Informational resources of Portal Cuba.cu. Description with metadata Dublin Core format

MSc. Isabel Daudinot Founier

Empresa de Tecnologías de la Información y Servicios Telemáticos Avanzados (CITMATEL). Cuba
Tel: +537 2043600
Fax: +537 204 8202
isabel@citmatel.inf.cu

Abstract:

This paper show the use of the Dublin Core (DC) format for the administration and description of the resources of information in the Portal Cuba.cu, created by the Enterprise of Information Technologies and Advanced Telematic Services (CITMATEL), with the subject to facilitate the retrieval of information. A guide was made based in DC format that permitted the description of the resources of information in the Portal.

Keywords:

Metadata, Dublin Core format, Description of information , Information retrieval, Portal Cuba.cu, Portal, Cuba

1. Introduction

The appearance and development of Internet has modified the more important aspects related with the creation, description, storage and the way of access to information resources. The use of metadata is one of the most promising alternatives for the description and organization of information resources in Internet; it also help in the localization and retrieval of these resources. Many institutions have identified this research field and already working in it.

CITMATEL, as well as others organizations involved in the generation, organization and retrieval of information in Internet is aware of the necessity of use standard and mechanisms that allow describe and organize its informational resources to facilitate not only the management, the searching process too, the localization and retrieval information.

Portal Cuba.cu, belonging to CITMATEL, offers access to a great quantity of growing current information about our country. For this reason it is of vital importance that the specialists that are dedicated to the processing and organization of these information resources, have the appropriate standardized tools that allow them to describe and to organize the information with homogeneity, clarity and the necessary speed. In this way, they can generate correct registrations that permit the users the access friendly to information resources.

2. Description of information resources of the Portal Cuba.cu

CITMATEL Web Division, creator and management of the Portal Cuba.cu, for the first time took the task of describing information resources of Portal, using a standard format of metadata, with the main objective of helping to increase its visibility and retrieval in Internet.

Knowing that in this Portal, the information is too varied and voluminous, we decided to use, the Dublin Core format, more adapted to the necessities of the Portal Cuba.cu, because of it flexibility and easy semantic.

Given the flexibility that the Dublin Core format allows, we only used those elements of the format responding to the necessities of the Portal. We inserted each description in the head of the code source of each resource of information.

Next, we show as an example, a description that we had included in the head of the code source of the one of services of the Portal.

<meta name="**DC.Title**" lang="es" content="Yellow Pages of the Portal Cuba.cu">
<meta name="**DC.Creator**" lang="es" content=" Enterprise of Informational Technologies and Advanced Telematic Services (CITMATEL)">
<meta name="**DC.Subject**" lang="es" content="Enterprises, Combined Enterprises, Mixed Enterprises, Industrial Enterprises, Yellow Pages, Cuba, Cuban Enterprises, Business, Cuban Institutions, Cuban Information">
<meta name="**DC.Description**" lang="es" content="This site offers information of the localization of Cuban institutions or foreigners, settled in the national territory, as well as of their products and services">
<meta name="**DC.Publisher**" lang="es" content="Enterprise of Informational Technologies and Advanced Telematic Services (CITMATEL)">
<meta name="**DC.Format**" lang="es" content="text/html">
<meta name="**DC.Date**" SCHEME="ISO8601" lang="es" content="1999-01-15">
<meta name="**DC.Identifier**" SCHEME="URL" lang="es"
content ="http://paginasamarillas.cuba.cu/">
<META name="**DC.Type**" lang="es" content="Home page">
<meta name="**DC.Language**" SCHEME="ISO 639" content="es">
<meta name="**DC.Coverage**" lang="es" content="Cuba">
<meta name="**DC.Right**" lang="es" content="Copyright© 1999-2005, CeniaInternet - CITMATEL">

The information resources that lodge in the Portal Cuba.cu grow quickly and the time that the specialists that describe resources have is very short.

For such reason to external creators of information resources that request lodging in the Portal Cuba.cu receive a document explaining the importance of the description; the document also explains how to code the resource using the format Dublin Core. They are in the obligation of giving each resource with the code of the format Dublin Core inserted in the head of the

code source. Before entering the resources in the database of the Portal, our specialists verify if the codes are correct; these data are used in the description of resources for the database of the Portal.

The Portal Cuba.cu offers the users the opportunity to recover the information hosting in the Portal through defined categories or using the internal search engine of the Portal (Catalejo). However, the existence of this search engine is not enough if the information is not properly organized to facilitate the search, the localization and the access to it.

For the standardization of the description the information resources of the database of the Portal, we made a guide based on the DC format. The guide is enlarged with some fields that are not part of this format in order to describe and control some necessary specific aspects for the administration of the database.

Next we show a brief description of the content of each field of this guide:

1. ID of registration
Automatic filling by the system

Definition of the field
Sequential unique number assigned and controlled automatically by the system to differentiate each registration of the database.

2. Title
Etiquette DC: "Title"
Obligatory filling

Definition of the field
Name assigned to the resource of information by the creator or editor. It should be concise, but informative, keeping the biggest relationship with the intellectual content of the resource.

3. Author
Etiquette DC: "Creator"
Obligatory filling
Repetitive

Definition of the field
Name of the author. If the author is an institution (corporate author), it is necessary to write the complete name of the institution, followed by the acronym in parenthesis.

4. URL or direction of the resource
Etiquette DC:"Identifier"
Obligatory filling

Definition of the field
Chain of signs (numbers), identifying the material in a unique way. The examples for the existent materials in the net include the URL (Uniform Resource Locator) and the URN (when they have been incorporated). This field registries the direction of the available resources of information in Internet.

5. Descriptors or keywords
Etiquette DC: "Subject"
Obligatory filling
Repetitive

Definition of the field
Theme of the material, expressed through keywords or sentences describing the content of the material.

6. Abstract or Description
Etiquette DC: "Description"
Obligatory filling.
Not repetitive

Definition of the field
This description must be as most informative as possible; it should offer enough information to allow the user to make a decision about the relevance of the resource, before accessing to it.

7. Language of the resource
Etiquette DC: "Language"
Obligatory filling

Definition of the field
This element identifies the language of the intellectual content of the resource of information. It goes in agreement with the standard ISO 639 (RFC 1766).

8. Date of publication of the resource in Internet
Etiquette DC: "Date"
Obligatory filling
Not repetitive

Definition of the field
Date associated with the creation or readiness of the material.
Registration comes according to the standard ISO 8601: 1988

The next fields or elements are not part of the format Dublin Core. We added them because of its importance for the administration and for the best handling of the database.

9. Category
Obligatory filling.
Repetitive

Definition of the field
Classification of the resource of information, according to the main thematic content.

10. Date of the last actualization (modification)
Obligatory filling.
Not repetitive.

Definition of the field
This element identifies the most recent date of modification of the resource of information. The date registration, according to the standard ISO 8601: 1988

11. Contact
Obligatory filling
Repetitive.

Definition of the field
Name and e-mail of the author (institution).

12. Notes
Optional filling

Definition of the field
Additional information of the resource, necessary for the management of the system; not concerns to the user. It is filled only when it is necessary.

All information resources of the database of the Portal Cuba.cu are described in correspondence with this guide.

Results of the work

1. The visits to the Portal increased significantly

2. The guide permitted:

- Describe the resources quickly and satisfy the recovery of information by the users.
- Establish the fields, its definition and its structure. Standard by this way the process of description of the informational resources of the Portal Cuba.cu. To guarantee the uniform presentation of metadata, a basic condition for retrieval information
- Define the retrieval information fields. This brought a bigger relevance and pertinence of the recovered information by the internal search engine of the Portal
- Determine the metadata for users. Every record returned by the search engine show the user the title, a summary and the address of the information resource

3. The results obtained with the application of this guide have allowed us to recommend the use of the format Dublin Core in the description of many other Cuban sites with a good acceptance of the creators and the visitors

References

1. Berners-Lee, T., *Metadata Architecture*, [Online] Available   at: http://www.w3.org/DesignIssues/Metadata.html.

2. Consorcio World Wide Web. 1999, *The global structure of an HTML document*, [Online] Available at: http://www.w3.org/TR/1999/REC-html401-19991224/struct/global.html.

3. Gradmann, S. 1998, 'Catalogación Versus Metadata: Vino Viejo en Odres Nuevos?', *64th IFLA General Conference August 16 - August 21,* [en   línea]   Disponible   en: http://www.ifla.org/IV/ifla64/007-126s.htm.

4. Méndez Rodríguez, E. M. & Merlo Vega, J. A. 2000, 'Localización, identicación y descripción de recursos web: tentativas hacia la standardalización', *VII Jornadas Españolas de Documentación*, Palacio Euskalduna, Argitalpen Zerbitzua, pp. 221-231.

5. Méndez Rodríguez, E. M. 2000, 'Localización, identificación y descripción de documentos web: tentativas hacia la Standardalización', *VII Jornadas Españolas de Documentación*, Palacio Euskalduna, Argitalpen Zerbitzua,   pp. 211-231.

6. Méndez Rodríguez, E. M. 2002*, Metadatos y recuperación de información: estandares, problemas y aplicabilidad en bibliotecas digitales*, Ediciones Trea, España

7. Merlo Vega, J. A..& Sorli Rojo, A. 2000, 'El uso de metainformación en los webs de las bibliotecas españolas', *VII Jornadas Españolas de Documentación*, Palacio Euskalduna, Argitalpen Zerbitzua, pp. 155-164.

8. Miller, Paul, Metadata for the masses: describes Dublin Core and means by wich it can be implemented, [Online] Available at: http://www.ariadne.ac.uk/issue5/metadata-masses/

9. Weibel, S.. 1997, *Dublin Core Metadata Element Set: reference description,* [Online] Available at: http://www.purl.org/metadata/dubli_core_elements.