A MULTILINGUAL VIRTUAL ENCYCLOPEDIA ON SUSTAINABLE DEVELOPMENT*

B. Kestemont[§], D. Le Roy[§§], P. Vanhaecke[§§] and W. Hecq[§]

[§] Center for Economic and Social Studies on the Environment, Université Libre de Bruxelles, Brussels, Belgium
[§§] Ecolas nv, Antwerpen, Belgium

ABSTRACT

A virtual (distributed) encyclopaedia on sustainable development, based on the WWW, would point to the best available and comparable data in a networked hypertext presentation, making the bridge between cultures, languages and disciplines. It would be presented from various perspectives and levels of aggregation. For each chapter or perspective, a panel of referees would guarantee the contents by a label. The hyperlinks would be internal (within the labelled chapters of the encyclopedia) as well as external (to other internet sources, points of view or candidate chapters or papers). A system of indexation would be realized through the use of common metadata headers, based on the html URCs and on most advanced environmental metadata standards, using multilingual thesaurus of general environmental descriptors, and tools (e.g. clickable maps, full text multilingual search, query by form), necessary to improve the extraction of information in various mirror servers.

1.0. INTRODUCTION

Sustainable development needs an improvement of information and data exchanges at all levels. Data sources covering a large range of aspects - environmental, economical, sociological and institutional - are needed to compile a set of reliable, relevant, generally accepted and harmonised indicators that

_____

* Presented at Eco-Informa '96, Lake Buena Vista, Florida, 4-7 November 1996.

would help the policy makers and the public lean towards a more sustainable development. A virtual encyclopedia would be a scaleable, evolutive, contradictory, multilingual and open base of knowledge around the concept of sustainable development, and a starting point to all valuable data sources around the world.

2.0. CONTENTS

A set of documents certified by referees in an 'intranet' system, accompanied with extern (not yet) certified documents, will be provided on the WWW. The standard metainformation included allow a building up of an index of the certified pages, and makes the Virtual Encyclopaedia searchable. This encyclopaedia will be a moving 'up to date' source of information and data on sustainable development and related areas. It has the advantages of a published multimedia encyclopaedia, but in addition, it can be built progressively, and updated part by part as the information becomes available: a continuous 'state of the sustainable development', with no limitation in new chapters provided by various collaborative partners in the world.

3 0 EDITING STANDARDS

The same behaviour of each of the chapters, a logo or a label will remind the reader that they are still in the Encyclopedia, and if the page or the author are certified or not.

4.0 METADATA STANDARDS

Metadata standards are crucial to make the Virtual Encyclopedia searchable, indexable and controllable through out the whole Internet. As illustrated in figure 1, the WWW pages indexing can be done automatically by search engines or manually, using metadata standards, with a higher degree of description quality. In the header of each digital source, metadata describe the source, as suggested in OCLC/NCSA metadata workshop (Weibel et al, 1995). The minimum number of metadata elements required to facilitate the discovery of documents in a networked environment are: Subject, title, author, publisher, other agent, date, object type, form, identifier, relation, source, language and coverage (spatial and temporal). This minimum set of metadata is not limitative (see CEC-DGI-EEA-TF, 1993), and the use of existing metadata is made possible

Figure 1. WWW Documents Indexing

by internal references to existing description standards, as shown in the example bellow.

Example of structured metadata in a html format:

```
<META NAME ="language" SCHEME = "ISO" VALUE="en">
<META NAME ="language" VALUE="English">
<META NAME="subject" SCHEME = "GMET" VALUE="environmental
information">
<META NAME="administrative-coverage"  VALUE="Europe">
<META NAME="resolution"  VALUE="country">
<META NAME="frequency"  UNITS  = "days" VALUE="7">
<META NAME="URI"  VALUE="http://www.ulb.ac.be/ceese/cds.html">
<META NAME="last-update" SCHEME = "ANSI X3.30-1985" VALUE="19960722">
<META NAME="form" SCHEME = "IMT" VALUE="html">
```

<META NAME="languages" SCHEME = "ISO" VALUE="en, fr, nl">
<META NAME="medium" VALUE="internet, WWW, html">
<META NAME="availability" VALUE="free">
<META NAME="size" UNITS = "records" VALUE="250">

5.0 INDEXES

The metadata concerning the Encyclopedia and related documents or sources is collected by standard questionnaires and by computer assisted indexing (figure 2):
-questionnaires ensure quality descriptions of selected sources, including the not yet networked ones (note: the sources are best described by their authors);
-harvester (search engine) scans the headers and contents of documents of the whole internet to check the sources and identify possible new ones.
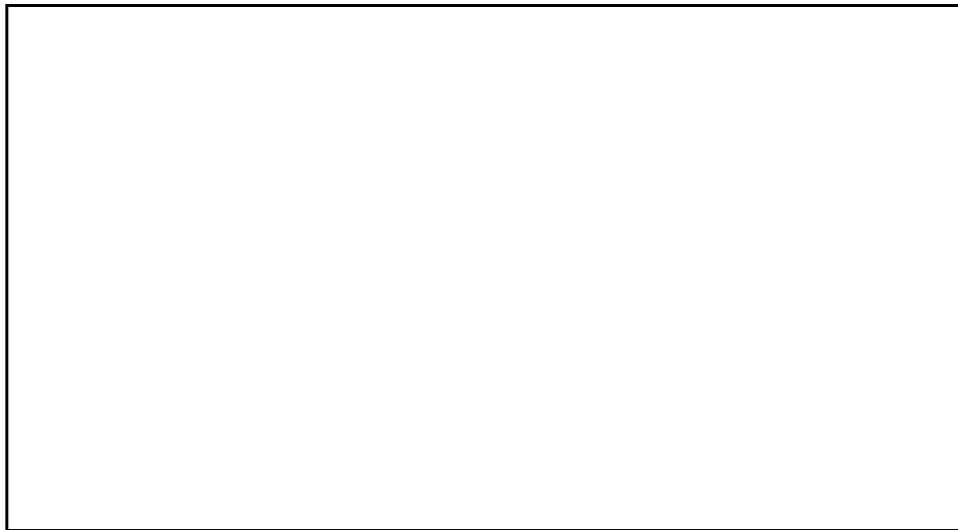
Figure 2. Metadata harvest

6.0 LANGUAGES

The tools to allow multilingual search of data sources can be:
-multilingual indexing, using agreed multilingual thesaurus like the INFOTERRA thesaurus (UNEP-INFOTERRA, 1990) or the Multilingual Environmental Thesaurus developed for the European Environment Agency (NBOI, 1995);
-graphic interfaces, namely clickable geographic interfaces, or even clickable charts, images or other objects with no language dependency.

On the other hand, the most read documents will be translated to key national or international languages. 'Cultural' translations and summaries for different target users could also be useful.

7.0 ORGANISATION

The organisation would follow the principle of a Special Interest Network or SIN as defined by Green (1994). The following elements are required:
-editors promoting limited editing and metadata standards and criteria;
-automatic and manual indexers;
-authors and referees on specific subject;
-translators and journalists for a multicultural approach.

8.0 CONCLUSION

A WWW Virtual Encyclopedia on sustainable development would mainly need a strong index system. The standardisation of metainformation on Internet documents would improve their searchability for the benefit of all potential users. The use of standard multilingual descriptors adapted to the themes of environment and sustainable development would make the link between users and data of different language or cultural origin. A core set of quality documents, certified by referees and labelled, provided in different languages, would be the starting point for an Encyclopedia open to any other networked resources.

9.0 REFERENCES

CEC-DGXI-EEA-TF, ''Catalogue of Datasources for the Environment: Analysis and Suggestions for a meta-data system and service for the European Environment Agency'', CEC DGXI and NRT-Nordic Council of Ministers, p. 185, 1993.

D.G. Green, 'A Web of SINs - the nature and organization of Special Interest Networks', http:www.csu.edu.au/sin/sin.htm, School of Environment and Information Sciences, Charles Sturt University, Albury, Australia, 1994.

NBOI, ''Multilingual Environmental Thesaurus'', The Netherlands Agency for Research Information, Amsterdam, 1995.

NSSDC, National Space Science Data Centre, ''Directory Interchange Format Manual'', WWW, version 4.1, NASA, USA, 1993-1996.

UNEP-INFOTERRA, ''Thesaurus of Environmental terms'', United nations Environment Programme, Nairobi, Kenya, 1990.

S. Weibel, J. Godby, E. Miller, and R. Daniel, ''OCLC/NCSA Metadata Workshop Report'', Dublin, NCSA, http://www.oclc.org:5046/conferences/metadata/diblin_core_report.htm, (18/4/96), 1995