



HAL
open science

Digital thesaurus: metadata resources for marine biology

Enrique Wulff-Barreiro

► To cite this version:

Enrique Wulff-Barreiro. Digital thesaurus: metadata resources for marine biology. 27th Annual Conference of the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC) and the 9th Conference of the European Association of Aquatic Sciences (EURASLIC), Oct 2001, Brest, France. hal-01234426

HAL Id: hal-01234426

<https://hal.science/hal-01234426>

Submitted on 26 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DIGITAL THESAURUS: METADATA RESOURCES FOR MARINE BIOLOGY

SUMMARY

The current status of the discipline of Marine Sciences, in Spain, is claimed to be bibliographically represented as complete according to the "repertoire" of actions, each reader has at his disposal when meeting the strategies and tools targeted to retrieve information after metadata. To keep abreast of the Spanish thesaurus regarding the bibliographic databases publicly available within the Web, a benchmarking search has been conducted. The four universities with Ms.D. studies on marine sciences, the Spanish Council for Scientific Research (CSIC), the Spanish Institute for the Oceanography (IEO), and the Spanish Network for the R&D (Network IRIS) have been focused on. The assessment is performed with the vision that there is no fixed, universal technology transfer process common to all institutions. Financial metrics have been discarded as particular disadvantaged for interdisciplinary research fields. Instead of this, a helpful reference point has resulted from the thematic groups oriented framing production of externalities. The study indicates the advantages of being able to mount the results of the metadata work on the internet. It presents the thesaurus of Icmán (Marine Sciences Institute of Andalusia), as a digital publication available to make sense of the web-based contents on the intranet of the library.

INTRODUCTION

These are two arguments favoring the domestic development of advanced digital thesaurus and metadata in marine science:

1) protection of domestic learning; a research strategy decision means that, in an information-equivalent translation, it is required a thorough breakdown of each field for the purpose of a systematic identification of all relevant potential bibliographical material.

2) the gaining of technical knowledge is the major benefit expected of the collaboration between science and industry. In a country like Spain only 2 international and 17 national recognised patents have been introduced in the period 82/97. So the use of a proprietary search engine (thesaurus) serves to enhance the business stature by discovering the dependence, complementarity, disjunction, and contravalance relationships between the elements of a relevant problem. The return to fashion of the thesaurus is then the reappearance of a content-rich searchable tool, presently available on the Web. Indexing and classification efforts of many companies to meet the challenges in the modern knowledge economy are a fascinating and exciting topic in business area today [1]. The activities of research and business convergence have been the subject guideline of the Canary Island sited VIIth national congress in aquaculture [2].

RESULTS AND DISCUSSION

MsD studies, in marine sciences, in Spain. Metadata.

Education in Sea Sciences in Spain with the goal of ultimately result in the availability of a Ms.D. mainly emerge out of the campuses of Alicante, Cádiz, Las Palmas de Gran Canaria, and Vigo. A major role play Cádiz and Las Palmas because of the presence there of specific Faculties of Sea Sciences.

Searching for data on the different metadata facilities available we have estimated that:

Las Palmas de Gran Canaria (Canary Islands), Faculty of Sea Sciences, located at the address URL <http://azuaje.ulpgc.es/centros/CMAR>, presents the 'lowest' profile with no metadata (and no specific program on the Ms.D. courses).

Present faculty of sea sciences of the university of Cadiz, URL: http://www.uca.es/facultad/ciencias_mar, do not yet possess more than the Metaname values: "Generator" and "ProgId". These meta-information has been written with Label Microsoft Front Page 4.0.

Vigo (Galician Region) Faculty of Sciences offered a Ms.D. in Sea Sciences, through a webpage, URL: <http://www.uvigo.es/w/centros/centros.gl.html>, where the program of the Ms.D. courses is implemented in an area written with Label Microsoft Front Page 3.0. The Metadata value "Generator" is displayed.

Available from 27-12-93, the web page of the Ms.D. in Sea Sciences offered by the Faculty of Sciences (University of Alicante), at the URL <http://www.ua.es/centros/ciencias/index.html>, do not use metadata but it illustrates the contents of the degree courses.

The RedIRIS (the Spanish academic internet network) resource in marine sciences, the Marinet Forum (http://tierra.rediris.es/marinet/marinet_english.html), offers the best meta-information set directly affecting the research agenda of those concerned with events in the area. Moderated by Dr. Somoza from the Spanish Geological and Minery Institute (Madrid) its competitive strategy offers to the browsers and search robots 13 keywords. Its metaname values exemplify the IRIS activities into virtual communities hosted at their servers and the percolation policy of support to meta-information. The range of values are: "Content-Type", "keywords", "description", CVU.Code, CVU.Name, Classification, DC.Description, DC.Format, DC.Subject, DC.Title, Description, Generator.

Digital thesaurus online.-

Only two digital thesaurus are concerned with marine sciences in Spain. Both come from the council for scientific research (CSIC) web resource. They are the Thesaurus of Animal Biology [10], and the Glossary of Aquaculture Terminology [4].

The specification of XML version 1.0 was released in february 1998, and during the three years it has passed since then it seems that only the CSIC digital thesaurus (both previous to 1996) are available through the net. The poster will assess then only with regard to document type html public written thesaurus.

Thesaurus of Animal Biology. A project once presented at the 46th Conference of the FID (International Federation for Information and Documentation), in Madrid 1992 [3]. Elaborated by the Scientific Information and Documentation Center (CINDOC, Madrid), it includes KwoC searches and Hierarchical-familial searches. It is written only in Spanish with a support DocType html Public, and no metadata. In internet it is sited at: http://pci204.cindoc.csic.es/tesauros/Biol_Ani/Biol_Ani.htm.

Glossary of Aquaculture Terminology. Under the conduction of Ms. Roldán, at the head of the aquaculture documentation unit (CINDOC, Madrid), this thesaurus was finished up in 1996. It includes KWOC searches. Some attention is shown on metadata, and a metaname value, "Content-type", is available with the subsequent title of the thesaurus. Available in Italian, Latin, English, French, German and Spanish. With the support of DocType html Public. In internet it is sited at: <http://pci204.cindoc.csic.es/tesauros/AcuiTerm/GIACuicu.htm>.

The digital thesaurus of Icmán (Marine Sciences Institute of Andalusia) (CSIC, Puerto Real).

A digital publication available to make sense of the web-based contents on the intranet of the library, of the Marine Sciences Institute of Andalusia.

To gain additional insights on the suitability of the web model to solve the librarian tasks the graphs resulting to distinctly index each paper extracted from the curriculum vitae of the researchers working in the laboratories were designed. It resulted in a thesaurus with a index of 318 keywords, and an annual index with 11 entries. The thesaurus is supported by the need to organise 253 papers. Indicating several instrumental or conceptual use all the thesaurus entries as much as all the papers set is provided with metadata.

The meta-information available for the query string syntax of the digital thesaurus, allow for several metaname Dublin Core values. DC.Title, creator, subject, description, publisher, date, language, rights allow the user to define the service information universe by setting mandatory keywords and by specifying an annual or a subject-oriented filter. An specific access to the digital objects of the library of icman is previewed from RedIris (R+D Spanish academic network), at www.rediris.es/si/iris-index/coord/contact.es.html, able to be searched by www.rediris.es/busquedas.

CONCLUSION

Lack of metadata standardization and very slow use of the Dublin Core standard seem the two meanings attached to the hope of increased validation of these digital resources on the web (34.2% of the servers contain metadata, and 0.3% of the sites used Dublin Core, according to the source [5]). This is obviously negative because metadata assumes a process more efficient than the brute-force search or time-consuming manual classification prevalent between general-purpose Internet engines [6].

Marine sciences, in the Spanish cyberspace, represent the rise of a new scientific community and only one virtual community in the field in question is able to be institutionally supported. This VC is present-day best internet implemented web resource in the discipline of marine sciences. And it is allocated to the R+D Spanish academic network (RedIris). Absences and failures in the Spanish electronic campus may be interestingly due to moral hazard circumstances. Because it seem that some tools suited to the needs of online publishing offices are available but that they are not used. Organizational changes to become familiar with the software are required.

The next generation of technologies attached to the production of digital thesaurus, the XML model is a remaining a tool on the supply side which is not making performances in Spain. This versatile technology have arrived after the production of the two digital thesaurus nowadays at work in the country.

The lack of mechanisms to access, store and classify data makes it difficult to reuse significant amount of it. And electronic records in marine sciences and in Spain seem allow to setup the display of the guidance and recommended procedures of the taskforce sited at IRIS-Index (the R+D Spanish academic network, RedIris, group inside the European pilot CHIC).

References.-

- [1] EUSIDIC 2001 Annual Conference. Session B: Knowledge management. 22-23 March 2001. Lille (France).
- [2] VIIth National Congress in Aquaculture. "Convergence between research and business: a challenge for the XXI century". Las Palmas de Gran Canaria, 1999 28-30 April. [<http://neptuno.iccm.rcanaria.es/nacional7.htm>].
- [3] Rey Rocha, J. "Development of a thesaurus on zoology, from the ICYT bibliographic database." International Federation for Information and Documentation 46th Conference and Congress. Madrid, 22-30 October 1992. Book of abstracts. 158-159.
- [4] Thesaurus of Animal Biology. Madrid, CSIC, 1994 [http://pci204.cindoc.csic.es/tesauros/Biol_Ani/Biol_Ani.htm]; and Glossary of Aquaculture Terminology. Madrid, CSIC, 1996 [<http://pci204.cindoc.csic.es/tesauros/AcuiTerm/GIACuicu.htm>].
- [5] Lawrence, S.; Giles, C.L. "Accessibility of information on the web". Nature, 400, 1999, 8 July, 107-109.
- [6] Cuenca, P.; Sosa, V. "Experiences in the use of metadata for web publishing". Electronic Publishing 99.



Enrique WULFF-BARREIRO
Marine Sciences Institute of Andalusia,
Spanish Council for Scientific Research. ICMAN (CSIC)
11510 Puerto Real (Cádiz). Spain.
E-mail: enrique.wulff@icman.csic.es

DEFINITION AND METHODOLOGY

The trade in technical knowledge constitutes a reference point for all digital libraries initiatives. It is why a benchmarking perspective has been adopted, in order to study and test the exploitation of digital thesaurus and metadata in the industrial/commercial/learned Spanish environment. The four campuses where MsD studies, in marine sciences, are taught in Spain (Alicante and Vigo inside the faculty of sciences, and Las Palmas de Gran Canaria and Cádiz inside the faculty of sciences of the sea), have been considered to elaborate the potential differences between their strategies of introduction of control software within their web pages, linking together their components or metadata, and of patterned electronic objects able to secure controlled access system to the internal content of the web resource or digital thesaurus.

Furthermore attested to the importance of the Aquatic Sciences and Fisheries Abstracts database (ASFA 1: Biological sciences & living resources) a recovery of the Spanish URL available for the years 1999-2000 has been performed. The nature and degree of the indexation practices have been appreciated as a critical variable in the choice of digital technology. 20 Web resources appear to break the bottleneck by dipping freely into the pool of marine sciences technological knowledge associated with internet. By this procedure the Spanish Institute for the Oceanography (IEO), and the Spanish Council for Scientific Research (CSIC) have been incorporated to the set of computational settings to be handled.

Requiring another approach the major axis of Spanish activity into automatic indexation and metadata, the in the net indexing experience task-force (IRIS-INDEX, <http://www.rediris.es/si/iris-index>), has been added to the analysis to support the seek functionality.

The second level of high level elements concerned, the digital thesaurus, at the present state of art in Spain bears an explicit extraction from the council for scientific research (CSIC) web resource. Thus the elements referred as Thesaurus of Animal Biology [3], and Glossary of Aquaculture Terminology [4], as two remarkably success histories, are kept for the time of analysis in this presentation.