

BRAZILIAN ELECTRONIC THESES AND DISSERTATIONS CONSORTIUM

Carlos Henrique Marcondes
Universidade Federal Fluminense/Depto. de Ciência da Informação
Project consultant
marcondes@alternex.com.br

Luis Fernando Sayão
Centro de Informações Nucleares/Comissão Nacional de Energia Nuclear
Project consultant
lsayao@cnen.gov.br

Ricardo Triska
Universidade Federal do Paraná/Depto. de Ciência e Gestão da Informação
Project manager - IBICT
ricardotriska@hotmail.com

Ana M B Pavani
Pontifícia Universidade Católica do Rio de Janeiro/Depto. de Engenharia Elétrica
Project partner
apavani@lambda.ele.puc-rio.br

Abstract

This paper relates and analyzes the experience of various Brazilian institutions organized in a consortium with the aim of providing integration and unified access, via a web gateway, to the full-text electronic theses and dissertations already available in different Brazilian universities. The increase and impact of academic free electronic publishing to scientific communication and to the development of science is discussed. The results of a pilot experience, developed under the coordination of IBICT as well as technical and technological aspects of the information system model adopted, are presented. Aspects such as metadata element set and format, protocols for automatic gathering of metadata, union catalog database, user interface and integration with other Brazilian STI systems are analyzed. As a conclusion, future developments and enlargement of the Brazilian Electronic Theses Consortia experience are discussed.

Key-words: electronic theses, electronic dissertations, electronic publishing, digital libraries, interoperability, metadata, science and technology information, science and technology management, Brazil

1. INTRODUCTION

This paper relates and discusses the recent experience of different Brazilian institutions, coordinated by IBICT – Brazilian Institute for Scientific and Technological Information – organized in a consortium to provide unified access through a single web gateway to full-text electronic theses and dissertations available today in different Brazilian universities. The Brazilian Consortium of Electronic Theses and Dissertations is the first subproject, available online in an experimental basis, of the Brazilian Digital Library in Science and Technology project – BDL. BDL has as its main objective the full-text publishing in digital format of contents of academic interest. A secondary

objective is to assure the interoperability of heterogeneous systems on the Internet to be integrated through a single web gateway.

For a long time, in the context of the so called “information overload”, science and technology information systems were built based on the paper document and referential information technological paradigm. These systems aimed to providing access to paper documents stored in a library or in a clearing house. Referential data bases made known “gray” literature while interlibrary loan mechanisms supplied copies of the documents. These mechanisms were expensive, slow, complex and required strong articulation among partners to be effective. **An expressive set of methodologies, technologies, professional training and science and technology information systems were built around the paper document paradigm. Não entendi esta frase!**

This paradigm has been changing with the rise of electronic publishing of academic documents directly on the Internet. The potential visibility of electronic academic publishing and the immediate access to full-text contrast with enormous quantity of information available on the Internet. This has turned information retrieval into the main task to the use of information in different activities (Marcondes, 2001). As a consequence, science and technology information systems face new challenges: interoperability, cooperation and unified access to different digital libraries. In order to achieve maximum visibility it is essential to assure access to academic documents, completing the knowledge transfer cycle.

Brazil has created a graduate culture unique among developing nations. It began with the National Graduate Plan in the seventies. The result of this plan is a set of graduate courses, funding agencies and periodic assessment that smoothly operate. In this context, theses and dissertations are a relevant result. They yield the results of scientific research as well as they indicate how these results come from: different areas of knowledge, geographic regions of the country, universities, etc. Even investments of funding agencies can be examined.

The first attempt to integrate Brazilian theses and dissertations in an unified data base was SITE - Theses System. This union catalog is coordinated by IBICT and began its online operation in 1996. Currently, it holds more than 140,000 records of theses and dissertations from 15 universities; a web interface provides access to it. As an expected follow on of this service, IBICT, in the scope of BDL project (Marcondes, 2001), proposed to various partners the development of a union catalog of the electronic theses and dissertations. Electronic theses and dissertations were available in many universities; the same happened with their catalogs. The union catalog is to be accessed through a unified web gateway. Isolated experiences of Brazilian institutions, regarding digital libraries and electronic publishing, are reported in a special number of *Ciência da Informação* (Information Science) Journal, volume 30, number 3, oct/dec 2001 (<http://www.ibict.br/cionline/300301/>).

This paper addresses and comments this experience. Section 2 analyzes the changes in science and technology information systems caused by the Internet and the new possibilities opened by electronic publishing, with a special focus on research results and scientific communications. Section 3 describes the standards and the technological and methodological options adopted by the consortium, specially those related to a common metadata interchange format and to the procedures for automatic gathering of metadata (to be included in common metadata database). Section 4

addresses the results of the pilot-project in the overall context of Brazilian Digital Library. Last but not least, section 5 presents future possibilities of enlargement of consortium to include other universities and to develop new services to add value to the theses and dissertation union catalog.

2. ELECTRONIC PUBLISHING PARADIGM

Information Science emerges around the “information overload” and the need of optimizing science communications flux as an aid to science development. Specialized information systems emerge in this context too. Information needed to scientific research is mainly constituted by non-conventional documents, those not found in the editorial market. They are also known as “gray literature” and contain research reports, congress papers, preprints, theses and dissertations. All of them report freshly obtained research results.

Currently, the communication of results of scientific research and development results is dispersed in two axes: spatially - scientists in different geographies get results and publish them in different communication vehicles, and temporally – this is done in different times. The role of information systems is to oppose to information dispersion, providing a focus of concentration to the communication of research results. Specialized science and technology information systems were mainly built around the task of providing access to this kind of literature, as NTIS, Dissertation Abstracts, Chemical Abstracts, INIS, ETDE, AGRIS and many others. These systems were built around the management, methodological and technological paradigm of referential information as a step to access copies of paper documents.

The widespread use of the Internet and of the mechanisms of direct publishing on the web have established a new paradigm to scientific communication. An ever-increasing number of academic journals and other scientific communication mechanisms, as preprints of papers, are migrating to electronic versions. A vivid debate on the problem of free access to academic articles is taking place between the scientific community and the scientific editors. The scientific community considers academic electronic publications as a means to increase visibility, to accelerate the development of science and to widely disseminate the results of research, considered as a humankind heritage (Harnard). According to this, charge of fees by scientific international editors hinders the free flux of research results and development of science, serving just commercial interests. Another fundamental question to the academic community, relative to information systems, is the visibility of electronic academic publishing: according to NEC Research Institute, electronically published articles are far more cited (157%) than paper published versions.

A pioneering initiative concerning a preprint electronic repository is ArXiv, at the Los Alamos National Laboratory, created in 1991 by American Physicist Paul Ginsparg (Van de Sompel, 2000). After this initial experience, the international scientific community has been developing a practical alternative for free publishing and access to academic works available directly on the Internet. As a consequence of this debate, new initiatives as PUBMED central (<http://www.pubmedcentral.nih.gov/>), Public Library of Science (<http://www.publiclibraryofscience.org/>) and OpenArchives Initiative (<http://www.openarchives.org/>) are arising. The dimensions of OpenArchives

Initiative worldwide can be evaluated through an extensive list in <http://www.osti.gov/eprints/ppnbrowse.html>.

Together with OpenArchives Initiative, there has been the development of various standards, technologies and methodologies concerning web publishing, metadata element sets and automatic exchange of metadata between digital libraries.

Besides preprints archives and electronic journals, theses and dissertations are one of the most common electronic publications. Theses and dissertations are important not only to an individual author, who looks for maximum visibility and scientific impact of his/her work, but also to advisors, academic departments, post-graduate programs and universities. There are initiatives of electronic theses and dissertations data banks in various countries, university consortia and isolated universities (see <http://www.ndltd.org/members/index.htm>). UNESCO has been supporting, since 1999, initiatives of staff training and building electronic libraries of theses and dissertations. Those initiatives are today consolidated around NDLTD – Networked Digital Library of Theses and Dissertations (<http://www.ndltd.org>). Currently, there are more than 120 university members of NDLTD (Suleman, 2001a and 2001b) worldwide. Among these, 7 (Virginia Tech University, CalTech, MIT, USA, Humboldt, Dresden and Duinsburg, German), plus PhysNet, contribute to the union catalog. NDLTD recommends a common metadata format, the ETD-ms (electronic theses and dissertations metadata set), available in <http://www.ndltd.org/standards/metadata/ETD-ms-v1.00.html>.

Electronic publishing on the Internet is a guarantee of enhanced visibility and easy accessibility to research results, once they are made available through appropriate digital libraries.

3. BRAZILIAN ETD CONSORTIUM MODEL OF INTEROPERABILITY

The Brazilian Digital Library is a long-term initiative of IBICT with the objective of providing the country scientific community with facilities for electronic publishing and access to full-text resources. Besides encouraging electronic publishing, BDL aims at integrating into a single web gateway heterogeneous information resources already operational. Some of these resources are SCIELO (<http://www.scielo.br>), a Brazilian gateway of electronic full-text, ANAIS, a gateway of electronic full-text proceedings of Brazilian events, PROSSIGA (<http://www.prossiga.br>), a gateway of Internet resources of interest to science and technology in different subject areas, various university OPAC's, some ETD data bases, Brazilian National Library OPAC, various open archives, etc.

To achieve integration between those resources in a single web gateway, BDL proposal will use a hybrid model of interoperability. BDL web gateway will be a Z39.50 client, accessing various Z39.50 servers of different Brazilian information systems. A special Z39.50 server will be the one at IBICT that will serve a common database formed with metadata gathered from other Brazilian information systems. This common metadata database will provide unified access via the same web gateway to information resources not served by Z39.50. OpenArchives Metadata Harvesting Protocol will be used to gather metadata from various sources to the common metadata database. The BDL project model of interoperability is showed in figure 1 and figure 2. Due to the importance of theses and dissertations to scientific communication and the emergence

of various isolated universities ETD databases, IBICT invited some partners to create an ETD union catalog as a first project of BDL.

Brazilian ETD Consortium aims to develop a cooperative union catalog. This union catalog will hold only metadata from ETD; full-text electronic documents will be stored in individual ETD databases of the members. Metadata will be centralized in a common data base at the coordinator institution site – IBICT . *ETD will be accessed through a single web gateway, providing unified access and allowing the user not to access each university ETD site one by one.*

The union catalog was created to support different kind of users and their specific information needs. The first group is made up of researchers and the graduate students who are interested in specific subject queries and to whom the union catalog of ETD is a knowledge transfer mechanism. For this type of users, the union catalog can be an important tool cooperation because it will allow the identification of research partners (or advisors) and of research groups working in the related subjects. This will help avoid duplication of efforts since a more effective cooperation can be established.

The ETD union catalog addresses the needs of the science and technology planners and managers too. This kind of user is interested in great sets of theses and dissertations – by author, by knowledge area, by geographic area, by institution, by funding agency, that could indicate trends and help decision making. ETD union catalog enables precise analysis and views of the graduate effort in Brazil, its needs, its development trends, its relations with economic and strategic sectors and its coordination with government priorities. Providing all this information side by side to compare them can reveal new relations and multiply its information potential, as says Latour (2000).

In a historical and social perspective, the ETD union catalog will become a cultural digital heritage of post-graduation in Brazil, a “datawarehouse” of multiple meanings, whose interpretation will depend on the viewpoint of those who analyze it.

Today, making its information available on the Internet is not the best way for a digital library to get maximum visibility. There is so much information on the Internet that finding what is needed may become quite a task and can be very time consuming too. Similar to conventional libraries in past, digital libraries are faced with the same challenge: collaboration and cooperation to guarantee maximum visibility of their holdings. Interoperability means the facility of different library catalogs being accessed simultaneously, not requiring that the user search one by one. Interoperability is a major research and technical objective. It requires a substantial amount of technology, standardization and protocols in order to achieve interoperability among different and heterogeneous electronic archives and digital libraries. Standards such as MARC and Z39.50 are mechanisms oriented to achieve interoperability between automated library catalogs.

The ETD databases in different Brazilian universities were developed according to different methodologies and they use different software and data formats. They were conceived to serve each university needs and not from a cooperative perspective. For this reason, to achieve interoperability among these systems, through a union catalog accessed by a single web gateway, it required an intense effort of standardization. This

work was developed in two directions: the first was the adoption of a protocol for the automatic gathering of metadata from each individual ETD database to a common metadata database where a user can submit queries; and the second was the definition of a common format to exchange metadata.

The unified web gateway gives users access to a common metadata database, containing metadata gathered from each individual university. Original full-text ETD is locally stored, as previously mentioned. The common metadata database is the mechanism to achieve integrated access to members' ETD collections. Each ETD record has a link to full-text electronic thesis or dissertation.

Brazilian science and technology information systems are experienced in maintaining cooperative union catalogs. Some examples are: Brazilian universities and IBICT operate SITE; different biomedical institutions, coordinated by BIREME, developed LILACS (Latin America and Caribbean Literature in Health Science) database; Brazilian universities, under coordination of FGV (Fundação Getúlio Vargas), maintain CALCO union catalog. Brazilian institutions that maintain cooperative information systems are conscious of how hard cooperation is. It requires both from the host and the member institutions. One of the objectives of the consortium is to minimize the managerial effort of each member. This was the main reason for the adoption of an automatic metadata gathering protocol, the OpenArchives Metadata Harvesting Protocol (<http://www.openarchives.org/OAI/openarchivesprotocol.htm>).

OAI MHP is a simple protocol that provides non-immediate interoperability between eprints repositories, digital libraries or any networked server that intends to make metadata of documents visible to be gathered by an external crawler. OpenArchives Initiative refuses the option of Z39.50 protocol due to its extreme complexity and to its characteristic of being computer resources intensive. (Troll, 2001).

In the OAI protocol conception there exist two kinds of partners. The first is made of institutions maintaining servers with electronic documents repositories and providing facilities to authors in publishing, storing and recovering electronic documents; they are called data providers. The second are the institutions that gather metadata about electronic documents stored in one or more data providers to develop value-added services with them, the service providers. Examples of value-added services would provide unified access to metadata of different repositories, as do Arc - Cross Archive Searching Service (<http://arc.cs.odu.edu/>) -, NDLTD gateway and BDL gateway; or the development of a qualified database on a specific subject with metadata collected from different sources, or publish a full-text journal with peer-review articles with metadata collect from different data providers.

The exchange of messages between data providers and service providers is unidirectional – service providers call data providers who answer sending metadata. Service providers calls are made using HTTP protocol (<http://www.ietf.org/rfc/rfc2616.txt>) GET or POST methods. Calls are answered by data providers sending metadata coded in XML (<http://www.w3.org/XML/>). OAI MHP determines Dublin Core Metadata Element Set (<http://purl.org/DC/documents/rec-dces-19990702.htm>) as the minimum mandatory metadata set to be supported by data providers. However, a data provider can offer its metadata according to other metadata sets, as for example, MARC, or Brazilian ETD metadata set. Interaction between

service providers and data providers using OAI MHP can be tested through OAI Repository Explorer, an interactive web interface developed with the aim to test repositories compliance with OAI MHP (<http://rocky.dlib.vt.edu/oai/cgi-bin/Explorer/oai1.0/testoai>). Slowly, digital library and digital publishing communities have become aware of the importance and the potential of OAI MHP for interoperability. The protocol has been receiving crescent attention from these communities due to its technological and conceptual simplicity. There are still proposals of extensions to OAI MHP, addressing questions like interfaces between digital libraries modules (Suleman, 2001c). A more technical view of OAI MHP can be found in Lynch (2001) and Warner (2001).

With the adoption of OAI MHP by the consortium, each university member will be a data provider, exposing metadata about its ETD collection. The metadata will be “visible” to a crawler operated by consortium coordinating institution – IBICT – that periodically will proceed automatic gathering of metadata to a common metadata database. OAI MHP will have a crescent impact in cooperation mechanisms between digital libraries and electronic archives. The use of OAI MHP by the BDL project will enable integration and unified access to metadata of electronic documents from different sources (Marcondes, 2001b). This option, the same adopted by NDLTD (Suleman, 2001b), will allow the integration of the consortium this international initiative.

The adoption of a common metadata set is a fundamental step to achieve interoperability between different ETD databases. The consortium adopted the Dublin Core Metadata Element Set (1999), extended to hold specific theses and dissertations information. The adopted format is fully compliant with NDLTD’s ETD-ms. Metadata is coded in XML, used as file format to exchange metadata, as established by OAI MHP. Besides typical bibliographic information relative to a thesis or dissertation, Brazilian ETD format includes complete author, advisor and committee-members information, including institutional affiliation, complete and citation names, full institution information including graduate programs, universities, funding agencies.

The extensive information set adopted by the consortium will enable new value-added services too. Together with other Brazilian Information systems partners, the consortium intends to provide full integration through cross-linking to other Brazilian information systems that maintain information relative to authors, relating them to different roles they play in ETD's - authors or advisors or committee-members. Initial partners will be SCIELO and CVLattes (curriculum vitae of Brazilian researchers, maintained by CNPq), which hold information on researchers or personal authors of Brazilian Journal articles. This will enable linking an advisor in an ETD database to his/her curriculum vitae information stored in CVLattes and/or his/her academic production stored in SCIELO. It will be possible to follow a link from SCIELO or CVLattes to ETD common database too.

There is an intention of achieving crescent semantic interoperability by adopting a common subject classification - the CNPq subject area classification, largely adopted by Brazilian scientific community.

4. PILOT-PROJECT EVALUATION

The Brazilian ETD Consortium developed a pilot ETD project as part of Brazilian Digital Library Project. This pilot-project included ETD's from the following institutions: BIREME/ENSP/FIOCRUZ, PUC-Rio, UFSC and USP. Three other universities already started negotiations to join the consortium: UNICAMP, UNESP and UnB. The main objective of this pilot-project was evaluated common ETD exchange format, focusing on the difficulties on generating metadata from individual databases to fill all the fields.

Initially, a great effort was dedicated to achieve consensus concerning the common metadata file format, with elements of each ETD coded in XML. An XML DTD was developed too. Based on this DTD, metadata of ETD from different Consortia members were sent to IBICT, forming the common metadata database. A web gateway was developed to provide integrated access to consortium (<http://www.ibict.br>). The Brazilian ETD pilot project encompassed 1047 electronic theses and dissertations, as of December 2001.

A preliminary evaluation of ETD metadata transferred to IBICT during the pilot-project shows the necessity of a great effort of revision of metadata exchange format to establish standards for some elements. Another important effort is to encourage institutions to fill out all information elements recommended by the standard format in order to develop a value-added data base and achieve different cross searching queries in the ETD union catalog. Lack of information relative to non-mandatory elements agreed upon in the metadata exchange format may increase as soon as more universities join the consortium. It seems that, after the pilot, more universities will want to be included in the consortium.

During the pilot-project, transfer of metadata from the members to IBICT was performed via FTP protocol; OIA MHP was not used. Consortium members that sent ETD metadata to pilot-project are experienced institutions, with expertise in using information technologies applied to information systems: BIREME – Latin American and Caribbean Information Center on Health Science (<http://www.bireme.br>) - a leader information center in Health Science, member of Pan-American Health Organization and World Health Organization; Stela group from UFSC (<http://www.stela.eps.ufsc.br>), conceived and implemented CVLattes system to CNPq (<http://www.cnpq.br>); and LAMBDA – Library, Museums and Archives Automation Laboratory (<http://www.maxwell.lambda.ele.puc-rio.br>), from PUC-Rio.

As more universities join the consortium, through their library systems, as USP, UNICAMP, UNESP, UNB, their library system staff will not have the same expertise relative to information technologies as BIREME, Stela Group or LAMBDA Laboratory. As the consortium grows bigger, more institutions with less experienced teams will outnumber the original members. This scenario points to the necessity of promptly adopting OAI MHP to automatically gather metadata with a minimum operational effort by the institutions. Different methodologies relative to information technologies to manage ETD data banks were used by institutions that participated in pilot-project. From those, only USP adopted one compliant with OAI MHP, the methodology of Digital Library Research Laboratory from Virginia Tech. University, USA (www.dlib.vt.edu). IBICT developed a methodology to manage electronic archives, based in Eprint software (<http://www.eprints.org>), used in a experience of open archives

with Brazilian Genetic Society (Triska, 2001). This methodology is compliant with OIA MHP and can be easily adapted to manage an ETD data bank. All other methodologies adopted in pilot-project are proprietary ones, not compliant with OAI MHP and must be adapted to full compliance with OAI MHP. Fortunately, Open Archives Initiative site has several links to various free software tools that can be used to develop servers and crawlers compliant with OAI MHP.

5. CONCLUSIONS AND OPEN QUESTIONS

After the pilot-project, the main objective of Brazilian ETD Consortium is the institutional consolidation of the initiative and its adoption by all Brazilian universities. IBICT developed a formal agreement document for universities to join the Consortium. This document will be discussed in next Consortium meeting. Other universities like UNICAMP, UNESP and UnB expressed the wish to join the Consortium immediately. The next Consortium meeting will establish a steering committee too. The steering committee will develop technical, organizational and political guidelines. All those questions must be institutionalized too.

There is also the need of solving technical questions that will permit Brazilian ETD Consortium to meet its full objectives and enable all planned services be operational.

The Consortium must be easily manageable and must be extended to other universities without an excessive managerial overload either to members or to coordinating institution. To meet this objective, automatic metadata gathering using OAI MHP must be fully operational. The use of OAI MHP is the main technological challenge to the Consortium after the pilot-project. The use of this protocol is a pioneering experience to Brazilian information systems and its dissemination will permit an integration of Brazilian information systems to international information flux. The use of OAI MHP by the Consortium will enable full integration to NDLTD, achieving maximum visibility to Brazilian electronic theses and dissertations, one of the objectives of the project.

The analysis of the metadata files transferred to IBICT by Consortium members (during the pilot-project) indicated the lack of many non mandatory fields of the metadata set. The complete filling of most metadata elements is essential for the union catalog to meet all its information potential. To meet this objective it is necessary to integrate the ETD union catalog with SITE, the referential bibliographic theses and dissertations data base maintained by IBICT too. As mentioned previously, SITE holds more than 140,000 records of theses and dissertations. This will result in a single theses and dissertation database, some of them full-text ETD, whose number will increase over the time.

Maximum semantic interoperability, an important specification too, will be met through generalized use of CAPES/CNPq Classification for indexing ETD's. Another initiative towards this objective is the creation of a cooperative Authority server to support all science and technology information systems community in Brazil.

As stressed before, the majority of institutions that participated in the pilot-project were experienced in information systems technology. They had the necessary expertise and technical resources to face the technical challenges of the pilot-project and immediately joined after the invitation by IBICT. However, for the consortium to meet the objective of including ETD's from all universities, even those with little experience and expertise in information technology-based systems, it is necessary help these universities in developing their own OAI MHP compliant ETD data bases and have their ETD metadata harvested to the common metadata data base. To meet this objective, the Consortium plans to develop a complete kit of an ETD data base site, including a web server, and a training program. It is also expected that this will be the seed to develop other electronic documents repositories that fully compliant with OAI MHP.

It is necessary to integrated Brazilian ETD union catalog with other important Brazilian science and technology information systems, as CVLattes/CNPq, SCIELO/BIREME, ANAIS/CIN-CNEN (electronic proceedings of Brazilian meetings in science and technology) and DATACAPES (information on Brazilian graduate programs), through the creation of a "link server" (Santana, 2001) to cross-link all these systems. This will enable easy navigation through all of them, achieving cross access through all individual works of a person, as an author of a journal article, of a research report or of a conference paper, or even as an advisor of a thesis.

Brazilian CNPq is developing an joint initiative with other Latin American and Caribbean countries with the aim of creating an integrated researcher curriculum vitae system, the CVLAC (De los Rios, 2001). There is also a demand from CNPq to integrate the Brazilian ETD union catalog with similar initiatives of other Portuguese speaking countries in order to develop a Portuguese language ETD union catalog (with Minho University or the National Library, Portugal). There is also a similar demand that encompass various Latin American and Caribbean countries.

The expressed wish of various Brazilian universities to join the Consortium indicates the real necessity of an union catalog of theses and dissertations to the Brazilian academic community. The proposed system has some interesting features: it provides unified access to Brazilian full-text theses and dissertations through a single web gateway and uses an exchange metadata format based on XML for the automatic gathering of metadata through OAI MHP. An increased visibility of Brazilian theses and dissertations, in the national context and worldwide, will be achieved by the Consortium web gateway. It is expected that this will result in more quality of Brazilian research effort.

BIBLIOGRAPHY AND SOURCES OF INFORMATION

DE LOS RIOS, Rebeca, SANTANA, Paulo Henrique de Assis. El espacio virtual de intercambio de información sobre recursos humanos em Ciência Y Tecnologia de América Latina y el Caribe: del CV Lattes al CvLAC. **Ciência da Informação**, Brasília, v.30, n.3, p. 42-47, set./dez. 2001.

DUBLIN CORE METADATA ELEMENTS SET, version 1.1: reference description. Dublin Core Initiative, 1999. Access in <http://dublincore.org/documents/1999/07/02/dces/>, visited in 06/06/2001.

ETD-ms: an interoperability metadata standard for electronic theses and dissertations – version 1.00. Access in <<http://www.ndltd.org/standards/metadata/ETD-ms-v1.00.html>>, visited in 22/11/2001.

GINSPARG, P. Winners and losers in the global research village. In: CONFERENCE ON ELECTRONIC PUBLISHING IN SCIENCE, 1996, Paris. **Proceedings...** Access in <<http://xxx.lanl.gov/blurb/pg96unesco.html>>, visited in 05 out. 2001.

HARNARD, Stevan. The self-archiving initiative. Nature Webdebates. Access in <<http://www.nature.com/nature/debates/e-access/Articles/harnad.html>>, visited in 03 nov. 2001.

LATOUR, Bruno. Redes que a razão desconhece: laboratórios, bibliotecas, coleções. In: BARATIN, Marc, JACOB, Christian. O poder das bibliotecas: a memória dos livros no ocidente. Rio de Janeiro: Ed. UFRJ, 2000. p. 21-44.

LAWRENCE, Steve. Free online availability substantially increases a paper's impact. Access in <http://www.nature.com/nature/debates/e-access/Articles/lawrence.html>, visited in 10 jun. 2001.

LYNCH, Clyfford. Metadata Harvesting and the Open Archives Initiative. **ARL Bimonthly Report**, n. 217, Aug. 2001, p.1-9. Access in <<http://www.arl.org/newsltr/217/mhp.html>>, visited in 18 nov. 2001.

MARCONDES, Carlos Henrique. Representação e economia da informação. **Ciência da Informação**, Brasília, v.30, n.1, p. 61-70, jan./abr. 2001.

MARCONDES, Carlos Henrique, SAYÃO, Luis Fernando. Integração e interoperabilidade no acesso a recursos informacionais em C&T: a proposta da Biblioteca Digital Brasileira. **Ciência da Informação**, Brasília, v.30, n.3, p. 24-33, set./dez. 2001.

ODLYSKO, Andrew. The Economics of Electronic Journals. **The Journal of electronic Publishing**, v.4, n.1, Sep. 1998. Access in <<http://www.press.umich.edu/jep/04-01/odlyzko.html>>, visited in 03 nov. 2001

SANTANA, Paulo Henrique de Assis, PACKER, Abel Laerte, BARRETO, Marcia Yamataka. Servidor de enlaces: motivação e metodologia. **Ciência da Informação**, Brasília, v.30, n.3, p. 48-55, set./dez. 2001

SULEMAN, Hussein, FOX, Edward A. A framework for building open digital libraries. **D-Lib Magazine**, v. 7, n. 12, Dec.. 2001. Access in <<http://www.dlib.org/dlib/december01/suleman/12suleman.html>>, visited in 18 dez. 2001.

SULEMAN, Hussein, et al. Networked digital library of theses and dissertations: bringing the gap for global access: part 1: mission and progress. **D-Lib Magazine**, v. 7, n. 9, Sept. 2001. Access in <<http://www.dlib.org/dlib/september01/suleman/09suleman-pt1.html>>, visited in 19 set. 2001.

SULEMAN, Hussein, et al. Networked digital library of theses and dissertations: bringing the gap for global access: part 2: services and research. **D-Lib Magazine**, v. 7, n. 9, Sept. 2001. Access in <<http://www.dlib.org/dlib/september01/suleman/09suleman-pt2.html>>, visited in 19 set. 2001.

TRISKA, Ricardo, CAFÉ, Ligia. Arquivos abertos: subprojeto da Biblioteca Digital Brasileira. **Ciência da Informação**, Brasília, v.30, n.3, p. 92-96, set./dez. 2001

TROLL, Denise; MOEN, Bill. Report to the DLF on the Z39.50 Implementers' Group. DLF : 2001. Access in <<http://www.diglib.org/architectures/zig0012.htm>>, visited in 22 ago. 2001.

VAN De SOMPEL, Herbert, Lagoze, Carl. The Santa Fe Convention of the Open Archives Initiative. **Dlib Magazine**, v.6, n.2, february 2000. Access in <<http://www.dlib.org/dlib/february00/vandesompel-oai/vandesompel-oai.html>>, visited in 15 ago. 2001.

WARNER, Simeon. Exposing and Harvesting Metadata Using the OAI Metadata Harvesting Protocol: A Tutorial, **High Energy Physics Libraries Webzine**, v.4, June 2001. Access in <<http://library.cern.ch/HEPLW/4/papers/3/>>, visited in 18 nov. 2001.

Acho que foi dada uma boa ênfase na visibilidade internacional - acredito que deveríamos mencionar que estamos prevendo metadados multi-língües para garantir a busca e a recuperação for a dos países de língua portuguesa!!!!!!!

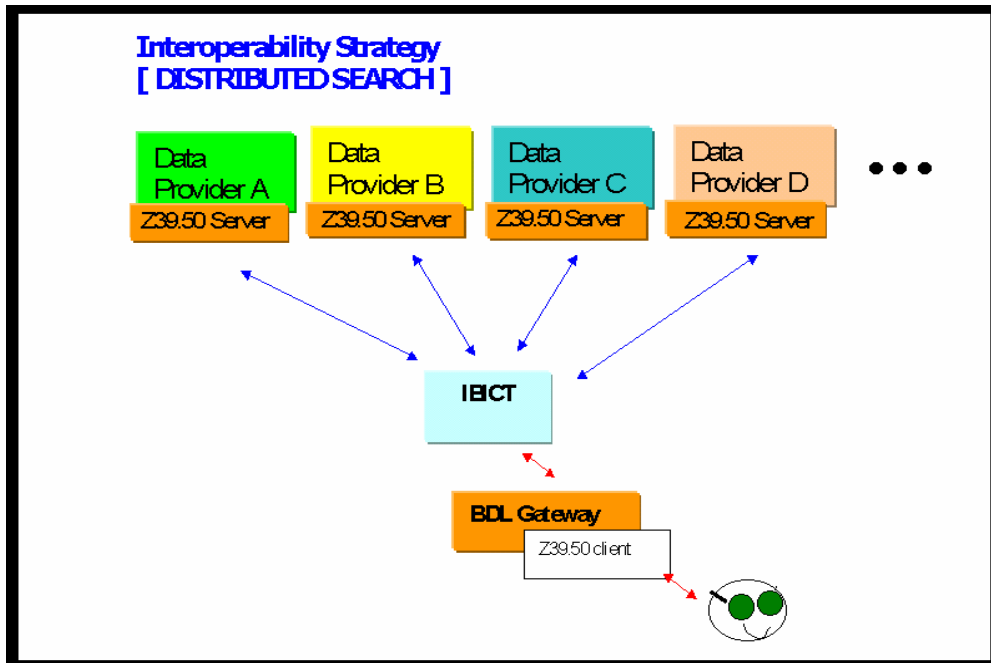


Figure 1– Interoperability Strategy – Distributed Search

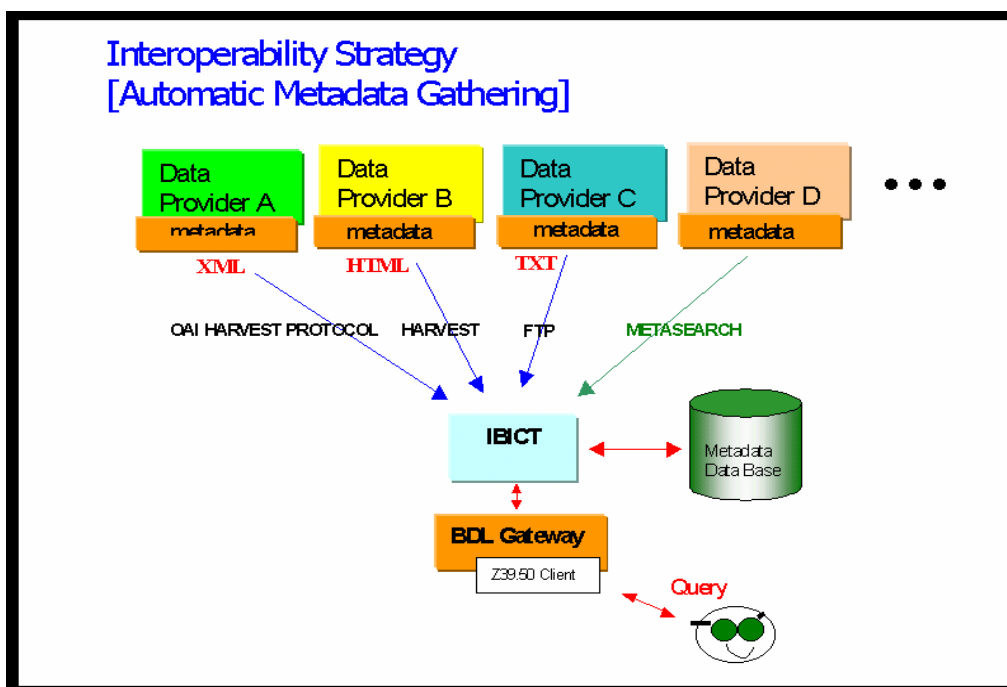


Figure 2 - Interoperability Strategy – Automatic Metadata Harvesting