

JAQUELINE COSTA ALVES

**ORGANIZAÇÃO E REPRESENTAÇÃO DA
INFORMAÇÃO NA BIBLIOTECA DIGITAL DE TESES E
DISSERTAÇÕES DA UNIVERSIDADE DO ESTADO DE
SANTA CATARINA - UDESC: UMA ANÁLISE FOCADA
EM METADADOS SOB A LUZ DO PADRÃO MTD-BR**

Florianópolis, 2009

JAQUELINE COSTA ALVES

**ORGANIZAÇÃO E REPRESENTAÇÃO DA
INFORMAÇÃO NA BIBLIOTECA DIGITAL DE TESES E
DISSERTAÇÕES DA UNIVERSIDADE DO ESTADO DE
SANTA CATARINA - UDESC: UMA ANÁLISE FOCADA
EM METADADOS SOB A LUZ DO PADRÃO MTD-BR**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Mestre, sob a orientação da Professora Doutora Lígia Maria Arruda Café

Florianópolis, 2009

Catálogo na fonte pela Biblioteca Universitária da
Universidade Federal de Santa Catarina

A474o Alves, Jaqueline Costa

Organização e representação da informação na biblioteca digital de teses e dissertações da Universidade do Estado de Santa Catarina - UDESC [dissertação] : uma análise focada em metadados sob a luz do padrão MTB-BR / Jaqueline Costa Alves ; orientadora, Lígia Maria Arruda Café. - Florianópolis, SC, 2009.
161 f.: il., tabs.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências da Educação. Programa de Pós-graduação em Ciência da Informação.

Inclui bibliografia

1. Biblioteca Digital de Teses e Dissertações.
2. Ciência da informação. 3. Recuperação da informação.
4. Metadados. I. Café, Lígia Maria Arruda. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Ciência da Informação. III. Título.

CDU 02

JAQUELINE COSTA ALVES

**ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO NA
BIBLIOTECA DIGITAL DE TESES E DISSERTAÇÕES DA
UNIVERSIDADE DO ESTADO DE SANTA CATARINA - UDESC:
UMA ANÁLISE FOCADA EM METADADOS SOB A LUZ DO
PADRÃO MTD-BR**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação do Centro de Ciências da Educação da Universidade Federal de Santa Catarina em cumprimento a requisito parcial para a obtenção do título de Mestre em Ciência da Informação.

APROVADA PELA COMISSÃO EXAMINADORA

EM FLORIANÓPOLIS, 03 DE JULHO DE 2009

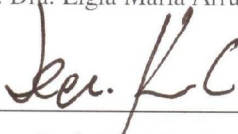


Profa. Dra. Lígia Maria Arruda Café - Coordenadora do Curso

BANCA EXAMINADORA



Profa. Dra. Lígia Maria Arruda Café – PGCIN/UFSC (Orientadora)



Prof. Dr. Hélio Kuramoto – IBICT (Examinador)



Prof. Dr. Raimundo Nonato Macedo dos Santos – PPGCIN/UFSC (Examinador)

Profa. Dra. Miriam Vieira da Cunha – PPGCIN/UFSC (Suplente)

Dedico este trabalho, com a mais profunda admiração e amor, a minha mãe Luiza, este grande exemplo de vida, determinação, dedicação e amor, em quem me espelho todos os dias.

A minha filha Mell, o mais perfeito, lindo e esperado presente que ganhei em toda a minha vida.

Ao meu pai José Luiz e aos meus irmãos Giovanni, Jacson e Daniel (um por todos e todos por um!!)

Aos vós Mário (in memoriam) e Abílio (in memoriam), por sempre acreditarem e torcerem por mim.

AGRADECIMENTOS

Agradeço a Deus pelo dom da vida e pela oportunidade de realizar mais este sonho.

A Luiza, minha amorosa e dedicada mãe, que esteve sempre ao meu lado me incentivando e apoiando. Por ter abdicado de seus interesses pessoais e profissionais, para dedicar-se totalmente a mim e a Mell durante o tempo que dela precisamos e por estar sempre pronta a me apoiar em todos os momentos de minha vida, mesmo quando não entende os motivos que me impulsionam a tomar determinadas atitudes. Mãe, muito obrigada! Sem sua ajuda eu não teria forças para alcançar esse meu objetivo!

A Mell, minha amada filha, que nasceu no início do desenvolvimento deste trabalho e que, com seu olhar e sorriso, repletos de amor e felicidade, é capaz de me fazer esquecer qualquer dificuldade ou tristeza, no momento em que estamos juntas.

Ao José Luiz, meu querido pai, que se fez presente durante toda a minha trajetória e, em especial, durante o desenvolvimento deste trabalho, demonstrando sua preocupação e prestando auxílio em todos os momentos que dele precisei.

Ao Giovanni, ao Jacson e ao Daniel, meus queridos irmãos (vítimas da minha mania de perfeição) que, mesmo morando longe, me incentivaram e apoiaram em todos os momentos.

A profa. Dra. Lígia Café, minha orientadora, por sua amizade, dedicação e estímulo. Foi um presente especial ter uma profissional tão competente como minha condutora nesta jornada. Obrigada por não me deixar perder o “foco”!

Ao prof. Dr. Hélio Kuramoto e ao prof. Dr. Raimundo Nonato Macedo dos Santos, por aceitarem o convite para contribuir com este trabalho e pelas valiosas sugestões.

Aos professores do PGCIN, pelo respeito e pela capacidade de estimular uma nova forma de reflexão sobre determinadas questões.

Aos colegas da turma 2007 do mestrado, por fazerem parte desta história. Foi muito bom conhecer cada um de vocês!

Aos colegas da UDESC, por entenderem e apoiarem minha vontade de realizar esta conquista.

A Salete, Mário, Bia, seu Pedro, Val e Iury, amigos que estão sempre dispostos a ajudar. Obrigada!

Ao Ricardo, meu querido amigo, pela força nas últimas horas.

A todos que, de uma forma direta ou indireta, sentiram vontade de compartilhar comigo esta vitória. Muito obrigada!

*Quando a gente acha que tem todas as
respostas, vem a vida e muda todas as
perguntas ...*

Luís Fernando Veríssimo

RESUMO

ALVES, Jaqueline Costa. **Organização e representação da informação na Biblioteca Digital de Teses e Dissertações da Universidade do Estado de Santa Catarina – UDESC: Uma análise focada em metadados sob a luz do padrão MTD-BR.** 2009. 152f. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação. Universidade Federal de Santa Catarina, Florianópolis, 2009.

Juntamente com a utilização crescente da *world wide web* (www), surgem dificuldades relacionadas com o aumento do volume de informações disponibilizadas sem uma forma estruturada de representação, o que dificulta o acesso e a recuperação da informação. Inseridas neste contexto, as bibliotecas digitais dependem da padronização da descrição física e de conteúdo de seus itens para garantir a interoperabilidade. Acredita-se que a adoção de metadados para a descrição de documentos eletrônicos possa contribuir para sanar estas dificuldades. No entanto, o conteúdo desses metadados deve obedecer a critérios cuja finalidade é tornar consistente a representação dos objetos incluídos nos repositórios, melhorando assim a interoperabilidade com outras bibliotecas digitais. Nesta realidade, encontra-se esta pesquisa cujo objetivo é analisar o conteúdo dos metadados das dissertações eletrônicas da Universidade do Estado de Santa Catarina – UDESC, disponibilizados para a Biblioteca Digital de Teses e Dissertações – BDTD, a fim de verificar se os mesmos estão sendo informados de acordo com o padrão MTD-BR e, desta forma, contribuir para a localização e recuperação das dissertações eletrônicas desta instituição. A abordagem geral da pesquisa foi exploratória, sendo analisados de forma qualitativa e quantitativa todos os metadados das dissertações eletrônicas da UDESC disponibilizados para a BDTD. A abordagem qualitativa foi realizada utilizando-se como parâmetro as recomendações do padrão MTD-BR no que se refere ao conteúdo dos elementos de metadados e seus respectivos atributos, enquanto a abordagem quantitativa teve como objetivo comparar a quantidade de informações corretas e incorretas de metadados. Os resultados quantitativos foram apresentados de forma descritiva e, com o auxílio de tabelas. Verificou-se que a maioria dos metadados de preenchimento obrigatório são registrados de acordo com o padrão MTD-BR. Entretanto, observou-se que os elementos de preenchimento facultativo, muitas vezes, deixaram de ser informados ou foram informados incorretamente, o que não implicou na exclusão do documento do processo de *harvesting* mas, dependendo da busca realizada, pode impossibilitar a recuperação do documento na BDTD. Os resultados observados induziram a apresentação de sugestões para a melhoria do conteúdo dos metadados informados pela UDESC para a BDTD. Conclui-se que a biblioteca digital da UDESC, apesar de ter uma boa qualidade no preenchimento dos metadados, pode ainda aprimorar este procedimento em favor da interoperabilidade e da recuperação da informação.

Palavras-chave: Metadados, Dublin Core, ETD-MS, MTD-BR, BDTD, NDLTD.

ABSTRACT

ALVES, Jaqueline Costa. **Organização e representação da informação na Biblioteca Digital de Teses e Dissertações da Universidade do Estado de Santa Catarina – UDESC: Uma análise focada em metadados sob a luz do padrão MTD-BR.** 2009. 152f. Dissertation (Master in Information Science) - Postgraduate Program in Science Information – Federal University of Santa Catarina, 2009.

Together with the increasing use of world wide web (www), difficulties arise related to the increased volume of information available without a structured way of representation, making it difficult to access and retrieval of information. Inserted in this context, digital libraries depend on the standardization of physical description and content of your items to ensure interoperability. It's believed that the adoption of metadata for description of electronic documents can help solve these problems. However, the content of metadata should conform to criteria whose purpose is to make a consistent representation of the objects included in the repositories, thus improving interoperability with other digital libraries. In fact, this research is aimed at analyzing the content of metadata for electronic theses at the University of the State of Santa Catarina - UDESC, available for the Digital Library of Theses and Dissertations - BDTD to verify whether they are informed in accordance with the standard MTD-BR and thus contribute to the location and retrieval of electronic dissertations that institution. The general approach of the research was exploratory, and analyzed qualitatively and quantitatively in all metadata of the electronic dissertations UDESC available for BDTD. A qualitative approach was performed using a parameter as the recommendations of the standard MTD-BR regarding the content of metadata elements and their attributes, while the quantitative approach aimed to compare the amount of correct information and incorrect metadata. The quantitative results were presented in a descriptive way, and with the aid of tables. It was found that most of the metadata required are registered according to the standard MTD-BR. However, it was observed that the elements of Optional often no longer be informed or have been incorrectly informed, which is not involved in the exclusion of the document but the process of harvesting, depending on the search carried out, may preclude the recovery of the document in BDTD. The results observed led to the submission of suggestions for improving the content of the metadata information for the UDESC BDTD. It is concluded that the library's digital UDESC, despite having a good quality in fulfillment of the metadata, you can still improve this procedure in favor of interoperability and retrieval of information.

Keywords: Metadata, Dublin Core, ETD-MS, MTD-BR, BDTD, NDLTD.

RESUMEN

ALVES, Jaqueline Costa. **Organização e representação da informação na Biblioteca Digital de Teses e Dissertações da Universidade do Estado de Santa Catarina – UDESC: Uma análise focada em metadados sob a luz do padrão MTD-BR.** 2009. 152f. Dissertação (Maestría en Ciencia de la Información) - Programa de Postgrado em Ciências de la Información – Universidad Federal de Santa Catarina, 2009.

Junto con el aumento del uso de World Wide Web (WWW), surgen dificultades relacionadas con el aumento del volumen de información disponible sin una forma estructurada de la representación, lo que dificulta el acceso y la recuperación de la información. Inserta en este contexto, las bibliotecas digitales dependen de la normalización de la descripción física y el contenido de sus artículos para garantizar la interoperabilidad. Se cree que la aprobación de los metadatos para la descripción de los documentos electrónicos pueden ayudar a resolver estos problemas. Sin embargo, el contenido de los metadatos debe ser conforme a criterios cuyo objetivo es lograr una representación coherente de los objetos incluidos en los repositorios, por lo tanto, mejorar la interoperabilidad con otras bibliotecas digitales. De hecho, esta investigación tiene por objeto analizar el contenido de los metadatos de tesis electrónicas en la Universidad del Estado de Santa Catarina - UDESC, disponible para la Biblioteca Digital de Tesis y Disertaciones - BDTD para verificar si se ha informado de conformidad con la norma MTD-BR y, por tanto, contribuir a la localización y recuperación de tesis electrónicas esa institución. El enfoque general de la investigación fue exploratorio, y se analizaron cualitativa y cuantitativamente en todos los metadatos de las tesis electrónicas UDESC disponibles para BDTD. Un enfoque cualitativo se realizó con un parámetro como las recomendaciones de la norma-BR MTD en relación con el contenido de elementos de metadatos y sus atributos, mientras que el enfoque cuantitativo para comparar la cantidad de información correcta e incorrecta de metadatos. Los resultados cuantitativos se presentan en forma descriptiva, y con la ayuda de tablas. Se constató que la mayoría de los metadatos necesarios están registradas de acuerdo a la norma MTD-BR. Sin embargo, se observó que los elementos de Facultativo a menudo ya no ser informado o han sido mal informados, que no participa en la exclusión del documento, pero el proceso de recolección, en función de la búsqueda realizada, puede impedir la recuperación del documento en BDTD. Los resultados observados llevaron a la presentación de sugerencias para mejorar el contenido de los metadatos para la información UDESC BDTD. Se concluye que la biblioteca digital de UDESC, a pesar de tener una buena calidad en el cumplimiento de los metadatos, aún puede mejorar este procedimiento en favor de la interoperabilidad y la recuperación de la información.

Palabras clave: Metadatos, Dublin Core, ETD-MS, MTD-BR, BDTD, NDLTD.

LISTA DE FIGURAS

Figura 1 - Integração dos níveis de abrangência da BDTD	40
Figura 2 – Integração dos metadados utilizados pela BDTD	61
Figura 3 – Interação entre Provedor de Serviços e Provedor de Dados, por meio do OAI-PMH	66
Figura 4 – Esquema Funcional do Modelo Open Archives	67
Figura 5 - Tela de busca de Teses e Dissertações da NDLTD.....	77
Figura 6 - Recuperação na base de dados da BDTD por meio do elemento NumeroChamada (3.4).....	81
Figura 7 – Título informado nos idiomas português (pt) e inglês (en)	82
Figura 8 – Título informado 02 vezes em português	83
Figura 9 – Título informado 05 vezes em português	83
A repetição de uma mesma informação não prejudica a recuperação, no entanto pode causar ruídos na compreensão do registro pelo usuário.	84
Figura 10 – Recuperação, na base de dados da UDESC, da dissertação com o elemento Titulo (4) não informado.....	84
Figura 11 – Resultado da busca, feita na base de dados da BDTD, da dissertação com o elemento Titulo (4) não informado	85
Figura 12 – Busca, na Biblioteca Digital da UDESC, da dissertação para a qual não foi informado o elemento URL (5.1)	87
Entretanto, pode-se verificar na figura abaixo que a mesma não pode ser recuperada na base de dados da BDTD. Este problema também se repetirá nas buscas realizadas pela NDLTD.	87
Figura 13 – Busca, na base de dados da BDTD, da dissertação para a qual não foi informado o elemento URL (5.1)	88
Figura 14 – Descrição da tabela Arquivos da base de dados da biblioteca digital de Teses e Dissertações da UDESC	89
Figura 15 – Atributo Idioma do elemento Legenda (5.2) apresentado de acordo com o padrão adotado	90
Figura 16 – Busca, na base de dados da BDTD, da dissertação para a qual o local de defesa foi informado incorretamente	97
Figura 17 – Recuperação de dissertação na BDTD com o elemento DataDefesa incorreto, realizada no mês de julho de 2008	99
Figura 18 – Recuperação de dissertação na BDTD com o elemento DataDefesa correto, realizada no mês de janeiro de 2009	100
Figura 19 - Recuperação da dissertação na biblioteca de teses e dissertações eletrônicas da UDESC, realizada em fevereiro de 2009	101
Este fato sugere que haja provavelmente um procedimento de correção dos metadados realizada pelo provedor de serviço BDTD.	101
Figura 20 - Recuperação da dissertação na biblioteca de teses e dissertações eletrônicas da UDESC, realizada em fevereiro de 2009	102
Figura 21 - Recuperação da dissertação na BDTD, realizada em fevereiro de 2009 ...	103
Figura 22 - Recuperação da dissertação na biblioteca de teses e dissertações eletrônicas da UDESC, realizada em fevereiro de 2009	104
Figura 23 - Recuperação da dissertação na BDTD, realizada em fevereiro de 2009 ...	105
Figura 24 – Recuperação, na base de dados da BDTD, da dissertação cujo nome da autora tem uma vírgula ao lado do nome.....	107
Figura 25 – Recuperação, na base de dados da NDLTD, da dissertação cujo nome da autora tem uma vírgula ao lado do nome.....	108

Figura 26 – busca das dissertações da UDESC na BDTD, no período de 1997 a 2007	122
Figura 27 – total de dissertações da UDESC na BDTD, no período de 1997 a 2007...	122
Figura 28 – Tela onde são informados os metadados referentes ao autor da tese/dissertação	124
Figura 29 – Tela onde são informados os metadados do contribuidor da tese/dissertação	125
Figura 30 - Tela onde são informados os metadados da agência de fomento	126
Figura 31 - Tela onde são informados os metadados da tese ou dissertação	127
Figura 32 - Tela onde são informados os metadados da tese ou dissertação	128
Figura 33 - Tela com os metadados da tese ou dissertação que foram informados.....	129

LISTA DE TABELAS

Tabela 1 – Componentes básicos <i>Dublin Core</i>	51
Tabela 2 – Componentes básicos ETD-MS	55
Tabela 3 – Componentes básicos MTD-BR	57
Tabela 4 - Tabela comparativa dos elementos do MTD-BR, ETD-MS e DC	61
Tabela 5 – Elementos para descrição física e descrição de conteúdo do documento informados pela UDESC para a BDTD	74
Tabela 6 – Descrição dos elementos Sigla (1.1), DataAtualizacao (1.2), IdentificacaoDocumento (1.3) e Tipo (1.4) do elemento Controle (1)	76
Tabela 7 – Descrição dos elementos Nome (2.1), Sigla (2.2) e URL (2.3) do elemento BibliotecaDigital (2)	78
Tabela 8 – Descrição dos elementos Nome (2.4.1), Sigla (2.4.2), Pais (2.4.3), UF (2.4.4), CNPJ (2.4.5) e URL (2.4.6) do elemento ProvedorServico (2.4)	79
Tabela 9 – Descrição dos elementos Nome (3.1), Sigla (3.2), URL (3.3) e NumeroChamada (3.4) do elemento BibliotecaDepositaria (3)	80
Tabela 10 – Descrição do elemento Titulo (4) e do atributo Idioma deste elemento	82
Tabela 11 – Descrição dos elementos URL (5.1), Legenda (5.2), NivelAcesso (5.3) e seus respectivos atributos do elemento Arquivo (5)	86
Tabela 12 – Descrição do elemento Idioma (6)	91
Tabela 13 – Descrição do elemento Grau (7)	92
Tabela 14 – Descrição do elemento Titulacao (8)	92
Tabela 15 – Descrição do elemento Resumo (9) e do atributo Idioma deste elemento ..	93
Tabela 16 – Descrição do elemento Cobertura (10) e do atributo Idioma deste elemento	94
Tabela 17 – Descrição do elemento Assunto (11) e dos atributos Idioma e Esquema deste elemento.....	95
Tabela 18 – Descrição dos elementos Cidade (12.1), UF (12.2) e Pais (12.3) do elemento LocalDefesa (12)	96
Tabela 19 – Descrição do elemento DataDefesa (13).....	98
Tabela 20 – Descrição dos elementos Nome (14.1), Citação (14.2), Lattes (14.3), CPF (14.4) e Afiliacao (14.5) do elemento Autor (14).....	106
Tabela 21 – Descrição dos elementos Nome (14.5.1), Sigla (14.5.2), Pais (14.5.3), UF (14.5.4), CNPJ (14.5.5) e URL (14.5.6) do elemento Afiliacao (14.5)	110
Tabela 22 – Descrição do atributo Papel do elemento Contribuidor (15)	111
Tabela 23 – Descrição dos elementos Nome (15.1), Citação (15.2), Lattes (15.3) e CPF (15.4) do elemento Contribuidor (15).....	112
Tabela 24 – Descrição dos elementos Nome (15.5.1), Sigla (15.5.2), Pais (15.5.3), UF (15.5.4), CNPJ (15.5.5) e URL (15.5.6) do elemento Afiliacao (15.5)	114
Tabela 25 – Descrição dos elementos Nome (16.1), Sigla (16.2), Pais (16.3), UF (16.4), CNPJ (16.5) e URL (16.6) do elemento InstituicaoDefesa (16).....	116
Tabela 26 – Descrição dos elementos Nome (16.7.1) e Area (16.7.2) do elemento Programa (16.7)	117
Tabela 27 – Descrição dos elementos Nome (17.1), Sigla (17.2), Pais (17.3), UF (17.4), CNPJ (17.5) e URL (17.6) do elemento AgenciaFomento (17)	118
Tabela 28 – Descrição do elemento Direitos (18) e de seu atributo Idioma.....	119
Tabela 29 – Descrição do elemento Extensão (19) e de seu atributo Namespace.....	120
Tabela 30 – Elementos de preenchimento obrigatório que deixaram de ser informados	121

LISTA DE ABREVIATURAS

BDTD.....	Biblioteca Digital de Teses e Dissertações
COMUT	Sistema de Comutação Bibliográfica
CTC.....	Conselho Técnico Consultivo
DC.....	Dublin Core
DCMI.....	Dublin Core Metadata Initiative
ETD-MS.....	Metadata Standard for Electronic Theses and Dissertations
ETDs	Eletronic Theses and Dissertations
FGDC.....	Federal Geographic Data Committee
FTP.....	File Transfer Protocol
HTML	HyperText Markup Language
IAFA/ WHOIS++	Internet Anonymous FTP Archive with transfer protocol
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
IES	Instituição de Ensino Superior
LDs.....	Linguagens Documentárias
MARC.....	Machine Readable Catalogue
MTD-BR.....	Padrão Brasileiro de Metadados de Teses e Dissertações
NDLTD	Networked Digital Library of Theses and Dissertations
OAI	Open Archives Initiative
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OCLC.....	Online Computer Library Center
PUC-RIO	Pontifícia Universidade Católica do Rio de Janeiro
RDF.....	Resource Description Framework
SAIF.....	Spatial Archive and Interchange Format
SGML	Standard Generalized Markup Language
SNCT&I.....	Sistema Nacional de Ciência, Tecnologia e Inovação
SRI	Sistema de Recuperação da Informação
SRI.....	Sistemas de Recuperação da Informação
TDEs	Teses e Dissertações Eletrônicas
TEI	Text Encoding Initiative
TIC	Tecnologia da Informação e comunicação
TICs	Tecnologias da Informação e comunicação
UDESC	Universidade do Estado de santa Catarina

USPUniversidade de São Paulo

UNESCOOrganização das Nações Unidas para Educação, Ciência e Cultura

WWWWorld Wide Web

XML.....Extensible Markup Language

SUMÁRIO

1	Introdução	18
1.1	Justificativa	20
1.2	Objetivos	25
1.2.1	Objetivo Geral:	26
1.2.2	Objetivos específicos:	26
2	Referencial Teórico.....	27
2.1	Ciência da Informação e Organização da Informação	27
2.2	Bibliotecas Digitais de Teses e Dissertações	31
2.2.1	<i>Networked Digital Library of Theses and Dissertations (NDLTD)</i>	37
2.2.2	Biblioteca Digital de Teses e Dissertações (BDTD)	38
2.3	Metadados	41
2.3.1	Estrutura	43
2.3.2	Interoperabilidade	45
2.3.3	Padrões de Metadados	48
2.3.3.1	Padrão de Metadados <i>Dublin Core (DC)</i>	49
2.3.3.1.1	Breve histórico do <i>Dublin Core</i>	50
2.3.3.1.2	Elementos do <i>Dublin Core</i>	51
2.3.3.1.3	Atributos do <i>Dublin Core</i>	52
2.3.3.1.4	Qualificadores do <i>Dublin Core</i>	53
2.3.3.2	Padrão de Metadados ETD-MS	53
2.3.3.2.1	Breve histórico do ETD-MS	54
2.3.3.2.2	Elementos do ETD-MS	54
2.3.3.2.3	Atributos do ETD-MS	55
2.3.3.2.4	Qualificadores Globais do ETD-MS	56
2.3.3.3	Padrão de Metadados MTD-BR	56
2.3.3.3.1	Breve histórico do MTD-BR	56
2.3.3.3.2	Elementos do MTD-BR	57
2.3.3.3.3	Atributos do MTD-BR	60
2.3.4	Comparação dos padrões MTD-BR, ETD-MS e <i>Dublin Core</i>	60
2.4	Protocolo OAI-PMH (<i>Open Archives Initiative Protocol for Metadata Harvesting</i>)	63
2.4.1	Breve histórico do OAI-PMH	64
2.4.2	<i>Harvesting</i>	65
3	Procedimentos Metodológicos.....	68
4	Análise dos metadados informados pela UDESC para a BDTD	69
5	Conclusão e Recomendações para trabalhos futuros	130
5.1	Conclusão.....	130
5.2	Recomendações para trabalhos futuros.....	131
	REFERÊNCIAS.....	133
	ANEXOS	140

1 Introdução

Atualmente, a informação é essencial para a realização de qualquer tarefa. Por esta razão, observa-se uma grande evolução nas pesquisas sobre tratamento e recuperação da informação (RI), possibilitando assim, o desenvolvimento de novas tecnologias para torná-los cada vez mais eficientes. Neste sentido, a Ciência da Informação tem um papel estratégico a exercer, considerando a importância da organização e representação da informação para a recuperação desta. Sendo assim, Dias (2001) faz a seguinte reflexão:

“Se adotarmos o pressuposto de que a questão básica da ciência da informação é a de acesso à informação, ou facilitação desse acesso, isso implica em trazer para o primeiro plano a importância de sistemas de informação e de sistemas de recuperação da informação, cujos objetivos são os de facilitar o acesso à informação necessitada por uma determinada comunidade de usuários”. (DIAS, 2001, p. 3)

A RI permite aos usuários que documentos sejam localizados por meio de consultas realizadas, normalmente, por palavras-chaves. Baseando-se nesta consulta, os documentos que possuem, literalmente, alguma palavra contida na representação dos mesmos, são recuperados. Entretanto, documentos relevantes podem deixar de ser recuperados ou, o documento recuperado e apresentado para o usuário pode não ser o que melhor atende sua necessidade, uma vez que diferentes fatores influenciam na RI. É nesse sentido que Baeza-Yates e Ribeiro Netto (1999) relatam que os documentos recuperados, freqüentemente, não são suficientes para satisfazer a necessidade de informação dos usuários. Acerca dos fatores que afetam a qualidade da RI, Chen (1999) destaca os seguintes: variações nas estruturas e formatos das base de dados, diferentes formas de documentos disponibilizados (texto, áudio e vídeo) e abundância de conteúdos multilingües nas aplicações da Web. Trata-se, portanto, de uma falha no processo de organização da informação, o que provoca os prejuízos na recuperação relatados acima.

As medidas mais utilizadas para avaliar a eficiência na recuperação da informação textual são a precisão e a revocação (*precision and recall*). Sendo revocação, a relação entre o número de documentos relevantes recuperados pelo sistema de recuperação da informação (SRI) e o número total de documentos relevantes contidos no sistema, dito de outra forma, a revocação mede a capacidade que um SRI possui de recuperar o máximo de documentos que puder sobre a consulta que lhe foi formulada pelo usuário. Enquanto precisão é a relação entre o número de documentos relevantes recuperados pelo SRI e o número total de documentos recuperados pelo sistema, ou seja, avalia, dentre os documentos recuperados, quais correspondem realmente ao que o formulador da consulta está buscando. Ambas se baseiam na noção de relevância, que diz respeito a probabilidade dos documentos recuperados coincidirem com aquilo que o usuário está buscando. Foskett (1973) apresentou precisão e revocação, especificidade e exaustividade, facilidade de uso, tempo de resposta e tipo de pesquisa (interativa e heurística) como características importantes dos SRIs.

A exaustividade e a especificidade são princípios, ou seja, orientações para a representação do conteúdo dos documentos. A exaustividade é a denominação atribuída a extensão da representação temática de um documento. Lancaster (2004) menciona que “a medida que se aumenta a extensão da representação também se aumenta a recuperabilidade do item”. Isto significa que quanto mais for elevado o número de pontos de acesso maior as chances do usuário encontrar o documento desejado. No entanto, sabe-se que este princípio deve ser utilizado com cautela uma vez que pode afetar igualmente a precisão da recuperação.

A especificidade é o segundo princípio do processo de representação de assuntos. Lancaster (2004) diz ser esta a diretriz “mais importante da indexação de assunto”. De acordo com o autor, a especificidade é aquela em que “um tópico deve ser indexado sob o termo mais específico que o abranja completamente” (LANCASTER, 2004). A maior ou menor adoção destes princípios terá influência direta com as medidas de recuperabilidade dos SRIs já mencionadas: a revocação e a precisão. Foskett (1973) lembra que estas medidas são inversamente proporcionais, pois conforme se aumenta o número de itens em um resultado de pesquisa, diminui-se a precisão do conjunto de itens recuperados; e o contrário também é verdadeiro, quanto mais se refinam as estratégias de busca, mais diminui o número de itens recuperados, aumentando a precisão da recuperação.

Assim, enquanto a exaustividade e a especificidade estão relacionadas com a organização da informação, a revocação e a precisão se referem a recuperação, sendo que o processo de recuperação depende da qualidade do tratamento dado aos documentos. Esta qualidade está associada igualmente a adoção de padrões, principalmente quando se trata de ambientes disponibilizados na *web* em que o processo de recuperação da informação está cada vez mais sendo executado de forma automática.

De acordo com Souza e Alvarenga (2004), a padronização da descrição da informação é essencial. Segundo estes autores:

“[...] é necessária uma padronização de tecnologias, de linguagens e de metadados descritivos, de forma que todos os usuários da *web* obedeçam a determinadas regras comuns e compartilhadas sobre como armazenar dados e descrever a informação armazenada e que esta possa ser “consumida” por outros usuários humanos ou não, de maneira automática e não ambígua.” (SOUZA e ALVARENGA, 2004, p. 134)

Diante do exposto, analisa-se nesta pesquisa a aplicação do padrão MTD-BR de metadados no contexto do tratamento de documentos da Biblioteca Digital de Teses e Dissertações - BDTD. Para isso, estuda-se os padrões de metadados *Dublin Core*, MTD-BR e ETD-MS, bem como examina-se a consistência do conteúdo dos metadados das dissertações eletrônicas disponibilizados pela Universidade do Estado de Santa Catarina – UDESC para a BDTD. Acredita-se que esta pesquisa possa vir a contribuir para que a recuperação das dissertações eletrônicas da UDESC na BDTD torne-se mais eficiente, assim como aprimorar a interoperabilidade entre bibliotecas operantes neste repositório digital.

1.1 Justificativa

Kent (1972) pontua que o desenvolvimento e a utilização eficiente de qualquer sistema deve basear-se numa compreensão clara dos materiais e dos problemas que tal sistema tem de enfrentar. O mesmo autor ressalta que:

“A experiência na recuperação da informação mostra que, em geral, deve-se manusear os seguintes tipos de material:

- (1) Arquivos grandes, capazes de conter muitos documentos.
- (2) Materiais complexos, geralmente abrangendo vários assuntos de interesse potencial.
- (3) Informações diversas, exigindo quase sempre a seleção entre vários documentos e a subsequente correlação”. (KENT, 1972, p. 24)

Além disso, é necessário levar em consideração que um SRI pode ser utilizado para finalidades distintas, por usuários distintos e em locais distintos. Neste sentido, Salton (1975) sinaliza que um SRI pode ser usado para uma variedade de finalidades bastante diferentes, muitas vezes por um grande grupo usuário localizado a grandes distâncias da estação central de arquivos.

Para Choo (2006), recuperar uma informação é disponibilizá-la ao usuário/consulente, que a solicita por necessidades espontâneas e/ou induzidas, objetivando construir significado, produzir novo conhecimento e tomar decisões, sejam administrativas, sejam pessoais. É importante ressaltar que a RI engloba operações consecutivas para a localização de documentos que contenham uma informação necessária, assim como para a recuperação dos mesmos. Seu propósito é satisfazer uma necessidade de informação do usuário, a qual é expressa em uma consulta que procura encontrar respostas consideráveis. Entretanto, esta consulta pode não expressar, da melhor forma, a necessidade de informação, e pode conter palavras com erros ortográficos ou estar sendo usada de forma inadequada.

Segundo Araújo (1995), os SRIs, entre outras funções, objetivam dar acesso às informações potencialmente contidas em documentos neles registrados e serão usados indistintamente. Estes sistemas devem tentar recuperar documentos relevantes para satisfazer a necessidade do usuário, baseado na informação fornecida pela consulta. Uma grande dificuldade na RI é saber quando uma informação é relevante ou não para a necessidade do usuário. Baeza-Yates e Ribeiro Netto (1999) enfatizam que esse conceito é básico na RI, pois o seu objetivo primordial consiste em recuperar todas as informações relevantes e a menor quantidade possível de informações irrelevantes. Lancaster (2004) define um documento pertinente (útil) como aquele que contribui para satisfazer alguma necessidade de informação. Neste sentido, Balby (2002) defende que

somente a pessoa que expressa a necessidade de informação poderia decidir sobre o que é ou não pertinente.

Os primeiros sistemas de busca booleana foram SRIs baseados em palavras-chave. Esses sistemas recebem uma consulta e devolvem as informações que contenham literalmente alguma palavra desta. A linguagem como ela é usada em textos reais possui várias características complexas e pouco previsíveis, por exemplo, exceções, que podem levar os SRIs baseados em palavras-chave a confundir textos relevantes e irrelevantes. Para Mauldin (1991), as mais comuns destas características são sinonímia e polissemia. Segundo (Cintra et al. 2002), a sinonímia ocorre quando existe a relação de equivalência entre, ao menos, duas palavras e a polissemia, quando uma palavra comporta mais de um significado.

Um SRI apresenta dois processos principais bem definidos e diferenciados: o processo de indexação, no qual a coleção de informações é indexada para permitir uma recuperação posterior eficiente; e o processo de recuperação no qual, dada uma consulta, as informações mais fortemente relacionadas a ela são recuperadas. Antes da indexação, é importante lembrar que para que o assunto de um documento seja definido em termos adequados, é necessário que, primeiramente, sejam extraídos os conceitos que nele estão contidos. Dahlberg (1978), em sua Teoria do Conceito, define conceito como a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo lingüístico. Ainda segundo esta autora, esse símbolo pode ser verbal ou não-verbal, ou seja, pode ser formado de sinais ou conjunto de sinais independentes das palavras.

Resgatando o que foi dito na seção anterior, a exaustividade e a especificidade são duas variáveis que interferem na escolha de conceitos. Estas duas variáveis dependem da política de indexação adotada pelo SRI, que deve obedecer critérios pré-estabelecidos, levando em consideração os objetivos do sistema, os recursos disponíveis e características dos usuários. A exaustividade está relacionada à possibilidade de indexar o documento em profundidade, ou seja, além de indexar o assunto principal, o SRI indexa também, os assuntos secundários. A especificidade é a capacidade do SRI de permitir precisão na especificação do assunto de um documento. Os fatores de exaustividade e especificidade influenciam em todo o processo de RI, existindo uma ligação desses com as medidas de precisão e revocação.

Pode-se ressaltar que sistemas de busca booleana baseados em palavras-chave são muito utilizados na atualidade, devido à sua fácil aplicabilidade. Mas, de acordo

com Campos e Bax (2000), estes processos estão longe da perfeição e dependem, em grande parte, do preparo prévio dos documentos a serem indexados, uma tarefa para especialistas. Atualmente, o uso cada vez mais freqüente de novas tecnologias contribui para o surgimento de mudanças neste sentido. Bräscher (2002) ressalta que:

“As tecnologias da informação vêm provocando mudanças profundas nos processos tradicionais de comunicação científica, quase que eliminando o espaço de tempo entre a produção e a disseminação dos textos científicos. Essas mudanças, conseqüentemente, afetam os processos de tratamento da informação utilizados pela Ciência da Informação”.

(BRÄSCHER, 2002, p. 1)

De acordo com Castro e Santos (2007), no cenário atual, marcado principalmente pelo uso intensivo de tecnologias, em especial, as Tecnologias da Informação e Comunicação - TICs, é possível verificar uma mudança de paradigmas nos ambientes informacionais, tais como as bibliotecas, caracterizadas como espaços de preservação do conhecimento acumulado pela humanidade.

Inseridas neste contexto, encontram-se as bibliotecas digitais que, para alcançarem seus objetivos, precisam organizar o seu acervo seguindo uma padronização para a descrição das informações, a fim de possibilitar o acesso e a recuperação eficientes. Além disso, é importante destacar que o desenvolvimento e a utilização de novas tecnologias que considerem a informação como um conjunto de conceitos são essenciais para a evolução das bibliotecas, principalmente as digitais. Castro e Santos (2007) ressaltam que:

“A evolução das bibliotecas tem merecido destaque ao longo destes anos, no que se refere ao desenvolvimento e ao uso de tecnologias principalmente as da informação e comunicação que potencializam e geram recursos de acesso e formatos de intercâmbio para o tratamento das informações bibliográficas.”

(CASTRO e SANTOS, 2007, p. 14)

O uso das TICs cada vez mais freqüente nos últimos anos contribuíram para o aumento das informações em formato digital e para a utilização de redes de comunicação. Estes fatos facilitaram e incentivaram a troca de informações em diversos formatos, influenciando diretamente na história das bibliotecas digitais. Além disso, permitiram que fossem ampliadas as possibilidades de descrever, armazenar, recuperar, preservar e utilizar a informação. Entretanto, tem como uma das conseqüências o aumento do volume de informações oriundas de diferentes idiomas e culturas, que é disponibilizado na *web* em formato digital não padronizado, o que traz como conseqüência sérios problemas de interoperabilidade e recuperação. E, sendo a informação essencial para o desenvolvimento de qualquer organização, em especial, para as bibliotecas digitais, é necessário encontrar soluções para as dificuldades em recuperar informações relevantes. Dias (2001) ressalta que:

“[...] mesmo no caso de acervos cuja íntegra dos textos se encontra disponível em forma eletrônica, continua havendo a necessidade de um mecanismo que permita aos usuários identificar material de interesse, selecionar material que melhor atenda suas necessidades, encontrar grupos de documentos similares e localizar esse material”. (DIAS, 2001, p. 7)

As bibliotecas digitais utilizam metadados para padronizar as formas de representação e garantir interoperabilidade entre os sistemas. Para Alves (2005):

“[...] metadados são conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informacional de um recurso que pode estar em meio eletrônico ou não. Já os formatos de metadados, também chamados de padrões de metadados, são estruturas padronizadas para a representação do conteúdo informacional que será representado pelo conjunto de dados-atributos (metadados).” (ALVES, 2005, p. 115)

É importante destacar que a utilização de metadados, além de possibilitar a padronização das formas de representação do conteúdo informacional, tem como

objetivo garantir a interoperabilidade entre sistemas. De acordo com Castro e Santos (2007):

“Um dos objetivos do uso de metadados no contexto da *web* e das bibliotecas digitais é permitir não só descrever documentos eletrônicos e informações em geral, possibilitando sua avaliação de relevância por usuários humanos, mas também permitir agenciar computadores e programas especiais, robôs e agentes de *software*, para que eles compreendam os metadados associados a documentos e possam então recuperá-los, avaliar sua relevância e manipulá-los com mais eficiência”. (CASTRO e SANTOS, 2007, p. 16)

Entretanto, para que a utilização de metadados para a recuperação das teses e dissertações da BDTD atinja plenamente seus objetivos, é fundamental que os mesmos sejam informados corretamente, de acordo com o padrão MTD-BR. Neste sentido, Rosetto (2003) ressalta que o tratamento da informação, de acordo com padrões internacionais para o armazenamento em bases de dados e disponíveis em ambientes digitais, continua sendo imprescindível e importante para a recuperação adequada.

Sendo assim, os metadados passam a merecer cada vez mais a atenção da comunidade científica, uma vez que serão os responsáveis por facilitar a busca e a recuperação da informação, além de garantir interoperabilidade entre diversos sistemas e plataformas computacionais, em ambientes informacionais digitais.

Diante do exposto, acredita-se que pesquisas que reflitam sobre a organização e representação da informação em ambientes digitais contribuam para a eficiência da recuperação.

Em decorrência destes fatores que influenciam muito no desenvolvimento das organizações, realizar uma pesquisa sobre a padronização do conteúdo dos metadados disponibilizados pela UDESC para a BDTD, vem ressaltar a importância destes como meio facilitador de localização da informação e de interoperabilidade, contribuindo para o aperfeiçoamento da performance do SRI desta biblioteca digital.

1.2 Objetivos

Levando em consideração que:

- as bibliotecas digitais são instrumentos que visam a integração social e intelectual dos usuários,
- essa integração só é possível por meio de um tratamento eficiente do conteúdo físico e temático;
- esses tratamentos dependem da adoção de normas, padrões, formatos e protocolos que assegurem o acesso da informação e a comunicação entre os sistemas (interoperabilidade);
- a inconsistência no preenchimento dos metadados prejudica a circulação da comunicação científica e tecnológica.

Esta pesquisa propõe os seguintes objetivos:

1.2.1 Objetivo Geral:

Analisar o conteúdo dos metadados das dissertações eletrônicas da UDESC, disponibilizados para a BDTD, a fim de verificar se os mesmos estão sendo informados de acordo com o padrão MTD-BR e, desta forma, contribuir para uma maior e melhor visibilidade da produção científica e tecnológica.

1.2.2 Objetivos específicos:

- Levantar os metadados descritivos alimentados pela UDESC e disponibilizados para a BDTD;
- Organizar esses metadados a fim de identificar quais se destinam à descrição física e quais se destinam à descrição de conteúdo do documento;
- Examinar o conteúdo desses metadados com base nas regras estabelecidas pelos modelos adotados para os elementos do padrão MTD-BR de metadados e para seus respectivos atributos;
- Propor recomendações que objetivem a melhoria do conteúdo dos metadados informados pela UDESC para a BDTD.

2 Referencial Teórico

A seguir será apresentado o referencial teórico que fundamentou a pesquisa e serviu de base para a análise dos dados.

2.1 Ciência da Informação e Organização da Informação

A área da Ciência da Informação é nova; teve o início de sua organização como ciência, no início da década de 60. Barreto (2008) relata que:

“Alguns estudos indicam o começo da ciência da informação relacionando-o a uma pequena reunião realizada em 1961 e de novo em 1962 no *Georgia Institute of Technology* no Estado da Georgia, nos EUA. Esta Conferência chamou-se "*Conferences on training science information specialists, october 12-13, 1961 [and] April 12-13, 1962*". Agregou um total de cerca de 60 pessoas, somando os dois anos de sua realização. A maioria dos participantes foram docentes e bibliotecários da própria universidade americana sede do evento. A reunião da Geórgia tratou, particularmente, do treinamento de especialistas da informação e unicamente no contexto dos EUA é o que revela as parcas 100 laudas de seus Anais.” (BARRETO, 2008, p.8)

A origem da Ciência da Informação está diretamente relacionada com a explosão informacional. De acordo com Le Coadic (1996), o fato do crescimento do volume de informação foi primordial para o surgimento e o desenvolvimento da Ciência da Informação. Embora tenha se iniciado de forma modesta, sua importância para a sociedade em geral tem sido demonstrada desde então. Seus principais objetivos são estudar as propriedades gerais da informação e analisar os processos de construção, comunicação e uso da mesma. De acordo com Novellino (1996):

“A Ciência da Informação é uma disciplina voltada para o estudo de fenômenos subjacentes à produção, circulação e uso da informação. O estudo desses fenômenos tem como finalidade

possibilitar a criação de instrumentos e o estabelecimento de metodologias que viabilizem a transferência de informações”. (NOVELLINO, 1996, p. 37)

Mesmo tendo uma forte associação com a Biblioteconomia, é uma ciência de abrangência interdisciplinar que se interliga a diversas áreas do saber. Segundo Saracevic (1996), a Ciência da Informação estabelece relações com diversas áreas ou campos do conhecimento, entre as áreas estão: a Biblioteconomia, a Ciência da Computação, as Ciências Cognitivas e a Comunicação. Para Le Coadic (1996):

“A ciência da informação é uma dessas novas interdisciplinas, um desses novos campos de conhecimentos onde colaboram entre si, principalmente, a psicologia, a lingüística, a sociologia, a informática, a matemática, a lógica, a estatística, a eletrônica, a economia, o direito, a filosofia, a política e as telecomunicações”. (LE COADIC, 1996, p.22)

A percepção de Pinheiro (1997) sobre o assunto é que:

“[...] a Ciência da Informação tem seu próprio estatuto científico, como ciência social que é, portanto, interdisciplinar por natureza, e apresenta interfaces com a Biblioteconomia, Ciência da Computação, Ciência Cognitiva, Sociologia da Ciência e Comunicação, entre outras áreas, e suas raízes, em princípio, vêm da bifurcação da Documentação/Bibliografia e da Recuperação da Informação.” (PINHEIRO, 1997, p. 1)

Saracevic (1996) afirma que:

"a ciência da informação é um campo dedicado às questões científicas e à prática profissional voltadas para os problemas de efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual

do uso e das necessidades de informação. No tratamento destas questões são consideradas de particular interesse as vantagens das modernas tecnologias informacionais". (SARACEVIC, 1996, p. 47)

A definição de Ciência da Informação de Saracevic (1996) demonstra a evolução desta área, ressaltando sua relação com a comunicação humana e a importância da contribuição oferecida pelas novas tecnologias informacionais. O que vem ao encontro da seguinte afirmação de Lima (2003):

“O campo da ciência da informação (CI) tem consolidado seus pontos de vistas e valores desde que ela foi reconhecida como uma disciplina social e científica nos anos 60 e, mais recentemente, como uma área interdisciplinar que reúne a biblioteconomia, a ciência cognitiva (CC), a ciência da computação e a comunicação, com íntima dependência dos processos da comunicação humanos e da tecnologia no seu contexto contemporâneo.” (LIMA, 2003, p. 86)

Lima (2006) pontua que o grande volume de informações, gerado no crescente número de áreas do conhecimento, a partir da revolução técnico-científica ocorrida após a Segunda Guerra Mundial, passou a demandar um nível maior de organização. A organização da informação assume neste contexto um importante papel no processo de criação do conhecimento. Para Svenonius (2001), o objetivo principal da organização da informação é agrupar informações semelhantes e separar as informações que são diferentes. Estando relacionada aos documentos, a organização da informação procura ordená-los para torná-los disponíveis. O empenho para realizar esta tarefa se concentra em estabelecer procedimentos para localizar a informação. É pelo processo de tratamento da informação que estes procedimentos são reconhecidos. Silva (2007) pondera que o tratamento da informação não pode alterar o documento, porém pode criar novas informações a partir dele. O mesmo autor afirma que o processo de tratamento da informação é feito a partir de metodologias formalizadas de reconhecimentos temáticos e descritivos do conhecimento inscrito num suporte documental. A organização da informação é realizada por meio dos procedimentos para

localização desta. Entretanto, Silva (2007) ressalta que “a organização da informação não se limita a um conjunto de procedimentos, mas se realiza através deles a partir do tratamento da informação, que é orientado por recursos¹ bem definidos, adotados por instituições ou ambientes específicos destas instituições”. Lima (2006) ressalta que os profissionais da CI buscam trabalhar com a organização e acesso aos sistemas de informação, habilitando-se para a classificação, indexação e busca em sistemas informatizados.

Barreto (2008) ressalta que a partir de 1990, após a internet, a informação assumiu um novo status, principalmente com a sua interface gráfica a *World Wide Web* (www). Assim que se iniciou o uso da internet, os sites costumavam apresentar poucas páginas; o que possibilitava que apenas um responsável pela atualização dos mesmos os mantivesse atualizados, com facilidade. Entretanto, com o aumento do volume de acessos à internet, o volume de informações disponibilizadas na *web* tomou proporções muito grandes; o que faz com que mesmo uma equipe de profissionais encontre dificuldades para publicar todo o volume de informações produzido diariamente, atualizar o conteúdo de documentos já publicados e excluir os documentos desnecessários; o que contribui para dificultar o acesso à informação. Souza e Alvarenga (2004), sinalizam que:

“Embora tenha sido projetada para possibilitar o fácil acesso, intercâmbio e a recuperação de informações, a *Web* foi implementada de forma descentralizada e quase anárquica; cresceu de maneira exponencial e caótica e se apresenta hoje como um imenso repositório de documentos que deixa muito a desejar quando precisamos recuperar aquilo de que temos necessidade. Não há nenhuma estratégia abrangente e satisfatória para a indexação dos documentos nela contidos, e a recuperação das informações, possível por meio dos "motores de busca" (*search engines*), é baseada primariamente em palavras-chave contidas no texto dos documentos originais, o que é muito pouco eficaz.” (SOUZA e ALVARENGA, 2004, p. 3)

¹ Entenda-se por recursos, os processos, métodos, e produtos como a indexação, catalogação, classificação, tesouro, vocabulários, índices, listas, dicionários, etc.

Diante desta realidade, torna-se essencial o desenvolvimento de ações direcionadas à organização e ao acesso da informação na *web*. Neste sentido, Souza et al. (2000) chama atenção para o seguinte:

“O inquestionável crescimento da Internet e o incessante aumento no volume de recursos eletrônicos disponíveis na *web* vêm demandando ações voltadas à organização e ao acesso à informação eletrônica, haja visto o inexpressivo uso de padrões de tratamento e classificação da informação, o que torna extremamente difícil o seu manuseio eficaz.” (SOUZA et al., 2000, p. 2)

Sendo a informação essencial para o desenvolvimento das organizações, estas têm se dedicado intensamente ao desenvolvimento e uso de padrões que possibilitem a organização e o acesso à informação disponibilizada na *web*. Inseridas neste contexto, as bibliotecas digitais, que são ambientes propícios para a recuperação de informações, utilizam metadados para padronizar as formas de representação e possibilitar a interoperabilidade entre os sistemas, tornando eficientes os processos de recuperação e acesso à informação. Na literatura, a definição mais utilizada para metadados, é dado sobre dado. Segundo Milstead e Feldman (1999), metadados descrevem os atributos e o conteúdo de um documento original, e se usados de forma efetiva, viabilizam o acesso à informação precisa. De acordo com Souza et al. (2000), a representação da informação por metadados é uma necessidade e visa criar uma estrutura de descrição padronizada de documentos eletrônicos, para tornar mais efetiva a recuperação de informação em redes eletrônicas. No capítulo 2.3, Metadados, será dada ênfase a este assunto.

2.2 Bibliotecas Digitais de Teses e Dissertações

A geração de conhecimento, assim como a transmissão deste, no âmbito das universidades, se dá por meio da informação científica. Neste contexto, a biblioteca assume a importante responsabilidade de satisfazer a necessidade informacional dos acadêmicos, assim como, de gerir a produção intelectual da instituição.

Os serviços oferecidos pela biblioteca, principalmente o intercâmbio de acervos entre as instituições, viabilizam a criação de uma rede dinâmica para a informação

trafegar e, desta forma, contribuem para a expansão do fluxo da comunicação científica. Lima (2004) enfatiza que:

“As bibliotecas não são apenas lugares onde se preserva a memória e se possibilita a continuidade de uma sociedade, mas principalmente lugares onde ocorrem rupturas com a tradição. A história das bibliotecas e a história de como elas classificam os conhecimentos e conseqüentemente as ciências é também a história do que uma sociedade, as instâncias de poder e um meio intelectual decidem transmitir”. (LIMA, 2004, p. 46)

Atualmente, a produção científica das universidades está passando a ser armazenada em formato digital, em ambientes chamados de repositórios institucionais, nos quais os pesquisadores podem disponibilizar seus trabalhos. Desta forma, contribuem para que as pesquisas realizadas na sua instituição possam ser acessadas com mais frequência, colaborando para a promoção da instituição a qual pertencem, além de garantir a preservação da produção científica.

Na biblioteca tradicional, o acesso aos exemplares das teses e dissertações pode ser feito pelo contato com a instituição que armazena o exemplar desejado ou por meio do Sistema de Comutação Bibliográfica – COMUT, que permite a solicitação das cópias que interessam. Com o surgimento das bibliotecas digitais e suas tecnologias, novas formas de representação de forma e conteúdo têm surgido, contribuindo para o aumento da eficiência do processo de recuperação da informação. Entretanto, muitas das funções básicas de ambas, continuam sendo as mesmas. Dias (2001) destaca que vamos encontrar nos diversos tipos de bibliotecas e bibliografias as mesmas funções básicas: a função de selecionar os materiais de informação, de aplicar-lhes um tratamento, a função de dialogar com o usuário em torno de uma consulta para melhor defini-la, entre tantas outras funções.

O desenvolvimento constante de novas TICs e o uso cada vez mais freqüente das redes de comunicação têm colaborado para o surgimento de diversas mudanças no cenário atual da informação. Estas mudanças facilitam muito o acesso e uso da informação, uma vez que, encurtando distâncias e diminuindo consideravelmente a necessidade de tempo para as buscas, permitem que qualquer indivíduo, dos mais diversos locais, possam acessar e usar a informação no momento que dela precisam.

A disponibilização de novos produtos e serviços assegurado pela evolução das TIC's tem possibilitado novas formas de busca, acesso e disseminação da informação científica. Fato este que contribui para que os usuários da informação tornem-se cada vez mais independentes para selecionarem a informação que buscam. O novo contexto da informação permite que diversas bases de dados eletrônicas possam ser acessadas via internet, tornando-se assim muito úteis para pesquisas científicas. Convém destacar que as redes eletrônicas tornam-se extremamente importantes para a evolução da diversificação na transmissão da informação científica de uma determinada área da ciência para outra. Entretanto, deve ser levado em consideração que nem todos os indivíduos têm acesso às facilidades oferecidas pelas TICs. Dias (2001) sinaliza que embora pudéssemos aceitar a informação, defendida por muitos, de que hoje em dia todo o conhecimento está disponível na internet, coisa muito diferente é essas informações estarem disponíveis para toda a população.

As bibliotecas universitárias brasileiras, cujos catálogos já haviam sido informatizados para garantir o acesso em rede local, podem disponibilizá-los, atualmente, na grande rede mundial denominada internet. O desenvolvimento e o uso de novas tecnologias, que possibilitam novas formas de acesso e formatos para intercâmbio, têm contribuído para a evolução das bibliotecas ao longo do tempo.

Diante das novas possibilidades oferecidas pelas TIC's, as bibliotecas digitais estão tornando cada vez mais forte a idéia de preservação, conservação e, principalmente, acesso à informação. Para Borgman (1996), a definição de bibliotecas digitais pode ser feita por duas idéias:

1. Bibliotecas Digitais são um conjunto de fontes eletrônicas e serviços técnicos associados para a criação, pesquisa e uso da informação, que possibilitam uma extensão e um aumento do armazenamento da informação e dos sistemas de recuperação, manipulando dados digitais em qualquer formato (texto, imagens, sons; imagens dinâmicas e estáticas) em redes distribuídas de trabalho. O conteúdo das bibliotecas digitais inclui dados e metadados que descrevem vários aspectos do dado (representação, criador, dono, direitos de reprodução) e metadados que consistem em ligações ou relacionamentos com outros dados ou metadados, sejam esses externos ou internos à biblioteca digital.

2. Bibliotecas Digitais são construídas, coletadas e organizadas, por (e para) uma comunidade de usuários, e suas capacidades funcionais satisfazem as necessidades de informação e usos dessa comunidade. Nelas os indivíduos e grupos interagem uns com os outros, usando sistemas e fontes de dados, informação e conhecimento. Neste sentido, representa extensão, aumento e integração de uma variedade de instituições de informação, como locais físicos, onde recursos são selecionados, coletados, organizados, preservados, e acessados em função de uma comunidade de usuários. Estas instituições de informação incluem, entre outras, bibliotecas, museus, arquivos e escolas, mas as bibliotecas digitais também se estendem e servem outros ambientes de comunidades, incluindo salas de aulas, escritórios, laboratórios, casas e espaços públicos. (BORGMAN, 1996, p.6)

Sayão (2007), ressalta que:

“Uma biblioteca digital – no seu sentido pleno - não é meramente um repositório ou uma coleção de informações em formato digital; também não é somente uma tecnologia ou um conjunto de tecnologias que se pode avaliar isoladamente. Antes disso, é um sistema aberto, de múltiplas interligações e múltiplos subsistemas, envolvendo um ambiente organizacional, profissionais especializados provenientes de diversas áreas, recursos informacionais, usuários claramente definidos, tecnologia de informação, procedimentos, padrões e protocolos e, não menos importante, compromissos de longo prazo.” (SAYÃO, 2007, p. 19)

Neste contexto, o mesmo autor afirma que as normas, padrões, formatos e protocolos cumprem um papel de fundamental importância, já que estabelecem as regras pelas quais os objetos são descritos, identificados e preservados, seus dados são armazenados, e os sistemas aos quais estão inseridos se comunicam.

Mas, mesmo diante de tantas facilidades oferecidas pelo uso de bibliotecas digitais, é necessário levar em consideração que existem algumas desvantagens. Como por exemplo: a obsolescência dos equipamentos e sistemas utilizados que, assim como as interfaces pouco amigáveis, podem dificultar o acesso e a inexistência de infra-estruturas necessárias que, assim como a complexidade dos sistemas utilizados, podem levar à info-exclusão. Lembremos também, que a interação social, que ocorre nas bibliotecas tradicionais, não ocorrerá da mesma forma nas bibliotecas digitais, embora ambas prestem serviços para usuários de diversos perfis. Neste sentido, Rasche e Varvakis (2006) ressaltam que:

“No atendimento ao público, há a convivência com o heterogêneo, da dona de casa ao intelectual, em tese todos têm o direito de acesso à biblioteca. Da busca pelo silêncio, de um lugar para ficar só com suas próprias divagações à busca por alguém para conversar sobre o livro lido, as necessidades e as expectativas são diferentes”. (RASCHE e VARVAKIS, 2006, p. 127)

Entretanto, Dias (2001) defende que as bibliotecas digitais parecem estar se firmando como a expressão que significaria, no contexto digital, um conjunto de artefatos, conhecimento, práticas e uma comunidade, que engendra compromissos realísticos assumidos por profissionais da informação, analistas de sistemas e usuários.

A principal diferença entre a biblioteca tradicional e a biblioteca digital é a ampliação das formas de tratamento, acesso e recuperação das informações. Cunha (1999) sinaliza que a futura biblioteca coexistirá em um ambiente onde os usuários estarão conectados a uma ampla variedade de recursos informacionais que muitas bibliotecas não poderão prover, a menos que se conectem a tais recursos. Mas, Dias (2001) ressalta que, em ambos os contextos, o tratamento da informação permanece como uma atividade essencial, pois cumpre função fundamental no trabalho de facilitar o acesso à informação.

Muitas universidades brasileiras oferecem, além dos cursos de graduação, cursos de pós-graduação que, abrangem os níveis de mestrado e doutorado. Estes cursos têm como finalidade principal capacitar professores para o ensino superior e formar pesquisadores e profissionais qualificados. Os programas de pós-graduação têm se

tornado responsáveis por grande parte da produção científica brasileira; as teses e dissertações geradas nestes programas estão contribuindo, cada vez mais, para a evolução do conhecimento científico. De acordo com Campello (2000), no Brasil, o termo dissertação está associado ao grau ou título de mestre, e o termo tese ao grau de doutor.

A organização das Bibliotecas Digitais de Teses e Dissertações pode ser considerada uma tendência mundial para o desenvolvimento e construção de projetos para bibliotecas integradas e sistemas de informação. A disponibilização de teses e dissertações em formato eletrônico tem como principal objetivo, permitir a representação (abstração) do modo físico/conhecimento físico. São esses aspectos que justificam o conceito/concepção de “representação”. Por exemplo: se um documento que armazena uma tese ou dissertação, em formato eletrônico, não for representado corretamente ou, se for representado por meio de um formato digital que não é o mais apropriado, este documento não será recuperado. A criação das teses e dissertações eletrônicas permite que os acadêmicos se tornem cada vez mais atuantes na produção de conhecimento, facilita o compartilhamento de pesquisas e atende a demanda informacional das universidades.

Uma das grandes vantagens oferecidas pelas teses e dissertações em formato eletrônico é a redução de tempo e de custo para a realização das buscas de documentos possibilitada pela centralização e extensa divulgação dos trabalhos científicos. Esta vantagem torna possível a execução de consultas simultâneas dos mais diversos locais, facilitando o acesso à informação e contribuindo para o aumento da visibilidade das pesquisas científicas. Outro benefício que merece destaque refere-se à preservação, uma vez que torna-se mais fácil e prático o armazenamento dos documentos, em meio digital.

No Brasil, as bibliotecas digitais de teses e dissertações ocupam cada vez mais o seu espaço com o propósito de disponibilizar acervos em formatos digitais, visando à conservação e a disponibilização de seus conteúdos e assegurando o compartilhamento da informação para os mais diversos locais do mundo. Também estão sendo realizados muitos encontros que abordam temas relacionados as bibliotecas digitais, fato este que demonstra o interesse dos profissionais brasileiros pelo assunto.

2.2.1 *Networked Digital Library of Theses and Dissertations (NDLTD)*

A *Networked Digital Library of Theses e Dissertations* (NDLTD) é uma organização internacional que se dedica a promover a adoção, criação, uso, disseminação e preservação das cópias das teses e dissertações publicadas e tem como objetivo principal o desenvolvimento de uma rede mundial de teses e dissertações digitais. Para atingir seus objetivos, encarrega-se de patrocinar e incentivar as instituições interessadas em contribuir para a evolução das teses e dissertações eletrônicas, realizando estudos e atividades sobre o assunto.

Os países que participam da NDLTD, visando contribuir para a ampliação do desenvolvimento das bibliotecas digitais, se propõem a compartilhar experiências, resultantes de pesquisas por eles realizadas, com os países em desenvolvimento, a fim de que as mesmas sirvam-lhes de base. (NETWORKED, 2008)

Desde a sua criação em 1996, a NDLTD tem trabalhado para melhorar a educação, aumentar a disponibilidade do aluno para pesquisa, capacitar os estudantes e as universidades e reduzir os custos da apresentação e manipulação de teses e dissertações eletrônicas (ETDs).

A fim de melhor servir os seus membros, em maio de 2003, foi devidamente formada como uma corporação mundial para fins educativos e de caridade. É atualmente liderada por um Conselho de Administração, em colaboração com os membros de várias comissões adicionais sobre os objetivos da organização.

O projeto NDLTD se propõe a disponibilizar coleções individuais de ETDs, arquivadas em instituições distintas, para estudantes e pesquisadores que procuram teses e dissertações em uma biblioteca digital. Apóia as bibliotecas digitais e promove encontros anuais, possibilitando a integração das universidades por meio de uma cooperação entre as mesmas. Também assegura aos estudantes a possibilidade de aprender pela prática e estimula a participação dos mesmos na comunicação dos resultados das pesquisas realizadas, uma vez que possibilita que eles próprios possam desenvolvê-las por um processo simples e objetivo.

A UNESCO (Organização das Nações Unidas para Educação, Ciência e Cultura), assim como outros grupos auxiliam neste processo, incentivando novos projetos na área. (NETWORKED, 2008)

Para tornar as ETDs de uma instituição acessíveis, primeiramente, o responsável pelas mesmas precisa disponibilizar seus metadados. Sendo assim, as instituições que

produzem e armazenam ETDs disponibilizam metadados, como título, autor, etc, para a NDLTD. Desta forma, é possível apresentá-las para os pesquisadores como uma única coleção e, quando um original relevante é encontrado, o pesquisador é direcionado à instituição que armazena este documento em formato digital.

O formato de metadados utilizado é o ETD-MS, um conjunto de padrões internacional para os elementos de metadados usados para a descrição de teses ou dissertações eletrônicas. Será dada ênfase a este assunto no capítulo 2.3.3.2 Padrão de Metadados ETD-MS.

Esta possibilidade de tornar metadados disponíveis para auxiliar na pesquisa é sustentada pela iniciativa dos arquivos abertos (*Open Archives Initiative - OAI*). A coleta dos metadados é feita pela utilização do protocolo OAI-PMH que torna possível a coleta de metadados de seus locais de origem para disponibilizá-los na NDLTD.

Os provedores de dados² da NDLTD são as chamadas instituições cooperantes³.

2.2.2 Biblioteca Digital de Teses e Dissertações (BDTD)

Trata-se de um projeto cujos principais objetivos são integrar os sistemas de informação de teses e dissertações das diversas Instituições de Ensino Superior (IES) brasileiras e estimular o registro e a publicação de teses e dissertações em formato digital.

O modelo da BDTD integra as iniciativas de registro bibliográfico e de publicação eletrônica de teses e dissertações existentes nos acervos das IES brasileiras. Ao integrar essas duas iniciativas, o IBICT amplia a abrangência da BDTD e disponibiliza para os usuários um catálogo nacional de teses e dissertações em texto integral e referencial provenientes das IES, possibilitando uma forma única de busca e acesso a estes documentos.

O desenvolvimento da BDTD ocorre no âmbito do programa da Biblioteca Digital Brasileira, que, portanto, vale-se das soluções tecnológicas que têm sido adotadas naquele programa. Segundo Baptista et al. (2007), para a implantação da BDTD, o IBICT elaborou e implantou o Padrão Brasileiro de Metadados para Teses e Dissertações (MTD-BR), o qual é um superconjunto do padrão ETD-MS, que por sua

² Os provedores de dados são sistemas que utilizam o protocolo OAI-PMH como um meio de expor metadados. Esses provedores também podem (mas não necessariamente) oferecer acesso gratuito a textos completos e a outros recursos. Fonte: < <http://clube-oai.incubadora.fapesp.br/portal/antigo/provedores>>

³ No anexo 1, encontra-se a lista das IES cooperantes da NDLTD.

vez é superconjunto do *Dublin Core*. A adoção do MTD-BR e a tecnologia de coleta automática de metadados (*harvesting*) foram particularmente importantes na implementação da BDTD. (INSTITUTO, 2008)

A BDTD se apóia em um modelo distribuído e utiliza as tecnologias de arquivos abertos, sendo as IES provedores de dados e o IBICT um agregador responsável pela coleta de metadados de teses e dissertações destes provedores de dados e pelo fornecimento de serviços de informação sobre esses metadados, expondo-os para serem coletados por outros provedores de serviços⁴. (BIBLIOTECA, 2008)

Para as IES que ainda não implementaram suas bibliotecas digitais de teses e dissertações, o IBICT oferece gratuitamente duas opções do Sistema de Publicação Eletrônica de Teses e Dissertações – TEDE:

- O TEDE simplificado que possibilita a publicação eletrônica da tese ou dissertação pela biblioteca da IES, mediante a autorização do autor da dissertação ou tese;
- O TEDE modular que necessita de uma infra-estrutura de integração entre o curso de pós-graduação, o autor da dissertação ou da tese e a biblioteca da IES no processo de publicação eletrônica, uma vez que o sistema possui funções específicas para cada etapa do processo de publicação das teses e dissertações.

Para as instituições que já implantaram suas bibliotecas digitais de teses e dissertações adotando tecnologias diferentes do sistema TEDE, o IBICT possibilita que se integrem à BDTD por meio da implementação do Protocolo OAI-PMH⁵ em seus repositórios de dados, expondo metadados no padrão MTD-BR. (INSTITUTO, 2008)

A BDTD se operacionaliza por meio da geração de uma base centralizada de metadados. Na fase inicial do projeto, as instituições consorciadas enviavam arquivos de metadados para alimentar a base centralizada. Posteriormente, foi implantado, no IBICT, o processo de coleta automática de metadados capaz de ir buscar os metadados diretamente nos servidores das instituições provedoras das informações. Em especial, participaram desta fase o IBICT (como provedor de serviços), a PUC-RIO e a USP,

⁴ Os provedores de serviços utilizam os metadados coletados via protocolo OAI-PMH como base para construção de novos serviços. Fonte: < <http://clube-oai.incubadora.fapesp.br/portal/antigo/provedores>>

⁵ O *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) é um protocolo que permite o compartilhamento de metadados, sem demandar custos elevados.

como fornecedoras de metadados (provedoras de dados) de teses e dissertações eletrônicas.

Assim como a BDTD utiliza a coleta de metadados nacionais para alimentar uma base centralizada, sistemas internacionais utilizam esta mesma base para realizarem coleta automática. Em particular, o IBICT promoveu a integração do Consórcio Brasileiro de Teses e Dissertações⁶ à NDLTD. Desta forma, a produção nacional de teses e dissertações é também disponibilizada internacionalmente. Esquemáticamente, a BDTD pode ser representada em três níveis de abrangência: local, nacional e internacional, conforme a figura abaixo:

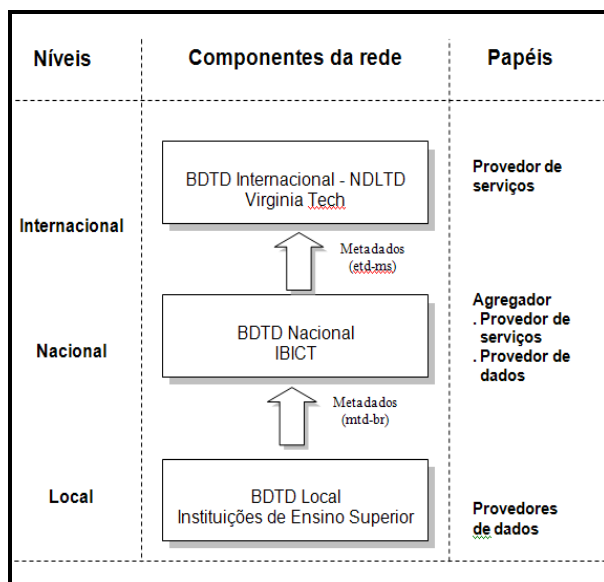


Figura 1 - Integração dos níveis de abrangência da BDTD

Fonte: (INSTITUTO, 2008)

Em termos gerais, o projeto da BDTD promove não só maior visibilidade da produção científica e tecnológica brasileira, em níveis nacional e internacional, mas também gera capacitação nacional nas tecnologias de informação e comunicação usadas para implementação de bibliotecas digitais.

Do ponto de vista do usuário final, ao acessar a BDTD, este pode realizar buscas de forma unificada (em uma única base de dados) às informações de teses e dissertações existentes nas diversas instituições consorciadas. As teses em meio eletrônico podem

⁶ O Consórcio Brasileiro de Teses e Dissertações é constituído pelas IES que mantêm bases de dados bibliográficas de teses e dissertações e/ou publicam esse tipo de documento em meio eletrônico. (IBICT, 2008)

ser recuperadas no servidor da instituição provedora da informação. (INSTITUTO, 2008)

As IES⁷ cooperam com a BDTD atuando como provedores de dados enquanto o IBICT atua como agregador, coletando os metadados das teses e dissertações dos provedores de dados, provendo serviços de informação sobre esses metadados e expondo-os para serem coletados por outros provedores de serviços.

2.3 Metadados

Embora, atualmente, o termo informação seja amplamente utilizado, se constituindo numa presença constante, tanto no vocabulário técnico-científico quanto na fala de uma grande parte das pessoas comuns, integrantes de todas as classes sociais, nada é menos seguro do que esse vocábulo possuir o mesmo significado para todos os que o utilizam ou dele ouvem falar. De acordo com Robredo (2003), a idéia mais apropriada sobre o significado da informação é a de documentação no seu mais amplo sentido, incluindo o registro e transmissão do conhecimento, o armazenamento, processamento, análise, organização e recuperação da informação registrada, e os processos e técnicas relacionados.

A recuperação da informação permite aos usuários que informações sejam localizadas por meio de consultas realizadas. Seu propósito é a satisfação de uma necessidade de informação do usuário, a qual é expressa em uma consulta que procura encontrar respostas consideráveis. Para Ferneda (2003):

“O processo de recuperação da informação consiste em identificar, no conjunto de informações (*corpus*) de um sistema, quais atendem à necessidade de informação do usuário. O usuário de um sistema de recuperação da informação está, portanto, interessado em recuperar “informação” sobre um determinado assunto e não recuperar dados que satisfazem sua expressão de busca. Essa característica é que diferencia os sistemas de recuperação da informação dos sistemas gerenciadores de banco de dados, estudados e implementados

⁷ No anexo 2, encontra-se a lista de IES cooperantes da BDTD.

desde o nascimento da ciência da computação”. (FERNEDA, 2003, p. 14)

A evolução que vem ocorrendo nas pesquisas sobre recuperação da informação tem possibilitado o desenvolvimento de novas tecnologias para acesso e recuperação das informações. Entretanto, muitos SRIs, apesar de serem muito utilizados atualmente, não estão aptos a recuperarem informações relevantes para aqueles que as buscam; retornando respostas, muitas vezes, que não satisfazem sua necessidade. O grande volume de informações disponibilizadas na *web* sem uma forma estruturada de representação contribui para dificultar a recuperação da informação. Cardoso (2006) destaca que a crescente complexidade dos objetos armazenados e o grande volume de dados exigem processos de recuperação cada vez mais sofisticados.

A organização da informação é essencial para que a recuperação relevante da mesma tenha êxito. Segundo Café e Bräscher (2008), a descrição da informação é fundamental na recuperação da informação, pois cumpre a função de possibilitar o acesso a itens específicos que se encontram armazenados. Neste sentido, McGarry (1999) sinaliza que a informação deve ser ordenada, estruturada ou contida de alguma forma, senão permanecerá amorfa e inutilizável. Dentre as ferramentas disponíveis atualmente para a padronização da descrição das informações, em rede eletrônica, encontram-se os metadados. A padronização permite facilitar o acesso e a recuperação da informação, assegurando aos usuários a fácil movimentação de dados entre diferentes sistemas e plataformas computacionais.

Para muitos autores, metadados significa “dados sobre dados”. Da mesma forma, Takahashi (2000) conceitua metadados como: “dados a respeito de outros dados, ou seja, qualquer dado usado para auxiliar na identificação, descrição e localização de informações. Trata-se, em outras palavras, de dados estruturados que descrevem as características de um recurso de informação”. Neste sentido, metadado refere-se a uma estrutura descritiva da informação sobre outro dado. Segundo Berners-Lee (1997), metadados são definidos como uma máquina para compreensão das informações sobre os recursos *web* ou outras coisas.

Os metadados têm como objetivo viabilizar a compreensão da informação por humanos e sistemas computacionais, ajudando na identificação, descrição, localização e gerenciamento de recursos da *web*. Segundo Alves e Souza (2007):

“Os elementos de metadados têm o propósito primário de descrever, identificar e definir um recurso de informação com o objetivo de modelar e filtrar o acesso. Os metadados são importantes na organização, gestão e recuperação da informação digital, principalmente. Nesse sentido, são adotados procedimentos técnicos de catalogação, indexação e categorização dos conteúdos informacionais, o que possibilita a integração de fontes diversificadas e heterogêneas de informação”. (ALVES, SOUZA, 2007, p. 22)

Trata-se de um assunto que sempre esteve presente em vários domínios de aplicação. Entretanto, sua utilização passou a aumentar significativamente com a evolução das plataformas computacionais, tornando-se popular com o surgimento dos sistemas de gerenciamento de banco de dados relacionais, que os adaptavam para descrever as estruturas desses sistemas.

Iannella e Waugh (1997) afirmam que os metadados assumem um importante papel, o de apoiar o uso de recursos eletrônicos e serviços. Os mesmos autores pontuam que, contudo, para um apoio efetivo e implantação de sistemas de metadados muitas questões necessitam ainda de ser abordadas.

2.3.1 Estrutura

Para que as informações disponibilizadas na web possam ser utilizadas de maneira correta, são necessárias informações sobre o seu contexto, ou seja, sobre sua organização e sobre seu significado. Estas informações podem ser utilizadas como base para a determinação de relacionamentos entre as informações e as características do mundo real descrito por elas. De acordo com a biblioteca digital de Oxford (OXFORD, 2008), os metadados podem ser classificados em três categorias:

- **Descritivos:** descrevem o conteúdo de um documento, suas configurações ou um outro recurso que facilite a identificação, pesquisa, intercâmbio e gerenciamento das informações;

- Administrativos: fornecem uma grande variedade da informação relacionada à exposição, uso, interpretação e gerenciamento dos repositórios de objetos. Como exemplos, pode-se citar o seu formato de armazenamento;
- Estruturais: apresentam objetos digitais ao usuário, descrevendo-o em termos de uso e navegação. Uma imagem associada, arquivos de áudio ou de vídeo, podem ser citados como exemplos.

Sayão (2007), esclarece o seguinte:

“Metadados podem descrever recursos informacionais em qualquer nível de agregação – uma coleção, um recurso simples, ou um elemento que faz parte de um outro recurso, como, por exemplo, uma fotografia inserida num artigo; pode ainda ser usado para descrever um trabalho, bem como uma manifestação ou uma expressão desse trabalho, por exemplo, um relatório, uma edição particular desse relatório, ou uma cópia específica da edição. Os metadados podem estar embutidos num objeto digital – inscritos na sua codificação, como é comum nos documentos HTML; ou podem estar armazenados separadamente, estruturados em base de dados e vinculados ao objeto que eles descrevem, facilitando a busca e a recuperação.”
(SAYÃO, 2007, p. 34)

Sendo assim, é necessário elaborar um modelo de metadado para descrever o contexto da informação, evitando a ambigüidade ou a redundância. É conveniente utilizar uma conceitualização de um domínio específico de problema ou ontologias que forneça um acordo comum de vocabulários para que os dados sejam referenciados. Desta forma, uma ontologia pode servir de base comum para a representação de dados e metadados.

Os metadados podem ser utilizados em áreas distintas como educação à distância, data warehouse, serviços disponibilizados na *web* e *web* semântica, por exemplo. Dentre as vantagens oferecidas pelos metadados, pode-se citar:

- Interoperabilidade entre objetos distribuídos em plataformas distintas;
- Troca padronizada de dados entre componentes distribuídos;
- Padronização de objetos de aprendizagem;
- Descrição de serviços e conteúdo multimídia;
- Representação de informações contextuais.

Os metadados podem ser representados no próprio recurso ou isoladamente. Quando a representação é feita no próprio recurso, os metadados podem estar inseridos em páginas *web* desenvolvidas em “*Standard Generalized Markup Language*” (SGML), em “*HyperText Markup Language*” (HTML) ou em “*Extensible Markup Language*” (XML). Essas linguagens de marcação têm como objetivos principais possibilitar maior interação entre humanos e computadores e assegurar a interoperabilidade entre as diversas plataformas computacionais.

2.3.2 Interoperabilidade

Atualmente a grande maioria das informações disponibilizadas na *web* é gerada de forma independente, resultando em heterogeneidade e, conseqüentemente, contribuindo muito para dificultar o acesso e o compartilhamento das mesmas.

Para sanar dificuldades relacionadas a heterogeneidade das informações disponibilizadas na *web* faz-se necessária uma estrutura que facilite o acesso a essas informações, assegurando interoperabilidade entre as bases de dados e as informações que estão sendo buscadas. Entretanto, as características complexas e pouco previsíveis da linguagem utilizada em textos reais, como a sinonímia e a polissemia, por exemplo, são obstáculos difíceis de serem vencidos para que isso se torne possível.

Neste contexto, os metadados surgem como alternativa para auxiliar na resolução do problema de heterogeneidade, uma vez que possibilitam o acesso de maneira uniforme as informações disponibilizadas e a definição de significado para a informação. A utilização de metadados para a descrição de recursos garante que estes sejam compreendidos por diversos *softwares*, possibilitando a troca e o compartilhamento de dados entre diferentes plataformas de *software* e *hardware*. De acordo com Alves e Souza (2007):

“A capacidade de bases de dados trocarem e compartilharem documentos, consultas e serviços, usando diferentes plataformas de *hardware* e *software*, estrutura de dados e interfaces, é chamada de interoperabilidade. Através dessa troca e compartilhamento são realizadas interações entre sistemas. Porém, as informações devem estar organizadas eficientemente para que essas interações aconteçam. É a principal característica para o sucesso dessas interações é a consistência, a qual é alcançada através do uso de padrões.” (ALVES, SOUZA, 2007, p. 4)

É possível descrever os dados seguindo um único padrão de metadados. Mas, como existem diversos padrões de metadados e sendo a internet uma rede tão ampla e rica em informações, é comum utilizar padrões distintos para a descrição dos dados.

Ao abordar a questão da interoperabilidade, é importante destacar que os metadados podem ser codificados de diversas formas mas, no contexto da *web*, a que permite descrever as informações para a troca de metadados, da maneira mais adequada é a linguagem XML, considerada uma grande evolução na internet. Trata-se de uma especificação técnica desenvolvida pelo W3C⁸ (*World Wide Web Consortium*), órgão responsável pelo desenvolvimento de tecnologias para a interoperabilidade por meio de especificações, diretrizes, *softwares* e ferramentas, para superar limitações do HTML, o padrão atual das páginas da *web*. É definida como o formato universal para dados estruturados na *web*; esses dados podem ser tabelas, desenhos ou parâmetros de configuração, por exemplo. A linguagem trata da definição de regras, a fim de permitir que documentos sejam descritos de tal forma que sejam facilmente acessados pelo computador. De acordo com Souza e Alvarenga (2004), o XML é uma recomendação formal do W3C e, em determinados aspectos, assemelha-se ao HTML. A principal diferença entre o XML e o HTML é o fato de utilizar *tags*⁹. Em ambas as linguagens, cada *tag* consiste em duas partes, uma que inicia e outra que fecha o comando. No entanto, em muitos casos, se uma *tag* é aberta no HTML e não é fechada, a página é

⁸ URL do *World Wide Web Consortium* - W3C: <www.w3.org>

⁹ *Tags* são comandos inseridos em um documento que especifica como o documento ou uma parte dele deve ser formatado. São utilizadas no código html de uma página para indicar ao *browser* como a mesma será visualizada na tela. As *tags* vêm sempre entre sinais de < e >. A cada nova versão de *browsers*, novas *tags* são acrescentadas ao código de programação html pelo *World Wide Web Consortium*.

Fonte: <<http://dicionario.babylon.com/Tag>>

exibida mesmo assim. Já no XML, se houver qualquer erro desse tipo, a aplicação não será executada. Com esse exemplo, é possível perceber que o HTML é uma linguagem mais flexível, enquanto o XML é altamente rígido. Isso pode até parecer uma desvantagem, mas se for, é compensada pela extensibilidade do XML. Enquanto no HTML, uma *tag* indica o início e o fim de um parágrafo, no XML, as *tags* são usadas para definir blocos de dados, podendo significar qualquer coisa que o programador desejar, como nome, idade, ou endereço, por exemplo. Desta forma, no XML, as *tags* podem significar o que o usuário quiser representar. Essa característica faz com que o XML seja considerado por muitos como uma linguagem capaz de gerar outras linguagens, visto que quem define os comandos e funções é o programador. A praticidade torna possível que um usuário crie uma coleção própria de *tags* e aplique-as nas páginas e documentos que desejar. Souza e Alvarenga (2004) ressaltam que:

“Os dados contidos nos documentos XML podem ser exibidos em uma infinidade de maneiras, dependendo do dispositivo em que serão manuseados (telas de computador, celulares, PDAs, etc.). Os documentos XML não contém, em si, as diretivas para exibição dos dados, e, para cada dispositivo-destinado específico, podemos realizar uma transformação do documento originalmente em XML para um documento passível de ser exibido ao usuário ou entendido e utilizado por outro dispositivo tecnológico”. (SOUZA e ALVARENGA, 2004, p. 135)

Como principais funções do XML, pode-se citar:

- Descrição de dados;
- Apresentação de dados em algum formato, como HTML;
- Transporte de dados;
- Troca de dados de forma transparente entre plataformas diferentes.

A independência de plataforma é outra grande vantagem do XML em relação ao HTML; a manipulação dos dados no XML não depende de plataforma, versões e *software*. Desta forma, possibilita que o cliente tenha acesso aos dados, facilmente.

A realização do processamento de um grande número de informações, no servidor, pode comprometer a performance do sistema como um todo. Uma das formas de otimizar o sistema e diminuir a carga do servidor é dividir o processamento com o

cliente *web*; o servidor passa a desempenhar, somente, a função de fornecer os dados armazenados, enquanto o processamento desses dados passa a ser elaborado no cliente. Para que o processamento ocorra desta forma, muitas vezes, faz-se necessária a troca contínua de informações entre cliente e servidor. Embora seja possível o processamento de alguns dados no cliente com *applets*¹⁰ em Java nas conexões estabelecidas com HTML, após o envio dos dados, a conexão é desfeita e é necessária a realização de uma nova conexão a cada troca de informação. Sendo assim, a possibilidade de processamento no cliente, garantida pelo XML também merece destaque, pois, torna possível que seja estabelecida uma conexão constante com o servidor para haver a troca de dados contínua entre cliente e servidor. Além de permitir o processamento de dados complexos no próprio cliente.

2.3.3 Padrões de Metadados

Para satisfazer as necessidades dos usuários, de acordo com o cenário de aplicação no qual estes se encontram inseridos, são desenvolvidos e utilizados diversos padrões de metadados. Alves (2005, p. 115) esclarece:

“[...] formatos de metadados, também chamados de padrões de metadados, são estruturas padronizadas para a representação do conteúdo informacional que será representado pelo conjunto de dados-atributos (metadados). Em outras palavras, os formatos ou padrões de metadados podem ser considerados como formas de representação de um item documentário”. (ALVES, 2005, p.115)

Dentre os padrões de metadados utilizados, pode-se citar:

- DC (*Dublin Core*): catalogação de documentos eletrônicos na web;
- FGDC (*Federal Geographic Data Committee*): descrição de dados geoespaciais;

¹⁰ *Applets* são aplicativos pequenos, de escopo limitado que em geral, são utilizados em sites da Internet. Fonte: < <http://dicionario.babylon.com/applet>>

- IAFA/ WHOIS++ (*Internet Anonymous FTP Archive with transfer protocol*): descrição do conteúdo e serviços disponíveis em arquivos FTP (File Transfer Protocol);
- MARC (*Machine Readable Catalogue*): catalogação bibliográfica;
- SAIF (*Spatial Archive and Interchange Format*): compartilhamento de dados espaciais e espaço temporal.
- TEI (*Text Encoding Initiative*): representação de materiais textuais em formato eletrônico;
- ETD-MS (*An Interoperability Metadata Standard for Electronic Theses and Dissertations*): padrão de metadados para descrição de teses e dissertações adotado pela NDLTD;
- MTD-BR: padrão brasileiro de metadados para descrição de teses e dissertações adotado pela BDTD (Biblioteca Digital Brasileira de Teses e Dissertações).

Convém ressaltar que, neste trabalho, os padrões de metadados DC, MTD-BR e ETD-MS merecerão mais atenção, uma vez que trata-se de padrões relacionados direta ou indiretamente com a estrutura da BDTD.

2.3.3.1 Padrão de Metadados *Dublin Core* (DC)

Dublin Core pode ser definido como um conjunto de elementos de metadados planejado para facilitar a descrição de recursos eletrônicos. É um formato menos estruturado e mais flexível, que adota a sintaxe do *Resource Description Framework* (RDF) para assegurar a interoperabilidade entre outros formatos. Para Weibel (1997), as características do *Dublin Core*, um proeminente candidato a descrição dos recursos eletrônicos, dividem-se em várias categorias:

- Simplicidade: proporciona um conjunto de elementos para a descrição que pode ser usado por usuários especialistas em catalogação e por usuários leigos. Sua ênfase está na semântica, uma vez que a maior parte dos seus elementos tem uma semântica comumente entendida, descartando a necessidade de uma sintaxe rígida;

- Interoperabilidade semântica: promove o entendimento comum entre os demais padrões de metadados;
- Consenso internacional: reconhecimento da cobertura internacional do escopo do recurso;
- Flexibilidade: permite a codificação de semânticas mais elaboradas em padrões mais sofisticados;
- Modularidade de metadados na *web*: embora dirija-se especificamente à descoberta de recursos, pode-se imaginar funcionalidades distintas que atendem a outros objetivos (termos e condições, arquivamento de gestão, administração de metadados e muitos outros).

O padrão *Dublin Core* possui dois níveis: o simples e o qualificado. O *Dublin Core Simple* é composto por quinze elementos enquanto o *Dublin Core Qualificado* inclui três elementos adicionais (Audiência, Proveniência e Detentor de Direitos), assim como um grupo de refinamentos de elementos (qualificadores), que aperfeiçoam a semântica dos elementos a fim de torná-los úteis nas descobertas de recursos.

Alves e Souza (2007) destacam que o *Dublin Core* vem sendo aplicado internacionalmente por diversas instituições de renome.

2.3.3.1.1 Breve histórico do *Dublin Core*

Visando adaptar-se ao contexto atual, no qual a produção, a organização e a recuperação da informação na *web* são cada vez mais levadas em consideração e contribuir para que as buscas realizadas tornem-se mais rápidas e eficientes, várias iniciativas estão sendo conduzidas para discutir a questão e propor padrões de descrição de recursos informacionais, como é o caso do *Dublin Core Metadata Initiative* (DCMI)¹¹.

O DCMI é uma organização dedicada a promover e difundir a adoção da norma de interoperacionalidade dos padrões de metadados e, ainda, o desenvolvimento de

¹¹ URL do *Dublin Core Metadata Initiative* (DCMI): <<http://dublincore.org/>>

vocabulários especializados dos metadados para descrever recursos que permitam o uso de sistemas mais inteligentes. Teve início em 1995 com um workshop em Dublin, Ohio que reuniu bibliotecários, pesquisadores, fornecedores de conteúdos e especialistas em computação para desenvolver uma norma para a pesquisa de recursos. Conforme Rosetto e Nogueira (2002), foi estabelecido por consenso de um grupo internacional e interdisciplinar de profissionais bibliotecários, analistas de sistemas, lingüistas, museólogos, entre outros. O DC original começou com um pequeno conjunto de elementos que imediatamente despertou o interesse dos fornecedores de conteúdos em diversas áreas como a das artes, da educação, dos negócios e setores governamentais.

2.3.3.1.2 Elementos do *Dublin Core*

Possui um conjunto de 15 elementos básicos, apresentados na tabela abaixo, que podem ser implementados livremente para atender as necessidades de cada usuário. (DUBLIN, 2008)

Elementos	Descrição
Título	Nome dados aos recursos
Criador	Entidade originalmente responsável pela criação do conteúdo do recurso
Assunto	Tema do conteúdo do recurso
Descrição	Relato do conteúdo do recurso. Exemplos: sumário, resumo e texto livre
Publicador	Entidade responsável por tornar o recurso disponível
Colaborador	Entidade responsável pela contribuição intelectual ao conteúdo do recurso
Data	Data associada a um evento ou ciclo de vida do recurso
Tipo	Natureza do conteúdo do recurso. Exemplos: texto, imagem, som
Formato	Manifestação física ou digital do recurso. Exemplos: html, pdf, ppt
Identificador	Referência não-ambígua para o recurso dentro de determinado contexto
Fonte	Referência a um recurso do qual o presente é derivado
Idioma	Língua do conteúdo intelectual do recurso
Relação	Referência para um recurso relacionado
Cobertura	Extensão ou escopo do conteúdo do recurso; pode ser temporal e espacial
DireitosAutorais	Informação sobre os direitos assegurados dentro e sobre o recurso Elementos Metadados <i>Dublin Core</i>

Tabela 1 – Componentes básicos *Dublin Core*

Fonte: DUBLIN (2008)

Também é um formato padrão adotado para efetuar a interoperabilidade entre outros formatos. É utilizado para descrever uma variedade de recursos existentes na internet com o propósito de ser um meio de comunicação e de procura de informações.

2.3.3.1.3 Atributos do *Dublin Core*

Cada um dos elementos DC é descrito por dez atributos da norma ISO/IEC 11179¹². Estes consistem em:

- **Nome** - Etiqueta atribuída ao elemento de dado
- **Identificador** - Identificador único atribuído ao elemento de dado
- **Versão** - Versão do elemento de dado
- **Autoridade de Registo** - Entidade autorizada a registar o elemento de dado
- **Língua** - Linguagem na qual o elemento de dado é definido
- **Definição** - Uma afirmação que representa claramente o conceito e a natureza do elemento de dado
- **Obrigaçã**o - Indica se o elemento de dados é sempre obrigatório ou não (contém um valor)
- **Tipo dos Dados** - Indica os tipos de dados que podem ser representados no valor do elemento de dado
- **Maxima Ocorrência** - Indica qualquer limite à repetição do elemento de dado
- **Comentário** - Uma nota relativa à aplicação do elemento de dado

Dos dez atributos referidos, seis são comuns a todos os elementos "*Dublin Core*". São estes, com seus respectivos valores:

Versão : 1.1
 Autoridade de Registo : *Dublin Core Metadata Initiative*
 Língua : pt ("en" para a versão original)
 Obrigação..... : Opcional

¹² O padrão internacional ISO/IEC 11179 é constituído de várias partes relacionadas com padronização e especificação de elementos de dados. O conjunto completo é composto por seis partes inter-relacionadas, sendo cada parte focada em um aspecto de desenvolvimento e manutenção de elementos de dados.

Tipo dos Dados : Cadeia de caracteres

Máxima Ocorrência : Ilimitada

É importante destacar que os atributos mencionados acima fazem parte da definição formal dos elementos.

2.3.3.1.4 Qualificadores do *Dublin Core*

O *Dublin Core* pode usar apenas os 15 elementos básicos ou usar também elementos de refinamento, denominados Qualificadores *Dublin Core* (*Dublin Core Qualifiers*), para cada um destes 15 elementos.

Quando são utilizados apenas os elementos básicos, denomina-se *Dublin Core* Simples (*Simple Dublin Core Metadata*) e quando são utilizados três elementos adicionais (Audiência, Proveniência e Detentor de Direitos) e elementos de refinamento da semântica, denomina-se *Dublin Core* Qualificado (*Qualified Dublin Core*).

O objetivo do refinamento de elementos é tornar a especificação dos mesmos mais precisa. De acordo com o princípio guia para a qualificação de um elemento *Dublin Core* (*Dumb-Down Principle*), uma aplicação que não interprete um termo utilizado para refinamento de um determinado elemento deve ignorar o qualificador e tratar o valor do metadado como um elemento simples.

O refinamento de elementos que ocorre no *Dublin Core* Qualificado inclui um conjunto de esquemas de codificação recomendados e desenvolvidos para auxiliar na interpretação dos valores dos metadados. Para que este objetivo seja atingido, podem ser utilizados vocabulários controlados, notações formais ou regras de *parsing*¹³.

2.3.3.2 Padrão de Metadados ETD-MS

O ETD-MS pode ser definido como um conjunto de padrões internacional para os elementos de metadados usados para descrever uma tese ou dissertação eletrônica.

Estes padrões de metadados procuram descrever o autor, o trabalho e o contexto no qual a pesquisa foi produzida, de tal forma que estas informações tornem-se úteis para aqueles que procuram pelas teses e/ou dissertações.

¹³ *Parsing* é o processo de analisar sintaticamente, ou seja, verificar se uma entrada de dados está escrita corretamente.

Sendo este o padrão internacional, todas as bibliotecas digitais de teses e dissertações que quiserem fazer parte da base cooperativa (*union catalog*) devem estar aderentes a ele.

2.3.3.2.1 Breve histórico do ETD-MS

A versão do ETD-MS datada de 06/05/2001 foi elaborada em resposta aos comentários feitos pelos usuários de normas da NDLTD.

A atualização de versão, que ocorreu em 01/10/2001, foi resultado de uma reunião de dois dias que ocorreu na OCLC, em Dublin Ohio, nos dias 9 e 10 de janeiro de 2001. (NETWORKED, 2008)

2.3.3.2.2 Elementos do ETD-MS

Utilizam a descrição de elementos comuns ao *Dublin Core* e de um elemento novo utilizado especificamente para teses e dissertações, o *thesis.degree*, cuja finalidade é indicar informações associadas ao trabalho. O ETD-MS possui um conjunto de 14 elementos básicos, apresentados abaixo: (NETWORKED, 2008)

Elementos	Descrição
<i>dc.title</i>	Nome dado aos recursos
<i>dc.creator</i>	Entidade originalmente responsável pela criação do conteúdo do recurso
<i>dc.subject</i>	Tema do conteúdo do recurso
<i>dc.description</i>	Relato do conteúdo do recurso. Exemplos: sumário, resumo e texto livre
<i>dc.publisher</i>	Entidade responsável por tornar o recurso disponível
<i>dc.contributor</i>	Entidade responsável pela contribuição intelectual ao conteúdo do recurso
<i>dc.date</i>	Data associada a um evento ou ciclo de vida do recurso
<i>dc.type</i>	Natureza do conteúdo do recurso. Exemplos: texto, imagem, som
<i>dc.format</i>	Manifestação física ou digital do recurso. Exemplos: html, pdf, ppt, gif
<i>dc.identifier</i>	Referência não-ambígua para o recurso dentro de determinado contexto
<i>dc.language</i>	Referência a um recurso do qual o presente é derivado

<i>dc.coverage</i>	Língua do conteúdo intelectual do recurso
<i>dc.rights</i>	Referência para um recurso relacionado
<i>thesis.degree</i>	Grau associado com o trabalho. (exemplo: Mestre em pesquisa de operações)

Tabela 2 – Componentes básicos ETD-MS

Fonte: NETWORKED (2008)

2.3.3.2.3 Atributos do ETD-MS

Alguns dos elementos ETD-MS possuem atributos que auxiliam na descrição. Estes são enumerados abaixo com suas respectivas finalidades: (NETWORKED, 2008)

- *dc.title.alternative* (elemento *dc.title*) – título alternativo da tese ou da dissertação;
- *dc.description.abstract* (elemento *dc.description*) – texto completo do resumo;
- *dc.description.note* (elemento *dc.description*) – informações adicionais a respeito da tese ou da dissertação. Exemplo: nota da aceitação do departamento;
- *dc.description.release* (elemento *dc.description*) – descrição da versão do trabalho. Deve somente ser usado para erratas, etc.;
- *dc.contributor.role* (elemento *dc.contributor*) – papel que a pessoa assumiu na criação ou na aprovação do trabalho. Exemplos: conselheiro, organizador, banca, etc.;
- *thesis.degree.name* (elemento *thesis.degree*) – nome do grau associado ao trabalho como aparece dentro do mesmo. Exemplo: mestre;
- *thesis.degree.level* (elemento *thesis.degree*) – nível de instrução associado ao original. Exemplos: bacharel, mestre, doutor;
- *thesis.degree.discipline* (elemento *thesis.degree*) – área de estudo do índice intelectual do original. Geralmente, será o nome de um programa ou de um departamento;
- *thesis.degree.grantor* (elemento *thesis.degree*) – instituição que concede o grau associado com o trabalho. Como outros nomes e instituições, este campo deve ser incorporado ao formulário de texto livre como aparece na página de título ou equivalente, com uma ligação a um registro da autoridade se disponível.

2.3.3.2.4 Qualificadores Globais do ETD-MS

Além dos atributos especificados para os diversos elementos, existem também atributos globais que podem ser especificados para qualquer elemento. Estes são numerados abaixo, com suas respectivas definições:

- *Lang* (língua): indica a língua na qual o valor de um campo é escrito. É inteiramente distinto do elemento *dc.language*, que indica a língua preliminar própria do trabalho;
- *Translated* (tradução de): é um elemento vazio que indica que o valor de um campo é uma tradução fornecida por alguém à exceção do autor. Para traduções fornecidas pelo autor é indicada simplesmente a própria língua (usando o atributo “*lang*”);
- *Scheme* (esquema): descrição do vocabulário ou do esquema usado para determinar os assuntos.

2.3.3.3 Padrão de Metadados MTD-BR

O Padrão Brasileiro de Metadados de Teses e Dissertações (MTD-BR) foi desenvolvido no âmbito do projeto Biblioteca Digital de Teses e Dissertações (BDTD), visando atender aos objetivos de geração de produtos e serviços de informação, com vistas à identificação e localização das teses e dissertações eletrônicas (TDEs), bem como permitir a coleta de informação para geração de indicadores e integração com outros repositórios nacionais de informação de ensino e pesquisa no país. (INSTITUTO, 2008)

É compatível com os padrões ETD-MS e *Dublin Core* e possibilita a integração dos registros de teses e dissertações com registros de outros repositórios brasileiros, como o Banco de Currículos da Plataforma Lattes.

2.3.3.3.1 Breve histórico do MTD-BR

A versão inicial do MTD-BR, utilizada para testes, foi implementada em 2001.

A primeira versão oficial, a 1.0, foi lançada em novembro de 2002.

A versão atual, a 2.0, foi aprovada e homologada, pelo Conselho Técnico Consultivo (CTC) da BDTD em março de 2005.

2.3.3.3.2 Elementos do MTD-BR

O padrão MTD-BR possui um conjunto de 19 elementos básicos. Os mesmos são descritos na tabela abaixo: (BIBLIOTECA, 2008)

Elementos	Descrição
Controle	Dados de controle do registro de Teses e Dissertações
BibliotecaDigital	Dados da Instituição responsável pela publicação digital da tese ou dissertação
BibliotecaDepositária	Dados da Biblioteca onde o exemplar físico da tese ou dissertação está arquivado
Título	Título da tese ou dissertação
Arquivo	Endereço eletrônico dos arquivos da tese ou dissertação na Biblioteca Digital
Idioma	Idioma da tese ou dissertação
Grau	Grau acadêmico associado à tese ou dissertação
Titulação	Nome do grau acadêmico associado à tese ou dissertação
Resumo	Resumo da tese ou dissertação
Cobertura	Escopo espacial ou temporal da tese ou dissertação
Assunto	Tópicos tratados na tese ou dissertação e a tabela de onde estes tópicos foram extraídos, quando for o caso
LocalDefesa	Local de defesa da tese ou dissertação
DataDefesa	Data em que foi defendida a tese ou dissertação
Autor	Dados do Autor da tese ou dissertação
Contribuidor	Dados do Contribuidor da tese ou dissertação e forma de participação (papel)
InstituiçãoDefesa	Dados da Instituição onde a tese foi defendida
AgênciaFomento	Dados da Agência de financiamento que apoiou financeiramente o autor
Direitos	Informa as condições de distribuição, reprodução e utilização da tese ou dissertação
Extensão	Reservado para a criação de sub-itens (subcampos) para uso específico

Tabela 3 – Componentes básicos MTD-BR

Fonte: Biblioteca Digital de Teses e Dissertações (BDTD)

Além dos elementos básicos, o padrão MTD-BR possui também elementos vinculados aos mesmos. Estes são listados abaixo, juntamente com seu elemento básico e sua descrição:

- Sigla (Controle) – sigla da instituição consorciada;
- DataAtualização (Controle) – data em que foi realizada a última atualização do registro da tese ou dissertação;
- IdentificaçãoDocumento (Controle) – código que identifica a tese ou dissertação na base de dados da biblioteca;
- Tipo (Controle) – tipo de informação;
- Nome (BibliotecaDigital) – nome da biblioteca digital;
- Sigla (BibliotecaDigital) – sigla da biblioteca digital;
- URL (BibliotecaDigital) – endereço eletrônico do site da biblioteca digital;
- ProvedorServiço (BibliotecaDigital) – instituição que abriga a biblioteca digital;
- Nome (ProvedorServiço – BibliotecaDigital) – nome da instituição;
- Sigla (ProvedorServiço – BibliotecaDigital) – sigla da instituição;
- País (ProvedorServiço – BibliotecaDigital) – país da instituição;
- UF (ProvedorServiço – BibliotecaDigital) – UF da instituição;
- CNPJ (ProvedorServiço – BibliotecaDigital) – CNPJ da instituição;
- URL (ProvedorServiço – BibliotecaDigital) – endereço eletrônico do site da instituição;
- Nome (BibliotecaDepositária) – nome da biblioteca depositária;
- Sigla (BibliotecaDepositária) – sigla da biblioteca depositária;
- URL (BibliotecaDepositária) – endereço eletrônico do site da biblioteca depositária;
- NúmeroChamada (BibliotecaDepositária) – código que identifica a tese ou dissertação no acervo da biblioteca depositária;
- URL (Arquivo) – endereço eletrônico do arquivo da tese ou dissertação e/ou página de apresentação dos metadados na biblioteca digital de origem;
- Legenda (Arquivo) - legenda para o arquivo, descrição de seu conteúdo;
- NivelAcesso (Arquivo) – restrição de acesso ao arquivo;
- Cidade (LocalDefesa) - cidade onde foi defendida a tese ou dissertação;
- UF (LocalDefesa) - UF da cidade onde foi defendida a tese ou dissertação;
- País (LocalDefesa) - país onde foi defendida a tese ou dissertação;

- Nome (Autor) – nome da pessoa física;
- Citação (Autor) - forma pela qual a pessoa física será citada;
- Lattes (Autor) - endereço eletrônico do currículo da pessoa na Plataforma Lattes;
- CPF (Autor) – CPF da pessoa;
- Afiliação (Autor) - instituição à qual a Pessoa é afiliada;
- Nome (Afiliação – Autor) – nome da instituição;
- Sigla (Afiliação – Autor) – sigla da instituição;
- País (Afiliação – Autor) – país da instituição;
- UF (Afiliação – Autor) – UF da instituição;
- CNPJ (Afiliação – Autor) – CNPJ da instituição;
- URL (Afiliação – Autor) – endereço eletrônico do site da instituição;
- Nome (Contribuidor) – nome da pessoa física;
- Citação (Contribuidor) – forma pela qual a pessoa deseja ser citada;
- Lattes (Contribuidor) – endereço eletrônico do currículo da pessoa na Plataforma Lattes;
- CPF (Contribuidor) – CPF da pessoa;
- Afiliação (Contribuidor) - instituição à qual a Pessoa é afiliada;
- Nome (Afiliação – Contribuidor) – nome da instituição;
- Sigla (Afiliação – Contribuidor) – sigla da instituição;
- País (Afiliação – Contribuidor) – país da instituição;
- UF (Afiliação – Contribuidor) – UF da instituição;
- CNPJ (Afiliação – Contribuidor) – CNPJ da instituição;
- URL (Afiliação – Contribuidor) – endereço eletrônico do site da instituição;
- Nome (InstituiçãoDefesa) - nome da instituição;
- Sigla (InstituiçãoDefesa) - sigla da instituição;
- País (InstituiçãoDefesa) – país da instituição;
- UF (InstituiçãoDefesa) – UF da instituição;
- CNPJ (InstituiçãoDefesa) – CNPJ da instituição;
- URL (InstituiçãoDefesa) – endereço eletrônico do site da instituição;
- Programa (InstituiçãoDefesa) - programa de pós-graduação onde a tese ou dissertação foi defendida;
- Nome (Programa – InstituiçãoDefesa) – nome do programa de pós-graduação;
- Área (Programa – InstituiçãoDefesa) – área de concentração do programa de pós-graduação;

- Nome (AgênciaFomento) – nome da instituição;
- Sigla (AgênciaFomento) – sigla da instituição;
- País (AgênciaFomento) – país da instituição;
- UF (AgênciaFomento) – UF da instituição;
- CNPJ (AgênciaFomento) – CNPJ da instituição;
- URL (AgênciaFomento) – endereço eletrônico do site da instituição;

2.3.3.3.3 Atributos do MTD-BR

O atributo Formato, do elemento URL (vinculado ao elemento Arquivo) é utilizado somente para as teses ou dissertações eletrônicas a fim de especificar o(s) formato(s) eletrônico(s) no(s) qual(is) a tese está disponível.

O atributo Idioma, dos elementos Legenda (vinculado ao elemento Arquivo), Título (elemento básico), Resumo (elemento básico), Cobertura (elemento básico), Assunto (elemento básico), Direitos (elemento básico), segue a Norma ISO 639¹⁴ tanto para 2 (dois) ou 3 (três) caracteres, que estabelece que o idioma português deve ser identificado como “pt” ou “por” (ambos em minúsculas).

O atributo Esquema, do elemento Assunto (elemento básico) é o instrumento de indexação utilizado pela biblioteca depositária.

O atributo Papel, do elemento Contribuidor (elemento básico), é o tipo de contribuição e forma de participação na geração/aprovação da tese ou dissertação. Pode ser Orientador, Co-Orientador ou Membro da Banca.

É importante destacar que todos os atributos adotados pelo padrão MTD-BR são facultativos¹⁵.

2.3.4 Comparação dos padrões MTD-BR, ETD-MS e Dublin Core

A relação entre os padrões de metadados DC, ETD-MS e MTD-BR pode ser visualizada na figura abaixo:

¹⁴ A ISO 639 - *Codes for the representation of names of languages* é uma norma técnica da ISO para especificação de nomes de idiomas. Fonte: <<http://www.oasis-open.org/cover/iso639a.html>>

¹⁵ No anexo 3, encontra-se a listagem completa dos elementos do padrão MTD-BR, com seus respectivos atributos.

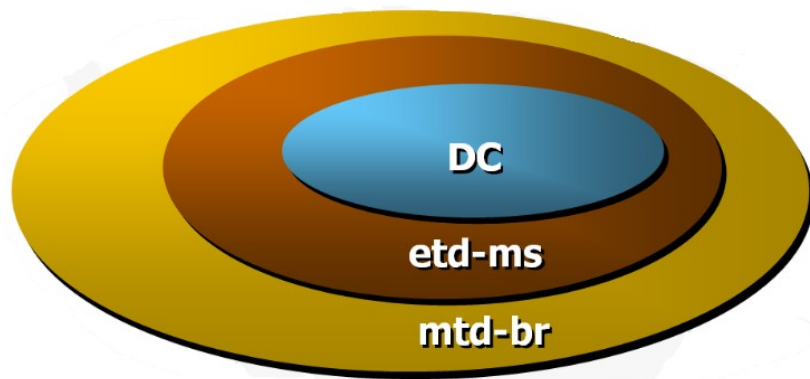


Figura 2 – Integração dos metadados utilizados pela BDTD

Fonte: (INSTITUTO, 2008)

Com o intuito de aprofundar nosso conhecimento sobre o padrão de metadados MTD-BR, julgamos conveniente fazermos uma comparação entre este e os padrões ETD-MS e *Dublin Core*, que serviram de base para as pesquisas realizadas pelo IBICT para o desenvolvimento do padrão adotado pela BDTD. Na tabela abaixo é apresentada uma comparação entre os elementos dos padrões MTD-BR, ETD-MS e *Dublin Core*:

MTD-BR	ETD-MS	<i>Dublin Core</i>
Controle	<i>dc.title</i>	Título
Biblioteca Digital	<i>dc.creator</i>	Criador
Biblioteca Depositaria	<i>dc.subject</i>	Assunto
Título	<i>dc.description</i>	Descricao
Arquivo	<i>dc.publisher</i>	Publicador
Idioma	<i>dc.contributor</i>	Colaborador
Grau	<i>dc.date</i>	Data
Titulacao	<i>dc.type</i>	Tipo
Resumo	<i>dc.format</i>	Formato
Cobertura	<i>dc.identifier</i>	Identificador
Assunto	<i>dc.language</i>	Fonte
LocalDefesa	<i>dc.coverage</i>	Idioma
DataDefesa	<i>dc.rights</i>	Relacao
Autor	<i>thesis.degree</i>	Cobertura
Contribuidor		DireitosAutorais
InstituicaoDefesa		
AgenciaFomento		
Direitos		
Extensao		

Tabela 4 - Tabela comparativa dos elementos do MTD-BR, ETD-MS e DC

Por meio desta comparação pode-se constatar que:

- 1) o padrão ETD-MS possui 13 elementos comuns ao padrão *Dublin Core*. Os elementos do DC não comuns ao ETD-MS são o Fonte e o Relacao. Para o padrão ETD-MS foi criado o elemento *thesis.degree*, cuja finalidade é indicar informações associadas ao trabalho. Convém lembrar que este elemento possui os atributos *thesis.degree.name*, *thesis.degree.level*, *thesis.degree.discipline* e *thesis.degree.grantor*, cujas finalidades são, respectivamente, indicar o nome do grau associado ao trabalho como aparece dentro do mesmo, indicar o nível de instrução associado ao original, indicar a área de estudo do índice intelectual do original e indicar a instituição que concede o grau associado com o trabalho.
- 2) o padrão ETD-MS possui somente elementos para a descrição dos recursos digitais.
- 3) o padrão MTD-BR adotou 13 elementos do padrão DC, que também são comuns ao ETD-MS. Os elementos do DC não adotados pelo MTD-BR são os mesmos não utilizados pelo ETD-MS, o Fonte e o Relacao. A descrição, neste padrão, foi detalhada incluindo autor, contribuidor, instituição, informações referentes as bibliotecas e aos programas de pós-graduação.
- 4) a maioria dos elementos criados para o padrão MTD-BR se destina a metadados administrativos.
- 5) o padrão *Dublin Core* possui apenas 1 elemento que se destina ao metadado administrativo, o elemento Fonte, cuja finalidade é indicar um recurso do qual o presente é derivado.
- 6) embora o elemento Relacao do *Dublin Core*, que representa a relação de um documento digital com outros, não tenha sido incluído no padrão MTD-BR, foram inclusos os elementos URL e Assunto, que possibilitam estabelecer uma relação entre uma tese ou dissertação com uma determinada universidade, agência de fomento ou biblioteca. Esses elementos tornam possível que uma tese/dissertação se relacione com outras teses ou dissertações, com trabalhos do

mesmo autor e com trabalhos das instituições que contribuíram para sua produção.

Diante do exposto, observa-se que o padrão MTD-BR é o que melhor agrupa os elementos necessários para que a BDTD possa desempenhar suas tarefas e atingir seus objetivos, uma vez que é composto por elementos descritivos, estruturais e administrativos, além de conter elementos específicos de interesse nacional e contemplar os padrões ETD-MS e DC. Lourenço (2005) defende que:

“[...] o padrão MTD-BR, além de conter os elementos descritivos necessários para o tratamento da informação de teses e dissertações eletrônicas, ele também contém elementos estruturais que irão garantir uma melhor recuperação da informação em uma biblioteca digital e elementos administrativos, que permitiram um melhor gerenciamento e organização desse acervo digital.” (LOURENÇO, 2005, p. 100)

A mesma autora defende que a presença de metadados administrativos e estruturais, além dos descritivos no padrão de metadados adotado pela biblioteca digital, é fator essencial para que essa biblioteca digital atinja seus propósitos e metas.

2.4 Protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*)

O *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) é um protocolo que permite o compartilhamento de metadados. Trata-se de uma solução que possibilita a troca de informações sem demandar custos elevados.

O protocolo OAI-PMH fornece uma estrutura de interoperabilidade independente da aplicação baseada na coleta dos metadados. Há duas classes de participantes na estrutura do OAI-PMH:

- Os provedores ou repositórios de dados que administram os sistemas que suportam o OAI-PMH como meio de expor metadados;

- Os provedores de serviços que utilizam metadados coletados pelo OAI-PMH. (OPEN, 2008)

De acordo com Baptista et al. (2007), os provedores de dados devem ser dotados, no mínimo, das seguintes funcionalidades:

- mecanismos de submissão para o auto-arquivamento¹⁶ dos trabalhos ou *papers*;
- sistema de armazenamento a longo prazo;
- mecanismos de exposição de metadados do arquivo para facilitar a sua coleta por terceiros, ou provedores de serviços.

É importante ressaltar que algumas instituições, assim como a BDTD, podem ser tanto provedores de dados como provedores de serviços, ou seja, exportam metadados em algumas situações e utilizam metadados coletados em outras.

Convém destacar que o OAI-PMH é um protocolo baseado em um modelo de coleta de metadados e não um protocolo de pesquisa.

2.4.1 Breve histórico do OAI-PMH

Em 1999, foi promovido um encontro em Santa Fé, Novo México, que teve como objetivo principal definir especificações técnicas e princípios administrativos para se estabelecer um alto nível de interoperabilidade. Este encontro deu origem a fundação da *Open Archives Initiative* (OAI). A OAI visa facilitar a disseminação eficiente de conteúdo, desenvolve e promove padrões de interoperabilidade entre repositórios digitais e sua base é o protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*), que permite que os participantes da Iniciativa possam compartilhar seus metadados.

A OAI especificou o protocolo OAI-PMH em 2001 e o mesmo foi testado durante os anos de 2001 e 2002. A versão inicial do OAI-PMH, a 1.0, disponibilizada em janeiro de 2001, introduziu o conjunto de elementos *Dublin Core* como uma base

¹⁶ De acordo com Baptista et al. (2007), auto-arquivamento ou *self-archiving* é um mecanismo que permite aos autores submeterem ou depositarem, eles mesmos, os seus trabalhos ou *papers* em um repositório digital.

para interoperabilidade dos metadados. E a versão atual, a 2.0, em junho de 2002. Com o lançamento da versão 2.0 do OAI-PMH, foi finalizado o período de experimentação inicial do protocolo. Neste período, foram testadas, durante 18 meses, a funcionalidade do protocolo e a idéia geral da coleta de metadados. A versão 2.0 tornou-se estável após pouco tempo de experimentação.

As mudanças estruturais, assim como a evolução do protocolo, demonstram o interesse da OAI em alcançar um nível de interoperabilidade que permita a exposição de metadados de uma forma global.

2.4.2 *Harvesting*

O OAI-PMH é um protocolo modelo de interoperabilidade baseado no processo de coleta automática de metadados (*metadata harvesting*). Operando sobre o protocolo HTTP¹⁷, os provedores de serviço enviam solicitações de metadados aos provedores de dados. Estes respondem disponibilizando metadados estruturados em registros XML, obedecendo a um determinado padrão de metadados. (INSTITUTO, 2008)

A aplicação responsável pela busca dos metadados é denominada *harvester* e a política adotada para a coleta automática de metadados segue uma definição lógica para estabelecer o que deverá ser coletado, por exemplo, se a coleta será por data ou por tipo de informação. É possível também especificar uma escala de tempo para a coleta dos registros. Os limites desta escala são inclusivos, isto é, deve ser especificado um limite a ser interpretado como “superior ou igual” ou “inferior ou igual” a um determinado argumento. Conseqüentemente, o argumento deve ser “superior ou igual” ou “inferior ou igual” ao limite. Caso contrário, o repositório emitirá um erro. (OPEN, 2008)

As operações de *harvesting* do OAI-PMH são baseadas em seis comandos transmitidos aos repositórios de dados, denominados verbos. Estes são enumerados abaixo, com suas respectivas funções:

- *Identify*: recupera dados administrativos sobre o repositório de dados;
- *ListMetadataFormats*: recupera os formatos de metadados disponibilizados no repositório de dados;
- *GetRecord*: recupera um único registro do repositório de dados;

¹⁷ O Protocolo de transferência de Hipertexto é um protocolo de comunicação utilizado para transferir dados através de intranets e da www.

- *ListRecords*: lista todos os registros de um repositório;
- *ListIdentifiers*: lista os identificadores (cabecinhos) de todos os registros de um repositório;
- *ListSets*: lista a estrutura do conjunto, também conhecida como árvore de assuntos ou hierarquia do repositório.

De acordo com Marcondes e Sayão (2002), a interação entre provedor de serviços e provedor de dados usando o OAI-PMH pode ser visualizada na figura abaixo:

Interação Service Providers e Data Providers

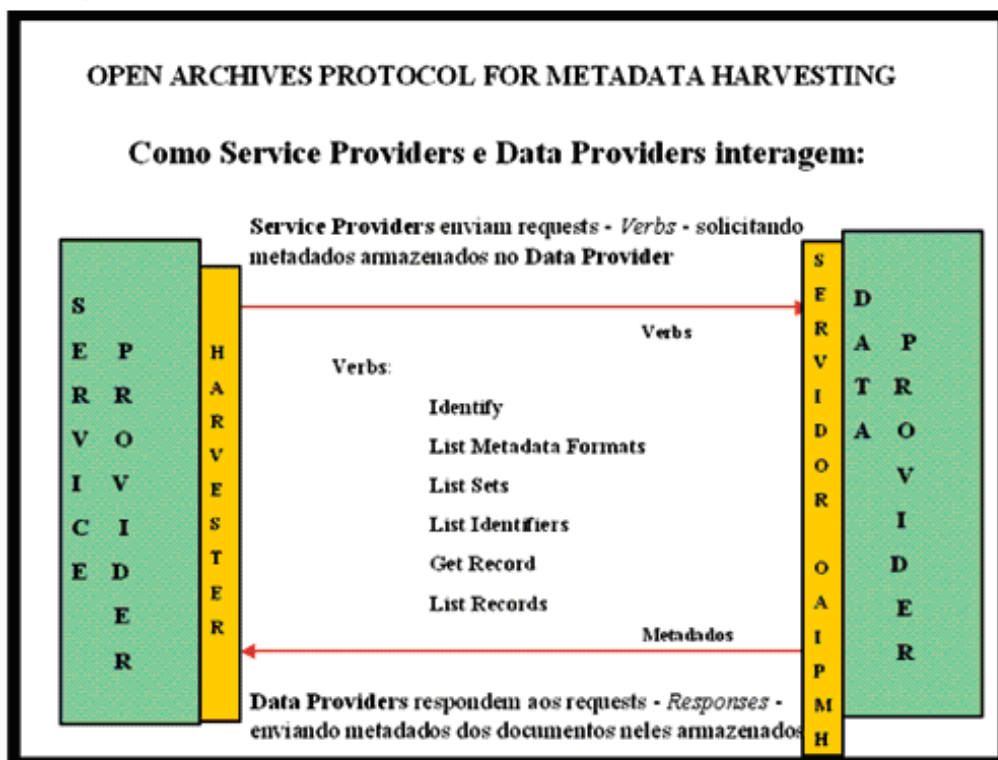


Figura 3 – Interação entre Provedor de Serviços e Provedor de Dados, por meio do OAI-PMH

Fonte: Marcondes e Sayão (2002)

Os provedores de dados são sistemas que utilizam o protocolo OAI-PMH como um meio para expor metadados e os provedores de serviços utilizam os metadados, coletados via protocolo OAI-PMH, como base para a construção de novos serviços. Tendo a implementação da BDTD como exemplo, na qual as IES atuam como provedores de dados e o IBICT e a NDLTD, que implementam os serviços com valor

agregado a partir dos dados coletados junto aos repositórios digitais, como provedores de serviços, pode-se visualizar a interação entre os mesmos por meio do protocolo OAI-PMH na figura abaixo:

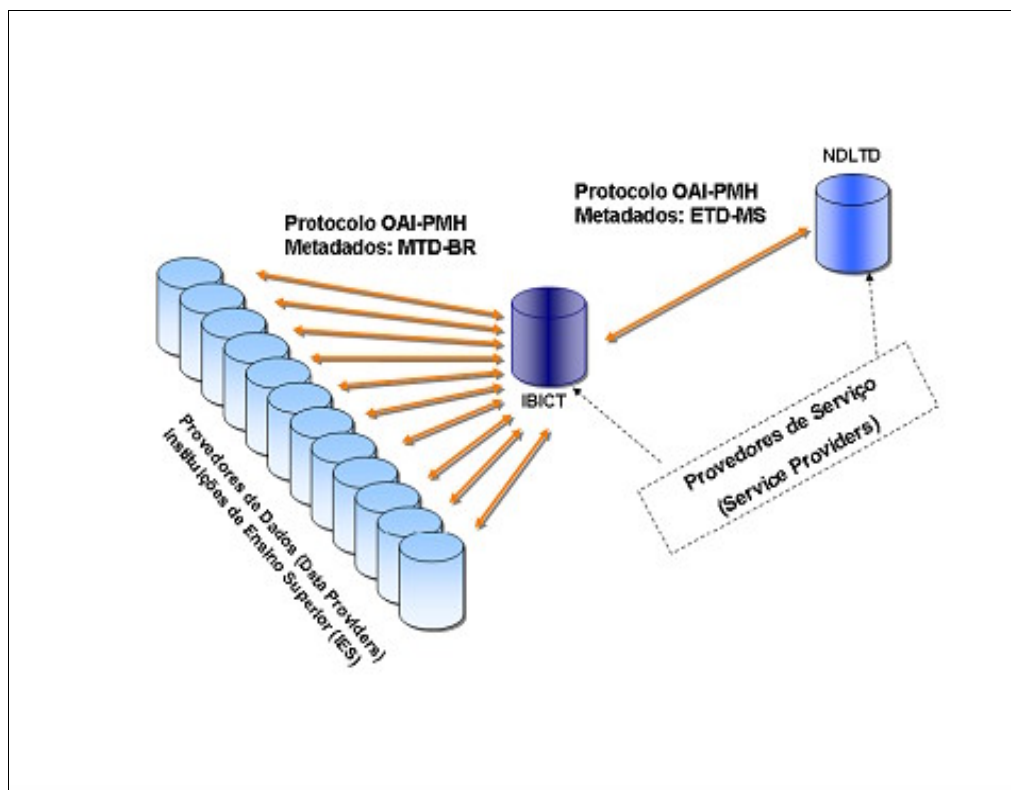


Figura 4 – Esquema Funcional do Modelo Open Archives

Fonte: Baptista et. al (2007)

Pode-se observar claramente a interação, possibilitada pelo protocolo OAI-PMH, que ocorre entre as IES, o IBICT e a NDLTD. As IES atuando como provedores de dados, disponibilizando metadados para o IBICT. Este atuando como provedor de serviços, disponibilizando serviços com valor agregado a partir dos metadados coletados das IES e como provedor de dados para a NDLTD, disponibilizando metadados. A NDLTD atuando como provedor de serviços, disponibilizando serviços com valor agregado a partir dos metadados coletados do IBICT.

A seguir, apresentaremos a caracterização da pesquisa e os procedimentos metodológicos utilizados no processo de coleta e análise dos dados. Procedimentos estes que têm como objetivo garantir coerência e validade científica para a pesquisa.

3 Procedimentos Metodológicos

A abordagem geral da pesquisa foi exploratória, uma vez que, de acordo com Gil (1991), estas pesquisas têm como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a construir hipóteses.

A estratégia de pesquisa foi a sintética de caso ou estudo de caso, pois, segundo Gil (1991), o estudo de caso é caracterizado pelo estudo profundo e exaustivo de um ou poucos objetos, de maneira que permita o seu detalhado conhecimento, tarefa praticamente impossível para outras estratégias de pesquisa.

A população da pesquisa foi constituída pelas dissertações eletrônicas da UDESC disponibilizadas na BDTD. Não foram inclusas teses pois, a UDESC, apesar de atualmente oferecer cursos de pós-graduação com grau de mestre e de doutor, este último teve seu início no ano de 2008, não tendo, portanto, teses defendidas.

O universo da pesquisa está restrito as dissertações defendidas nesta instituição, no período de 1997 a 2007, totalizando 451 dissertações. Destas, 443 estão disponibilizadas na BDTD e 08, apesar de não terem sido processadas pelo *harvesting* e, portanto, constarem somente na BDTD da UDESC, também foram consideradas no corpus de análise.

O método de coleta de dados foi o de registros institucionais (ou análise documental).

Foram analisados de forma qualitativa e quantitativa todos os metadados das dissertações eletrônicas da UDESC disponibilizados para a BDTD.

A abordagem qualitativa foi realizada utilizando-se como parâmetro as recomendações do padrão MTD-BR no que se refere ao conteúdo dos elementos de metadados e seus respectivos atributos.

A abordagem quantitativa teve como objetivo comparar a quantidade de informações corretas e incorretas de metadados. Os resultados quantitativos foram apresentados de forma descritiva com o auxílio de tabelas. O exame destas tabelas resultou em recomendações para a melhoria do conteúdo dos metadados informados pela UDESC para a BDTD.

4 Análise dos metadados informados pela UDESC para a BDTD

Para que se obtenha a informação confiável, de fácil acesso, com um tempo de resposta reduzido e com um custo acessível, é essencial uma adequada gestão de informações, assim como de seus fluxos. Com o objetivo de garantir estas características, o tratamento físico e a Análise Documentária assumem um papel fundamental, pois, pretendem analisar, traduzir e representar a forma e o conteúdo dos documentos com a finalidade de auxiliar a recuperação de informações.

Segundo Kobashi (1994), a circulação de informações e documentos, tal como é compreendida no âmbito da Ciência da Informação, comporta as seguintes operações básicas: a) coleta de documentos; b) tratamento de documentos e c) difusão dos documentos. O tratamento é a fase que se dedica a descrever o documento sob dois enfoques: forma e conteúdo. Neste sentido, Kobashi (1994) defende que:

“O tratamento documentário do suporte material, objeto da Representação Descritiva ou Catalogação, visa elaborar a descrição normalizada de aspectos físicos do documento (nome da obra, do autor, local de publicação, ano de publicação, editora, entre outros dados). Tal modalidade de tratamento permite o acesso físico ao documento.” (KOBASHI, 1994, p. 20)

A mesma autora afirma que o tratamento documentário do conteúdo – a Análise Documentária – visa, por outro lado, elaborar representações condensadas daquilo que é dito em um dado texto, sendo o resumo e o índice, as representações documentárias típicas.

Para que tanto o tratamento físico como o de conteúdo seja feito de forma consistente, é necessária a utilização de padrões. Estes, garantirão a interoperabilidade dos SRIs. No caso da BDTD, este padrão é definido pelo MTD-BR (anexo 3) que se compõe de três partes:

- 1) Lista de Elementos do Schema;
- 2) Padrões adotados para o preenchimento de elementos específicos, baseando-se em tabelas elaboradas e utilizadas por organizações nacionais e/ou

internacionais, como Normas ISO ou listas normalizadas de siglas de bibliotecas depositárias utilizadas no CCN/COMUT¹⁸, por exemplo.

3) Padrões adotados para os atributos com a finalidade de identificar alguns elementos. Convém ressaltar que os atributos do MTD-BR são elementos facultativos adotados no Schema.

O uso adequado do MTD-BR garantirá a circulação das teses e dissertações no meio acadêmico. Tomando por base a distinção entre forma e conteúdo citada por Kobashi (1994), examinamos os elementos do padrão MTD-BR de metadados, a fim de identificar quais se destinam à descrição física e quais se destinam à descrição de conteúdo do documento. Constatou-se que, dos 67 elementos analisados, 61 destinam-se à descrição física, 03 destinam-se à descrição de conteúdo do documento. Três não foram informados pela UDESC, portanto não constam da tabela abaixo. Os elementos que deixaram de ser informados são: Cobertura (10), que indica o escopo espacial ou temporal da dissertação, Area (16.7.2), que informa a área de concentração do Programa de Pós-graduação em que a dissertação foi defendida e Extensão (19), reservado para a criação de subitens (sub-campos) para uso específico. A tabela a seguir mostra estes dois conjuntos de elementos:

Descrição física	Descrição de conteúdo
1 Controle	
1.1 Sigla (da instituição consorciada)	
1.2 DataAtualização (do registro da dissertação)	
1.3 IdentificacaoDocumento (na base de dados da biblioteca)	

¹⁸ O Catálogo Coletivo Nacional de Publicações Seriadas (CCN), coordenado pelo IBICT, é uma rede cooperativa de unidades de informação de instituições localizadas no Brasil. Possibilita o acesso a publicações periódicas científicas e técnicas e reúne informações de centenas de catálogos, produzidos pelas principais bibliotecas do país, em um único catálogo nacional de acesso público. O Programa de Comutação Bibliográfica (COMUT) foi criado em 1980 pelo Ministério da Educação, por meio da CAPES, visando dotar o País de um mecanismo eficiente de acesso à informação. Permite às comunidades acadêmica e de pesquisa o acesso a documentos em todas as áreas do conhecimento (através de cópias de artigos de revistas técnico-científicas, teses e anais de congressos), exclusivamente para fins acadêmicos e de pesquisa, respeitando-se rigorosamente a Lei de Direitos Autorais. Fonte: <<http://www.ibict.br>>

1.4 Tipo (da fonte de informação)	
2 BibliotecaDigital	
2.1 Nome (da biblioteca digital)	
2.2 Sigla (da biblioteca digital)	
2.3 URL (da biblioteca digital)	
2.4 ProvedorServico	
2.4.1 Nome (da instituição que abriga a biblioteca digital)	
2.4.2 Sigla (da instituição que abriga a biblioteca digital)	
2.4.3 Pais (da instituição que abriga a biblioteca digital)	
2.4.4 (da instituição que abriga a biblioteca digital)	
2.4.5 CNPJ (da instituição que abriga a biblioteca digital)	
2.4.6 URL (da instituição que abriga a biblioteca digital)	
3 BibliotecaDepositaria	
3.1 Nome (da biblioteca depositária)	
3.2 Sigla (da biblioteca depositária)	
3.3 URL (da biblioteca depositária)	
3.4 NumeroChamada (da biblioteca depositária)	
	4 Titulo (da dissertação)
5 Arquivo	
5.1 URL (da dissertação e/ou página de apresentação dos metadados na biblioteca digital de origem)	
5.2 Legenda (para descrição do conteúdo do arquivo)	
5.3 NivelAcesso (restrição de acesso ao arquivo)	

6 Idioma (da dissertação)	
7 Grau (grau acadêmico associado à dissertação)	
8 Titulacao (nome do grau acadêmico associado à dissertação)	
	9 Resumo (da dissertação)
	11 Assunto (tratado na dissertação)
12 LocalDefesa	
12.1 Cidade (onde foi defendida a dissertação)	
12.2 UF (onde foi defendida a dissertação)	
12.3 Pais (onde foi defendida a dissertação)	
13 DataDefesa (da dissertação)	
14 Autor	
14.1 Nome (do autor da dissertação)	
14.2 Citacao (forma pela qual o autor da dissertação será citado)	
14.3 Lattes (endereço eletrônico do currículo do autor na plataforma Lattes)	
14.4 CPF (do autor)	
14.5.1 Afiliacao (Instituição à qual o autor é afiliado)	
14.5.2 Sigla (da Instituição à qual o autor é afiliado)	
14.5.3 Pais (da Instituição à qual o autor é afiliado)	
14.5.4 UF (da Instituição à qual o autor é afiliado)	
14.5.5 CNPJ (da Instituição à qual o autor é afiliado)	
14.5.6 URL (da Instituição à qual o autor é afiliado)	

15 Contribuidor	
15.1 Nome (do contribuidor)	
15.2 Citacao (forma pela qual o contribuidor deseja ser citado)	
15.3 Lattes (endereço eletrônico do currículo do contribuidor na plataforma Lattes)	
15.4 CPF (do contribuidor)	
15.5.1 Nome (da Instituição à qual o contribuidor é afiliado)	
15.5.2 Sigla (da Instituição à qual o contribuidor é afiliado)	
15.5.3 Pais (da Instituição à qual o contribuidor é afiliado)	
15.5.4 UF (da Instituição à qual o contribuidor é afiliado)	
15.5.5 CNPJ (da Instituição à qual o contribuidor é afiliado)	
15.5.6 URL (da Instituição à qual o contribuidor é afiliado)	
16 InstituicaoDefesa	
16.1 Nome (da Instituição onde a dissertação foi defendida)	
16.2 Sigla (da Instituição onde a dissertação foi defendida)	
16.3 Pais (da Instituição onde a dissertação foi defendida)	
16.4 UF (da Instituição onde a dissertação foi defendida)	
16.5 CNPJ (da Instituição onde a dissertação foi defendida)	
16.6 URL (da Instituição onde a dissertação foi defendida)	

16.7.1 Nome (do programa de pós-graduação onde a dissertação foi defendida)	
17 AgenciaFomento	
17.1 Nome (da Instituição que apoiou financeiramente o autor)	
17.2 Sigla (da Instituição que apoiou financeiramente o autor)	
17.3 País (da Instituição que apoiou financeiramente o autor)	
17.4 UF (da Instituição que apoiou financeiramente o autor)	
17.5 CNPJ (da Instituição que apoiou financeiramente o autor)	
17.6 URL (da Instituição que apoiou financeiramente o autor)	
18 Direitos (informa as condições de distribuição, reprodução e utilização da dissertação)	

Tabela 5 – Elementos para descrição física e descrição de conteúdo do documento informados pela UDESC para a BDTD

Tendo como base os padrões do MTD-BR adotados para os metadados e para seus respectivos atributos, será apresentada a seguir a análise do conteúdo dos metadados informados pela UDESC para a BDTD. Ressalta-se que para cada metadado analisado, apresenta-se uma tabela que sintetiza os seguintes pontos:

- Elemento: informa o nome do elemento analisado;
- Descrição: informa a descrição do elemento analisado;
- Preenchimento: informa se o preenchimento do elemento analisado é obrigatório ou facultativo;
- Padrão: informa se foi adotado algum padrão para o preenchimento do elemento analisado;

- Total de elementos: informa o total de elementos analisados;
- Corretos: informa o número de elementos analisados informados corretamente;
- Incorretos: informa o número de elementos analisados que foram informados incorretamente;
- Não informados: informa o número de elementos analisados que deixaram de ser informados.

É importante destacar que quando o preenchimento do elemento analisado tiver sido feito corretamente, não será apresentada nenhuma discussão a respeito. Entretanto, serão tecidos alguns comentários sobre a consequência do preenchimento incorreto dos elementos, assim como da falta de preenchimento dos mesmos.

1. Controle: o elemento Controle (1) indica dados de controle do registro de Teses e Dissertações. Trata-se de um elemento de preenchimento obrigatório que se subdivide em quatro (04) elementos, Sigla (1.1), DataAtualizacao (1.2), IdentificacaoDocumento (1.3) e Tipo (1.4). Pode-se observar, na tabela abaixo, a descrição dos mesmos na biblioteca digital de teses e dissertações da UDESC:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
1.1 Sigla	indica a sigla da Instituição Consorciada	Obrigatório	Não adotado	451	451	0	0
1.2 DataAtualizacao	indica a data da última atualização do registro da dissertação	Obrigatório	Protocolo OAI ¹⁹	451	450	0	1
1.3 Identificação Documento	indica o código que identifica a dissertação na base de dados da biblioteca	Obrigatório	Não adotado	451	451	0	0

¹⁹ O padrão DateTime (XML) adotado pelo protocolo OAI é AAAA-MM-DDThh:mm:ss, sendo as especificações de hora, minuto e segundo facultativas para o protocolo OAI embora sejam importantes para a coleta automática de dados (*harvesting*). Fonte: <<http://www.w3.org/TR/xmlschema-2/#dateTime>>

1.4 Tipo	indica o tipo da fonte de informação	Obrigatório	DCMI Type Vocabulary ²⁰	895	894	1	0
---------------------	--------------------------------------	-------------	------------------------------------	------------	------------	----------	----------

Tabela 6 – Descrição dos elementos Sigla (1.1), DataAtualizacao (1.2), IdentificacaoDocumento (1.3) e Tipo (1.4) do elemento Controle (1)

O elemento DataAtualizacao (1.2) deixou de ser informado para uma das dissertações e, por ser de preenchimento obrigatório, a falta desta informação resultou na exclusão da respectiva dissertação do processo de *harvesting*.

Deve-se ressaltar que uma dissertação eletrônica pode possuir somente um elemento Arquivo (5) ou um para cada capítulo. Por esta razão, o número de elementos Tipo (1.4) é superior ao número de dissertações defendidas entre 1997 e 2007, que é 451. Dos 895 elementos Tipo (1.4) analisados, constatou-se que para uma das dissertações, mesmo tendo sido informado de acordo com o padrão especificado, o elemento foi informado como “som”, quando o correto seria “Texto” ou “Tese ou Dissertação Eletrônica”, por ser a fonte de informação, um texto e não um som. O preenchimento incorreto do elemento não resultou na exclusão desta dissertação do processo de *harvesting* porque está de acordo com os tipos especificados no padrão para este elemento. E, como a BDTD não disponibiliza a busca por meio do Tipo, o preenchimento incorreto deste elemento não terá conseqüências na recuperação desta dissertação na base da BDTD. Entretanto, se considerarmos a busca por outros repositórios, o preenchimento incorreto deste metadado ocasionará problema de interoperabilidade. Este é o caso da NDLTD. Esta rede disponibiliza a busca por meio do elemento Tipo, conforme pode-se observar na figura abaixo:

²⁰ O DCMI Type Vocabulary oferece um modo geral, uma lista de termos que podem ser usados para determinar a natureza de um recurso. Fonte: <<http://dublincore.org/documents/dcmi-type-vocabulary/>>

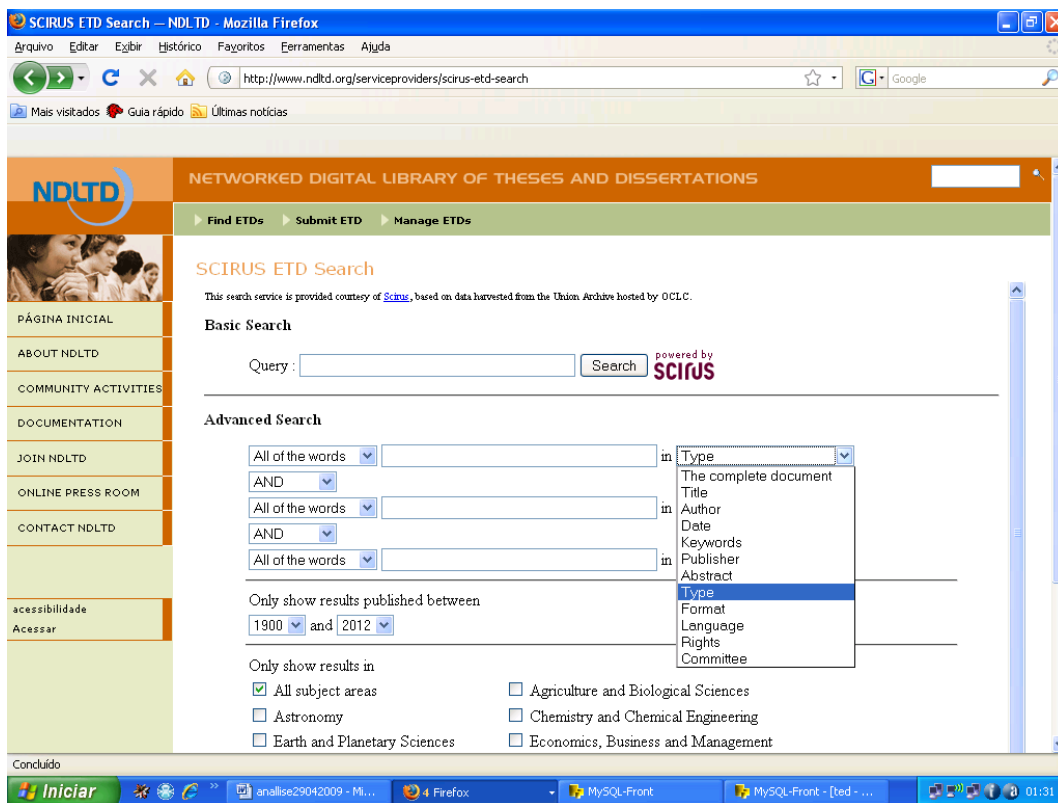


Figura 5 - Tela de busca de Teses e Dissertações da NDLTD

Sendo assim, a recuperação desta dissertação na base de dados da NDLTD por meio da busca por Tipo (*Type*) não será possível, impossibilitando o fluxo deste documento em âmbito internacional.

2. BibliotecaDigital: o conteúdo do elemento BibliotecaDigital (2) indica o nome da instituição responsável pela publicação digital da tese ou dissertação. Trata-se de um elemento cujo preenchimento é facultativo e que se subdivide em quatro (04) elementos, Nome (2.1), Sigla (2.2), URL (2.3) e ProvedorServico (2.4). É importante mencionar que, quando um elemento é de preenchimento facultativo, mesmo sendo composto por elementos de preenchimento obrigatório, o mesmo pode ou não ser informado. Entretanto, se um elemento de preenchimento facultativo for informado, é necessário que sejam informados todos os elementos que o compõem e são de preenchimento obrigatório. Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
2.1 Nome	indica o nome da biblioteca digital	Obrigatório	Não adotado	1	1	0	0
2.2 Sigla	indica a sigla da biblioteca digital	Obrigatório	Não adotado	1	1	0	0
2.3 URL	indica o endereço eletrônico do site da biblioteca digital	Obrigatório	URI ²¹	1	1	0	0

Tabela 7 – Descrição dos elementos Nome (2.1), Sigla (2.2) e URL (2.3) do elemento BibliotecaDigital (2)

Convém esclarecer que a instituição responsável pela publicação digital da tese ou dissertação, ou seja, a Biblioteca Digital é a mesma para todas as dissertações e, por este motivo, os dados da mesma foram registrados em uma tabela, juntamente com seu respectivo código. Sendo assim, os elementos foram informados somente uma vez e, são buscados por meio de seu respectivo código, para cada uma das dissertações. Este procedimento ajuda a manter a consistência da base de dados, evitando o preenchimento incorreto e, finalmente, aumentando a qualidade da interoperabilidade.

2.4 ProvedorServico: o conteúdo do elemento ProvedorServico (2.4) indica o nome da instituição que abriga a biblioteca digital. Trata-se de um elemento cujo preenchimento é facultativo e que se subdivide em seis (06) elementos, Nome (2.4.1), Sigla (2.4.2), Pais (2.4.3), UF (2.4.4), CNPJ (2.4.5) e URL (2.4.6). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
2.4.1 Nome	indica o nome da instituição que abriga a biblioteca di-	Obrigatório	Não adotado	1	1	0	0

²¹ URI - *Uniform Resource Identifiers* proporciona um meio simples e extensível para a identificação de um recurso. Fonte: <<http://www.rfc.sunsite.dk/rfc/rfc2396.html>>

	gital						
2.4.2 Sigla	indica a sigla da instituição que abriga a biblioteca digital	Facultativo	Não adotado	1	1	0	0
2.4.3 País	indica o país da instituição que abriga a biblioteca digital	Facultativo	ISO 3166 ²²	1	1	0	0
2.4.4 UF	indica a unidade federativa da instituição que abriga a biblioteca digital	Facultativo	Unidade da Federação ²³	1	1	0	0
2.4.5 CNPJ	indica o CNPJ da instituição que abriga a biblioteca digital	Facultativo	MF 14 dígitos ²⁴	1	0	1	0
2.4.6 URL	indica o endereço eletrônico da instituição que abriga a biblioteca digital	Facultativo	URI	1	1	0	0

Tabela 8 – Descrição dos elementos Nome (2.4.1), Sigla (2.4.2), País (2.4.3), UF (2.4.4), CNPJ (2.4.5) e URL (2.4.6) do elemento ProvedorServico (2.4)

Verifica-se que o elemento CNPJ (2.4.5) foi informado incorretamente. Isto deve-se ao fato de que, para o preenchimento deste código, deve-se consultar uma tabela que, atualmente, registra o CNPJ da UDESC incorretamente. Conseqüentemente,

²² A ISO 3166 - *Codes for the representation of names of countries* - Código para representação de nomes de cidades - é uma norma técnica da ISO para especificação de nomes de cidades. Fonte:

<<http://www.oasis-open.org/cover/country3166.html>>

²³ Tabela de Unidades da Federação do IBGE. Fonte:

<<http://www.ibge.gov.br/cidadesat/topwindow.htm?1>>

²⁴ O padrão MF 14 dígitos especifica que número do CNPJ deve ser informado incluindo-se os 14 dígitos numéricos, sem traços e sem pontos.

toda vez que este código é solicitado, ele será informado de forma errada. Por ser um campo facultativo, o preenchimento incorreto do mesmo não implica em exclusão das respectivas dissertações do processo de *harvesting*. Também não terá conseqüências na busca das dissertações, uma vez que não é disponibilizado busca por meio deste elemento na BDTD e nem na NDLTD. No entanto, trata-se de um dado sem utilidade por estar preenchido incorretamente.

3. BibliotecaDepositaria: o conteúdo do elemento BibliotecaDepositaria (3) indica a biblioteca onde o exemplar físico da tese ou dissertação está arquivado. Trata-se de um elemento cujo preenchimento é facultativo e que se subdivide em quatro (04) elementos, Nome (3.1), Sigla (3.2), URL (3.3) e NumeroChamada (3.4). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
3.1 Nome	indica o nome da biblioteca depositária	Facultativo	Não adotado	451	451	0	0
3.2 Sigla	indica a sigla da biblioteca depositária	Obrigatório	Cadastro CCN/COMUT	451	451	0	0
3.3 URL	indica o endereço eletrônico da biblioteca depositária	Facultativo	URI	451	451	0	0
3.4 NumeroChamada	indica o código que identifica a dissertação no acervo da biblioteca depositária	Facultativo	Não adotado	451	433	0	18

Tabela 9 – Descrição dos elementos Nome (3.1), Sigla (3.2), URL (3.3) e NumeroChamada (3.4) do elemento BibliotecaDepositaria (3)

Ao examinarmos a tabela BibliotecaDepositaria (anexo 4), observou-se que os elementos Nome (3.1), Sigla (3.2) e URL (3.3) foram informados apenas 1 vez, juntamente com o código da biblioteca depositária. O código da biblioteca depositária

também é registrado juntamente com o cadastro dos programas de pós-graduação, estabelecendo desta forma, uma relação entre a tabela que armazena os dados referentes a biblioteca depositária e a tabela que armazena os dados referentes aos programas de pós-graduação.

O elemento NumeroChamada (3.4) deixou de ser informado para 18 dissertações e, por ser um elemento de preenchimento facultativo, as mesmas não foram excluídas do processo de *harvesting*. Mas, como a BDTD disponibiliza a busca pelo número da chamada, a recuperação da informação será prejudicada. Assim, conforme pode-se observar na figura abaixo, as dissertações para as quais não foram informados o elemento NumeroChamada não poderão ser recuperadas por meio deste tipo de busca.

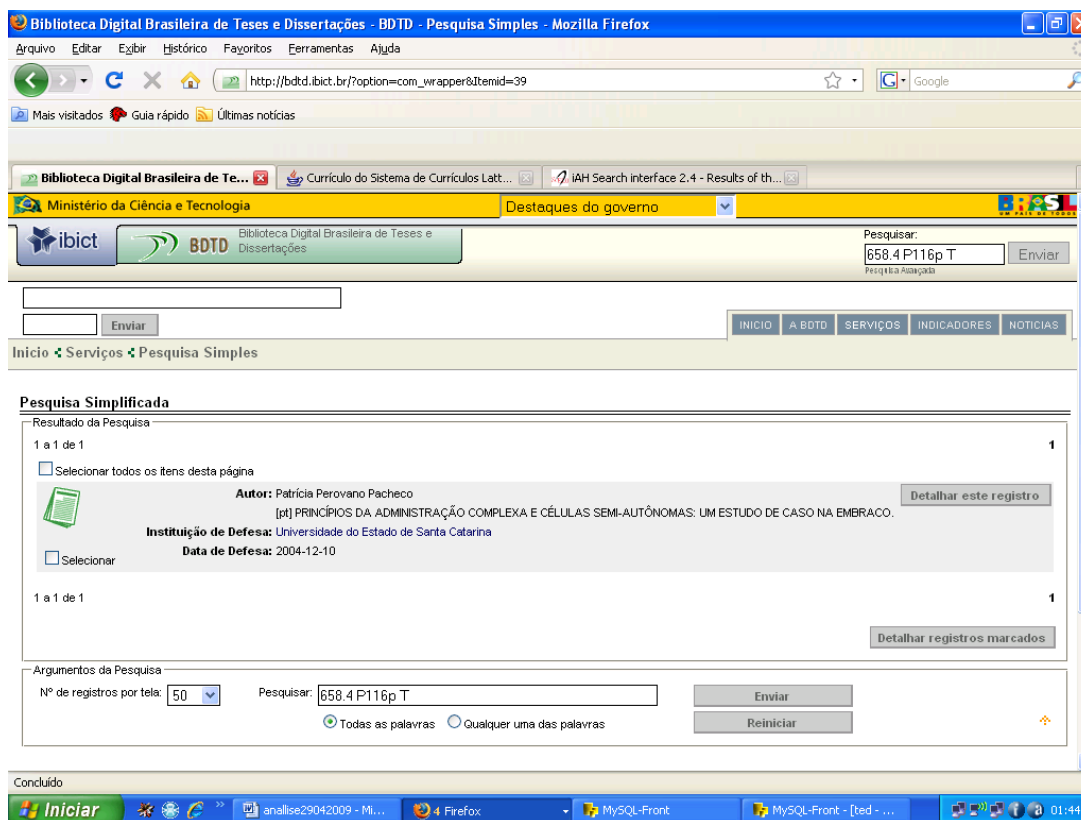


Figura 6 - Recuperação na base de dados da BDTD por meio do elemento NumeroChamada (3.4)

4. Título: o conteúdo do elemento Título (4) indica o título da tese ou dissertação. Trata-se de um elemento de preenchimento obrigatório e que possui o atributo Idioma, cuja finalidade é registrar o Idioma do título informado. Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
4 Título	indica o título da dissertação	Obrigatório	Não adotado	615	613	1	1
Atributo Idioma	indica o idioma no qual foi informado o título da dissertação	Facultativo	Norma ISO 639	615	610	0	5

Tabela 10 – Descrição do elemento Título (4) e do atributo Idioma deste elemento

Deve-se ressaltar que o título da dissertação deve ser, obrigatoriamente, informado no idioma português. Entretanto, o mesmo pode ser informado ou não em outros idiomas. E, por este motivo, o número de elementos Título (4) é superior ao número de dissertações. Convém sinalizar que, para três dissertações, houve repetição da informação do título. Na tela mostrada abaixo, pode-se visualizar o título informado corretamente nos idiomas português (pt) e inglês (en) enquanto nas duas telas seguintes (figuras 8 e 9), pode-se observar a repetição incorreta do idioma em português (pt).

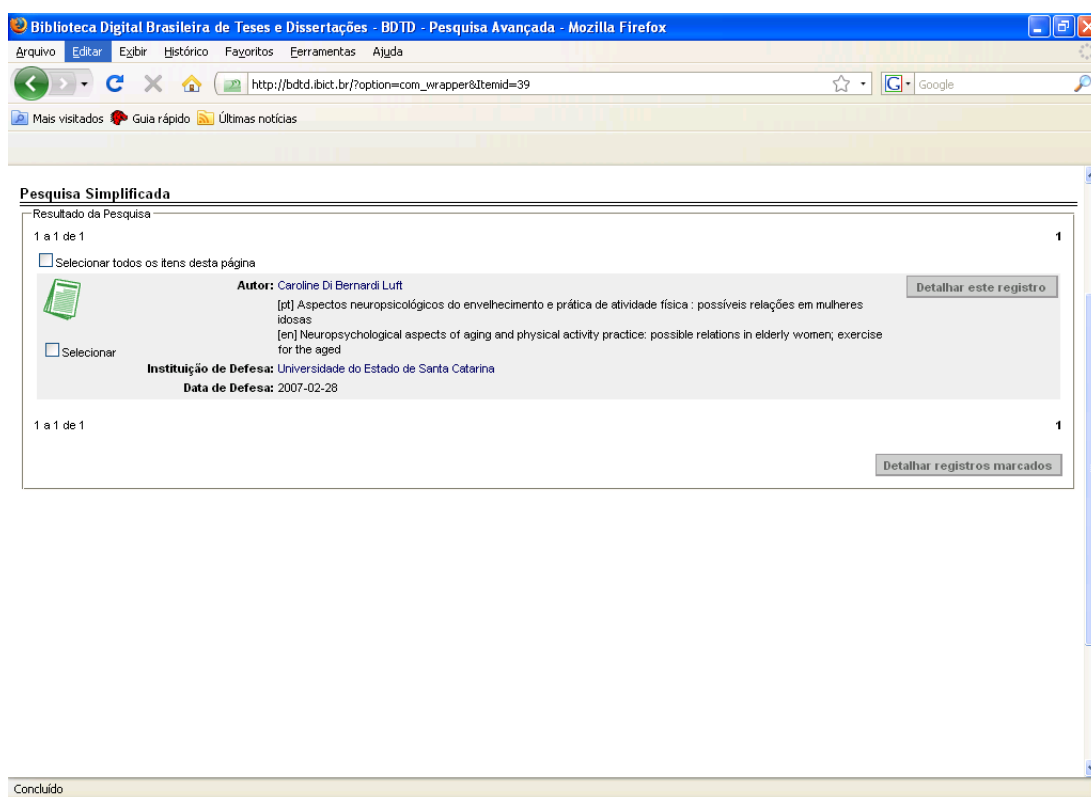


Figura 7 – Título informado nos idiomas português (pt) e inglês (en)

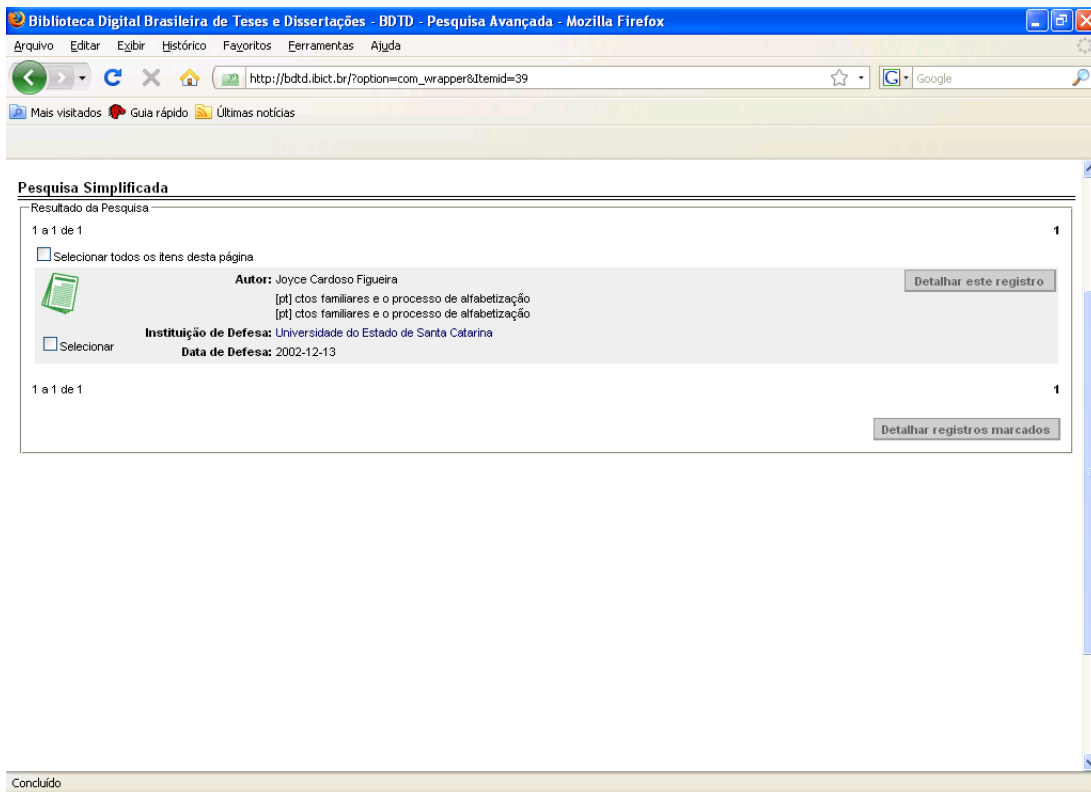


Figura 8 – Título informado 02 vezes em português

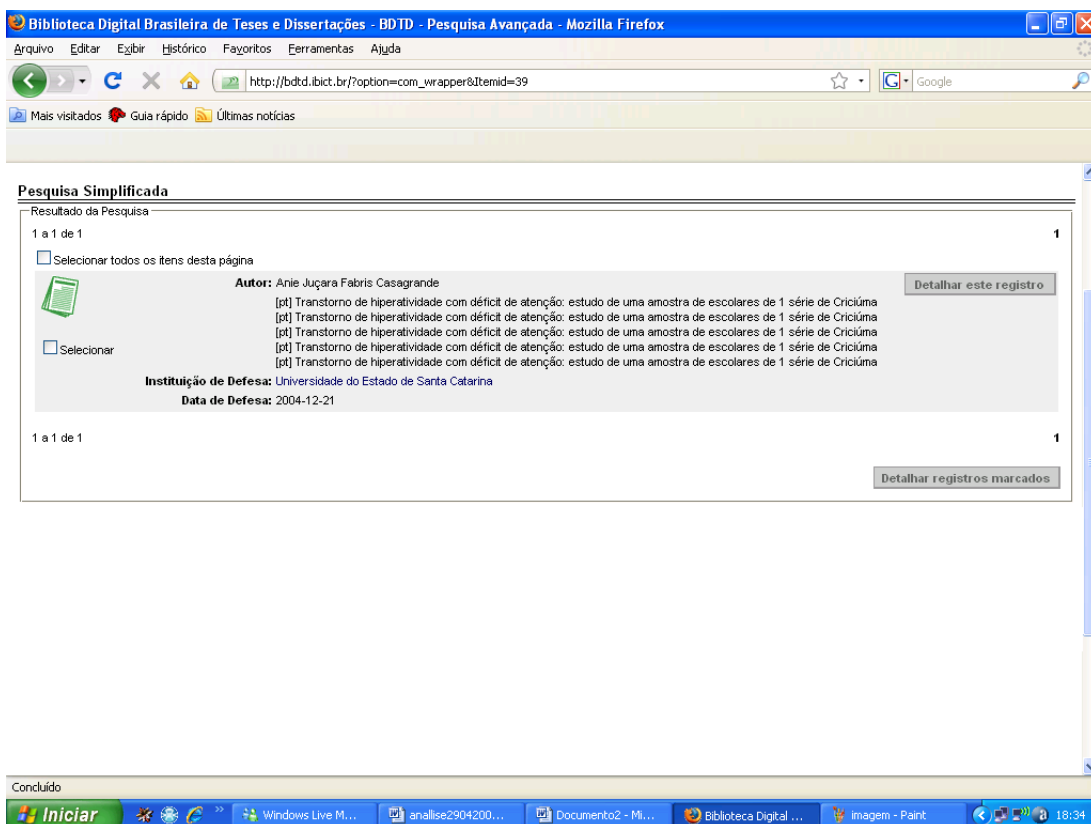


Figura 9 – Título informado 05 vezes em português

A repetição de uma mesma informação não prejudica a recuperação, no entanto pode causar ruídos na compreensão do registro pelo usuário.

Ainda sobre o elemento Título (4), observou-se que o mesmo deixou de ser informado para uma dissertação e foi informado, incorretamente, para outra. Por tratar-se de um elemento de preenchimento obrigatório, a falta de preenchimento implica na exclusão da respectiva dissertação da base de dados da BDTD e, conseqüentemente, da base de dados da NDLTD. Na figura abaixo, pode-se verificar que a dissertação cujo título deixou de ser registrado encontra-se na base de dados da biblioteca da UDESC.

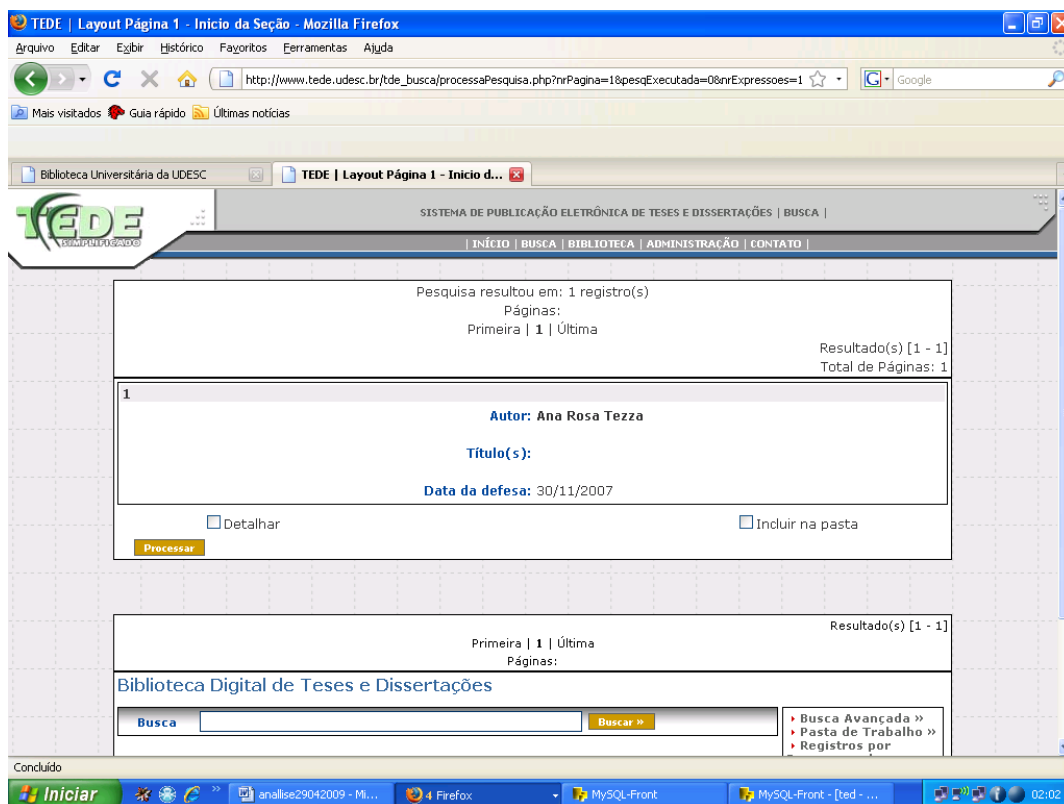


Figura 10 – Recuperação, na base de dados da UDESC, da dissertação com o elemento Título (4) não informado

Entretanto, seus metadados não estão na base de dados da BDTD, conforme pode-se visualizar na figura abaixo:

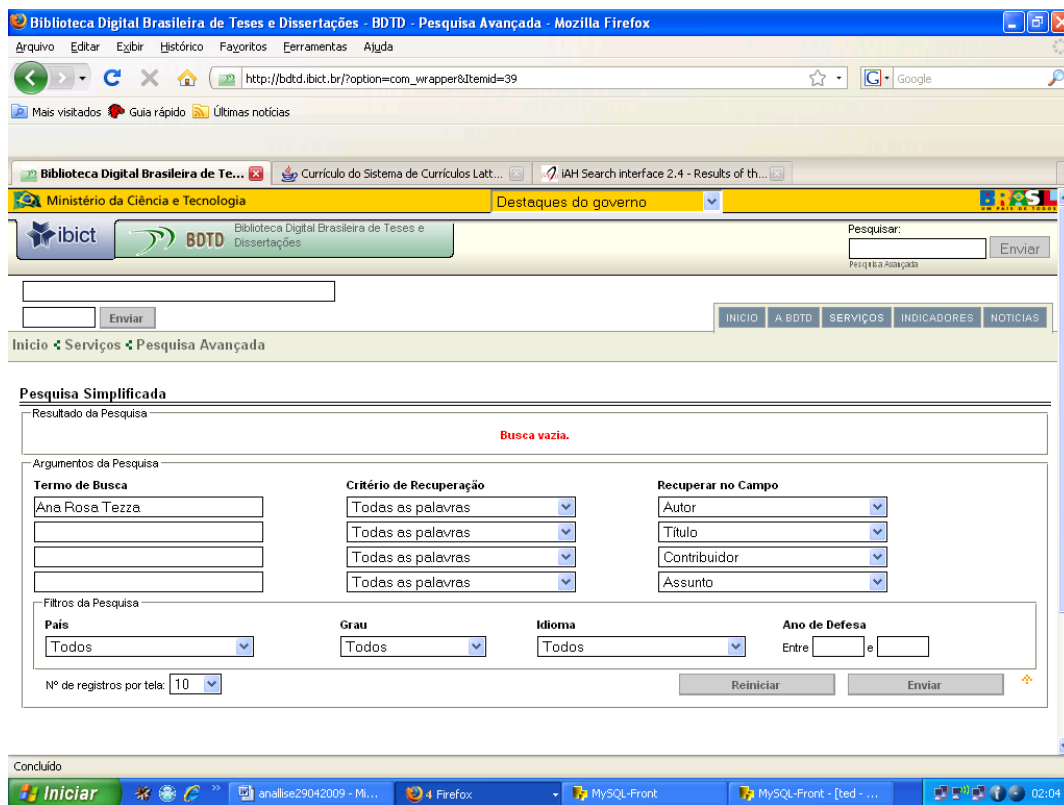


Figura11 – Resultado da busca, feita na base de dados da BDTD, da dissertação com o elemento Título (4) não informado

O atributo Idioma do elemento Título (4) deixou de ser informado para 5 dissertações. Por não ser de preenchimento obrigatório, a falta de informação não ocasionou a exclusão das respectivas dissertações do processo de *harvesting*. Entretanto, impossibilitou a busca por meio do idioma na base de dados da BDTD e da ND LTD.

5. Arquivo: o conteúdo do elemento Arquivo (5) indica o endereço eletrônico dos arquivos da tese ou dissertação. Trata-se de um elemento cujo preenchimento é facultativo e que se subdivide em três (03) elementos, URL (5.1), Legenda (5.2), NivelAcesso (5.3) e seus respectivos atributos. Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
5.1 URL	indica o ende-	Obrigatório	URI	1042	1041	0	1

	reço eletrônico do arquivo ou dos metadados da dissertação na biblioteca de origem						
Atributo Formato ²⁵	indica o formato eletrônico no qual a dissertação está disponível	Facultativo	Não adotado	1042	1041	0	1
Atributo Idioma	indica o(s) idioma(s) no(s) qual(is) o(s) arquivo(s) da dissertação está disponível	Facultativo	URI	1042	0	0	1042
5.2 Legenda	indica a legenda do arquivo da dissertação	Facultativo	Não adotado	1042	1040	0	2
Atributo Idioma	indica o(s) idioma(s) no(s) qual(is) o arquivo vinculado a Legenda está disponível	Facultativo	Não adotado	1042	0	1042	0
5.3 NivelAcesso	indica se o acesso à dissertação poderá ser público ou não	Obrigatório	Público Ou Restrito	1042	1039	3	0

Tabela 11 – Descrição dos elementos URL (5.1), Legenda (5.2), NivelAcesso (5.3) e seus respectivos atributos do elemento Arquivo (5)

²⁵ O atributo Formato é utilizado para especificar o(s) formato(s) eletrônico(s) no(s) qual(is) a tese ou dissertação eletrônica está disponível. Fonte: < <http://www.iana.org/assignments/media-types> >

O elemento URL (5.1) e seu atributo Formato deixaram de ser informados para uma das dissertações. Por tratar-se de um elemento de preenchimento obrigatório, a falta desta informação resultou na exclusão da respectiva dissertação do processo de *harvesting*. Na tela abaixo pode-se visualizar esta dissertação na base de dados da UDESC.

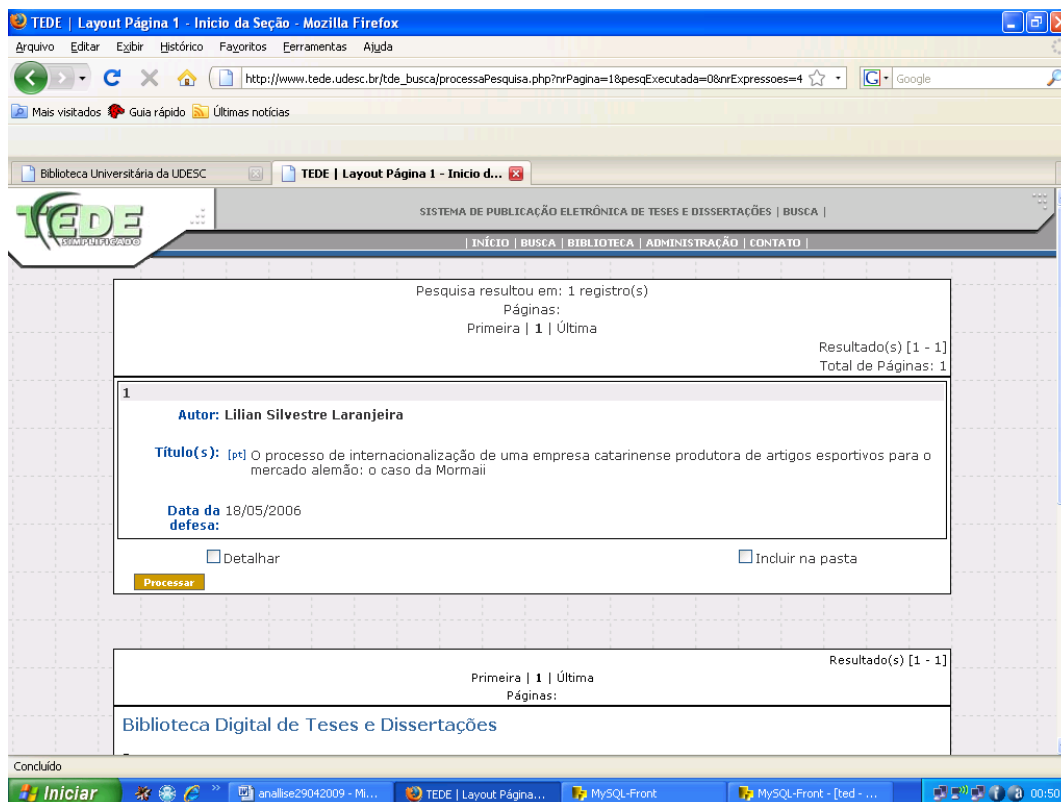


Figura 12 – Busca, na Biblioteca Digital da UDESC, da dissertação para a qual não foi informado o elemento URL (5.1)

Entretanto, pode-se verificar na figura abaixo que a mesma não pode ser recuperada na base de dados da BDTD. Este problema também se repetirá nas buscas realizadas pela NDLTD.

The screenshot shows the BDTD search interface in a Mozilla Firefox browser. The address bar shows the URL: http://bdt.d.ibict.br/?option=com_wrapper&Itemid=39. The page header includes the logo of the Ministério da Ciência e Tecnologia and the BDTD logo. The search bar contains the text "Pesquisar:" and "Pesquisa Avançada". Below the search bar, there are navigation links: "INICIO", "A BDTD", "SERVIÇOS", "INDICADORES", and "NOTÍCIAS". The main content area is titled "Pesquisa Simplificada" and shows the search results. The search criteria are: "Termo de Busca" (Lilian Silvestre Laranjeira, UDESC), "Critério de Recuperação" (Todas as palavras), and "Recuperar no Campo" (Autor, Instituição de Defesa, Contribuidor, Assunto). The filters are: "País" (Todos), "Grau" (Todos), "Idioma" (Todos), and "Ano de Defesa" (Entre e). The number of records per page is set to 10. The search results are empty, displaying "Busca vazia." at the bottom of the results area.

Figura 13 – Busca, na base de dados da BDTD, da dissertação para a qual não foi informado o elemento URL (5.1)

Observou-se que o atributo Idioma do elemento URL (5.1) deixou de ser informado para todas as dissertações. Analisamos a estrutura da tabela Arquivos da base de dados da biblioteca digital de Teses e Dissertações da UDESC, que armazena os metadados deste elemento e constatamos que a mesma possui campos para armazenar o código do arquivo (arCod), o identificador da dissertação (tsIdentificador), o formato do arquivo (arFormato), a URL (arURL) onde se encontra o arquivo e direitos (arDireitos) do mesmo. Entretanto, não possui um campo para armazenar o atributo Idioma do elemento URL. A descrição desta tabela pode ser visualizada na figura abaixo:

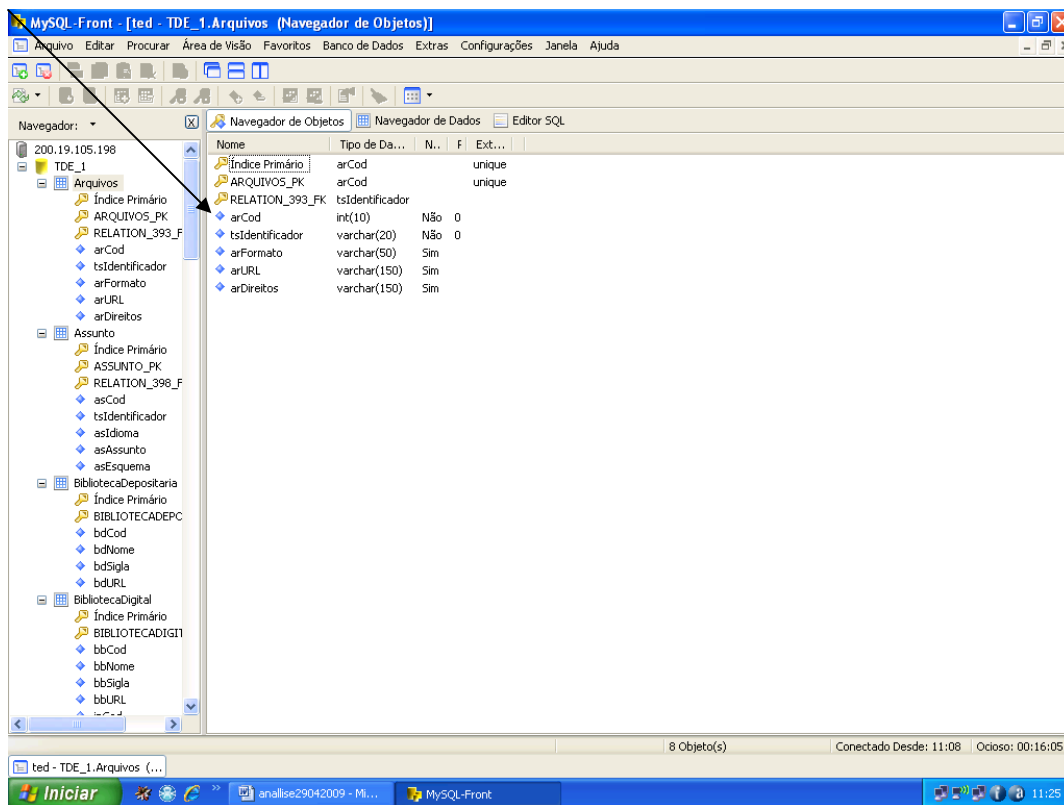


Figura 14 – Descrição da tabela Arquivos da base de dados da biblioteca digital de Teses e Dissertações da UDESC

É importante mencionar que todas as demais tabelas desta base de dados, que encontram-se no anexo 4, foram examinadas a fim de verificar se alguma delas armazena este atributo. No entanto, não foi localizado um campo para o armazenamento do referido atributo, o que impossibilita o preenchimento do mesmo para todas as dissertações. Este fato, nos induziu a refletir sobre este atributo e surgiram as seguintes questões:

- qual a utilidade do atributo Idioma para o elemento URL (5.1)?
- por que não tem um campo para seu armazenamento nas tabelas da biblioteca digital de teses e dissertações fornecidas pela BDTD?
- este atributo poderia ser descartado do padrão MTD-BR?

Por ser facultativo este atributo, o não preenchimento do mesmo não resultou na exclusão das respectivas dissertações do processo de *harvesting*. E, como não é disponibilizada a busca pelo mesmo na BDTD e nem na NDLTD, a falta de preenchimento não impossibilitou as buscas. Sendo assim, concluímos que este atributo

poderia ser dispensado do padrão MTD-BR, contribuindo para que o mesmo se torne mais “enxuto”.

Sobre o elemento Legenda (5.2), observou-se que dos 1042 atributos Idioma analisados, todos foram informados em letras maiúsculas (PT) e, conforme mencionado, anteriormente, o preenchimento deste atributo deve seguir a Norma ISO 639 tanto para 2 (dois) ou 3 (três) caracteres, que determina que o idioma português deve ser identificado como “pt” ou “por” (ambos em minúsculas). Desta forma, constatou-se que a adoção do padrão MTD-BR ocorreu nenhuma vez, durante o preenchimento dos 1042 atributos Idioma do elemento Legenda (5.2). Pode-se observar na figura abaixo que, na base de dados da BDTD, o atributo Idioma apresentado ao lado do elemento Legenda está de acordo com o padrão adotado, mesmo não tendo sido este informado corretamente.

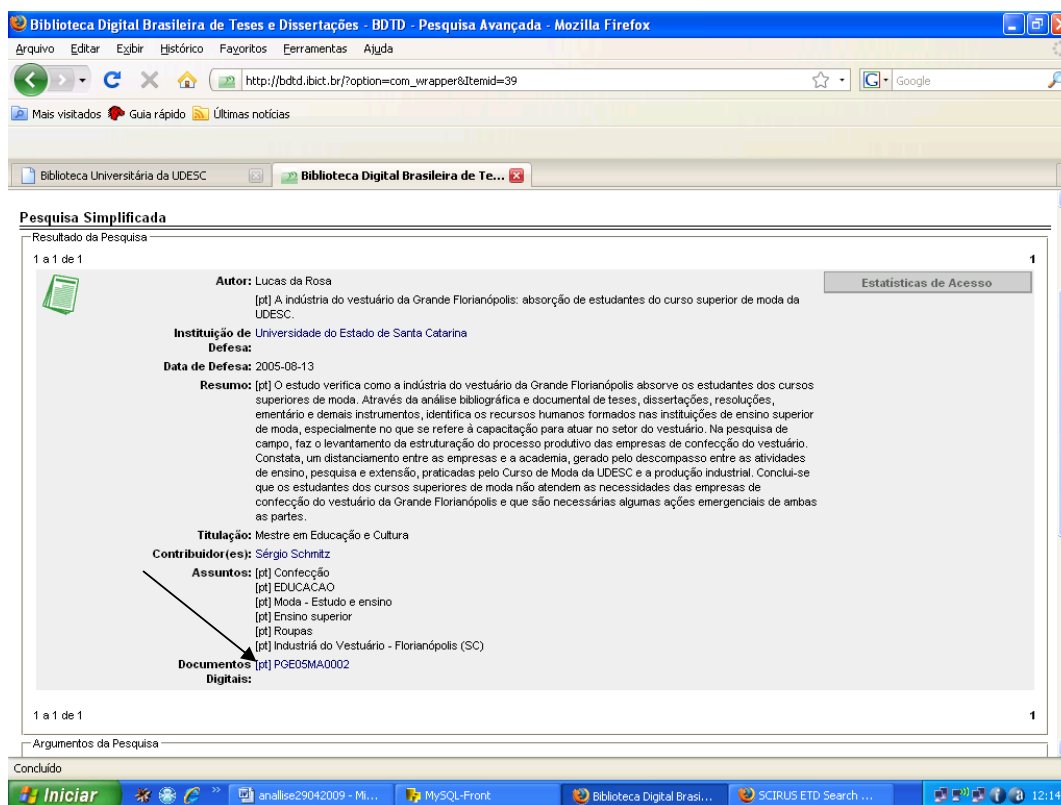


Figura 15 – Atributo Idioma do elemento Legenda (5.2) apresentado de acordo com o padrão adotado

Desta forma, pode-se subentender que existem duas possibilidades; o idioma apresentado ao lado da legenda é o da dissertação e não o da legenda, uma vez que a tabela Teses possui o campo tsIdioma que destina-se ao armazenamento do elemento

Idioma (6), que indica o idioma no qual a dissertação foi escrita, e tendo sido este informado corretamente para todas as dissertações ou o IBICT providenciou as correções necessárias no conteúdo deste elemento, da mesma forma que ocorreu com o elemento `tsDataDefesa(13)`, como mencionado anteriormente. Por ser facultativo o atributo Idioma do elemento Legenda (5.2), o não preenchimento do mesmo não resultou na exclusão das respectivas dissertações do processo de *harvesting*. E, como não é disponibilizada a busca pelo mesmo na BDTD e nem na ND LTD, as dissertações podem ser recuperadas em ambas as bases de dados. Sendo assim, acredita-se ser conveniente uma reflexão sobre este atributo também, uma vez que o mesmo não está sendo utilizado, poderia ser excluído do padrão MTD-BR, sem causar prejuízo.

Ao examinarmos o preenchimento do elemento `NivelAcesso (5.3)`, para todas as dissertações defendidas entre 1997 e 2007, observou-se que o mesmo foi informado 1042 vezes. Destes, 1039 indicam que o acesso poderá ser público, por meio do valor “Publico” e 03 indicam que o acesso não poderá ser público, por meio do valor “Retido”. Sendo assim, pode-se observar que a adoção do padrão MTD-BR, que estabelece que o valor informado para este elemento deve ser “Público” ou “Restrito”, ocorreu em todas as vezes que o conteúdo deste elemento devia indicar que o acesso poderá ser público. Entretanto, ocorreu nenhuma vez, quando devia indicar que o acesso não poderá ser público. O preenchimento incorreto deste elemento não ocasionou a exclusão das respectivas dissertações do processo de *harvesting*. Mas, impossibilitou a busca da mesma por meio da busca por nível de acesso na base de dados da BDTD e da ND LTD.

6. Idioma: o conteúdo do elemento Idioma (6) indica o idioma no qual a tese ou dissertação foi escrita. Trata-se de um elemento de preenchimento obrigatório e o padrão adotado para o seu preenchimento é a Norma ISO 639. Na tabela abaixo, pode-se observar que a descrição dos mesmos foi feita corretamente todas as vezes, favorecendo assim a circulação destes.

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
6 Idioma	indica o idioma da dissertação	Obrigatório	Norma ISO 639	451	451	0	0

Tabela 12 – Descrição do elemento Idioma (6)

7. Grau: o conteúdo do elemento Grau (7) indica o grau acadêmico associado à tese ou dissertação. Trata-se de um elemento de preenchimento é obrigatório e o padrão adotado para o seu preenchimento é o valor “Mestre” ou “Doutor”. Pode-se observar a descrição do mesmo na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
7 Grau	indica o grau acadêmico associado a dissertação	Obrigatório	Mestre ou Doutor	451	450	1	0

Tabela 13 – Descrição do elemento Grau (7)

Ao examinar o preenchimento deste elemento, observou-se que o mesmo foi informado para todas as dissertações defendidas entre 1997 e 2007 e, embora todos estejam de acordo com o padrão especificado, para uma dissertação, o grau informado foi “Doutor”, quando deveria ter sido “Mestre”, o que impossibilita a recuperação da mesma, quando é busca é feita por grau de Mestre.

8. Titulação: o conteúdo do elemento Titulação (8) indica o nome do grau acadêmico associado à tese ou dissertação. Trata-se de um elemento que deve ser, obrigatoriamente, informado e não é adotado padrão para o seu preenchimento. Na tabela abaixo, pode-se observar que a descrição do mesmo foi correta todas as vezes:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
8 Titulacao	indica o nome do grau acadêmico associado a dissertação	Obrigatório	Não adotado	451	451	0	0

Tabela 14 – Descrição do elemento Titulacao (8)

No entanto, por tratar-se de um campo utilizado para buscas, recomenda-se a adoção de uma tabela para evitar possíveis preenchimentos incorretos que possam surgir.

9. Resumo: o conteúdo do elemento Resumo (9) indica o resumo da tese ou dissertação. Trata-se de um elemento que deve ser, obrigatoriamente, informado e não é adotado padrão para o seu preenchimento. O elemento Resumo possui o atributo Idioma, que é facultativo, como mencionado anteriormente, e cuja finalidade é especificar o(s) idioma(s) no(s) qual(is) o(s) resumo está disponível. Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
9 Resumo	indica o resumo da dissertação	Obrigatório	Não adotado	832	830	0	2
Atributo Idioma	indica o idioma do resumo da dissertação	Facultativo	Norma ISO 639	832	733	0	99

Tabela 15 – Descrição do elemento Resumo (9) e do atributo Idioma deste elemento

Deve-se ressaltar que uma dissertação eletrônica pode possuir somente o resumo no idioma português (pt) ou nos idiomas português (pt) e inglês (en). Por esta razão, o número de elementos Resumo é superior ao número de dissertações defendidas entre 1997 e 2007, que é 451. Observou-se que o elemento Resumo (9) deixou de ser informado para duas dissertações, o que ocasionou a exclusão das mesmas do processo de *harvesting* e, conseqüentemente, das bases de dados da BDTD e da NDLTD. Destaca-se a grande importância deste elemento com item de descrição de conteúdo. É por meio dele que o usuário toma a decisão final da recuperação integral da dissertação, uma vez que o resumo funciona como elemento que complementa a descrição pelas palavras-chaves e o título. Desta forma, seu não preenchimento pode acarretar sérios problemas para a circulação dos documentos. O atributo Idioma do elemento Resumo (9) deixou de ser informado para 99 dissertações. Por ser facultativo, a falta de informação deste atributo não resultou na exclusão das respectivas dissertações do processo de *harvesting*. Entretanto, impossibilitou a recuperação das mesmas, por meio da busca por idioma tanto na base de dados da BDTD como na da NDLTD.

10. Cobertura: o conteúdo do elemento Cobertura (10) indica o escopo espacial ou temporal da tese ou dissertação. Trata-se de um elemento que deve ser, facultativamente, informado e não é adotado padrão para o seu preenchimento. O elemento Cobertura possui o atributo Idioma, cuja finalidade é especificar o(s) idioma(s) no(s) qual(is) a cobertura está disponível. Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
10 Cobertura	indica o escopo espacial ou temporal da dissertação	Facultativo	Não adotado	451	0	0	451
Atributo Idioma	indica o idioma no qual foi informada a cobertura	Facultativo	Norma ISO 639	451	0	0	451

Tabela 16 – Descrição do elemento Cobertura (10) e do atributo Idioma deste elemento

O elemento Cobertura (10) e seu respectivo atributo deixaram de ser informados para todas as dissertações eletrônicas da UDESC, defendidas entre 1997 e 2007. Por serem ambos facultativos, a falta de preenchimento não ocasionou a exclusão das respectivas dissertações do processo de *harvesting*. Entretanto, tornou inviável a recuperação das mesmas por meio da busca por cobertura na base de dados da BDTD e da NDLTD.

11. Assunto: o conteúdo do elemento Assunto (11) indica os tópicos tratados na tese ou dissertação e a tabela de onde estes tópicos foram extraídos, quando for o caso. O elemento Assunto possui o atributo Idioma, cuja finalidade é especificar o(s) idioma(s) no(s) qual(is) o(s) assunto(s) da tese ou dissertação eletrônica está(ão) disponível(is) e o atributo Esquema, cujo propósito é especificar o(s) instrumento(s) de indexação utilizado(s) pela biblioteca depositária. Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
11 Assunto	indica os tópicos tratados na dissertação e a tabela de onde os mesmos foram extraídos	Facultativo	Não adotado	3655	3653	0	2
Atributo Idioma	indica o(s) idioma no(s) qual(is) o(s) assunto(s) da dissertação eletrônica está disponível	Facultativo	Norma ISO 639	3655	3487	0	168
Atributo Esquema	indica o(s) instrumento(s) de indexação utilizado(s) pela biblioteca depositária	Facultativo	Não adotado	3655	3655	0	0

Tabela 17 – Descrição do elemento Assunto (11) e dos atributos Idioma e Esquema deste elemento

Deve-se ressaltar que uma dissertação eletrônica pode abranger muitos assuntos. Por esta razão, o número de elementos Assunto é superior ao número de dissertações defendidas entre 1997 e 2007. Ao examinar o preenchimento deste elemento, observou-se que o mesmo foi informado 3653 vezes e deixou de ser informado para duas dissertações, o que resultou na impossibilidade de recuperação das mesmas por meio da busca por assunto na base de dados da BDTD e da NDLTD, impactando na revocação e na precisão dos respectivos SRIs. Ainda sobre o preenchimento deste elemento, é importante mencionar que são inseridas as palavras-chaves descritas pelos autores, e também os assuntos padronizados pelo catálogo de Autoridades (termos autorizados de

nomes e assuntos). oferecido pela Rede Bibliodata²⁶ da Fundação Getúlio Vargas – FGV. O atributo Idioma deste elemento deixou de ser informado 168 vezes. Mas, por ser de preenchimento facultativo, não ocasionou exclusão das referidas dissertações do processo de *harvesting*. E, como não está disponível a busca pelo atributo Idioma do elemento Assunto (11) na base de dados da BDTD nem na da NDLTD, a recuperação destas dissertações continuou sendo possível em ambas as bases, por meio de outros tipos de busca.

12. LocalDefesa: o conteúdo do elemento LocalDefesa (12) indica o local da defesa da tese ou dissertação. Trata-se de um elemento cujo preenchimento é facultativo e que se subdivide em três (03) elementos, Cidade (12.1), UF (12.2) e País (12.3). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
12.1 Cidade	indica a cidade onde foi defendida a dissertação	Obrigatório	Não adotado	451	449	2	0
12.2 UF	indica a UF onde foi defendida a dissertação	Facultativo	Unidade Da Federação	451	444	2	5
12.3 País	indica o país onde foi defendida a dissertação	Facultativo	Norma ISO 3166	451	451	0	0

Tabela 18 – Descrição dos elementos Cidade (12.1), UF (12.2) e País (12.3) do elemento LocalDefesa (12)

Convém pontuar que, mesmo não sendo adotado um padrão para o preenchimento do campo Cidade (12.1), para 449 dissertações os mesmos foram informados corretamente. Entretanto, para duas destas, as cidades foram informadas

²⁶ A Rede Bibliodata é uma rede cooperativa de Bibliotecas brasileiras que tem seus acervos representados no Catálogo Coletivo Bibliodata, realizam a catalogação cooperativa e compartilham produtos e serviços, visando a redução dos custos, além de promover a difusão dos acervos bibliográficos de suas instituições. Fonte: < <http://www8.fgv.br/bibliodata/> >

incorretamente, uma foi informada como “Florinaópolis” e a outra, como “Florianópolis, SC”, ao invés de Florianópolis, o que inviabiliza a recuperação das mesmas por meio da busca por cidade na base de dados da BDTD e da NDLTD. Também observou-se que este elemento foi informado com e sem acento, demonstrando uma falta de padronização no preenchimento. Na figura abaixo, pode-se observar que para recuperar a dissertação por meio da busca pela cidade de defesa, é necessário que a mesma seja informada da mesma forma que se encontra na base de dados da BDTD. Sendo assim, foi adotada a opção incorreta “Florinaópolis” e as respectivas dissertações foram recuperadas.

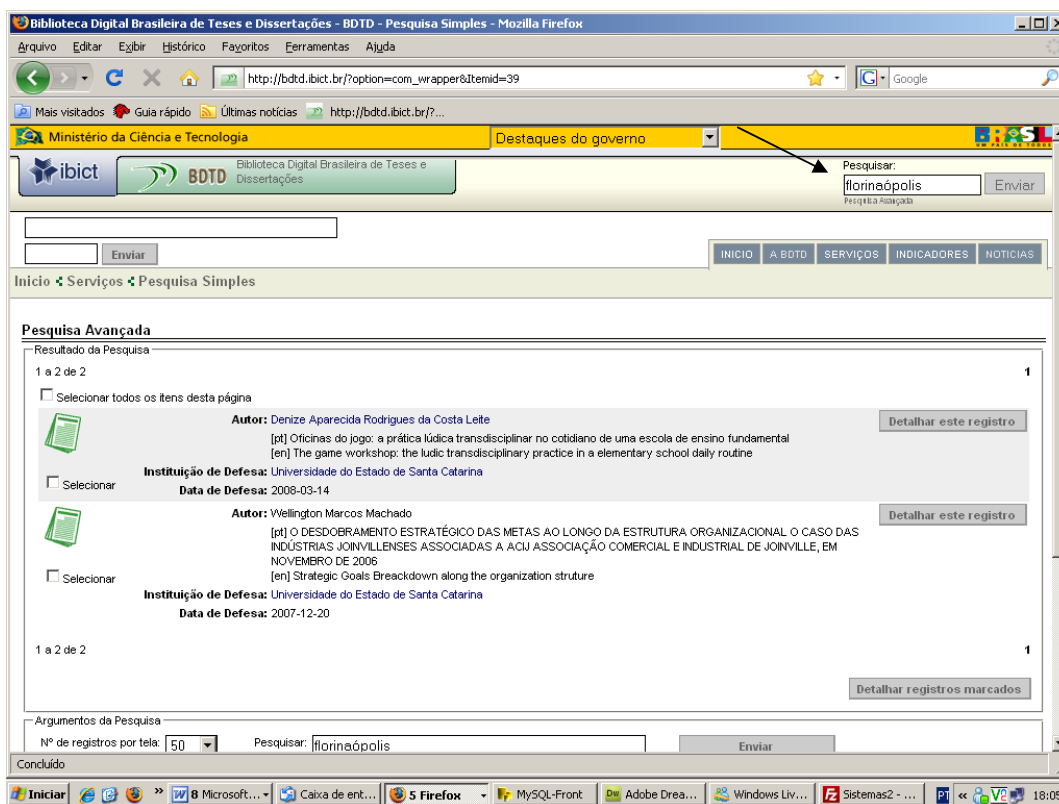


Figura 16 – Busca, na base de dados da BDTD, da dissertação para a qual o local de defesa foi informado incorretamente

Convém esclarecer que a busca recuperou duas dissertações com o local de defesa informado como “Florinaópolis” mas, uma destas foi defendida no ano de 2008 e, por este motivo, não faz parte da população desta pesquisa.

Sobre o preenchimento do elemento UF (12.2), observou-se que, mesmo sendo respeitado o padrão adotado, duas vezes o mesmo foi descrito incorretamente, uma vez que foi registrado “SP”, ao invés de SC. Informar incorretamente este elemento tem

como consequência a impossibilidade de recuperar as respectivas dissertações, por meio da busca por UF, na base de dados da BDTD e da NDLTD.

13. DataDefesa: o conteúdo do elemento DataDefesa (13) indica a data em que foi defendida a tese ou dissertação eletrônica. Este elemento deve ser, obrigatoriamente, informado e é adotado o padrão da ISO 8601²⁷. Pode-se observar a descrição do mesmo na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
13 DataDefesa	indica a data da defesa da dissertação eletrônica	Obrigatório	Norma ISO 8601	451	449	0	2

Tabela 19 – Descrição do elemento DataDefesa (13)

Apesar desta tabela mostrar que não houve preenchimento incorreto do elemento DataDefesa, é conveniente relatar um fato histórico a este respeito. Uma busca realizada no mês de julho de 2008, o ano de defesa de uma dissertação, foi inscrito como 2204, ao invés de 2004, impedindo que a respectiva dissertação fosse localizada e recuperada por meio da busca por período. Entretanto, a mesma podia ser localizada e recuperada por meio de outras formas de busca, conforme pode-se visualizar na figura abaixo:

²⁷ A ISO 8601 - *Representations of dates and times, second edition* – Representação de datas e horários- é uma norma técnica da ISO que especifica que a data deve ser informada nos seguintes formatos: ‘YYYY-MM-DD’ ou ‘YYYY-MM’ ou ‘YYYY’, onde YYYY = ano com 4 dígitos, MM = mês com dois dígitos e DD = dia com dois dígitos.

The screenshot shows the BDTD search interface. At the top, the browser title is 'Biblioteca Digital Brasileira de Teses e Dissertações - BDTD - Pesquisa Avançada - Mozilla Firefox'. The address bar shows the URL 'http://bdtbd2.ibict.br/?option=com_wrapper&Itemid=39'. The page header includes the logo of the Ministério da Ciência e Tecnologia and the BDTD logo. A search bar is visible with the text 'Pesquisar:' and a search button. Below the search bar, there are navigation links: 'INICIO', 'A BDTD', 'SERVIÇOS', 'INDICADORES', and 'NOTÍCIAS'. The main content area is titled 'Pesquisa Simplificada' and shows the search results. The first result is for a dissertation by Kasselandra Mattos Soares, with the title '[pt] História da matemática na formação de professores do ensino fundamental - (1 a 4 série)'. The search criteria are 'Instituição de Defesa: Universidade do Estado de Santa Catarina' and 'Data de Defesa: 2204-10-15'. The page also displays search filters and navigation options.

Figura 17 – Recuperação de dissertação na BDTD com o elemento DataDefesa incorreto, realizada no mês de julho de 2008

Após o mês de julho de 2008, observou-se uma correção realizada na data de defesa da dissertação mencionada acima, conforme pode-se observar nas figuras abaixo:

Biblioteca Digital Brasileira de Teses e Dissertações - BDTD - Pesquisa Avançada - Mozilla Firefox

Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://bdtb.ibict.br/?option=com_wrapper&Itemid=39

Ministério da Ciência e Tecnologia Destaques do governo

ibict BDTD Biblioteca Digital Brasileira de Teses e Dissertações

Pesquisar: Enviar

Enviar

INICIO A BDTD SERVIÇOS INDICADORES NOTÍCIAS

Início Serviços Pesquisa Avançada

Pesquisa Simplificada

Resultado da Pesquisa

1 a 1 de 1

Selecionar todos os itens desta página

Autor: Kasselandra Mattos Soares [Detalhar este registro](#)

[pt] História da matemática na formação de professores do ensino fundamental - (1 a 4 série)

Instituição de Defesa: Universidade do Estado de Santa Catarina

Selecionar **Data de Defesa:** 2004-10-15

1 a 1 de 1 [Detalhar registros marcados](#)

Argumentos da Pesquisa

Termo de Busca	Critério de Recuperação	Recuperar no Campo
<input type="text" value="kasselandra"/>	Todas as palavras	Autor
<input type="text" value="udesc"/>	Todas as palavras	Instituição de Defesa
<input type="text"/>	Todas as palavras	Contribuidor
<input type="text"/>	Todas as palavras	Assunto

Filtros da Pesquisa

Pais	Grav	Idioma	Ano de Defesa
Concluído			

Figura 18 – Recuperação de dissertação na BDTD com o elemento DataDefesa correto, realizada no mês de janeiro de 2009

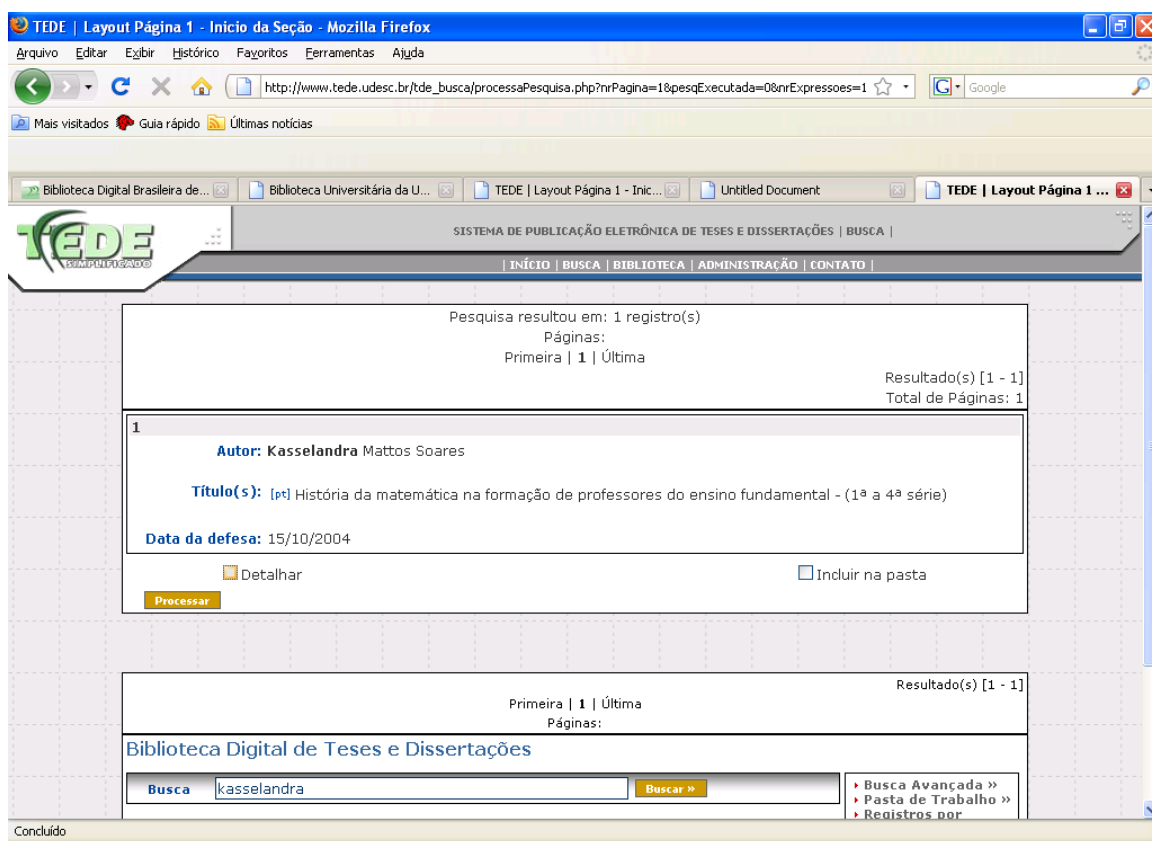


Figura 19 - Recuperação da dissertação na biblioteca de teses e dissertações eletrônicas da UDESC, realizada em fevereiro de 2009

Este fato sugere que haja provavelmente um procedimento de correção dos metadados realizada pelo provedor de serviço BDTD.

Além deste caso de data de defesa cadastrada incorretamente, mencionamos também, o caso das dissertações que ainda encontram-se sem data de defesa na biblioteca de teses e dissertações eletrônicas da UDESC atualmente, embora as correções tenham sido efetuadas na base de dados da BDTD, conforme pode-se observar nas figuras abaixo:

TEDE | Layout Página 1 - Início da Seção - Mozilla Firefox

Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://www.tede.udesc.br/tdc_busca/processaPesquisa.php?nrPagina=1&pesqExecutada=0&nrExpressoes=1

Mais visitados Guia rápido Últimas notícias HotMail gratuito Personalizar Links http://200.19.105.30... Windows Media Windows

Biblioteca Universitária da UDESC TEDE | Layout Página 1 - Início d...

Resultado(s) [1 - 1]
Total de Páginas: 1

1

Autor: Rogério Luiz Saturnino

Título(s): [pt] Feira de ciências e tecnologia: um espaço para a produção de conhecimento no âmbito da SOCIESC - A 34ª FECIETT / Rogério Luiz Saturnino

Data da defesa: 0000

Detalhar Incluir na pasta

Processar

Primeira | 1 | Última
Páginas:

Resultado(s) [1 - 1]

Biblioteca Digital de Teses e Dissertações

Busca saturnino Buscar »

- » Busca Avançada »
- » Pasta de Trabalho »
- » Registros por Programas de pós-graduação »
- » Ajuda ?

Registros por página 5

v52.04 | COPYRIGHT © desenvolvido para IE 5.0+

logo 1 logo 2 BDTD ibict

Concluído

Figura 20 - Recuperação da dissertação na biblioteca de teses e dissertações eletrônicas da UDESC, realizada em fevereiro de 2009

Biblioteca Digital Brasileira de Teses e Dissertações - BDTD - Pesquisa Avançada - Mozilla Firefox


Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://bdt.d.ibict.br/?option=com_wrapper&Itemid=39

Mais visitados Guia rápido Últimas notícias HotMail gratuito Personalizar Links http://200.19.105.30... Windows Media Windows

1 a 1 de 1

Selecionar todos os itens desta página

 **Autor:** Rogério Luiz Saturnino
[pt] Feira de ciências e tecnologia: um espaço para a produção de conhecimento no âmbito da SOCIESC - A 34 FECIETT / Rogério Luiz Saturnino

Selecionar **Instituição de Defesa:** Universidade do Estado de Santa Catarina
Data de Defesa: 2004-04-31

1 a 1 de 1

Detalhar este registro

Detalhar registros marcados

Argumentos da Pesquisa

Termo de Busca	Critério de Recuperação	Recuperar no Campo
saturnino	Todas as palavras	Autor
udesc	Todas as palavras	Instituição de Defesa
	Todas as palavras	Contribuidor
	Todas as palavras	Assunto

Filtros da Pesquisa

Pais	Grau	Idioma	Ano de Defesa
Todos	Todos	Todos	Entre e

10

REQUISITOS MÍNIMOS ELABORAÇÃO COPYLEFT

[Voltar]

Concluído

topo

Figura 21 - Recuperação da dissertação na BDTD, realizada em fevereiro de 2009

The screenshot shows a web browser window with the URL http://www.tede.udesc.br/tdc_busca/processaPesquisa.php?nrPagina=1&pesqExecutada=0&nrExpressoes=4f. The page title is "TEDE | Layout Página 1 - Início da Seção - Mozilla Firefox". The browser's address bar shows the URL and a search engine icon. The page content includes a navigation menu with links for "INÍCIO", "BUSCA", "BIBLIOTECA", "ADMINISTRAÇÃO", and "CONTATO". The main content area displays the search results for the query "Nelson de Santi Junior". The results show one record with the following details:

- Pesquisa resultou em:** 1 registro(s)
- Páginas:** Primeira | 1 | Última
- Resultado(s) [1 - 1]**
- Total de Páginas:** 1
- 1**
- Autor:** Nelson de Santi Junior
- Título(s):** [pt] Avaliação numérico-experimental da fratura dúctil em processos predominantementetrativos e compressivos
- Data da defesa:** 0000
- Detalhar
- Incluir na pasta
- Processar**

Below the search results, there is a section for "Biblioteca Digital de Teses e Dissertações" with a search bar labeled "Buscar por". The search bar contains the text "Nelson de Santi Junior" and a dropdown menu set to "Autor". A "Busca" button is visible to the right of the search bar. The status bar at the bottom of the browser window shows "Concluído".

Figura 22 - Recuperação da dissertação na biblioteca de teses e dissertações eletrônicas da UDESC, realizada em fevereiro de 2009

The screenshot shows the BDTD search interface. At the top, the browser window displays the URL http://bdtd.ibict.br/?option=com_wrapper&Itemid=39. The search results section shows a single entry for a dissertation by Nelson de Santi Junior, with the title "[pt] Avaliação numérico-experimental da fratura dúctil em processos predominantemente tratativos e compressivos", defended at Universidade do Estado de Santa Catarina on 2006-02-29. Below the search results, there are sections for "Argumentos da Pesquisa" (Search Arguments) and "Filtros da Pesquisa" (Search Filters). The "Argumentos da Pesquisa" section includes fields for "Termo de Busca" (Search Term) with the text "de santi" and "udesc", and "Critério de Recuperação" (Recovery Criteria) with dropdown menus for "Todas as palavras" (All words) and "Recuperar no Campo" (Recover in Field) with dropdown menus for "Autor" (Author), "Instituição de Defesa" (Defending Institution), "Contribuidor" (Contributor), and "Assunto" (Subject). The "Filtros da Pesquisa" section includes dropdown menus for "País" (Country) set to "Todos" (All), "Grau" (Degree) set to "Todos" (All), "Idioma" (Language) set to "Todos" (All), and "Ano de Defesa" (Year of Defense) with fields for "Entre" (Between) and "e" (and). At the bottom, there are buttons for "Reiniciar" (Reset) and "Enviar" (Submit), and a "Nº de registros por tela" (Number of records per page) dropdown set to "10".

Figura 23 - Recuperação da dissertação na BDTD, realizada em fevereiro de 2009

Percebe-se portanto que o IBICT efetivamente providenciou as correções necessárias no conteúdo do elemento DataDefesa, em sua base de dados, para as dissertações que estavam com este elemento registrado incorretamente. Desta forma, possibilitou a recuperação das mesmas por meio da busca por período. Entretanto, a necessidade deste retrabalho poderia ter sido evitada se uma verificação de datas fosse realizada no momento em que os metadados são informados pela UDESC. Diante do exposto, deve-se ressaltar a importância da verificação da data de defesa, via sistema, a fim de evitar cadastros com datas incorretas.

14. Autor: o conteúdo do elemento Autor (14) indica dados do autor da tese ou dissertação eletrônica. Trata-se de um elemento cujo preenchimento é obrigatório e que se subdivide em cinco (05) elementos, Nome (14.1), Citação (14.2), Lattes (14.3), CPF (14.4), Afiliação (14.5) e seus respectivos atributos. Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
14.1 Nome	indica o nome da pessoa física do autor	Obrigatório	Não adotado	451	450	1	0
14.2 Citação	indica a forma pela qual a pessoa física será citada	Facultativo	Não adotado	451	451	0	0
Atributo Norma	indica a(s) norma(s) utilizada(s) para informar a forma pela qual o autor será citado	Facultativo	Não adotado	0	0	0	451
14.3 Lattes	indica o endereço eletrônico do currículo do autor na Plataforma Lattes	Facultativo	URI	451	101	26	324
14.4 CPF	indica o CPF do autor da dissertação	Facultativo	MF 11 dígitos ²⁸	451	14	437	0
14.5 Afiliação	indica o(s) idioma(s) no(s) qual(is) a legenda está disponível	Facultativo	Não adotado	1045	0	1045	0

Tabela 20 – Descrição dos elementos Nome (14.1), Citação (14.2), Lattes (14.3), CPF (14.4) e Afiliação (14.5) do elemento Autor (14)

Convém pontuar que, mesmo não sendo adotado padrão para o preenchimento do elemento Nome (14.1), para 450 dissertações os mesmos foram inscritos corretamente. Entretanto, para uma destas, o nome foi registrado incorretamente, uma vez que foi incluída uma vírgula entre o nome e sobrenome da autora. Por tratar-se de um erro que não alterou o nome da autora, a recuperação da dissertação, por meio da

²⁸ O padrão MF 11 dígitos especifica que número do CPF deve ser informado incluindo-se os 11 dígitos numéricos, sem traços e sem pontos.

busca pelo nome do autor, continuou sendo possível na base de dados da BDTD e da NDLTD. Conforme pode-se observar nas telas abaixo:

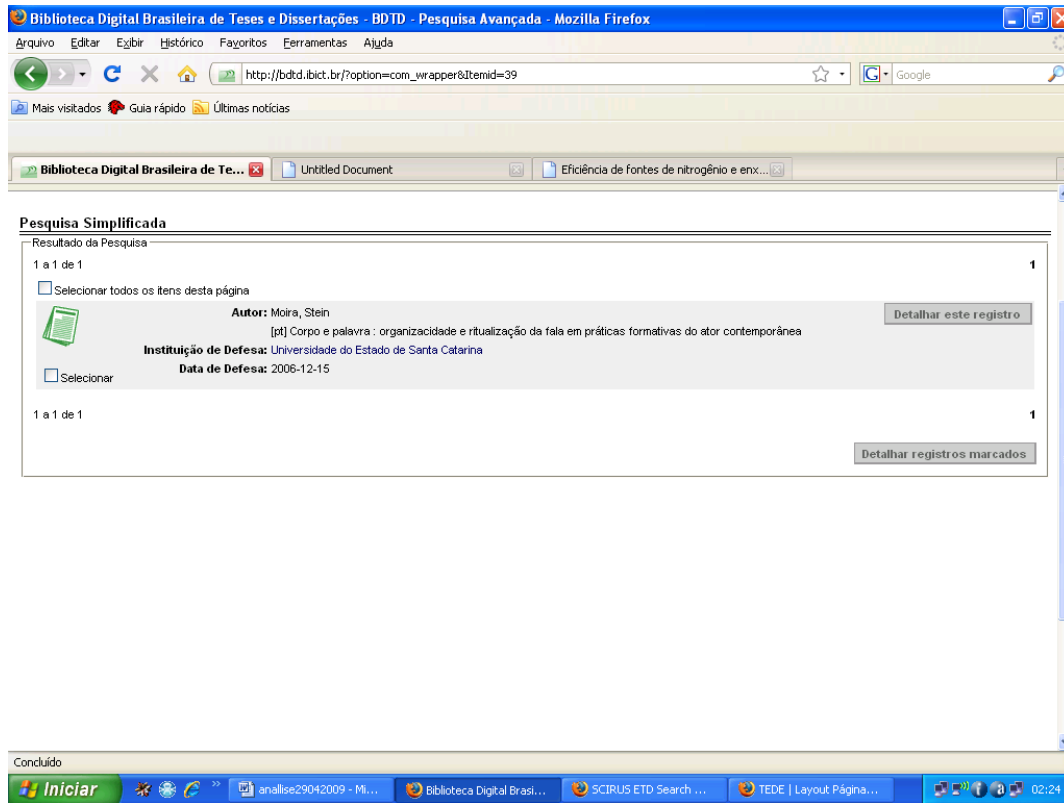


Figura 24 – Recuperação, na base de dados da BDTD, da dissertação cujo nome da autora tem uma vírgula ao lado do nome

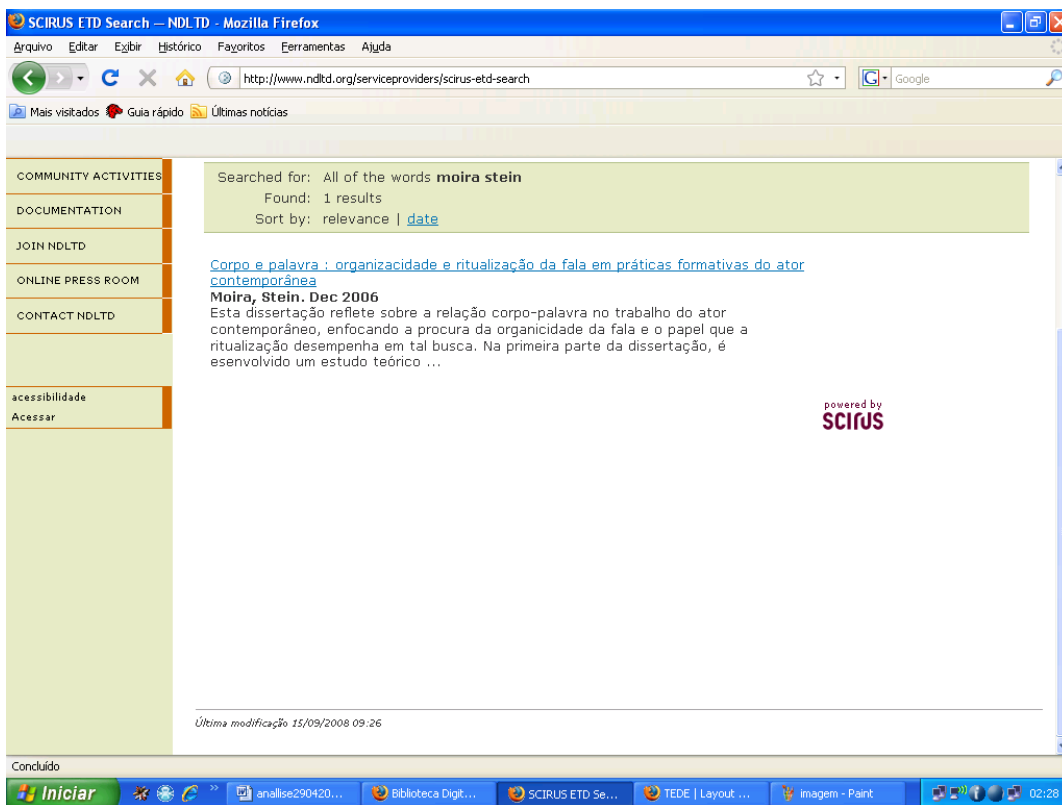


Figura 25 – Recuperação, na base de dados da NDLTD, da dissertação cujo nome da autora tem uma vírgula ao lado do nome

Mesmo que o elemento Citacao (14.2) tenha sido registrado corretamente todas as vezes, não foram identificados atributos Norma para este elemento, uma vez que o mesmo deixou de ser informado para as 451 dissertações defendidas entre 1997 e 2007. Por ser de preenchimento facultativo e não estar disponível a busca por meio do atributo Norma, a falta de preenchimento não impossibilitou a busca das respectivas dissertações na base de dados da BDTD e da NDLTD.

Ao examinar o preenchimento do elemento Lattes (14.3), observou-se que o mesmo foi registrado 101 vezes, corretamente e 26 vezes incorretamente. Neste último caso, embora o padrão adotado tenha sido respeitado, o endereço eletrônico informado direciona para um local que não é o currículo Lattes do autor. Este elemento ainda deixou de ser informado 324 vezes. Cabe destacar que o currículo Lattes possui links para outras bases como Scielo, por exemplo. A falta de informação deste elemento ou seu preenchimento incorreto impossibilita o acesso ao currículo Lattes do autor e, conseqüentemente, aos links para outros trabalhos do autor, assim como a recuperação da respectiva dissertação por meio da busca pelo Lattes na base de dados da BDTD.

Sobre o elemento CPF (14.4), observou-se que ele foi inscrito 14 vezes, corretamente e 437 vezes incorretamente. Para estes últimos, o padrão adotado foi totalmente desconsiderado, sendo observados os mais diversos formatos como com 10, 11, 12, 13 e 14 dígitos; com pontos, com hífen e com seqüência de zeros. O preenchimento incorreto do elemento CPF (14.4) inviabiliza a recuperação das respectivas dissertações na base de dados da BDTD e da NDLTD. Além disso, torna impossível sua utilização quando, futuramente, a BDTD venha a adotar este elemento de descrição única como item de ligação com outras bases de dados como o Lattes.

14.5 Afiliacao: o conteúdo do elemento Afiliacao (14.5) indica dados referentes a Instituição à qual o autor da tese ou dissertação eletrônica é afiliado. Trata-se de um elemento de preenchimento facultativo e que se subdivide em seis (06) elementos, Nome (14.5.1), Sigla (14.5.2), País (14.5.3), UF (14.5.4), CNPJ (14.5.5) e URL (14.5.6). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
14.5.1 Nome	indica o nome da Instituição à qual o autor da dissertação é afiliado	Obrigatório	Não adotado	451	241	0	210
14.5.2 Sigla	indica a sigla da Instituição à qual o autor da dissertação é afiliado	Facultativo	Não adotado	451	241	0	210
14.5.3 País	indica o país da Instituição à qual o autor da dissertação é afiliado	Facultativo	Norma ISO 3166	451	241	0	210
14.5.4 UF	indica a Unidade Federativa da Instituição à qual o autor da dis-	Facultativo	Unidade da Federação	451	241	0	210

	sertação é afiliado						
14.5.5 CNPJ	indica o CNPJ da Instituição à qual o autor da dissertação é afiliado	Facultativo	MF 14 dígitos	451	0	241	210
14.5.6 URL	indica o endereço eletrônico da Instituição à qual o autor da dissertação é afiliado	Facultativo	URI	451	241	0	210

Tabela 21 – Descrição dos elementos Nome (14.5.1), Sigla (14.5.2), País (14.5.3), UF (14.5.4), CNPJ (14.5.5) e URL (14.5.6) do elemento Afiliacao (14.5)

Ao examinar o preenchimento dos elementos Nome (14.5.1), Sigla (14.5.2), País (14.5.3), UF (14.5.4) e URL (14.5.6) observou-se que os mesmos foram informado 241 vezes corretamente e deixaram de ser informados 210 vezes. A falta de informação destes elementos resulta na impossibilidade de recuperar, na base de dados da BDTD e da NDLTD, as respectivas dissertações, por meio da busca pelo nome, sigla e URL da instituição à qual o autor da dissertação eletrônica é afiliado. A recuperação das dissertações por meio da busca pelo país ou pela UF da instituição à qual o autor da dissertação eletrônica é afiliado também não é possível, quando estes elementos deixam de ser informados. Entretanto, se os elementos país e UF da biblioteca digital responsável pela publicação digital da dissertação, do local de defesa da dissertação, da instituição de defesa da dissertação ou da agência de fomento que apoiou financeiramente o autor tiverem sido informados corretamente, a busca será feita por meio destes, uma vez que é pela busca simples que o país e a UF são informados para a busca. Sendo assim, as respectivas dissertações serão recuperadas na base de dados da BDTD e da NDLTD.

Sobre o preenchimento do elemento CNPJ (14.5.5), observou-se que o mesmo foi registrado 241 vezes incorretamente e deixou de ser informado 210 vezes. É importante destacar que a busca na base de dados da BDTD e da NDLTD não é realizada pelo CNPJ e, por este motivo, o preenchimento incorreto ou a falta de

informação deste elemento não resulta no processo de exclusão das respectivas dissertações do processo de *harvesting*.

15. Contribuidor: o conteúdo do elemento Contribuidor (15) indica dados do contribuidor da tese ou dissertação eletrônica. Trata-se de um elemento cujo preenchimento é obrigatório, que possui o atributo Papel, cuja finalidade é informar o tipo de contribuição e forma de participação na geração/aprovação da tese ou dissertação. A tabela de papéis é composta por três (03) destes, que são: Orientador, Co-Orientador e Membro da Banca. Pode-se observar a descrição do atributo Papel na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
Atributo Papel	indica o tipo de contribuição e forma de participação na geração/aprovação da dissertação	Facultativo	Orientador, Co-Orientador ou Membro da banca	451	450	0	1

Tabela 22 – Descrição do atributo Papel do elemento Contribuidor (15)

O elemento Contribuidor se subdivide em cinco (05) elementos, Nome (15.1), Citação (15.2), Lattes (15.3), CPF (15.4) e Afiliação (15.5). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
15.1 Nome	indica o nome da pessoa física do contribuidor	Obrigatório	Não adotado	451	450	0	1
15.2 Citação	indica a forma pela qual a pessoa física do contribuidor será citada	Facultativo	Não adotado	451	449	0	2

15.3 Lattes	indica o endereço eletrônico do currículo do contribuidor na Plataforma Lattes	Facultativo	URI	451	97	22	332
15.4 CPF	indica o CPF do contribuidor da dissertação	Facultativo	MF 11 dígitos	451	20	431	0

Tabela 23 – Descrição dos elementos Nome (15.1), Citação (15.2), Lattes (15.3) e CPF (15.4) do elemento Contribuidor (15)

Convém pontuar que, mesmo não sendo adotado padrão para o preenchimento do elemento Nome (15.1), o mesmo foi informado corretamente para todas 450 dissertações, deixando de ser inscrito para uma destas. A consequência da falta desta informação é a exclusão da respectiva dissertação do processo de *harvesting*, uma vez que o preenchimento do mesmo é obrigatório. Sendo assim, a mesma deixa de fazer parte da BDTD e da NDLTD.

Sobre o elemento Citação (15.2), observou-se que o mesmo foi informado corretamente para 449 dissertações defendidas entre 1997 e 2007, deixando de ser informado para duas destas. Por se de preenchimento facultativo, a respectiva dissertação não foi excluída do processo de *harvesting*. Entretanto, uma busca realizada pelo elemento Citação (15.2) não recuperaria a respectiva dissertação na base de dados da BDTD nem da NDLTD.

Ao examinar o preenchimento do elemento Lattes (15.3), observou-se que o mesmo foi registrado 97 vezes corretamente e 22 vezes incorretamente. Nos casos incorretos, embora o padrão adotado tenha sido respeitado, o endereço eletrônico informado direciona para um local que não é o currículo Lattes do contribuidor. Além disso, este elemento deixou de ser descrito 332 vezes. É importante ressaltar que o currículo Lattes possui links para acesso a outras bases como a base do CNPQ, por exemplo. E, a falta desta informação, prejudica o acesso a estes links pelo currículo Lattes.

Quanto ao preenchimento do elemento CPF (15.4), observou-se que o mesmo foi informado 20 vezes corretamente e 431 vezes incorretamente, tendo sido o padrão

totalmente desconsiderado. Nestes últimos foram observados os mais diversos formatos para este elemento, ou seja, o mesmo foi informado com 09, 11, 13 e 14 dígitos; com pontos, com hífen e com seqüência de zeros. A informação incorreta deste elemento não prejudica a busca na BDTD nem da NDLTD, uma vez que ambas não disponibilizam busca pelo CNPJ do contribuidor da dissertação.

15.5 Afiliacao: o conteúdo do elemento Afiliacao (15.5) indica dados referentes a Instituição à qual o contribuidor da tese ou dissertação eletrônica é afiliado. Trata-se de um elemento cujo preenchimento é facultativo e que se subdivide em seis (06) elementos, Nome (15.5.1), Sigla (15.5.2), Pais (15.5.3), UF (15.5.4), CNPJ (15.5.5) e URL (15.5.6). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchi- mento	Padrão	Total de elementos	Corretos	Incorretos	Não infor- mados
15.5.1 Nome	indica o nome da Instituição à qual o contribuidor da dissertação é afiliado	Obrigatório	Não adotado	451	279	0	172
15.5.2 Sigla	indica a sigla da Instituição à qual o contribuidor da dissertação é afiliado	Facultativo	Não adotado	451	279	0	172
15.5.3 Pais	indica o país da Instituição à qual o contribuidor da dissertação é afiliado	Facultativo	Norma ISO 3166	451	279	0	172
15.5.4 UF	indica a Unidade Federativa da Instituição à qual o contribuidor da disserta-	Facultativo	Unidade Da Federação	451	279	0	172

	ção é afiliado						
15.5.5 CNPJ	indica o CNPJ da Instituição à qual o contribuidor da dissertação é afiliado	Facultativo	MF 14 dígitos	451	0	279	172
15.5.6 URL	indica o endereço eletrônico da Instituição à qual o contribuidor da dissertação é afiliado	Facultativo	URI	451	279	0	172

Tabela 24 – Descrição dos elementos Nome (15.5.1), Sigla (15.5.2), País (15.5.3), UF (15.5.4), CNPJ (15.5.5) e URL (15.5.6) do elemento Afiliacao (15.5)

Ao examinar o preenchimento dos elementos Nome (15.5.1), Sigla (15.5.2), País (15.5.3), UF (15.5.4) e URL (15.5.6) observou-se que os mesmos foram informado 279 vezes corretamente e deixou de ser informado 172 vezes. A falta de informação destes elementos tem como consequência a impossibilidade de recuperação das respectivas dissertações, por meio da busca pelo nome, sigla e URL da instituição à qual o contribuidor da dissertação eletrônica é afiliado, na base de dados da BDTD e da NDLTD. A recuperação das dissertações por meio da busca pelo país ou pela UF da instituição à qual o autor da dissertação eletrônica é afiliado também não é possível, quando estes elementos deixam de ser informados. Mas, se os elementos país e UF da biblioteca digital responsável pela publicação digital da dissertação, do local de defesa da dissertação, da instituição de defesa da dissertação ou da agência de fomento que apoiou financeiramente o contribuidor tiverem sido informados corretamente, a busca será feita por meio destes, uma vez que é por meio da busca simples que o país e a UF são informados para a busca na base de dados da BDTD e da NDLTD. Sendo assim, as respectivas dissertações serão recuperadas em ambas as bases.

Sobre o preenchimento do elemento CNPJ (15.5.5), observou-se que o mesmo foi registrado 241 vezes incorretamente e deixou de ser informado 210 vezes. A informação incorreta ou a falta de preenchimento deste elemento não prejudica a busca

na BDTD nem da NDLTD, uma vez que nenhuma das duas disponibiliza busca pelo CNPJ da Instituição à qual contribuidor da dissertação é afiliado.

16. InstituicaoDefesa: o conteúdo do elemento InstituicaoDefesa (16) indica dados referentes a Instituição onde a tese ou dissertação eletrônica foi defendida. Trata-se de um elemento cujo preenchimento é obrigatório e que se subdivide em seis (06) elementos, Nome (16.1), Sigla (16.2), Pais (16.3), UF (16.4), CNPJ (16.5), URL (16.6) e Programa (16.7). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incor-retos	Não informados
16.1 Nome	indica o nome da Instituição onde a dissertação foi defendida	Obrigatório	Não adotado	451	451	0	0
16.2 Sigla	indica a sigla da Instituição onde a dissertação foi defendida	Facultativo	Não adotado	451	451	0	0
16.3 Pais	indica o país da Instituição onde a dissertação foi defendida	Facultativo	Norma ISO 3166	451	451	0	0
16.4 UF	indica a Unidade Federativa da Instituição onde a dissertação foi defendida	Facultativo	Unidade Da Federação	451	451	0	0
16.5 CNPJ	indica o CNPJ da Instituição onde a dissertação foi defendida	Facultativo	MF 14 dígitos	451	0	451	0
16.6	indica o ende-	Facultativo	URI	451	451	0	0

URL	reço eletrônico da Instituição onde a dissertação foi defendida						
------------	---	--	--	--	--	--	--

Tabela 25 – Descrição dos elementos Nome (16.1), Sigla (16.2), País (16.3), UF (16.4), CNPJ (16.5) e URL (16.6) do elemento InstituicaoDefesa (16)

Ao examinarmos o preenchimento do elemento CNPJ (16.5), observamos que o mesmo foi informado incorretamente para todas as dissertações. Convém esclarecer que a instituição de defesa é a mesma para todas as dissertações e, por este motivo, os dados da mesma foram informados em uma tabela, juntamente com seu respectivo código. Sendo assim, os elementos foram informados somente uma vez e, são buscados por meio de seu respectivo código, para cada uma das dissertações. Como o elemento CNPJ (16.5) foi registrado incorretamente, conseqüentemente, toda vez que o código da Instituição de defesa é inscrito para uma dissertação, o CNPJ buscado é o que se encontra na tabela, ou seja, o incorreto. Por ser um campo de preenchimento facultativo, o preenchimento incorreto do mesmo não implica em exclusão das respectivas dissertações do processo de *harvesting*. Também não terá conseqüências na busca das dissertações, uma vez que não é disponibilizado busca por meio deste elemento na BDTD e nem na NDLTD.

16.7 Programa: o conteúdo do elemento Programa (16.7) indica dados referentes ao programa de pós-graduação onde a tese ou dissertação eletrônica foi defendida. Trata-se de um elemento cujo preenchimento é facultativo e que se subdivide em dois (02) elementos, Nome (16.7.1) e Area (16.7.2). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
16.7.1 Nome	indica o nome do programa de pós-graduação onde a dissertação foi defendida	Obrigatório	Não adotado	451	451	0	0

16.7.2 Area	indica a área de concentração do programa de pós-graduação onde a dissertação foi defendida	Facultativo	Não adotado	451	0	0	451
------------------------	---	-------------	-------------	------------	----------	----------	------------

Tabela 26 – Descrição dos elementos Nome (16.7.1) e Area (16.7.2) do elemento Programa (16.7)

Ao examinar o preenchimento do elemento Área (16.7.2), observou-se que o mesmo deixou de ser informado para todas as dissertações defendidas entre 1997 e 2007. A falta de informação deste elemento implicará na impossibilidade de recuperação das respectivas dissertações, por meio da busca por área, na base de dados da BDTD e da NDLTD, uma vez que ambas disponibilizam a busca por área.

17. AgenciaFomento: o conteúdo do elemento AgenciaFomento (17) indica dados referentes a agência de fomento que apoiou financeiramente o autor da tese ou dissertação eletrônica. Trata-se de um elemento cujo preenchimento é facultativo e que se subdivide em seis (06) elementos, Nome (17.1), Sigla (17.2), Pais (17.3), UF (17.4), CNPJ (17.5) e URL (17.6). Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
17.1 Nome	indica o nome da agência de fomento que apoiou financeiramente o autor da dissertação	Obrigatório	Não adotado	451	451	0	0
17.2 Sigla	indica a sigla da agência de fomento que apoiou financeiramente o autor da dis-	Facultativo	Não adotado	451	451	0	0

	sertação						
17.3 Pais	indica o país da agência de fomento que apoiou financeiramente o autor da dissertação	Facultativo	Norma ISO 3166	451	451	0	0
17.4 UF	indica a Unidade Federativa da agência de fomento que apoiou financeiramente o autor da dissertação	Facultativo	Unidade Da Federação	451	451	0	0
17.5 CNPJ	indica o CNPJ da agência de fomento que apoiou financeiramente o autor da dissertação	Facultativo	MF 14 dígitos	451	0	451	0
17.6 URL	indica o endereço eletrônico da agência de fomento que apoiou financeiramente o autor da dissertação	Facultativo	URI	451	451	0	0

Tabela 27 – Descrição dos elementos Nome (17.1), Sigla (17.2), Pais (17.3), UF (17.4), CNPJ (17.5) e URL (17.6) do elemento AgenciaFomento (17)

Ao examinarmos o preenchimento do elemento CNPJ (17.5), observamos que o mesmo foi informado incorretamente para todas as dissertações. Deve-se esclarecer que os metadados referentes a agência de fomento foram informados em uma tabela, juntamente com seu respectivo código. Desta forma, os elementos foram informados somente uma vez e são buscados por meio de seu respectivo código, para cada uma das

dissertações. Como o elemento CNPJ (17.5) foi salvo incorretamente nesta tabela, conseqüentemente, toda vez que o código da agência de fomento é inscrito para uma dissertação, o CNPJ buscado é o que se encontra na tabela, ou seja, o incorreto. Por ser um campo de preenchimento facultativo, o preenchimento incorreto do mesmo não implica em exclusão das respectivas dissertações do processo de *harvesting*. Também não terá conseqüências na busca das respectivas dissertações, uma vez que não é disponibilizado busca por meio deste elemento na BDTD e nem na NDLTD.

18. Direitos: o conteúdo do elemento Direitos (18) indica as condições de distribuição, reprodução e utilização da tese ou dissertação. Trata-se de um elemento cujo preenchimento é facultativo e que possui o atributo Idioma, cuja finalidade é indicar o Idioma no qual os direitos foram informados. Pode-se observar a descrição do atributo Papel na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
18 Direitos	indica as condições de distribuição, reprodução e utilização da dissertação	Facultativo	Não adotado	451	450	0	1
Atributo Idioma	indica o idioma no qual os direitos foram informados	Facultativo	Norma ISO 639	451	450	0	1

Tabela 28 – Descrição do elemento Direitos (18) e de seu atributo Idioma

Ao examinar o preenchimento do elemento Direitos (18) e de seu atributo Idioma, observou-se que os mesmos foram registrados para 450 dissertações defendidas entre 1997 e 2007, deixando de ser informado para uma destas. A falta de informação do elemento Direitos (18) implicará na impossibilidade de recuperação da respectiva dissertação, por meio da busca por este elemento, na base de dados da BDTD e da NDLTD.

19. Extensao: o elemento Extensao (19) é reservado para a criação de sub-itens (subcampos) para uso específico. Trata-se de um elemento cujo preenchimento é facultativo e que possui o atributo Namespace. Pode-se observar a descrição dos mesmos na tabela abaixo:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
19 Extensao	reservado para a criação de sub-itens (subcampos) para uso específico	Facultativo	Não adotado	451	0	0	451
Atributo Namespace	indica o idioma no qual a extensão foi informada	Facultativo	Norma ISO 639	451	0	0	451

Tabela 29 – Descrição do elemento Extensão (19) e de seu atributo Namespace

Ao examinar o preenchimento do elemento 19 Extensao e de seu atributo Namespace, observou-se que os mesmos deixaram de ser informados para todas as dissertações defendidas entre 1997 e 2007. Como ambos são de preenchimento facultativo e se destinam a criação de sub-itens para uso específico da Instituição de defesa, a falta de informação dos mesmos não resultará no processo de exclusão das respectivas dissertações do processo de *harvesting*. Da mesma forma, não irá interferir na recuperação das mesmas na base de dados da BDTD e da NDLTD, uma vez que a busca por meio deste elemento não é disponibilizada.

Na tabela abaixo, pode-se observar todos os elementos de preenchimento obrigatório que deixaram de ser informados, excluindo as respectivas dissertações do processo de *harvesting*:

Elemento	Descrição	Preenchimento	Padrão	Total de elementos	Corretos	Incorretos	Não informados
1.2 DataAtualizacao	indica a data em que foi realizada a última atualização	Obrigatório	Protocolo OAI	451	450	0	1

	do registro da dissertação						
4 Título	indica o título da dissertação	Obrigatório	Não adotado	615	613	1	1
5.1 URL	indica o endereço eletrônico do arquivo ou dos metadados da dissertação na biblioteca de origem	Obrigatório	URI	1042	1041	0	1
9 Resumo	indica o resumo da dissertação	Obrigatório	Não adotado	832	830	0	2
13 DataDefesa	indica a data da defesa da dissertação eletrônica	Obrigatório	Norma ISO 8601	451	449	0	2
15.1 Nome (Contribuidor)	indica o nome da pessoa física do contribuidor	Obrigatório	Não adotado	451	450	0	1

Tabela 30 – Elementos de preenchimento obrigatório que deixaram de ser informados

Nas figuras abaixo, é possível observar que na base de dados da BDTD, o total de dissertações defendidas na UDESC, no período de 1997 a 2007, é 443:

Biblioteca Digital Brasileira de Teses e Dissertações - BDTD - Pesquisa Avançada - Mozilla Firefox

Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://bdtd.ibict.br/?option=com_wrapper&Itemid=39

Mais visitados Guia rápido Últimas notícias

Data de Defesa: 2005-12-09

Autor: Rogéria Rebelo Diegoli
[en] Solid instruction and sharpened education: the disciplinary strategies in the Santo Antônio - Blumenau - SC

REQUISITOS MÍNIMOS ELABORAÇÃO COPYLEFT

1 Selecionar (1932-1942)

Instituição de Defesa: Universidade do Estado de Santa Catarina
Data de Defesa: 2006-12-01

1 a 5 de 443 1|2|3|4|5|6|7|8|9|10|11|12|13|14|15|16|17|18|19|20|>>>>

Argumentos da Pesquisa

Termo de Busca: udesc

Critério de Recuperação: Todas as palavras

Recuperar no Campo: Instituição de Defesa

Filtros da Pesquisa: País: Todos, Grau: Todos, Idioma: Todos, Ano de Defesa: Entre 1997 e 2007

Nº de registros por tela: 5

Reiniciar Enviar

[Voltar]

topo

Concluído

Figura 26 – busca das dissertações da UDESC na BDTD, no período de 1997 a 2007

Biblioteca Digital Brasileira de Teses e Dissertações - BDTD - Pesquisa Avançada - Mozilla Firefox

Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://bdtd.ibict.br/?option=com_wrapper&Itemid=39

Mais visitados Guia rápido Últimas notícias

Envia

INICIO A BDTD SERVIÇOS INDICADORES NOTÍCIAS

Início Serviços Pesquisa Avançada

Pesquisa Avançada

Resultado da Pesquisa

1 a 5 de 443 1|2|3|4|5|6|7|8|9|10|11|12|13|14|15|16|17|18|19|20|>>>>

Selecionar todos os itens desta página

Selecionar

Autor: Jeannette Sofia Quidel Espinoza
[pt] Compósitos de matriz de alumínio reforçados com aluminetos de níquel: fabricação e caracterização mecânica e estrutural

Instituição de Defesa: Universidade do Estado de Santa Catarina
Data de Defesa: 2002-11-29

Selecionar

Autor: Diane Deuschle
[en] Iteration program to international customers as a strategy of relationship marketing: case Ciser
[pt] Programa de integração de clientes internacionais como estratégia de marketing de relacionamento : o caso ciser

Instituição de Defesa: Universidade do Estado de Santa Catarina
Data de Defesa: 2007-06-19

Selecionar

Autor: Roseli Maria Battistella
[pt] O jovem Brecht e Karl Valentin : a cena cômica na república de Weimar

Instituição de Defesa: Universidade do Estado de Santa Catarina
Data de Defesa: 2007-05-05

Selecionar

Autor: Mônica Teresinha Marçal
[en] The insertion of the women in the Catarinense College, 1967-73
[pt] A inserção das mulheres no Colégio Catarinense, 1967 - 73

Instituição de Defesa: Universidade do Estado de Santa Catarina
Data de Defesa: 2005-12-09

Selecionar

Autor: Rogéria Rebelo Diegoli
[en] Solid instruction and sharpened education: the disciplinary strategies in the Santo Antônio - Blumenau - SC

REQUISITOS MÍNIMOS ELABORAÇÃO COPYLEFT

1 Selecionar (1932-1942)

Concluído

Figura 27 – total de dissertações da UDESC na BDTD, no período de 1997 a 2007

Das 451 dissertações defendidas na UDESC no período de 1997 a 2007, 443 tiveram seus metadados coletados pela BDTD e, conseqüentemente, pela NDLTD. Mas, 8 delas foram excluídas do processo de *harvesting* por falta de informação de elementos obrigatórios.

A análise dos metadados informados pela UDESC para as dissertações defendidas entre 1997 e 2007, nos permitiu observar que:

- 1) os atributos dos elementos, cujo preenchimento é facultativo, muitas vezes, deixaram de ser informados ou foram informados incorretamente. Como por exemplo, o atributo Idioma, do elemento URL, que deixou de ser informado 1042 vezes, o atributo Norma do elemento Citacao, que deixou de ser informado para todas as dissertações e o atributo Idioma do elemento Legenda, que foi informado incorretamente para todos os arquivos;
- 2) os elementos cujo preenchimento é facultativo, muitas vezes, deixaram de ser informados. Como por exemplo, o elemento Cobertura, que deixou de ser informado para todas as dissertações, o elemento Lattes (do autor), que deixou de ser informado para 324 dissertações, o elemento Lattes (do contribuidor), que deixou de ser informado para 332 dissertações, o elemento Afilicao (do autor), que deixou de ser informado para 210 dissertações e o elemento Afilicao (do contribuidor), que deixou de ser informado para 172 dissertações;
- 3) o maior número de preenchimentos incorretos foram dos elementos Lattes, CPF e CNPJ;
- 4) o preenchimento dos elementos que devem ser obrigatoriamente informados, na grande maioria das vezes, foi feito corretamente. Entretanto, alguns destes deixaram de ser informados, implicando na exclusão das respectivas dissertações do processo de *harvesting*. Desta forma, as mesmas deixaram de fazer parte da base de dados da BDTD e da NDLTD.

Sendo assim, com o objetivo de melhorar a qualidade dos metadados informados pela UDESC para a BDTD, são feitas algumas sugestões a seguir.

Na figura abaixo são apresentados os metadados referentes ao autor da tese ou dissertação. Pode-se verificar que o campo CPF não apresenta um texto explicativo sobre como o mesmo deve ser informado, um exemplo de preenchimento ou alguma sugestão para auxiliar na utilização do formato especificado pelo padrão MTD-BR, como acontece, na mesma tela, com o campo e-mail.

Figura 28 – Tela onde são informados os metadados referentes ao autor da tese/dissertação

Para auxiliar na dificuldade relativa ao preenchimento do metadado CPF, sugerimos que seja incluso um texto explicativo sobre como informar o CPF, juntamente com um exemplo. Ou ainda, automatizar a consistência do preenchimento deste metadado, a fim de impedir que o mesmo seja registrado de maneira incorreta. É importante destacar que estas sugestões se aplicam a todas as telas em que são informados os CPFs e não somente na tela da figura acima.

Na figura abaixo são apresentados os metadados referentes ao contribuidor da tese ou dissertação. Pode-se perceber que o campo onde é informado o Lattes, mesmo sendo o seu título destacado em outra cor, não informa que o endereço pode ser buscado ao clicar sobre a palavra Lattes, não apresenta um texto explicativo sobre como o mesmo deve ser informado, um exemplo de preenchimento ou alguma sugestão para auxiliar no preenchimento correto.

Figura 29 – Tela onde são informados os metadados do contribuidor da tese/dissertação

Para auxiliar o preenchimento do metadado Lattes, sugerimos que seja incluso um texto explicativo sobre como informá-lo, juntamente com um exemplo, bem como uma explicação da importância de seu preenchimento e como proceder para localizar o endereço correto, que fica dentro do currículo e não na barra de endereços do site da plataforma Lattes.

Na figura abaixo, são informados os metadados referentes a agência de fomento. Pode-se observar que o campo onde é informado o CNPJ não apresenta um texto explicativo sobre como o mesmo deve ser registrado, um exemplo de preenchimento ou alguma sugestão para auxiliar na utilização do formato especificado pelo padrão MTD-BR.

Sistema de Publicação Eletrônica de Teses e Dissertações | Pós-Graduação | Edição de Cadastro de Alunos - Mozilla Firefox

Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://www.tede.udesc.br/tde_biblioteca/insrir_agfomento.php

Mais visitados Guia rápido Últimas notícias

Biblioteca Universitária da UDESC | Sistema de Publicação Eletrônica... | Plataforma Lattes

SISTEMA DE PUBLICAÇÃO ELETRÔNICA DE TESES E DISSERTAÇÕES | BIBLIOTECA | INÍCIO | BUSCA | BIBLIOTECA | ADMINISTRAÇÃO | CONTATO

Agência Fomento

Nome

Sigla

País

UF

CNPJ

URL

Inserir >>

v52.04 | COPYRIGHT © desenvolvido para IE 5.0+

logo 1 logo 2 BOTD ibict

http://www.tede.udesc.br/tde_biblioteca/insrir_agfomento.php

Figura 30 - Tela onde são informados os metadados da agência de fomento

Para auxiliar na dificuldade relativa ao preenchimento do metadado CNPJ, sugerimos que seja incluso um texto explicativo sobre como informá-lo, juntamente com um exemplo. Ou ainda, automatizar a consistência do preenchimento deste metadado, a fim de impedir que o mesmo seja registrado de maneira incorreta. É importante ressaltar que estas sugestões se aplicam a todas as telas em que são descritos os CNPJs e não somente na tela da figura acima.

Na figura abaixo, são apresentados os metadados referentes a localização da defesa da tese ou dissertação. Pode-se observar que o campo onde é informada a cidade não apresenta uma caixa de seleção, como acontece com os campos País e UF.

Sistema de Publicação Eletrônica de Teses e Dissertações | Pós-Graduação | Criação de TDE - Mozilla Firefox

Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://www.tede.udesc.br/tde_biblioteca/iniciar_tde.php

Mais visitados Guia rápido Últimas notícias

TEDE

SISTEMA DE PUBLICAÇÃO ELETRÔNICA DE TESES E DISSERTAÇÕES | BIBLIOTECA |

INÍCIO | BUSCA | BIBLIOTECA | ADMINISTRAÇÃO | CONTATO |

Iniciar TDE

Entre com as 3 primeiras letras do Nome do Autor

jaq

Nome

CPF 65003985053

Seu email pode ser disponibilizado na página de rosto? Sim Não

Número de chamada

Grau

Titulação

País

UF

Cidade

Data da defesa - - - 2009 (AAAA)

Idioma

Concluído

Figura 31 - Tela onde são informados os metadados da tese ou dissertação

Para auxiliar o preenchimento do metadado cidade de defesa, sugerimos que seja inclusa uma caixa de seleção com as cidades, a fim de evitar que sejam registradas cidades com erro de grafia e também para que as mesmas sigam um padrão. É importante apontar que esta sugestão se aplica a todas as telas em que são informadas as cidades e não somente na tela da figura acima.

Na figura abaixo, são mostrados a data de defesa, o idioma, as palavras-chave e o resumo da tese ou dissertação. Pode-se observar que o campo onde é informado o ano de defesa não apresenta uma caixa de seleção, como acontece com os campos Dia e Mês.

Sistema de Publicação Eletrônica de Teses e Dissertações | Pós-Graduação | Criação de TDE - Mozilla Firefox

Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://www.tede.udesc.br/tde_biblioteca/iniciar_tde.php

Mais visitados Guia rápido Últimas notícias

Data da defesa - - 2009 (AAAA)

Idioma Portugues

Título

Palavras-chave

Digite as palavras separadas por ponto-e-vírgula(;).
Exemplo: resina composta; sistemas lineares

Não utilize vírgula ou ponto final

Resumo

Concluído

Figura 32 - Tela onde são informados os metadados da tese ou dissertação

Observou-se também que não é feita uma verificação a fim de impedir que seja cadastrado um ano maior que o atual. Conforme pode-se observar na figura 21, o ano cadastrado foi 2204.

Autor	
Nome	Jaqueline (teste)
Citação	
Lattes	
CPF	65003985053
E-mail	jaquealves@ig.com.br
	Seu email pode ser disponibilizado na página de rosto? Não

Dados da TDE	
Grau	Mestre
Titulação	Mestre em Administração
Programa	Mestrado em Administração - Acadêmico
Área de Concentração	Administração

Contribuidor	
Contribuidor	Nenhum arquivo anexado.

Local de defesa	
Cidade	Florianópolis
UF	SC
País	Brasil
Data da defesa	01/12/2204
Idioma	Portugues

Figura 33 - Tela com os metadados da tese ou dissertação que foram informados

Para auxiliar o preenchimento do metadado ano de defesa, sugerimos que seja inclusa uma caixa de seleção com, pelo menos, dez anos anteriores ao atual, juntamente com o mesmo. Ou ainda, automatizar a verificação do ano de defesa, a fim de evitar cadastros com datas incorretas, o que pode interferir no resultado das buscas por período.

Além das sugestões acima, acredita-se ser conveniente conscientizar os responsáveis pelo preenchimento dos metadados, sobre a importância de informá-los corretamente, mesmo quando o preenchimento é facultativo.

5 Conclusão e Recomendações para trabalhos futuros

A seguir serão apresentadas algumas informações relevantes e considerações finais do trabalho de acordo com os objetivos propostos pela pesquisa, bem como sugestões de pesquisas posteriores que possam vir a desenvolver estudos relativos ao tema abordado.

5.1 Conclusão

A pesquisa bibliográfica sobre metadados realizada neste trabalho possibilitou a formação de um referencial teórico sobre o assunto, que contribui para evidenciar a importância do uso de metadados nos SRIs para que a busca e acesso tornem-se mais eficientes. Na realidade, o uso de metadados sociabiliza a informação uma vez que adota um padrão único comum a uma comunidade de usuários. Conseqüentemente, favoriza a circulação da informação. No caso da BDTD, aumenta a visibilidade da produção acadêmica em níveis nacional e internacional, otimiza o uso dos dados, maximiza a troca de informações e seu compartilhamento e, finalmente, influencia os fatores de impacto, favorecendo desta forma o progresso científico.

Esta pesquisa procurou propor recomendações para a melhoria do conteúdo dos metadados informados pela UDESC para a BDTD, no sentido de apresentar uma reflexão mais aprofundada sobre como os metadados estão sendo informados. A proposição das melhorias partiu da análise de todos os metadados informados pela UDESC e disponibilizados para a BDTD, de acordo com o padrão de metadados MTD-BR.

Para alcançar o objetivo proposto por este trabalho foi necessário estabelecer alguns objetivos específicos que direcionassem as proposições necessárias para a melhoria do conteúdo dos metadados informados pela UDESC para a BDTD.

O primeiro destes objetivos, levantamento dos metadados descritivos alimentados pela UDESC e disponibilizados para a BDTD, foi atingido por meio da análise da base de dados que armazena os metadados das dissertações eletrônicas da UDESC. Essa análise possibilitou que fossem filtrados todos os metadados das dissertações defendidas entre 1997 e 2007.

Por meio do segundo objetivo deste trabalho, procurou-se organizar esses metadados a fim de identificar quais se destinam à descrição física e quais se destinam à

descrição de conteúdo do documento. Dos 67 elementos analisados, foram identificados 03 que se destinam à descrição de conteúdo do documento, 61 que se destinam à descrição física do documento e 03 que não foram informados.

O terceiro objetivo, exame do conteúdo desses metadados com base nas regras estabelecidas pelos modelos adotados para os elementos do padrão MTD-BR de metadados e para seus respectivos atributos, nos permitiu verificar que, na grande maioria das vezes, os metadados são informados de acordo com o padrão MTD-BR. Entretanto, observou-se que os elementos de preenchimento facultativo, assim como os atributos dos elementos, deixaram de ser informados muitas vezes. Também observou-se que os atributos dos elementos foram informados incorretamente para algumas dissertações. O preenchimento incorreto dos atributos e dos elementos de preenchimento facultativo, bem como a falta de informação dos mesmos, não ocasiona a exclusão dos respectivos documentos do processo de *harvesting*. Mas, dependendo da busca realizada, pode tornar impossível a recuperação dos mesmos na BDTD e na NDLTD. Pode-se citar a busca pelo número de chamada da dissertação na BDTD como exemplo de recuperação impossível quando o elemento Numerochamada deixa de ser informado ou é informado incorretamente. Observou-se que o maior número de preenchimento incorreto dos elementos ocorreu para os elementos Lattes, CPF e CNPJ, o que resultou na impossibilidade de relacionamentos com outros sistemas de recuperação por meio destes elementos. O preenchimento dos elementos que devem ser obrigatoriamente informados, na grande maioria das vezes, foi feito corretamente. Embora, poucas vezes, alguns destes não tenham sido informados, resultando na exclusão das respectivas dissertações do processo de *harvesting*.

5.2 Recomendações para trabalhos futuros

As limitações desta pesquisa, aliadas a complexidade do assunto estudado, abrem precedentes para trabalhos futuros. Com o seu desenvolvimento, serão levantadas novas questões que poderão estimular o surgimento de novas formas de observar, investigar e entender; iniciando novos estudos, a partir dos resultados obtidos neste trabalho.

Uma sugestão de trabalho futuro seria verificar se todos os elementos e atributos do padrão MTD-BR são realmente de interesse ou se alguns destes poderiam ser excluídos. Pois, embora a comparação entre os padrões MTD-BR, ETD-MS e Dublin

Core, feita nesta pesquisa, tenha nos permitido concluir que o primeiro é o que melhor agrupa os elementos necessários para que a BDTD possa desempenhar suas tarefas e atingir seus objetivos, observou-se que o MTD-BR, além de possuir um número bem maior de elementos do que os outros dois, possui elementos que nos impulsionam a refletir sobre sua real necessidade. Nessa linha, citamos o caso do atributo Idioma do elemento URL (5.1), cuja finalidade é indicar o idioma do endereço eletrônico do arquivo ou dos metadados da dissertação na biblioteca de origem. Seria realmente necessário manter um atributo para indicar o idioma de um endereço eletrônico? Convém ressaltar que o número elevado de elementos do padrão MTD-BR pode dificultar ou, até mesmo, tornar inviável a sua adoção pelas IES. Portanto, acredita-se que seria uma grande contribuição o desenvolvimento de estudos que refletissem sobre a possibilidade de redução do número de elementos do padrão MTD-BR, sem causar prejuízos no agrupamento de elementos necessários para que a BDTD possa desempenhar suas tarefas e atingir seus objetivos. O resultado desse estudo viria certamente contribuir para uma maior e melhor aceitação do padrão pelas IES.

Sabe-se que as organizações podem utilizar os metadados para a) fazer o mapeamento de suas informações, b) auxiliar na criação de significado para a informação da qual dependem para atingir suas metas c) apoiar a tomada de decisões e d) facilitar a localização da informação. Diante desta constatação, acredita-se que seria de grande valia tratar da importância da alimentação correta dos metadados, o que traria benefícios para estudos bibliométricos e cientométricos que produzem índices importantes para a tomada de decisão por Agências de Fomento, além de aprimorar o Sistema Nacional de Ciência, Tecnologia e Inovação (SNCT&I), suprimindo suas necessidades de informação de forma mais eficiente.

Espera-se que este trabalho possa cooperar com a área da CI como uma referência conceitual, instigando novas reflexões, discussões e análises sobre um tema de elevada importância para a descrição dos recursos disponibilizados em meio eletrônico, que são os metadados.

REFERÊNCIAS

ALVES, M. D. R.; SOUZA, I. F. Estudo de Correspondência de Elementos Metadados: *Dublin Core* e MARC21. Revista Digital de Biblioteconomia e Ciência da Informação. ISSN: 1678-765X Vol. 4, n. 2, p. 20-38, 1º sem. Campinas, SP 2007. Disponível em: <<http://server01.bc.unicamp.br/seer/ojs/include/getdoc.php?id=389&article=94&mode=pdf>> Acesso em: maio de 2008.

ALVES, R. C. V. Web semântica: uma análise focada no uso de metadados. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.

ARAÚJO, V.M.R. Sistemas de informação: nova abordagem teórico-conceitual. Ciência da Informação. Brasília, v.24,n.1,p.54-76, jan/abr.1995.

BAEZA-YATES, R.; RIBEIRO-NETO, B. Modern Information Retrieval. New York: Addison Wesley, 1999.

BALBY, C. N. Estudos de uso de catálogos on-line (OPACs): revisão metodológica e aplicação da técnica de análise de *log* de transações a um OPC de biblioteca universitária brasileira. 2002. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes da Universidade de São Paulo, USP, São Paulo.

BAPTISTA, A. A.; COSTA, S. M.; KURAMOTO, H.; RODRIGUES, E. COMUNICAÇÃO CIENTÍFICA: o papel da *Open Archives Initiative* no contexto do Acesso Livre. Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf. ISSN: 1518-2924 Vol. 13, n. esp. 1º sem. Florianópolis, SC 2008. Disponível em: http://www.encontros-bibli.ufsc.br/bibesp/esp_06/bibesp_esp_06_apresentacao_esp_20071.pdf Acesso em: julho de 2008.

BAPTISTA, Dulce Maria. O IMPACTO DOS METADADOS NA REPRESENTAÇÃO DESCRITIVA Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis, v.12, n.2, p. 177-190, jul./dez., 2007.

BARRETO, A. de A. Uma quase história da ciência da informação. DataGramZero - Revista de Ciência da Informação - v.9 n.2 abr/08. Disponível em <http://www.dgz.org.br/abr08/F_I_art.htm> Acesso em: fevereiro de 2009.

BERNERS-LEE, T. Metadata architecture: documents, metadata and

link. 1997. Disponível em: <<http://www.w3.org/DesignIssues/Metadata.html>>
Acesso em: fevereiro de 2008.

BIBLIOTECA (2008) “Biblioteca Digital de Teses e Dissertações”, Disponível em:
<http://bdtd.ibict.br>. Acesso em: maio de 2008.

BORGMAN, C. L. Social Aspects of digital libraries. In: WORKSHOP ON SOCIAL ASPECTS OF DIGITAL LIBRARIES, 1996, Los Angeles. Final Report UCLA-NSF, 1996. Disponível em <<http://www.ifla.org/documents/libraries/net/ucladl01.pdf>>
Acesso em: abril de 2008.

BRÄSCHER, M. A Ambigüidade na Recuperação da Informação”. DataGramZero - Revista de Ciência da Informação - v.3 n.1, Fev 2002. Disponível em:
http://www.dgz.org.br/fev02/Art_05.htm Acesso em: outubro de 2008.

CABRAL, A. Informática e comunicação para uma sociedade democrática: uma abordagem sobre organizações, propostas e ideologias no cenário brasileiro. Anais do 24. Congresso Brasileiro de Ciências da Comunicação, Campo Grande/MS, setembro 2001 [cd-rom]. São Paulo: Intercom, 2001. São Paulo, SP
Disponível em: <<http://hdl.handle.net/1904/4846>> Acesso em: junho de 2008.

CAFÉ, L. Aplicação do modelo de predição sintático-semântica na construção de linguagens documentárias facetadas In.: CUNHA, M. V. D. C. ; SOUZA, F. D. C. D. (Org.). Comunicação, gestão e profissão: abordagens para o estudo da Ciência da Informação. Belo Horizonte, MG: Autêntica, 2006, p. 69-92

CAFÉ, L. BRÄSCHER, M. Organização da informação e bibliometria. Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf. ISSN: 1518-2924 Vol. 13, n. esp. 1º sem. Florianópolis, SC 2008. Disponível em:
<http://www.periodicos.ufsc.br/index.php/eb/article/view/1334/1032>
Acesso em: julho de 2008.

CAMPELLO, B.S. Teses e dissertações. In: CAMPELLO, B.S. ; CENDÓN, B.V. ; KREMER, J.M. (Org.). Fontes de informação para pesquisadores e profissionais. Belo Horizonte : Ed. UFMG, 2000. Cap. 9 (Aprender)

CAMPOS, F.; BAX, M. P. Como os Mecanismos de Busca da Web Indexam Páginas HTML. Minas Gerais, 2000. Disponível em:
<<http://www.bax.com.br/research/publications/ComoMaquinasBuscaIndexamPaginasWeb.pdf>>. Acesso em: setembro de 2007.

CARDOSO, O. N. P. Recuperação de Informação. INFOCOMP Journal of Computer Science. ISSN: 1807-4545 Ano 5, n. 4, dezembro Lavras, MG. 2006 Disponível em: <www.dcc.ufla.br/infocomp/artigos/v2.1/olinda.pdf> Acesso em: fevereiro de 2008.

CASTRO, F. F.; SANTOS, P. L. V. A. C. Os Metadados como instrumentos tecnológicos na padronização e potencialização dos recursos informacionais no âmbito das bibliotecas digitais na era da Web Semântica. Inf. & Soc.: Est., João Pessoa, v.17, n.2, p.12-21, maio/ago. 2007.

CHAGAS, M. A recuperação da informação em diferentes suportes textuais In.: CUNHA, M. V. D. C. ; SOUZA, F. D. C. D. (Org.). Comunicação, gestão e profissão: abordagens para o estudo da Ciência da Informação. Belo Horizonte, MG: Autêntica, 2006, p. 47-67

CHEN, H. Semantic research for digital libraries. D-Lib Magazine, v.5, n. 10 out.1999. Disponível em : <<http://www.dlib.org/dlib/october99/chen/10chen.html>. > Acesso em: outubro de 2008.

CHOO, C. W. A organização do conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões. São Paulo: Ed. Senac São Paulo, 2003.

CINTRA, A. M. et al. Para entender as linguagens documentárias. 2 ed. São Paulo: Polis, 2002. 92 p. (Coleção Palavra-Chave, 4).

CORREIA, A. E. G. C.; SILVA, E. L. As influências das tecnologias de informação e comunicação no processo de pesquisa científica: um estudo aplicado à UFPE In.: CUNHA, M. V. D. C. ; SOUZA, F. D. C. D. (Org.). Comunicação, gestão e profissão: abordagens para o estudo da Ciência da Informação. Belo Horizonte, MG: Autêntica, 2006, p. 9-32

CUNHA, M. B. Desafios na construção de uma biblioteca digital. Ci. Inf. ISSN: 0100-1965 Vol.28, n.3 Sept./Dec Brasília, DF 1999.

DAHLBERG, I. Teoria do conceito. Ciência da Informação. Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978.

DIAS, E. W. Contexto Digital e Tratamento da Informação. DataGramZero - Revista de Ciência da Informação - v.2 n.5, Out 2001. Disponível em: http://www.dgz.org.br/out01/Art_01.htm Acesso em: agosto de 2008.

DUBLIN (2008) “*Dublin Core Metadata Initiative*”, Disponível em:

<http://dublincore.org/> . Acesso em: maio de 2008.

FERNEDA, E. Recuperação da Informação: Análise sobre a Contribuição da Ciência da Computação para a Ciência da Informação. 2003. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes da Universidade de São Paulo, USP, São Paulo. Disponível em:

< <http://www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/publico/Tese.pdf>> Acesso em: janeiro de 2008.

FOSKETT, A. C. A abordagem temática da informação; tradução de Antonio Agenor Briquet de Lemos. São Paulo, Polígono; Brasília, Ed. Universidade de Brasília, 1973.

FOX, E. A.; McMILLAN, G.; SULEMAN, H.; GONÇALVES, M. A.; LUO, M.

Networked Digital Library of Theses and Dissertations (NDLTD). Capítulo XXX

Disponível em: <<http://pubs.cs.uct.ac.za/archive/00000015/01/ndltd12.pdf>> Acesso em: maio de 2008.

GIL, A.C. Como elaborar projetos de pesquisa. São Paulo: Atlas, 1991.

IANNELLA, R.; WAUGH, A. Metadata: Enabling the Internet. 1997. Disponível em: <<http://www.ifla.org/documents/libraries/cataloging/metadata/ianr1.pdf>> . Acesso em: julho de 2008.

INSTITUTO (2008) “Instituto Brasileiro de Informação e Tecnologia – IBICT”, Disponível em <http://btdt.ibict.br>. Acesso em: abril de 2008.

JACKSON, P.; MOULINIER, I. Natural Language Processing for Online Applications: Text Retrieval, Extraction, and Categorization. John Benjamins Publishing Co, 2002.

KENT, A. Manual da recuperação mecânica da informação. São Paulo: Polígono, 1972. 427 p. ISBN (Broch.) tradução de Beatriz Berrini

KOBASHI, N. Y. A elaboração de informações documentárias, em busca de uma metodologia. 1994. Tese (doutorado em Ciências da Comunicação) - Escola de Comunicações e Artes, Universidade de São Paulo, 1994.

LANCASTER, F. W. Indexação e resumos: teoria e prática . 2. ed. Brasília, DF: Lemos Informação e Comunicação, 2004. 452 p. ISBN 8585637242 (broch.)

LE COADIC, Y-F. A ciência da informação. Brasília: Briquet de Lemos, 1996. 119p.

LIMA, G.A.B. Interfaces entre a ciência da Informação e a ciência cognitiva. *Ciência da Informação* jan/abr. 2003, 32 (1), 77-87. Disponível em:
< <http://www.scielo.br/pdf/ci/v32n1/15975.pdf> >

LIMA, G. A. B. Organização da informação para sistemas de hipertextos. In.: NAVES, Madalena Martins Lopes; KURAMOTO, Hélio. (Org.). *Organização da informação: princípios e tendências*. Brasília, DF: Briquet de Lemos/Livros, 2006, p. 99-116

LIMA, V. M. A. Da classificação do conhecimento científico aos sistemas de recuperação de informação: enunciação de codificação e enunciação de decodificação da informação documentária. 2004. 156f. Tese (Doutorado em Ciências da Comunicação). Universidade de São Paulo. Escola de Comunicações e Artes. São Paulo, 2004.

LOURENÇO, C. de A., Análise do padrão brasileiro de metadados de teses e dissertações segundo o modelo entidade-relacionamento. 2005. 164 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais, UFMG, Minas Gerais.

MARCONDES, C. H.; SAYÃO, L. F. Documentos digitais e novas formas de cooperação entre sistemas de informação em c&t. *Ci. Inf.* Vol. 31 n.3 Brasília set./dez. 2002.

MAULDIN, M. *Conceptual Information Retrieval A case study in adaptive partial parsing*. Kluwer Academic Publishers, 1991.

McGARRY, K. *O contexto dinâmico da informação: uma análise introdutória*; Tradução de Helena Vilar de Lemos. – Brasília-DF: Briquet de Lemos/Livros, 1999.

MILSTEAD, J; FELDMAN, S. Metadata: Cataloging by any other name. *Online*, v. 23, n. 1, p. 24-31, 1999.

NETWORKED (2008) “Networked Digital Library of Theses and Dissertations”, Disponível em: <<http://www.ndltd.org>> Acesso em: maio de 2008.

NOVELLINO, M. S. F. Instrumentos e metodologias de representação da informação. **Informação & Informação**, Londrina, v.1, n.2, p. 37-45, jul./dez. 1996.

OHIRA, M. L. B.; PRADO, N. S. Bibliotecas virtuais e digitais: análise de artigos de periódicos brasileiros (1995/2000). *Revista Ciência da Informação*, v. 31, n. 1, p. 61–74, Janeiro/Abril 2002.

ONLINE (2008) “Online Computer Library Center”, Disponível em:
<<http://www.oclc.org>> Acesso em: junho de 2008.

OPEN (2008) “Open Archives Initiative”, Disponível em:
<<http://www.openarchives.org>> Acesso em: junho de 2008.

OXFORD (2008) “Oxford Digital Libra”, Disponível em:
<<http://www.odl.ox.ac.uk/metadata.htm>> Acesso em: setembro de 2007.

PINHEIRO, L. V. R. A Ciência da Informação entre sombra e luz: domínio epistemológico e campo interdisciplinar. 1997. Tese. (Doutorado em Comunicação e Cultura) - Universidade Federal do Rio de Janeiro, Escola de Comunicação, Rio de Janeiro, 1997.

RASCHE, F.; VARVAKIS, G. Bibliotecas públicas e seus serviços In.: CUNHA, M. V. D. C. ; SOUZA, F. D. C. D. (Org.). *Comunicação, gestão e profissão: abordagens para o estudo da Ciência da Informação*. Belo Horizonte, MG: Autêntica, 2006, p. 127-140

ROBREDO, J. Da Ciência da informação Revisitada aos Sistemas Humanos de Informação, 2003.

ROSETTO, M. Metadados e recuperação da informação: padrões para bibliotecas digitais. In: CIBERÉTICA: SIMPÓSIO INTERNACIONAL DE PROPRIEDADE INTELECTUAL, INFORMAÇÃO e ÉTICA, 2., 2003, Florianópolis. *Anais eletrônicos...* Disponível em: <<http://www.ciberetica.org.br/trabalhos/anais/58-87-p1-87.pdf>> Acesso em: fevereiro de 2008.

ROSETTO, M. NOGUEIRA, A. H. Aplicação de elementos metadados *Dublin Core* para descrição de dados bibliográficos on-line da Biblioteca Digital de Teses da USP. São Paulo, 2002.
Disponível em: <<http://www.sibi.ufrj.br/snbu/snbu2002/oralpdf/82.a.pdf>>
Acesso em: janeiro de 2008.

SALTON, G. *Dynamic information and library processing*. Englewood Cliffs: Prentice-Hall, 1975.

SANTOS, B. de S. Um discurso sobre as ciências. 9. ed. Porto : Afrontamento, 1997. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40141988000200007> Acesso em: fevereiro de 2009.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. Perspectiva em Ciência da Informação, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SAYÃO, L. F. PADRÕES PARA BIBLIOTECAS DIGITAIS ABERTAS E INTEROPERÁVEIS. Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf., Florianópolis, n. esp., 1º sem. 2007. Disponível em: <http://www.encontros-bibli.ufsc.br/bibesp/esp_06/bibesp_esp_06_sayao_esp_20071.pdf> Acesso em: julho de 2008.

SILVA, F. M. e. Organização da Informação em sistemas eletrônicos abertos de Informação Científica & Tecnológica: Análise da Plataforma Lattes. 2007 163 f. Tese (Doutorado em Ciência da Informação) – Departamento de Biblioteconomia e Documentação, Universidade de São Paulo, São Paulo, 2007.

SOUZA, M. I. F.; SANTOS, A. D. dos.; OLIVEIRA, M. J. de.; CINTRA, M. A. M. de U.; VENDRUSCULO, L. G. Informação para internet: uso de metadados e o padrão *Dublin Core* para catalogação de recursos eletrônicos na embrapa. In Proceedings XIX Congresso Brasileiro de Biblioteconomia e Documentação 1, Centro de eventos da PUCRS. 2000. Disponível em: <<http://dici.ibict.br/archive/00000702/01/T042.pdf>> Acesso em fevereiro de 2009.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas contribuições para a ciência da informação. Ci. Inf., Brasília, v. 33, n. 1, p. 132-141, jan./abril 2004. Disponível em: <<http://www.ibict.br/cionline/include/getdoc.php?id=356&article=71&mode=pdf>> Acesso em: maio de 2007.

SVENONIUS, E. Intellectual foundation of Information Organization. Cambridge: Mit Press, 2001

TAKAHASHI, T. (Org.). Sociedade da informação no Brasil: livro verde. Brasília: Ministério da Ciência e Tecnologia, 2000.

WEIBEL, S. The *Dublin Core*: A Simple Content Description Model for Electronic Resources. 1997. Disponível em: <<http://www.asis.org/Bulletin/Oct-97/weibel.htm>> Acesso em: julho de 2007.

ANEXOS

Anexo 1 – IES COOPERANTES DA NDLTD

IES dos EUA que cooperam com a NDLTD:

- Air University (Alabama)
- Baylor University
- Boston College
- Brigham Young University
- California Institute of Technology
- Clemson University
- College of William and Mary
- Concordia University (Illinois)
- Drexel University
- Duquesne University
- East Carolina University
- East Tennessee State University
- Florida Institute of Technology
- Florida International University
- Florida State University
- Georgetown University
- George Washington University
- John Hopkins University
- Louisiana State University
- Marshal University
- Massachusetts Institute of Technology
- Miami University of Ohio
- Michigan Tech
- Mississippi State University
- University of Central Florida
- University of Colorado
- University of Florida
- University of Georgia
- University of Hawaii at Manoa
- University of Illinois, Urbana-Champaign

- University of Iowa
- University of Kentucky
- University of Maine
- University of Missouri-Columbia
- University of Nevada, Las Vegas
- University of New Orleans
- University of North Texas
- University of Oklahoma
- University of Pittsburgh
- University of Rochester
- University of South Florida
- University of Central Florida
- University of Tennessee, Knoxville
- University of Tennessee, Memphis
- University of Texas at Austin
- University of Virginia
- University of West Florida
- University of Wisconsin, Madison
- Montana State University
- Naval Postgraduate School
- New Jersey Institute of Technology
- New Mexico Tech
- North Carolina State University
- Northwestern University
- Pennsylvania State University
- Regis University
- Rochester Institute of Technology
- Texas A&M University
- Vanderbilt University
- Virginia Commonwealth University
- Virginia Tech
- Wake Forest University
- West Virginia University
- Western Kentucky University

- Western Michigan University
- Worcester Polytechnic Institute
- Yale University

IES internacionais cooperantes da NDLTD:

- Aristotle University of Thessaloniki (Greece)
- Assumption University of Thailand (Thailand)
- Australian National University (Australia)
- Biblioteca de Catalunya (Spain)
- Centro University La Salle-UNILASALLE (Brazil)
- Chinese University of Hong Kong (Hong Kong)
- Chung Yuan Christian University (Taiwan)
- Chungnam National U., Dept of CS (S. Korea)
- City University, London (UK)
- Curtin University of Technology (Australia)
- Darmstadt University of Technology (Germany)
- Freie Universitat Berlin (Germany)
- Gerhard Mercator Universitat Duisburg (Germany)
- Griffith University (Australia)
- Gyeongsang National University, Chinju (Korea)
- Helsinki University of Technology (Finland)
- Humboldt-Universität zu Berlin (Germany)
- Indian Institute of Technology, Bombay (India)
- Lund University (Sweden)
- McGill University (Canada)
- Nanyang Technological University (Singapore)
- Naresuan University (Thailand)
- National Sun Yat-Sen University (Taiwan)
- National Taiwan Normal University (Taiwan)
- Universitat de Barcelona (Spain)
- Universitat de Girona (Spain)
- Universitat de Lleida (Spain)
- Universitat Oberta de Catalunya (Spain)

- Universitat Politecnica de Catalunya (Spain)
- Universitat Politecnica de Valencia (Spain)
- Universitat Pompeu Fabra (Spain)
- Universitat Rovira i Virgili (Spain)
- Universitat Sbibliothek Munchen (Germany)
- Universitè Laval (Quèbec, Canada)
- Universitè Lyon 2 (França)
- University Catolica de Brasilia-UCB (Brazil)
- University Catolica de Pernambuco-UNICAP (Brazil)
- University Catolica de Salvador-UCSAL (Brazil)
- University Catolica de Santos-UNISANTOS (Brazil)
- University Catolica Dom Bosco-UCDB (Brazil)
- University do Vale do Rio dos Sinos-UNISINOS (Brazil)
- University of Antioquia (Medellin, Colombia)
- University of Bergen (Norway)
- University of British Columbia (Canada)
- University of Edinburgh (UK)
- University of Glasgow (UK)
- University of Guelph (Ontario, Canada)
- University of Hong Kong (Hong Kong)
- University of Hyderabad (India)
- University of Melbourne (Australia)
- University of Mysore (India)
- National University of Singapore (Singapore)
- Pontificia Universidade Católica de Minas Gerais (Brazil)
- Pontificia Universidade Católica do Rio de Janeiro (Brazil)
- Pontificia University Catolica de Campinas-PUC/CAMPINAS (Brazil)
- Pontificia University Catolica do Parana-PUC/PR (Brazil)
- Pontificia University Catolica do Rio Grande do Sul-PUC/RS (Brazil)
- Rand Afrikaans University (South Africa)
- Rhodes University (South Africa)
- Shanghai Jiao Tong University (China)
- St. Petersburg State Technical U. (Russia)
- State University of Campinas (Brazil)

- UNESP – Universidade Estadual Paulista (Brazil)
- Universidad de Chile (Chile)
- Universidad de las Américas Puebla (México)
- Universidad Politecnica De Cartagena (Spain)
- Universidade Federal Fluminense (Brazil)
- Universitat Autònoma de Barcelona (Spain)
- Universitat d'Alacant (Spain)
- University of New Brunswick (Canada)
- University of New South Wales (Australia)
- University of Novi Sad (Yugoslavia)
- University of Pisa (Italy)
- University of Pretoria (S. Africa)
- University of Queensland (Australia)
- University of São Paulo (Brazil)
- University of Sydney (Australia)
- University of Tampere (Finland)
- University of the Free State (South Africa)
- University of Utrecht (Netherlands)
- University of Waterloo (Canada)
- University of Western Ontario (Canada)
- Uppsala University (Sweden)
- Wilfrid Laurier University (Canada)
- Xiamen University Library (China)
- Yupei University of Science and Technology (Taiwan)

Outras instituições cooperantes da NDLTD:

- British Library
- Cinemedia
- Coalition for Networked Information
- Committee on Institutional Cooperation
- Comunidade Virtual de Aprendizagem Da Rede de Instituições Católicas de

Ensino Superior

- Consorci de Biblioteques Univ. Catalunya

- Diplomatica.com
- Dissertatione Online
- Dissertation.com
- ETDweb
- Ibero-American Sci. & Tech. Ed. Cons. (ISTEC)
- Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)
- MathDISS International
- National Documentation Centre (NDC, Greece)
- National Library Of Canada
- National Library of Portugal
- Office of Scientific and Technical Information (Department of Energy)
- OhioLINK
- OCLC (Online Computer Library Center)²⁹
- Organization of American States (OAS)
- SOLINET
- Sudanese National Electronic Library (Sudan)
- UNESCO

²⁹ A OCLC é uma entidade sem fins lucrativos dedicada ao público para promover um maior acesso a informações. Mais de 60000 bibliotecas em 112 países e territórios ao redor do mundo utilizam os seus serviços. (ONLINE, 2008)

Anexo 2 – IES COOPERANTES DA BDTD

- Associação Educacional Nove de Julho – UNINOVE (SP)
- Comissão Nacional de Energia Nuclear – CNEN (RJ)
- Faculdade de Medicina de São José do Rio Preto – FAMERP (SP)
- Fundação Oswaldo Cruz – FIOCRUZ (RJ)
- Instituto Nacional de Pesquisas Espaciais INPE (SP)
- Instituto Nacional de Telecomunicações – INATEL (MG)
- Instituto Tecnológico de Aeronáutica – ITA (SP)
- Instituto de Pesquisas Tecnológicas – IPT (SP)
- Pontifícia Universidade Católica de Campinas - PUC-Campinas (SP)
- Pontifícia Universidade Católica de São Paulo PUC-SP (SP)
- Pontifícia Universidade Católica do Paraná – PUCPR (PR)
- Pontifícia Universidade Católica do Rio Grande do Sul PUCRS (RS)
- Pontifícia Universidade Católica do Rio de Janeiro PUC-Rio (RJ)
- Universidade Católica Dom Bosco – UCDB (MS)
- Universidade Católica de Brasília – UCB (DF)
- Universidade Católica de Pelotas – UCEPEL (RS)
- [Universidade Católica de Pernambuco - UNICAP](#) (PE)
- Universidade Católica de Santos – UNISANTOS (SP)
- Universidade Estadual Paulista "Júlio Mesquita Filho" – UNESP (SP)
- Universidade Estadual de Campinas – UNICAMP (SP)
- [Universidade Estadual de Londrina - UEL](#) (PR)
- Universidade Estadual do Oeste do Paraná – UNIOESTE (PR)
- Universidade Federal de Viçosa – UFV (MG)
- Universidade Federal Fluminense – UFF (RJ)
- Universidade Federal da Bahia – UFBA (BA)
- Universidade Federal da Paraíba – UFPB (PB)
- [Universidade Federal de Alagoas - UFAL](#) (AL)
- Universidade Federal de Goiás – UFG (GO)
- Universidade Federal de Lavras – UFLA (MG)
- Universidade Federal de Mato Grosso do Sul – UFMS (MS)
- Universidade Federal de Minas Gerais – UFMG (MG)
- Universidade Federal de Pelotas – UFPEL (RS)

- Universidade Federal de Pernambuco – UFPE (PE)
- Universidade Federal de Rondônia – UNIR (RO)
- Universidade Federal de Santa Catarina – UFSC (SC)
- Universidade Federal de Santa Maria – UFSM (RS)
- [Universidade Federal de Sergipe - UFS](#) (SE)
- Universidade Federal de São Carlos – UFSCar (SP)
- Universidade Federal de São Paulo – UNIFESP (SP)
- Universidade Federal de Uberlândia – UFU (MG)
- [Universidade Federal do Ceará - UFC](#) (CE)
- [Universidade Federal do Espírito Santo - UFES](#) (ES)
- Universidade Federal do Pará – UFPA (PA)
- Universidade Federal do Pará – UFPA (RS)
- Universidade Federal do Rio Grande do Norte – UFRN (RN)
- [Universidade Federal do Rio Grande do Sul UFRGS](#) (RS)
- [Universidade Federal do Rio de Janeiro - UFRJ](#) (RJ)
- Universidade Metodista de São Paulo (SP)
- [Universidade Presbiteriana Mackenzie](#) (SP)
- [Universidade Regional de Blumenau - FURB](#) (SC)
- [Universidade de Brasília - UnB](#) (DF)
- Universidade de Pernambuco – UPE (PE)
- Universidade de São Paulo USP (SP)
- Universidade de Taubaté – UNITAU (SP)
- Universidade do Estado de Santa Catarina – UDESC (SC)
- Universidade do Estado do Rio de Janeiro – UERJ (RJ)
- Universidade do Vale do Itajaí – UNIVALI (CE)
- Universidade do Vale do Rio dos Sinos – UNISINOS (RS)
- Universidade do Oeste Paulista – UNOESTE (SP)

Anexo 3 – Lista de elementos do padrão MTD-BR de metadados



MTD-BR Padrão Brasileiro de Metadados de Teses e Dissertações
 Lista de Elementos do Schema (versão 2)
 14/02/2005

Nº	NomedoElemento	Atributos	Conteúdo	R / NR	O / F	Padrão
1	<Controle>		Dados de controle do registro de Teses e Dissertações	NR	O	
1.1	<Sigla>		Sigla da Instituição Consorciada	NR	O	
1.2	<DataAtualizacao>		Data em que foi realizada a última atualização do registro da tese ou dissertação	NR	O	Protocolo OAI
1.3	<IdentificacaoDocumento>		Código que identifica a tese ou dissertação na base de dados da biblioteca	NR	O	--
1.4	<Tipo>		Tipo da fonte de informação.	R	O	DCMI Type Vocabulary
2	<BibliotecaDigital>		Nome da Instituição responsável pela publicação digital da tese ou dissertação	NR	F	
2.1	<Nome>		Nome da Biblioteca Digital	NR	O	--

Nº	NomedoElemento	Atributos	Conteúdo	R / NR	O / F	Padrão
2.2	<Sigla>		Sigla da Biblioteca Digital	NR	O	--
2.3	<URL>		Endereço eletrônico do site da Biblioteca Digital	NR	O	Padrão URI
2.4	<ProvedorServico>		Instituição que abriga a Biblioteca Digital	NR	F	--
2.4.1	<Nome>		Nome da Instituição	NR	O	--
2.4.2	<Sigla>		Sigla da Instituição	NR	F	--
2.4.3	<País>		País da Instituição	NR	F	ISO 3166
2.4.4	<UF>		UF da Instituição	NR	F	Unidade da Federação
2.4.5	<CNPJ>		CNPJ da Instituição	NR	F	Padrão MF 14 dígitos
2.4.6	<URL>		Endereço eletrônico do site da Instituição	NR	F	Padrão URI
3	<BibliotecaDepositaria>		Biblioteca onde o exemplar físico da tese ou dissertação está arquivado	NR	F	
3.1	<Nome>		Nome da Biblioteca Depositária	NR	F	--
3.2	<Sigla>		Sigla da Biblioteca Depositária	NR	O	Cadastro CCN/COMUT
3.3	<URL>		Endereço eletrônico do site da Biblioteca Depositária	NR	F	Padrão URI
3.4	<NumeroChamada>		Código que identifica a tese ou dissertação no acervo da Biblioteca Depositária	NR	F	--
4	<Titulo>	Idioma	Título da tese ou dissertação	R	O	--
5	<Arquivo>		Endereço eletrônico dos arquivos da tese ou dissertação na Biblioteca Digital	R	F	
5.1	<URL>	Formato Idioma	Endereço eletrônico do arquivo da tese ou dissertação e/ou página de apresentação dos metadados na biblioteca digital de origem	NR	O	Padrão URI

Nº	NomedoElemento	Atributos	Conteúdo	R / NR	O / F	Padrão
5.2	<Legenda>	Idioma	Legenda para o arquivo, descrição de seu conteúdo	R	F	--
5.3	<NivelAcesso>		Restrição de acesso ao arquivo	NR	O	[Restrito, Publico]
6	<Idioma>		Idioma da tese ou dissertação	NR	O	ISO 639
7	<Grau>		Grau acadêmico associado à tese ou dissertação	NR	O	Tabela de Grau
8	<Titulacao>		Nome do grau acadêmico associado à tese ou dissertação	NR	O	--
9	<Resumo>	Idioma	Resumo da tese ou dissertação	R	O	--
10	<Cobertura>	Idioma	Escopo espacial ou temporal da tese ou dissertação	R	F	--
11	<Assunto>	Idioma Esquema	Tópicos tratados na tese ou dissertação e a tabela de onde estes tópicos foram extraídos, quando for o caso	R	F	--
12	<LocalDefesa>		Local de defesa da tese ou dissertação	NR	F	--
12.1	<Cidade>		Cidade onde foi defendida a tese ou dissertação	NR	O	--
12.2	<UF>		UF da cidade onde foi defendida a tese ou dissertação	NR	F	Unidade da Federação
12.3	<País>		País onde foi defendida a tese ou dissertação	NR	O	ISO 3166
13	<DataDefesa>		Data em que foi defendida a tese ou dissertação	NR	O	ISO 8601
14	<Autor>		Autor da tese ou dissertação	R	O	--
14.1	<Nome>		Nome da Pessoa Física	NR	O	--
14.2	<Citacao>	Norma	Forma pela qual a Pessoa Física será citada em sob alguma <u>norma</u>	R	F	--
14.3	<Lattes>		Endereço eletrônico do currículo da Pessoa na Plataforma Lattes	NR	F	Padrão URI

Nº	NomedoElemento	Atributos	Conteúdo	R / NR	O / F	Padrão
14.4	<CPF>		CPF da Pessoa	NR	F	Padrão MF 11 dígitos
14.5	<Afiliacao>		Instituição à qual a Pessoa é afiliada	R	F	--
14.5.1	<Nome>		Nome da Instituição	NR	O	--
14.5.2	<Sigla>		Sigla da Instituição	NR	F	--
14.5.3	<País>		País da Instituição	NR	F	ISO 3166
14.5.4	<UF>		UF da Instituição	NR	F	Unidade da Federação
14.5.5	<CNPJ>		CNPJ da Instituição	NR	F	Padrão MF 14 dígitos
14.5.6	<URL>		Endereço eletrônico do site da Instituição	NR	F	Padrão URI
15	<Contribuidor>	Papel	Contribuidor da tese ou dissertação e forma de participação (papel)	R	O	--
15.1	<Nome>		Nome da Pessoa Física	NR	O	--
15.2	<Citacao>		Forma pela qual a Pessoa deseja ser citada	NR	F	--
15.3	<Lattes>		Endereço eletrônico do currículo da Pessoa na Plataforma Lattes	NR	F	Padrão URI
15.4	<CPF>		CPF da Pessoa	NR	F	Padrão MF 11 dígitos
15.5	<Afiliacao>		Instituição à qual a Pessoa é afiliada	R	F	--
15.5.1	<Nome>		Nome da Instituição	NR	O	--
15.5.2	<Sigla>		Sigla da Instituição	NR	F	--
15.5.3	<País>		País da Instituição	NR	F	ISO-3166
15.5.4	<UF>		UF da Instituição	NR	F	Unidade da Federação

Nº	NomedoElemento	Atributos	Conteúdo	R / NR	O / F	Padrão
15.5.5	<CNPJ>		CNPJ da Instituição	NR	F	14 dígitos
15.5.6	<URL>		Endereço eletrônico do site da Instituição	NR	F	Padrão URI
16	<InstituicaoDefesa>		Instituição onde a tese foi defendida	NR	O	
16.1	<Nome>		Nome da Instituição	NR	O	
16.2	<Sigla>		Sigla da Instituição	NR	F	
16.3	<País>		País da Instituição	NR	F	ISO 3166
16.4	<UF>		UF da Instituição	NR	F	Unidade da Federação
16.5	<CNPJ>		CNPJ da Instituição	NR	F	Padrão MF 14 dígitos
16.6	<URL>		Endereço eletrônico do site da Instituição	NR	F	Padrão URI
16.7	<Programa>		Programa de pós-graduação onde a tese ou dissertação foi defendida	R	F	
16.7.1	<Nome>		Nome do Programa de Pós-Graduação	NR	O	
16.7.2	<Area>		Área de concentração do Programa de Pós-Graduação	NR	F	
17	<AgenciaFomento>		Agencia de financiamento que apoiou financeiramente o autor	R	F	
17.1	<Nome>		Nome da Instituição	NR	O	-
17.2	<Sigla>		Sigla da Instituição	NR	F	-
17.3	<País>		País da Instituição	NR	F	ISO 3166
17.4	<UF>		UF da Instituição	NR	F	Unidade da Federação
17.5	<CNPJ>		CNPJ da Instituição	NR	F	Padrão MF 14 dígitos

Nº	NomedoElemento	Atributos	Conteúdo	R / NR	O / F	Padrão
17.6	<URL>		Endereço eletrônico do site da Instituição	NR	F	Padrão URI
18	<Direitos>	Idioma	Informa as condições de distribuição, reprodução e utilização da tese ou dissertação	R	F	
19	<Extensao>	Namespace	Reservado para a criação de sub-itens (subcampos) para uso específico	NR	F	

Legenda

R= Repetitivo

NR= Não Repetitivo

O= Obrigatório

F= Facultativo

Padrões adotados para os elementos

Elemento 1.2 <DataAtualização>

Padrão DateTime (XML) adotado pelo protocolo OAI.

AAAA-MM-DDThh:mm:ss - as especificações de hora, minuto e segundo são facultativas para o protocolo OAI embora sejam importantes para o processo de coleta automática de dados (harvesting).

Fonte: <<http://www.w3.org/TR/xmlschema-2/#dateTime> >

Elemento 1.4 <Tipo>

DCMI Type Vocabulary

English	Português
Collection	Coleção
Dataset	Conjunto de dados
Event	Evento/ocorrência
Image	Imagem
Interactive Resource	Recursos Interativos
Service	Serviço
Software	Software
Sound	Som
Text	Texto

Fonte: <<http://dublincore.org/documents/dcmi-type-vocabulary/>>

Termos adicionais:

Electronic Theses and Dissertation – Tese ou Dissertação Eletrônica
Printed Theses and Dissertation – Tese ou Dissertação Impressa

Elementos 2.3; 2.4.6; 3.3; 5.1;14.3; 14.5.6; 15.3; 15.5.6; 16.6; 17.6 <URL>

Padrão URI - Uniform Resource Identifiers

Fonte:< <http://rfc.sunsite.dk/rfc/rfc2396.html> >

Elementos 2.4.3; 12.2; 15.5.3; 16.3; 17.3 <País>

Norma ISO 3166

Fonte: <http://www.iso.ch/iso/en/Standards> >

Elementos 2.4.4; 12.2; 14.5.4; 15.5.4; 16.4; 17.4 <UF>

Unidade da Federação

Tabela de Unidades da Federação do IBGE

Elementos 2.4.5; 14.5.5; 15.5.5; 16.5; 17.5 <CNPJ>

O número do CNPJ deve ser informado incluindo-se os 14 dígitos, sem traços e sem pontos, conforme exemplo:
99999999999999

Elemento 5.3 <NivelAcesso>

Um dos dois valores

Restrito
Publico

Elemento 3.2 <Sigla>

Cadastro CCN/COMUT

Fonte: < <http://www.ibict.br> >

Elemento 6 <Idioma >

Norma ISO 639 tanto para 2 (dois) ou 3 (três) caracteres

Fonte: < <http://www.iso.ch/iso/en/Standards> >

Elemento 7 <Grau >

Tabela de Grau

Doutor
Mestre

Elemento 13 <DataDefesa>

Norma ISO 8601 Calendar Date ou no mínimo o ano da defesa, conforme as máscaras:

'YYYY-MM-DD' ou 'YYYY-MM' ou 'YYYY'

Fonte: < <http://www.iso.ch/iso/en/Standards> >

Elementos 14.4; 15.4 <CPF>

O número do CPF deve ser informado incluindo-se os 11 dígitos, sem traços e sem pontos.

Padrões adotados para os atributos

Todos os atributos adotados neste schema são facultativos

Formato (Atributo do elemento 5.1 <URL>)

O atributo formato é utilizado somente para as teses ou dissertações eletrônicas para especificar o(s) formato(s) eletrônico(s) no(s) qual(is) a tese está disponível.

Fonte: < <ftp://ftp.isi.edu/in-notes/iana/assignments/media-types/media-types> >

Idioma (Atributo dos elementos 5.2<Legenda>, 4 <Título>; 9 <Resumo>; 10 <Cobertura>; 11 <Assunto>; 18 <Direitos>)

Norma ISO 639 tanto para 2 (dois) ou 3 (três) caracteres

Fonte: < <http://www.iso.ch/iso/en/Standards> >

Esquema (Atributo do elemento 11 <Assunto>)

O atributo Esquema é o instrumento de indexação utilizado pela biblioteca depositária.

Papel (Atributo do elemento 15 <Contribuidor>)

O atributo Papel é o tipo de contribuição e forma de participação na geração/aprovação da tese ou dissertação.

Tabela de papéis

Orientador
Co-Orientador
Membro da Banca

Anexo 4 – TABELAS DAS QUAIS SÃO EXTRAÍDOS OS METADADOS DAS TESES E DISSERTAÇÕES ELETRÔNICAS DA UDESC

Tabela Arquivos:

Nome do campo	Conteúdo
arCod	Código do arquivo
tsIdentificador	Código que identifica a tese ou dissertação na base de dados da biblioteca
arFormato	Formato do arquivo
arURL	Endereço eletrônico onde se encontra o arquivo
arDireitos	Informação sobre Direitos

Tabela Assunto:

Nome do campo	Conteúdo
asCod	Código do assunto
tsIdentificador	Código que identifica a tese ou dissertação na base de dados da biblioteca
asIdioma	Idioma no qual o assunto foi informado
asAssunto	Assunto ao qual se refere a tese ou dissertação
asEsquema	Esquema do assunto

Tabela BibliotecaDepositaria:

Nome do campo	Conteúdo
bdCod	Código da instituição onde encontra-se a tese ou dissertação
bdNome	Nome da instituição onde encontra-se a tese ou dissertação
bdSigla	Sigla da instituição onde encontra-se a tese ou dissertação
bdURL	Endereço eletrônico da instituição onde encontra-se a tese ou dissertação

Tabela BibliotecaDigital:

Nome do campo	Conteúdo
bbCod	Código da biblioteca digital
bbNome	Nome da biblioteca digital
bbSigla	Sigla da biblioteca digital
bbURL	Endereço eletrônico do site da Biblioteca Digital

Tabela Direitos:

Nome do campo	Conteúdo
drCod	Código do direito
tsIdentificador	Código que identifica a tese ou dissertação na base de dados da biblioteca
drDireito	Texto que informa se o conteúdo do arquivo está liberado para acesso ou não
drIdioma	Idioma no qual o drDireito foi informado

Tabela Instituicao:

Nome do campo	Conteúdo
inCod	Código da Instituição de defesa
inNome	Nome da Instituição de defesa
inSigla	Sigla da Instituição de defesa
inPais	País da Instituição de defesa
inUF	Unidade Federativa da Instituição de defesa
inCNPJ	CNPJ da Instituição de defesa
inURL	Endereço eletrônico do site da Instituição de defesa

Tabela LegendaArquivo:

Nome do campo	Conteúdo
laCod	Código da legenda
arCod	Código do arquivo
laLegenda	Legenda para o arquivo, descrição de seu conteúdo
laIdioma	Idioma da legenda

Tabela OAIControle

Nome do campo	Conteúdo
tsIdentificador	Código que identifica a tese ou dissertação na base de dados da biblioteca

Tabela PFPJ:

Nome do campo	Conteúdo
pjCod	Código da pessoa jurídica
pfCod	Código da pessoa física
pfpjTipo	Tipo de pessoa física ou jurídica

Tabela PessoaFisica:

Nome do campo	Conteúdo
pfCod	Código da pessoa física
pfNome	Nome da pessoa física
pfCitacao	Forma pela qual a pessoa física será citada em sob alguma norma
pfCitacaoABNT	Forma pela qual a pessoa física será citada de acordo com a ABNT
pfLattes	Endereço do currículo Lattes da pessoa física
pfCPF	CPF da pessoa física
pfEmail	Email da pessoa física

Tabela PessoaJuridica:

Nome do campo	Conteúdo
pjCod	Código da pessoa jurídica
pjNome	Nome da pessoa jurídica
pjSigla	Sigla da pessoa jurídica

pjPais	País da pessoa jurídica
pjUF	Unidade Federativa da pessoa jurídica
pjCNPJ	CNPJ da pessoa jurídica
pjURL	Endereço eletrônico do site da pessoa jurídica

Tabela Programa:

Nome do campo	Conteúdo
prCod	Código do programa de pós-graduação
prNome	Nome do programa de pós-graduação
prArea	Área do programa de pós-graduação
bdCod	Código da biblioteca digital a qual pertence o programa de pós-graduação

Tabela Resumo:

Nome do campo	Conteúdo
rsCod	Código do resumo
tsIdentificador	Código que identifica a tese ou dissertação na base de dados da biblioteca
rsResumo	Resumo da tese ou dissertação
rsIdioma	Idioma no qual foi informado o resumo

Tabela TSPF:

Nome do campo	Conteúdo
pfCod	Código da pessoa física
prCod	Código do programa
tspfTipo	Tipo de pessoa física (Autor, co-autor ou orientador)
tspfDispEmail	Disponibilidade ou não do e-mail da pessoa física

Tabela Teses:

Nome do campo	Conteúdo
tsIdentificador	Código que identifica a tese ou dissertação na base de dados da biblioteca
tsNumeroChamada	Código que identifica a tese ou dissertação no acervo da biblioteca depositária
tsIdioma	Idioma no qual a tese ou dissertação foi escrita
tsGrau	Grau acadêmico associado à tese ou dissertação
tsTitulacao	Nome do grau acadêmico associado à tese ou dissertação
tsDataAtualizacao	Data em que foi realizada a última atualização do registro da tese ou dissertação
tsCidadeLocalDefesa	Cidade onde foi defendida a tese ou dissertação
tsUFLocalDefesa	Unidade Federativa onde foi defendida a tese ou dissertação
tsPaisLocalDefesa	País onde foi defendida a tese ou dissertação
tsDataDefesa	Data em que foi defendida a tese ou dissertação
tsDataFinalizacao	Data da finalização da tese ou dissertação
tsDataLiberacaoPos	Data de liberação da tese ou dissertação
cnCodPos	Sem conteúdo
tsDataRevisao	Data da revisão da tese ou dissertação

cnCodRevisor	Código do Revisor
tsDataCatalogacao	Data da catalogação da tese ou dissertação
cnCodCatalogador	Sem conteúdo
tsCampo1	Informações adicionais sobre a tese ou dissertação
tsCampo2	Informações adicionais sobre a tese ou dissertação
inCod	Código da Instituição
tsDataDiretorio	Data na qual foram informados os dados da teses ou dissertação

Tabela Tipo:

tpCod	Código do tipo do arquivo
tsIdentificador	Identificador da tese ou dissertação
tpTipo	Tipo do arquivo (Electronic Thesis and Dissertation, Texto, etc)

Tabela Titulo:

ttCod	Código do título da tese ou dissertação
tsIdentificador	Identificador da tese ou dissertação
ttTitulo	Título da tese ou dissertação
ttIdioma	Idioma no qual o título da tese ou dissertação foi informado