into account the physical limits of the hardware) |
| Processing of the metadata | Manual or semi automatic regarding to the skill of the administrator | Automatic   |
| Update frequency           | maximum once a week  | Automatic/daily   |
| Ease of use                | Good   | Very Good   |
| Search                     | Manual, textual  | Assisted by a search engine   |

**Table 1. Comparison of the two SMIS approaches**

### **3.3 Perspectives**

In continuation of our project, different tasks will be carried out to improve the performance of the system, in particular:

- creation of mirror sites; the role of a mirror site is to propose exactly the same information as the original site but on another server; the advantage of this process is to avoid the obstruction of the site and thus to ensure good performance in term of speed of access to the metadata;
- automation of creation procedures of thematic indexes to answer the requests of certain users; the success of some search engines such as Yahoo<sup>18</sup> show the preference that some users of the Web grant to thematic indexes, which satisfy this category of user and others who can not use the search engine due to the incompatibility of their browsers with the requirement of the web interface of the relational database;
- automation of metadata collection procedure, in particular by the use of indexing "robots" such as those used by the majority of the commercial search engines; these "robots" are in fact programs which traverse the Internet and automatically catalogue the information that they are charged to find;
- procedures of assistance for metadata creation allowing the users who wish to create their own metadata and to insert it in the information system to do so; these procedures must be developed in order to guarantee the integrity of the data and to avoid the problems of ill will or the errors of handling.

The goal of our work is to build a tool which at the end will be likely to function in a quasi autonomous way and require only one weak investment in human capital. Thereafter our work will be taken over and inserted within the information system for Sustainable Development that the OSTC are developing on the basis of the experience gained during our research program.

## **4. Conclusions**

The rise of information technologies are an opportunity and a determining factor for sustainable development. The processes of decision-making can be improved thanks to the speed of access to information necessary to the setting-up of effective and favourable policies to support the objectives defined in Agenda 21. These technologies allow more transparency and better information for citizens and people concerned with the decisions being made now or in the future.

From a world point of view, it is important to carry out standardisation's in order to ensure the quality of information and the transmission speed of this information. The advantage of making standardisation's also lies in the minimisation of cultural and linguistic barriers, in particular via the development of machine translation technologies.

However, information technologies are, of course only one tool of decision-making aid for Sustainable Development and should not be perceived as the solution for the problems we try to solve.

- 
- <sup>1</sup> UNEP & DPCSD 1992, "Agenda 21", United Nations Conference on Environment and Development, Rio de Janeiro 3 - 14 June 1992, at <http://kaos.erin.gov.au/portfolio/esd/nsesd/Agenda21.html>
- <sup>2</sup> UNEP & DPCSD 1992, "Agenda 21", United Nations Conference on Environment and Development, Rio de Janeiro 3 - 14 June 1992, Summary of the chapter 40 at <http://kaos.erin.gov.au/portfolio/esd/nsesd/a21summ.html#ch40>
- <sup>3</sup> History of the SSTC, summary, at [http://www.belspo.be/belspo/ostc/histo\\_uk.htm](http://www.belspo.be/belspo/ostc/histo_uk.htm)
- <sup>4</sup> Plan d'Appui Scientifique à une politique de Développement Durable, appel à propositions, at [http://www.belspo.be/belspo/ostc/announ/pages/gc\\_fr.htm](http://www.belspo.be/belspo/ostc/announ/pages/gc_fr.htm)
- <sup>5</sup> DGXIII - E/4 Second Metadata Workshop, 26 June 1998, at <http://www2.echo.lu/libraries/en/metadata2.html>
- <sup>6</sup> B. Kestemont and W. Hecq, IATAFI 96, Information technology tools for sustainable development, at <http://www.ulb.ac.be/ceese/PAPERS/IATAFI/iatafi.htm>
- <sup>7</sup> World Wide Web consortium, HTML 4.0 Specification, W3C Recommendation, revised on 24-Apr-1998, at <http://www.w3.org/TR/PR-html40/struct/global.html#h-7.4.4>
- <sup>8</sup> Dublin Core, at [http://purl.oclc.org/metadata/dublin\\_core/](http://purl.oclc.org/metadata/dublin_core/)
- <sup>9</sup> European Environment Agency, Catalog of Data Sources, at <http://www.mu.niedersachsen.de/cds/>
- <sup>10</sup> Matthew Stiff, Multilingual Issues, in DGXIII - E/4 Second Metadata Workshop, 26 June 1998, at <http://www2.echo.lu/libraries/en/metadata2.html>
- <sup>11</sup> Federal Geographic Data Committee, FGDC, at <http://fgdc.er.usgs.gov/>
- <sup>12</sup> U.S. Environmental Protection Agency (EPA), at <http://www.lbl.gov/~olken/epaintro.html>
- <sup>13</sup> International Organization for Standardization, at <http://www.iso.ch/>
- <sup>14</sup> Paul Miller, Controlled Vocabulary, in DGXIII - E/4 Second Metadata Workshop, 26 June 1998, at <http://www2.echo.lu/libraries/en/metadata2.html>
- <sup>15</sup> General European Multilingual Environmental Thesaurus (GEMET), at [http://www.mu.niedersachsen.de/cds/objectives\\_products.html](http://www.mu.niedersachsen.de/cds/objectives_products.html)
- <sup>16</sup> URL of the website: <http://www.ulb.ac.be/ceese/SMIS/themes.htm>
- <sup>17</sup> URL of the website: <http://oder.agr.kuleuven.ac.be/smis/>
- <sup>18</sup> <http://www.yahoo.com>