

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Silvia Maria Puentes Bentancourt

**A QUALIDADE NOS METADADOS DA REVISTA EM QUESTÃO:
ANÁLISE DAS RECOMENDAÇÕES DO DUBLIN CORE**

Porto Alegre, 2011

Silvia Maria Puentes Bentancourt

**A QUALIDADE NOS METADADOS DA REVISTA EM QUESTÃO:
ANÁLISE DAS RECOMENDAÇÕES DO DUBLIN CORE**

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção de título de Bacharel em Biblioteconomia da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul.

Orientador: Prof. Dr. Rafael Port da Rocha

Porto Alegre, 2011

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Carlos Alexandre Netto

Vice-Reitor: Prof. Dr. Rui Vicente Oppermann

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretor: Prof. Ricardo Schneiders da Silva

Vice-Diretor: Profª. Drª. Regina Helena van der Laan

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Chefe: Profª. Drª. Ana Maria Mielnickzuk de Moura

Chefe substituta: Profª. Drª. Sonia Elisa Caregnato

COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA

Coordenadora: Profª. Ms. Glória Isabel Ferreira Sattamini

Coordenadora Substituta: Profª. Drª. Samile Vanz

CIP- Brasil - Dados Internacionais de Catalogação na Publicação

B456 Bentancourt, Silvia Maria Puentes

A Qualidade nos Metadados da Revista Em Questão :
análise das recomendações do Dublin Core / Silvia Maria
Puentes Bentancourt. – 2011.

99 fl.

Monografia (graduação) - Universidade Federal do Rio
Grande do Sul, Faculdade de Biblioteconomia e
Comunicação, Departamento de Ciências da Informação,
2011.

Orientador: Prof. Dr. Rafael Port da Rocha.

1. Metadados 2. Dublin Core Metadata Initiative (DCMI)
3. Qualidade de dados 4. Revista Em Questão I. Rocha,
Rafael Port da II. Título.

CDU 001.103.2

Departamento de Ciências da Informação
Faculdade de biblioteconomia e Comunicação
Universidade Federal do Rio Grande do Sul
Rua Ramiro Barcelos, 2705
CEP 90035-007 Porto Alegre – RS
Fone: (51) 33085067
Fax: (51) 33085435
E-mail: fabico@ufrgs.br

Silvia Maria Puentes Bentancourt

**A QUALIDADE NOS METADADOS DA REVISTA EM QUESTÃO:
ANÁLISE DAS RECOMENDAÇÕES DO DUBLIN CORE**

Trabalho de Conclusão de Curso
apresentado como requisito parcial para
obtenção de título de Bacharel em
Biblioteconomia da Faculdade de
Biblioteconomia e Comunicação da
Universidade Federal do Rio Grande do
Sul.

Data aprovação: 08 de dezembro de 2011

Prof. Esp. Helen Rose Flores de Flores - UFRGS
Banca Examinadora

Ms. Ana Gabriela Clipes Ferreira - UFRGS
Banca examinadora

Prof. Dr. Rafael Port da Rocha - UFRGS
Prof. orientador

*À minha família, que de forma incondicional,
acompanham as minhas empreitadas sem tentar entender os motivos.*

AGRADECIMENTOS

Ao chegar ao fim deste trabalho, que representa o final de uma etapa, e tentar lembrar o seu percurso, percebo que ele é o somatório do trabalho e esforço de muitas pessoas. Pois, ninguém é uma ilha ..., alguém já falou isso antes¹.

Sendo assim, temos muitos a quem agradecer,
aos que com seus exemplos nos motivam a ir mais longe, e aqui um agradecimento especial à Prof Dr^a Ida Regina Stumpf, quem com generosidade permite um aprendizado sem fim; principalmente durante o período da bolsa de Iniciação Científica;
às colegas de bolsa Natália, Rosely e Camila;
ao Prof. Dr. Rafael Port da Rocha, meu paciente orientador, por sua atenção e precisão nos questionamentos;
a minha orientadora de estágio curricular obrigatório, Bibl. Ms. Sonia Regina Zanotto, com quem mantivemos diálogos produtivos sobre a Ciência da Informação e tantos outros assuntos correlatos;
aos professores que ampliam as perspectivas de atuação do profissional bibliotecário, destaco as professoras Dr^a Samile Vanz e Dr^a Sonia Caregnato;
àqueles colegas que nos acompanharam durante as aulas e que, com suas contribuições, propiciaram o debate de idéias nem sempre compartilhadas;
ao professores abertos a discussões e que permitem esses momentos, como o Prof. Dr. Valdir José Morigi;
aos amigos que estão prontos para compartilhar momentos de estudo e descontração, Gabriela, Fabiana, Ceres, Juliana, Fabrício e, inclusive, aos seus familiares, que me recebem de braços abertos;
àquela voz interior, que frequentemente não quero escutar, mas que está sempre aí, muito obrigada Deus;
à minha família, a quem dedico este trabalho.

¹ John Donne (1572-1631). Não resisti a um comentário bibliotecário.

A qualidade começa com a educação e termina com a educação.

Kaoru Ishikawa

RESUMO

O estudo dos metadados da Revista Em Questão investiga a qualidade dos mesmos pelo levantamento da não conformidade com as recomendações do *Dublin Core Metadata Initiative* sob a ótica das categorias da qualidade. Aborda as mudanças ocorridas no campo das Ciências da Informação pelas transformações tecnológicas, o uso de Repositórios Institucionais e Revistas Eletrônicas para a comunicação científica e o papel dos metadados na visibilidade das pesquisas científicas e o acesso e uso das suas informações. Estuda o padrão de metadados Dublin Core *Metadata Initiative* (DCMI), suas instruções e recomendações. Apresenta as dimensões da qualidade de dados, sejam inerentes aos dados, dependentes do sistema ou ambas; e suas categorias: Intrínseca, Acessibilidade, Contextual e Representacional. Para a pesquisa foram colhidos os metadados da Revista pelo ambiente FABRICO e analisados tal qual se apresentavam segundo as não conformidades do tipo: dados faltando, dados incorretos, dados confusos, dados insuficientes. Os resultados são apresentados por elemento do DCMI, em tabelas, junto aos índices fornecidos pelo FABRICO. Eles evidenciaram deficiências na política de metadados da Revista uma vez que não há padronização na adoção do DCMI, nem quanto à definição do elemento nem na instrução de uso. Foram vistos metadados com registros faltando, ambíguos e com incoerências no tratamento ao longo do tempo. Recomenda-se o estabelecimento de política de metadados alinhada com a estratégia da Revista para tornar seu conteúdo visível, acessível e pronto para uso.

Palavras-chave: Metadados. *Dublin Core Metadata Initiative* (DCMI). Qualidade de dados. Revista Em Questão.

RESUMEN

El estudio de los metadatos de la Revista Em Questão investiga la calidad de los mismos por el apunte de la no conformidad con las recomendaciones de el *Dublin Core Metadata Initiative* (DCMI) bajo la óptica de las categorías de la calidad. Aborda los cambios ocurridos en el campo de las Ciencias de la Información por las transformaciones tecnológicos, el uso de Repositórios Institucionales y Revistas Electrónicas para la comunicación científica y el papel de los metadatos en la visibilidad de las investigaciones científicas y el acceso y uso de sus informaciones. Estudia el padrón de metadatos *Dublin Core Metadata Initiative* (DCMI), sus instrucciones y recomendaciones. Presenta las dimensiones de la calidad de datos, sean inherentes a los datos, dependientes del sistema o ambos, y sus categorías: Intrínseca, Accesibilidad, Contextual y Representacional. Para el estudio fueron colectados los metadatos de la Revista por el ambiente FABRICO. Ellos evidenciaron deficiencias en la política de metadatos de la Revista una vez que no hay estandarización en la adopción de el DCMI, ni respecto a la definición de los elementos ni en la instrucción de uso. Fueron vistos metadatos con registros faltando, ambíguos y con incoherencias en el tratamiento a lo largo de el tiempo. Se recomienda establecer una política de metadatos alineada con la estrategia de la Revista para tornar su contenido visible, accesible y pronto para uso.

Palabras clave: Metadatos. *Dublin Core Metadata Initiative* (DCMI). Calidad de datos. Revista Em Questão.

LISTA DE TABELAS E QUADROS

Quadro 1 - Modelo Genérico de Qualidade de Dados	55
Quadro 2 – Folha de Verificação para Análise dos Metadados	60
Tabela 1 – Metadados do Elemento Título.....	62
Tabela 2 – Metadados do Elemento Criador.....	64
Quadro 3 – Elemento Título e Criador do artigo A Mediação das cartas dos leitores na mídia: mapas imaginários sobre Porto Alegre	66
Quadro 4 – Exemplos de falta de padronização no elemento Criador	66
Tabela 3 – Metadados do Elemento Criador.....	68
Tabela 4 – Metadados do Elemento Assunto.....	72
Quadro 5 – Exemplos de Representação de Assunto.....	74
Tabela 5 – Metadados do Elemento Descrição.....	76
Tabela 6 – Metadados do Elemento Cobertura.....	77
Tabela 7 – Metadados do Elemento Fonte	80
Tabela 8 – Metadados do Elemento Identificador	82
Tabela 9 – Metadados do Elemento Relação.....	83
Tabela 10 – Metadados do Elemento Formato	85
Tabela 11 – Metadados do Elemento Tipo.....	86
Tabela 12 – Metadados do Elemento Data	88
Tabela 13 – Metadados do Elemento Idioma.....	89
Tabela 14 – Metadados do Elemento Editor	90
Tabela 15 – Metadados do Elemento Direitos Autorais	91

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
AL	Acesso Livre
AO	<i>Open Access</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
DCMI	<i>Dublin Core Metadata Initiative</i>
DOI	<i>Digital Object Identifier</i>
FABICO	Faculdade de Biblioteconomia e Comunicação
ISBN	<i>International Standard Book Number</i>
ISO	<i>International Organization for Standardization</i>
MDR	<i>Metadata registries</i>
OAI PMH	<i>Open Archives Initiative Protocol for Metadata Harvesting</i>
ODLIS	<i>Online Dictionary for Library and Information Science</i>
OLAC	<i>Open Language Archives Community</i>
pdf	<i>Portable Document Format</i>
PPGCOM	Programa de Pós-Graduação em Comunicação e Informação
QAF	<i>Quality Assurance Framework</i>
RDF	<i>Resource Description Framework</i>
RFC	<i>Request for Comments</i>
RI	Repositórios Institucionais
SEER	Sistema Eletrônico de Editoração de Revistas
SGML	<i>Standard Generalized Markup Language</i>
TI	Tecnologias da Informação
UFRGS	Universidade Federal do Rio Grande do Sul
URI	<i>Uniform Resource Identifier</i>
URL	<i>Uniform Resource Locator</i>
XML	<i>eXtensible Markup Language</i>

SUMÁRIO

1	INTRODUÇÃO	14
2	OBJETIVOS	18
2.1	OBJETIVO GERAL	18
2.2	OBJETIVOS ESPECÍFICOS	18
3	REVISTA EM QUESTÃO	19
4	REFERENCIAL TEÓRICO	22
4.1	A CIÊNCIA DA INFORMAÇÃO EM AMBIENTE <i>WEB</i>	22
4.2	REVISTAS CIENTÍFICAS ELETRÔNICAS	30
4.3	METADADOS.....	34
4.4	<i>DUBLIN CORE METADATA INICIATIVE</i> (DCMI)	37
4.4.1	Título (<i>title</i>)	40
4.4.2	Criador (<i>creator</i>).....	40
4.4.3	Colaborador (<i>contributor</i>)	41
4.4.4	Assunto (<i>subject</i>).....	41
4.4.5	Descrição (<i>description</i>).....	42
4.4.6	Cobertura (<i>coverage</i>)	43
4.4.7	Fonte (<i>source</i>).....	44
4.4.8	Identificador (<i>identifier</i>).....	44
4.4.9	Relação (<i>relation</i>).....	45
4.4.10	Formato (<i>format</i>)	46
4.4.11	Tipo (<i>type</i>)	46
4.4.12	Data (<i>date</i>)	47
4.4.13	Idioma (<i>language</i>)	48
4.4.14	Editor (<i>publisher</i>).....	48
4.4.15	Direitos (<i>rights</i>).....	49
4.5	QUALIDADE DOS METADADOS	50
5	METODOLOGIA	58
6	RESULTADOS	61
6.1	TÍTULO (<i>title</i>).....	61
6.2	CRIADOR (<i>creator</i>)	64

6.3	COLABORADOR (<i>contributor</i>)	67
6.4	ASSUNTO (<i>subject</i>)	71
6.5	DESCRIÇÃO (<i>description</i>)	75
6.6	COBERTURA (<i>coverage</i>).....	77
6.7	FONTE (<i>source</i>).....	80
6.8	IDENTIFICADOR (<i>identifier</i>).....	82
6.9	RELAÇÃO (<i>relation</i>)	83
6.10	FORMATO (<i>format</i>).....	84
6.11	TIPO (<i>type</i>).....	86
6.12	DATA (<i>date</i>)	87
6.13	IDIOMA (<i>language</i>)	89
6.14	EDITOR (<i>publisher</i>).....	90
6.15	DIREITOS AUTORAIS (<i>rights</i>)	91
7	CONCLUSÕES	93
	REFERÊNCIAS	95

1 INTRODUÇÃO

As mudanças tecnológicas levaram a mudanças de hábitos e costumes que resultaram em transformações sociais, entre elas, a relação com a informação. Num novo paradigma esta tornou-se mais importante do que o seu suporte, o seu acesso e assimilação passaram a significar conhecimento e poder. A tecnologia também favoreceu a disseminação, a produção científica tornou-se acelerada, acompanhada de uma proliferação de informação desordenada, complexa e nem sempre confiável. A falta de uma infra-estrutura preparada, com conhecimento de tratamento da informação, pode levar a uma inacessibilidade da informação desejada mesmo esta estando armazenada.

Desta forma, pelo crescimento do número e diversidade de informação disponível, sua gestão, manutenção e recuperação cresceram em importância. O tratamento da informação passou a interessar não somente aos profissionais da informação como também aos seus usuários, ao perceberem o valor da organização da informação para que esteja disponível e acessível, em quantidade e qualidade pretendida.

A sociedade mudou e provocou mudanças nas bibliotecas e, por conseqüência, no mercado de trabalho do bibliotecário. Agora demandam, aos centros de informação e aos seus profissionais, habilidades e conhecimentos antes despercebidos ou desnecessários. Além disso, as unidades de informação reconheceram a impossibilidade de conter sob sua responsabilidade toda a informação produzida, donde se torna imperativo prestar outros serviços e trabalhar em parcerias, com intercâmbio de informações. Esse contexto exige que a biblioteca cumpra suas funções educativa, cultural, recreativa e informacional, além de possibilitar aos seus usuários fundamentação para serem cidadãos ativos. O bibliotecário passou a ser um mediador atento ao intercâmbio de informação que, embora sempre tenha existido, aumentaram as suas possibilidades pela evolução dos meios de comunicação.

A comunicação científica passou a beneficiar-se de diversos recursos, como os repositórios de periódicos editados em centros de pesquisa vinculados a universidades e a outras instituições similares. O movimento do Acesso Livre expandiu ainda mais as opções, onde o importante é compartilhar para propiciar o crescimento conjunto.

Mas o movimento trouxe também imposições técnicas para que os seus propósitos pudessem ser alcançados. A interoperabilidade das diferentes bases de dados é uma destas, onde as bases precisam disponibilizar seus metadados para informar quais recursos informacionais quer compartilhar e de que forma.

Assim, recuperação da informação armazenada numa base de dados depende dos metadados disponibilizados para colheita. No entanto, só conhecer os metadados não garante o acesso à informação disponibilizada. Ao longo de todo o canal de informação a comunicação deverá ocorrer sem ruídos, ou seja, os dados transmitidos precisarão estar corretos para permitir acessar o que se pretende, no momento necessário, de forma completa e no formato certo. Portanto, a qualidade dos metadados disponibilizados para acessar a informação é fundamental para que o processo se desenvolva a contento. Ela precisa ser garantida, seja através da definição de Política de Metadados, Normas de Procedimentos, treinamentos, avaliações, entre outros.

Embora já existam padrões de metadados, todos mantêm um grau de liberdade que torna necessário fazer opções no momento de uma aplicação específica. Também conhecer a forma como os dados são captados pelos motores de busca é importante porque é deles que depende a oportunidade de recuperar os documentos de interesse dos usuários. Cabe ressaltar que a responsabilidade de manter metadados corretos é de quem os inseriu e depende da capacidade do profissional que os digita.

O principal problema é como captar o mundo real e representá-lo por uma linguagem artificial que, todavia, permita sua recuperação através de mecanismos de busca. As dificuldades são ainda maiores pelo envolvimento de variáveis diversas relacionadas com a natureza da informação, representação e armazenamento, programas de computadores, sistemas de busca, comportamento dos usuários, por citar alguns.

A partir de leituras sobre Metadados percebeu-se a diversidade de enfoques que o tema permite. Isto também possibilita inferir o interesse despertado por pesquisadores de diversas áreas do conhecimento, como das Ciências da Informação, Análise de Sistemas, Informática, Linguística, entre outras. Sendo assim, também as soluções para a melhoria da qualidade dos metadados são de origens variadas e que, mediante esforços conjuntos, procura-se o estabelecimento

de normas e diretrizes para uso geral, como as da *International Organization for Standardization* (ISO).

No entanto, a Biblioteconomia é a área do conhecimento que tem como objeto de estudo a representação da informação. Nesse sentido, acredita-se que seus conhecimentos sobre a representação de documentos de forma qualificada e que permitam a sua recuperação por parte dos usuários são de fundamental importância para a definição de quais metadados são necessários e o seu grau de especificidade.

O interesse da autora pelo assunto decorreu do auxílio que as informações contidas nos metadados propiciam nos procedimentos de busca, principalmente aqueles relacionados com a recuperação da informação na rede mundial de computadores, doravante apenas *Web*. Assim, o propósito desta investigação é levantar a situação atual dos metadados disponibilizados na Revista *Em Questão*, publicação editada pela Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul (FABICO/UFRGS), e analisá-los sobre os aspectos da adequação a padrões internacionais e dimensões de qualidade.

Deste modo, o objeto geral de estudo é a Revista *Em Questão*. Para unidades de análise, foram selecionados todos metadados disponibilizados na *Web* referentes às publicações de 2003 a 2010, totalizando 17 edições. A pesquisa aborda o universo dos dados.

A opção pela Revista foi motivada pela facilidade da coleta dos metadados, por ser uma publicação da mesma faculdade em que este estudo é proposto e, portanto, viabiliza a apresentação dos resultados diretamente aos seus editores. Especificamente os metadados foram observados com o propósito de identificar a sua aderência às recomendações de melhores práticas do *Dublin Core Metadata Initiative* (DCMI) e a qualidade foi verificada pela não conformidade com diretrizes e normas, onde se discriminou a origem das falhas por dados faltando, incorretos, confusos ou insuficientes.

Assim, o problema de pesquisa é conhecer se os metadados da Revista *Em Questão* seguem as recomendações de melhores práticas do *Dublin Core Metadata Initiative* (DCMI) e assinalar a qualidade pela ocorrência das possíveis incorreções nos metadados disponíveis para colheita. Os metadados foram analisados em conjunto e por elemento que descrevem, através de uma análise horizontal.

Outro motivo para este trabalho foi conhecer o ambiente FABRICO² (ROCHA, *on-line*) e verificar suas possibilidades para a colheita e análise de metadados. O FABRICO foi desenvolvido pelo Prof. Dr. Rafael Port da Rocha, da mesma faculdade onde é editada a Revista. O ambiente FABRICO além de colher os metadados disponibilizados pelo sistema de acesso aberto, agrega os mesmos por elemento e por recurso, apresenta o conteúdo dos registros em forma de relatório ou nuvem; faz cálculos estatísticos básicos e organiza os recursos por frequência.

Trata-se de uma pesquisa exploratória, quanto aos seus objetivos, onde os dados são analisados de forma qualitativa e que buscou subsídios para os seus procedimentos em bibliografia pertinente. A fim de efetuar essa análise, foi necessário entender a Ciência da Informação, a gestão da informação, as revistas científicas e a interoperabilidade no ambiente *Web*, além da qualidade de dados e metadados.

Sendo assim, o trabalho está dividido em seis seções. A primeira traz a introdução, a segunda os objetivos do trabalho e a terceira apresenta a Revista Em Questão. A quarta seção corresponde à revisão da literatura e está subdividida em uma apresentação do ambiente globalizado das Ciências da Informação; as revistas científicas; a qualidade dos metadados, suas contribuições e problemas decorrentes da qualidade, padronização do DCMI. Na quinta são apresentados e discutidos os resultados da pesquisa e na sexta as conclusões do estudo.

Espera-se que este trabalho contribua na formação mais consistente da autora no tema, na familiaridade com o ambiente FABRICO e que colabore com a equipe da Revista Em Questão pelos resultados apontados no estudo.

Salienta-se que, quanto à grafia dos termos utilizados em língua estrangeira, estes estão escritos em itálico e alguns não foram traduzidos devido a sua difusão e uso frequente no original. Também em itálico encontram-se os exemplos no corpo do texto para destacá-los.

² <http://www6.ufrgs.br/fabrico/f2/notes/tipos.php>

2 OBJETIVOS

Os objetivos da pesquisa servem para ilustrar o que se pretende estudar. O objetivo geral relaciona-se com o assunto escolhido e os específicos com as limitações do mesmo.

2.1 OBJETIVO GERAL

Investigar a qualidade dos metadados da Revista Em Questão, mediante o levantamento das não conformidades com as recomendações de DCMI sob a ótica das categorias da qualidade.

2.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos são aqueles que detalham o objetivo geral, logo, definem e caracterizam as etapas ou fases da pesquisa. Neste caso, os objetivos específicos são:

- a) levantar critérios para observar a qualidade de metadados;
- b) conhecer as práticas de uso de *Dublin Core Metadata Initiative* (DCMI);
- c) examinar os metadados conforme as recomendações de melhores práticas do DCMI;
- d) assinalar as inconformidades nos metadados; e
- e) propor recomendações para a adequação dos metadados.

3 REVISTA EM QUESTÃO

Em Questão é uma Revista da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul (FABICO /UFRGS), que surgiu em substituição da Revista de Biblioteconomia & Comunicação. A Revista de Biblioteconomia & Comunicação originou-se do trabalho experimental dos alunos da disciplina de Projeto Experimental em Jornalismo II, em 1986, com o propósito de publicar matérias de interesse da Biblioteconomia e Comunicação a fim de discutir sobre as tendências e divulgar os estudos realizados por professores e alunos da Faculdade. Circulou com publicações irregulares e com interrupções nos anos de 1991 a 1993, 1995, 1997 a 1999, num total de oito volumes.

Em 2003 surgiu a Revista Em Questão com o mesmo objetivo de tornar pública a produção intelectual de professores, alunos e técnicos-administrativos da Faculdade e outros especialistas, segundo seu edital de jan./jul. 2003. Também procurou incentivar o desenvolvimento de massa crítica focando como público alvo a população discente e docente da universidade acreditando com isto transformar a sociedade. Passa a denominar seus campos de interesse como de Informação e Comunicação e não mais Biblioteconomia e Comunicação e transformou-se em uma publicação semestral.

Os documentos devem ser originais, no tema ou na abordagem, em português, inglês e espanhol e categorizados por artigos, entrevistas, resenhas, e outros que possam ser considerados adequados. O processo de avaliação segue o procedimento de avaliação por pares, onde conta com um Conselho Editorial além de consultores e especialistas para este propósito. A editora é a própria universidade e participa do Programa de Apoio à Edição de Periódicos da UFRGS.

A primeira edição de 2004 foi lançada concomitantemente *on-line* e em formato impresso, o que ocorre ainda hoje, sendo que o formato impresso é utilizado apenas para permuta entre bibliotecas. A disponibilização *on-line* é através do Sistema Eletrônico de Editoração de Revistas (SEER) e em 2007 foram incluídos todos os metadados das edições a partir de 2003, o que gerou maior visibilidade pela disponibilidade para colheita dos seus metadados.

Num processo de melhoria constante, o Conselho Editorial procura acompanhar as orientações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), especificamente no que diz respeito ao *Qualis*, sobre

periódicos científicos. “Qualis é o conjunto de procedimentos utilizados pela Capes para estratificação da qualidade da produção intelectual dos programas de pós-graduação” (CAPES, *on-line*). A Revista foi avaliada pela *Qualis* no triênio 2007-2009, com estrato B2³ nas áreas da Ciência Social Aplicada e Interdisciplinar.

A Revista participa da iniciativa do Acesso Livre disponibilizando seu conteúdo integral, gratuito e de forma imediata. A Revista *Em Questão* é indexada em bases de dados nacionais e internacionais.

Apresentam-se a seguir os dados da Revista, resumidamente:

Título Atual: *Em Questão*.

Título Anterior: Revista de Biblioteconomia & Comunicação

Editor: Universidade Federal do Rio Grande do Sul (UFRGS).

Editor Chefe: Prof. Dr. Valdir José Morigi (gestão 2009/2011).

Local de Publicação: Porto Alegre, RS, Brasil.

Início da publicação: 2003 – volume 9, substitui a Revista de Biblioteconomia & Comunicação, 1986 a 2000, volumes 1 a 8.

ISSN: Versão Impressa: 1807-8893.

E-ISSN: Versão Eletrônica: 1808-5245.

Qualis: B2.

Periodicidade: semestral.

Idioma: português, espanhol e inglês.

Seções principais: editorial, artigos inéditos e de revisão, entrevistas e resenhas.

Formato: impresso no período de 1986 a 2000 e também eletrônico a partir de 2003.

Disponível em: <<http://seer.ufrgs.br/EmQuestao/index>>.

Indexação:

Bases Nacionais:

- BASEBRES (Base Referencial de Revistas de Biblioteconomia e Ciência da Informação);
- BRAPCI (Base Brasileira de Periódicos em Ciência da Informação);
- Portal de Informação Científica e Tecnológica/Prossiga/IBICT;

³ Webqualis. Disponível em:
<<http://qualis.capes.gov.br/webqualis/ConsultaListaCompletaPeriodicos.faces>> Acesso em: ago. 2011

- Portal de Periódicos da UFRGS;
- Portcom/Intercom (Portal de Livre Acesso a Red Iberoamericana de revistas de comunicación y cultura);
- Portal de Periódicos da CAPES; e
- Univerciência (Portal de Revistas Eletrônicas em Ciências da Comunicação).

Bases Internacionais:

- DOAJ (*Directory of Open Access Journals*);
- LAPTOC (*Latin American Periodicals Tables of Contents*);
- Latindex (*Sistema Regional de Información en línea para revistas científicas de America Latina, el Caribe, España y Portugal*); Oásis - Ibict (*Open Access and Scholarly Information System*); e
- Red Iberoamericana de Revistas de Comunicación y Cultura.

4 REFERENCIAL TEÓRICO

O referencial teórico da pesquisa inicia com a contextualização da Ciência da Informação na área das Ciências Sociais Aplicadas e nas implicações que as mudanças tecnológicas demandaram. A seguir aborda as revistas científicas, sua função e publicação em ambiente *Web*. Na terceira subseção é focado o aspecto da padronização de metadados, especificamente o padrão *Dublin Core Metadata Initiative* (DCMI). A última está dedicada à qualidade dos dados e metadados, onde se apresentam estudos realizados para sua verificação e garantia de atendimento aos requisitos estabelecidos.

4.1 A CIÊNCIA DA INFORMAÇÃO EM AMBIENTE *WEB*

Numa nova ordem mundial, onde a globalização e as tecnologias de informação são uma característica marcante, ganha força e forma a Ciência da Informação. Ela é classificada pela Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) na área de Ciências Sociais Aplicadas 1, juntamente com Museologia e Comunicação, e é composta por:

Teoria da Informação;
Teoria Geral da Informação;
Processos da Comunicação;
Representação da Informação;
Biblioteconomia;
Teoria da Classificação;
Métodos Quantitativos, Bibliometria;
Técnicas de Recuperação da Informação;
Processos de Disseminação da Informação;
Arquivologia; e
Organização de Arquivos (CAPES, 2009, *on-line*)⁴.

Então, a partir dessa lista, pode-se dizer que o profissional da informação é aquele que participa dos “processos de geração, disseminação, recuperação, gerenciamento, conservação e utilização da informação” (MENDONÇA, 2006, p. 2). A continuação, a mesma autora relaciona os profissionais que tem a informação como seu objeto de trabalho, quais sejam “bibliotecários, arquivistas,

⁴ Tabela de Áreas de Conhecimento. [documento *on-line*, sem paginação]. Disponível em: <<http://www.capes.gov.br/avaliacao/tabela-de-areas-de-conhecimento>> Acesso em: 27 jun. 2011

documentalistas, administradores, analista de sistemas, comunicadores, jornalistas, além dos profissionais ligados à informática” (MENDONÇA, 2006, p. 2). Nessa mesma direção, Alvarenga (2003) acrescenta ainda linguistas e produtores comerciais e institucionais de bases de dados documentais. Logo, deriva que o processo da informação é composto por fases que requerem conhecimentos e instrumentos especializados que envolvem esses profissionais que atuam de forma coordenada ou independente.

Cabe ao bibliotecário lidar com esse contexto e socializar a informação. Tornando-se o mediador entre as pessoas e o conhecimento registrado, o que resulta em desenvolvimento social, cultural, científico e tecnológico, além da possibilidade da formação de cidadãos, considerando-se o termo em toda sua amplitude. Assim, a sua atuação aparecerá qualificando de maneira positiva ou negativa o processo de informação.

O processo de informação consiste em fazer com que uma mensagem atinja o destino. No caso dos documentos disponibilizados para consulta em geral, e na *Web* de forma particular, são fundamentais a categorização; indexação; recuperação de informação; e, mais recentemente, a interação do homem e a máquina. As tecnologias de informação favoreceram muito o processo que, aliado aos conhecimentos da biblioteconomia, psicologia e linguística possibilitaram agilizar a transferência do conhecimento organizado (FOSKETT⁵, 1980 *apud* LIMA, 2003).

A Comunicação Científica é um caso particular no qual a informação que flui é resultado de pesquisas na busca de novos conhecimentos. Meadows (1999, p. 161) salienta que a "realização de pesquisas e a comunicação de seus resultados são atividades inseparáveis." Valério (2008, p. 161) salienta que "Meadows inaugurou um novo ciclo de destaque para a comunicação de ciência, ao enfatizar a importância da comunicação, e ao deixar claro que a comunicação é tão vital quanto a própria pesquisa."

A comunicação científica pode ocorrer de duas formas, a formal e a informal. A formal caracteriza-se pela disponibilidade da informação por longo período a um público amplo, já a informal é o inverso, efêmera e acessível a um público restrito (MEADOWS, 1999). São considerados canais formais os periódicos e os livros, e os informais são os congressos, correspondência pessoal e comunicações orais.

⁵ FOSKETT, D. J. et. al. A ciência da informação como disciplina emergente: implicações educacionais. **Ciência da Informação ou informática?**. Rio de Janeiro: Calunga, 1980. p. 53-69.

O Fluxo da Informação Científica normalmente é representado por um modelo, que mostra o caminho de uma pesquisa desde a produção, passando pela publicação, até a utilização das informações da mesma, que é constatada pelos estudos de citações. Assim, a Comunicação Científica é definida por Garvey e Griffith⁶ (1979, *apud* VALERIO, 2008, p. 161) como o conjunto de atividades associadas à produção, disseminação e uso da informação.

Weitzel (2006, p. 85) aponta que “dentre as principais questões em discussão na área da comunicação científica em ambiente eletrônico destaca-se a mudança estrutural do fluxo de comunicação científica e a conseqüente fragilidade dos domínios formal e informal.” A importância da comunicação científica formal já é notória. Pavan (2008) salienta que a publicação é tão importante quanto o intenso trabalho de análise do objeto de estudo e exige a redação de um texto primoroso, não apenas uma mera exposição de idéias. Ela permite a presteza na divulgação de resultados, a perpetuação da informação, a ampliação da divulgação e a consulta por reiteradas vezes os mesmos dados inalterados. Ela ocorre uma vez que os textos escritos e publicados veiculam estudos de interesse para a coletividade.

Embora o sistema de comunicação científica sempre tenha ocorrido, foi no século XVII que, devido aos periódicos científicos, se delineou o modelo moderno de comunicação científica (LARA, 2006, p. 395). Pavan (2008, p. 45) também chama a atenção para a evolução dos suportes para a escrita e afirma que estes facilitaram a troca de informações entre os cientistas, além de permitir que “os dados não fossem esquecidos ou perdidos, como ocorria com a transmissão oral.” Nessa afirmação está presente a idéia do uso e reuso da informação, onde outros cientistas apreendem o conhecimento registrado para fomentar novos conhecimentos.

[...] existe um mito no qual o discurso científico se perde na *internet*, tanto no sentido figurativo diante dos discursos possíveis na *internet*, quanto literalmente, pela ausência de políticas de preservação digital – aspectos que implicam na legitimidade e persistência, quesitos fundamentais para a constituição e recuperação da memória científica e básica para estabelecer a comunicação e desenvolvimento científico (WEITZEL, 2006, p. 86).

⁶ GARVEY, W. D.; GRIFFITH, B. C. *Communication*: the essence of Science. Apêndice A, B. In: GARVEY, W. D. **Communication: the essence of Science**. Oxford: Pergamon Press, 1979. p. 299 Disponível em: <<http://global-reach.biz/globstats/evol.html>>. Acesso em: jun.2007.

Portanto, apenas o registro não garante a disseminação e o uso, independente do suporte, a indexação dos documentos é necessária para a recuperação. Trata-se de um “processo intelectual que envolve [...] a compreensão do texto e a composição da representação do documento.” (LIMA, 2003, p. 83) Podem-se identificar três etapas:

(1) análise do documento e estabelecimento do seu assunto; (2) identificação dos principais conceitos do documento; (3) tradução destes conceitos em termos de uma linguagem de indexação. Mediante o uso de uma estrutura da linguagem própria são estabelecidas categorias. (LIMA, 2003, p. 83)

A afirmação acima trata da representação dos assuntos abordados no documento. A organização do conhecimento de qualquer área está extremamente relacionada com a forma de como representá-la e, por sua vez, de como recuperar, portanto, a estruturação do conhecimento facilita a recuperação da informação (LIMA, 2003). Porém, para que exista a recuperação da informação é preciso que as estruturas de conhecimento do autor do texto, do indexador e a do usuário estejam em sintonia, ou seja, aquilo que é escrito (texto) foi representado por termos passíveis de ser capturados pelos motores de busca segundo uma estratégia definida pelo usuário. Como na comunicação científica a troca de informações ocorre entre os membros com interesses afins, é de se esperar que o problema de linguagens para representar o assunto seja minimizado.

No entanto, no contexto deste estudo, ambiente *Web*, são necessários mais elementos do que apenas o assunto para propiciar a recuperação e possível uso, também precisa-se da localização do documento, permissão de acesso, direitos de uso, entre outros. Nesse ambiente, onde a informação é abundante, Valério (2008, p. 161) distingue que “enquanto a comunicação científica é a forma de estabelecer o diálogo com o público da comunidade científica – comunicação entre os pares -, a divulgação científica visa à comunicação para o público diversificado, fora da comunidade científica.” Meadows (1999) afirma que a forma como o cientista divulga informações depende basicamente de três elementos: veículo empregado, natureza das informações e público-alvo.

Outra característica da comunicação científica é a visibilidade, que para Packer e Menegheni (2006, p. 237) “representa a capacidade de exposição que uma fonte ou fluxo de informação possui de, por um lado, influenciar seu público alvo e,

por outro, ser acessada em resposta a uma demanda de informação.” Os autores esclarecem ainda, que ela se aplica aos diferentes componentes e processos que dão forma ao conjunto da comunicação científica como os índices e bases de dados bibliográficas, publicações em geral, autores, instituições de fomento e de pesquisa, entre outros.

Nos modelos emergentes da comunicação científica, segundo Pavan (2008, p. 46) há uma forte presença da tecnologia e “reflete os novos anseios das comunidades científicas, como o livre acesso ao conhecimento científico.” O volume de produção científica disponível torna necessária a seleção do material. Nesse contexto, para embasar novos conhecimentos, uma solução são os índices e os artigos de revisão que, por sua vez, aumentam a visibilidade das pesquisas e seus pesquisadores.

Para Packer e Meneghini (2006, p. 239) “ocorre, portanto em duas dimensões principais: ser referência (de qualidade e credibilidade) no âmbito de uma disciplina ou área temática e ser indexado em índices de prestígio internacional e nacional.” No que tange à primeira, trata-se de sua condição de fonte de informação, enquanto que a segunda, trata-se da sua condição de ser acessado. Um reforço é a afirmação de Ziman (1979, p. 83): “[...] uma grande descoberta científica não passa a existir, apenas, por força da autoridade moral ou do talento literário do seu criador, e sim pelo seu reconhecimento e sua apropriação por toda a comunidade científica”. Disto decorre a importância do uso, apropriação e geração de novos conhecimentos para dar início a um novo ciclo do fluxo.

Nesse contexto de mudanças tecnológicas e no sentido da visibilidade das pesquisas surgiram os Repositórios Institucionais (RI). Estes

[...] abrigam a produção intelectual de dada instituição, oportunizando a recuperação de uma informação específica, abrindo espaço para a divulgação de obras consideradas - até a chegada e permanência do suporte eletrônico se propagar, como literatura cinzenta (DZIEKANIAK, 2010, p. 54).

Portanto, os Repositórios Institucionais dizem respeito a instituições que tem o intuito de divulgar sua produção agilmente, como afirma Sarmiento e Souza⁷ (2004, p. 143 *apud* DZIEKANIAK, 2010, p. 52)

tornar o trabalho disponível de forma rápida, favorecendo o acesso democrático e gratuito das publicações eletrônicas, e enfraquecendo, com isso, o monopólio das grandes editoras científicas que até recentemente detinham em seu poder os direitos de publicação.

Sayão e Marcondes (2008) também atentam para o crescente custo dos periódicos comerciais e para as práticas restritivas de publicação. Nesse ambiente, as ferramentas da Tecnologia da Informação vão produzindo mudanças ao longo do processo de comunicação científica e trazendo agora à discussão as publicações eletrônicas e as editoras científicas. Neste ponto, abriram-se duas alternativas: as assinaturas de publicações científicas com alto valor agregado e direcionadas a bibliotecas e instituições de pesquisa; e o pensamento de um acesso livre aos documentos na *Web*.

O Acesso Livre é a

disponibilização livre na *Internet* de literatura de caráter acadêmico ou científico (em particular os artigos de revistas científicas), permitindo a qualquer utilizador ler, baixar arquivos, copiar, distribuir, imprimir, pesquisar, indexar, fazer links ou referenciar o texto integral dos documentos. (LARA, 2006, p. 389).

Na literatura é encontrada indistintamente a expressão Acesso Livre ou Acesso Aberto. Kuramoto (2010, *on-line*) faz um levantamento da utilização de ambas e conclui que

a palavra aberta em tradução a *open* não significa que aquilo que se acessa possa ser reutilizado e distribuído. Enquanto que a palavra *livre* exprime a liberdade e, que a pesar de uma informação não ser de sua propriedade, ela pode ser reutilizada e distribuída, que é o espírito do *Open Access*.

⁷ SARMENTO Ee SOUZA, M. F. Mudanças no processo de comunicação científica: a alternativa dos repositórios institucionais. inIn: VIDOTTI, S. A. B. G. (Org.) **Tecnologia e Conteúdos Informativos**. São Paulo: Polis, 2004.

O Acesso Livre iniciou com o repositório de *arxiv* que, em 1991, foi criado para armazenar os artigos de física que ainda não haviam sido avaliados por seus pares e, portanto, não publicados. No Brasil, em 2005 foi elaborada a Declaração de Apoio ao Acesso Aberto à Literatura Científica – Carta de São Paulo⁸, onde o grupo de professores, pesquisadores, bibliotecários, alunos, cidadãos e representantes de organizações da sociedade civil manifestaram seu apoio ao acesso aberto da literatura científica. Nesse documento abordam os direitos autorais, reconhecem a contribuição do acesso livre para a qualidade da pesquisa e promoção do conhecimento e estabelecem uma série de recomendações para assegurar o acesso à literatura e documentação científica. Desde 2007 tramita na Câmara dos Deputados o Projeto de Lei de nº 1.120/2007⁹ que obriga as instituições públicas de ensino superior a construir repositórios institucionais para o depósito da produção técnico-científica do seu corpo docente e discente, que tem autoria de Rodrigo Rollemberg. De acordo com o artigo 1 § 1º os repositórios deverão ser compatíveis com padrões de interoperabilidade internacionalmente adotados.

Então, o movimento do Acesso Livre (AL) ou *Open Access* (AO) surgiu num contexto em que, ao mesmo tempo, ocorria o desenvolvimento das Tecnologias de Informação e o aumento nos custos de produção de periódicos impressos. O AL envolve, além de questões políticas ou de normalização, questões relacionadas a cultura e costumes da comunidade científica onde, novamente, a questão tecnológica se faz presente. O Acesso Livre propiciou o fortalecimento dos portais institucionais e dos repositórios. Desta forma, as universidades adotaram os repositórios para dispor sua produção e como meio de comunicação científica ao disponibilizar suas publicações. Estas publicações, disponíveis para a comunidade acadêmica, tornaram-se mais visíveis e converteram-se em uma questão fundamental para o desenvolvimento e a maturidade da pesquisa científica no Brasil (MARCONDES; SAYÃO, 2001, p. 25)

Apesar das mudanças, a prática da revisão por pares permaneceu, embora agora também conte com o auto-arquivamento, que é quando o próprio autor publica seu trabalho. Mas o Acesso Livre pode ir além da comunicação científica, é uma ferramenta que permite, inclusive, a divulgação científica. Com o autor dispondo

⁸ Documento disponível em: <http://www.acessoaberto.org/carta_de_sao_paulo_acesso_aberto.htm>

⁹ Documento disponível em:
<<http://www.camara.gov.br/proposicoesWeb/fichadetramitacao?idProposicao=352237>>

suas publicações diretamente na base de dados, além da divulgação das pesquisas, há uma economia de tempo, esforço e custo. Pela auto-publicação, o autor também passa a cumprir uma das funções do editor. Mais ainda, há repositórios que incorporam em seus acervos textos recomendados pelos leitores, nesse caso, também estes desempenham, em parte, o papel de editores.

Com as instituições preservando sua produção intelectual pelo RI aliado ao reconhecimento de que a informação precisa estar acessível a quem dela recorra, começou a ser freqüente também a troca de informações entre os RIs, por via eletrônica, a respeito dos dados do documento (DZIEKANIAK, 2010, p. 54).

A partir daí surgiu a necessidade da interoperabilidade entre os repositórios, onde O Dicionário de Ciência da Informação ODLIS (REITZ, 2010) define interoperabilidade como a capacidade de um hardware ou software comunicar-se e trabalhar com outro sistema de troca de dados, geralmente entre sistemas diferentes, concebidos e fabricados por fornecedores diferentes.

A interoperabilidade, que foi reconhecida pelo movimento do Acesso Livre, ocorre mediante a troca de metadados que representam as informações contidas ou sobre os documentos disponibilizados. Onde, os metadados existem para permitir interoperabilidade e precisam ser de fácil compreensão, precisos, completos, entre outros.

Em 1999 foi criado o *Open Archive Initiative (OAI)* para promover e desenvolver a interatividade e também foi responsável pela criação do *Open Archives Initiative Protocol for Metadata Harvesting (OAI – PMH)*. Esse protocolo foi desenvolvido com o propósito de distribuir e coletar metadados, com foco em descritores de documentos.

A demanda seguinte foi o desenvolvimento de padrões de metadados e protocolos para o recolhimento dos mesmos. Nesse sentido, os padrões mais utilizados para troca de dados e metadados na *Web* são o HTML, o SGML e o XML. Pois, não somente através do uso de metadados se garante a comunicação dos dados, é necessário também o uso de ontologias, de tecnologias, como o RDF, e de linguagens de marcação, como o XML, que disponibilizam modelos e sintaxe capazes de permitir a interoperabilidade de informações.

A insuficiência de padrões revela uma falta de clareza quanto às estratégias dos RI e das bases de dados em geral, que embora pretendam tornar acessíveis seus recursos, acabam por dificultar o seu acesso. Os principais problemas que

atrapalham a integração dos sistemas são: o uso de uma grande diversidade de esquemas de metadados; mesmos recursos armazenados em diversos locais; duplicação de esforços para gestão de documentos; identificação e validação dos dados na origem.

4.2 REVISTAS CIENTÍFICAS ELETRÔNICAS

O surgimento das revistas científicas normalmente está ligado a grupos de pesquisadores, que remonta ao Século XVII com a criação das sociedades e academias (GONÇALVES; RAMOS; CASTRO, 2006). No Brasil isto também ocorreu, onde elas surgiram e se fortaleceram junto aos programas de pós-graduação e iniciativas de qualificação acadêmica. Pois elas “representam, de alguma forma, um processo de finalização da pesquisa. É quando ela se converte em produto científico, indicando que o processo chegou a um grau de maturidade que permite expressá-lo” (STUMPF, 2003).

Meadows (1999) afirma que o periódico científico se constitui no canal de comunicação formal da ciência no qual os cientistas divulgam suas descobertas e promovem sua imagem. No mesmo sentido, Ziman (1979) leciona que as revistas científicas também permitem a ascensão do cientista trazendo promoção, reconhecimento e poder no seu meio. Assim, o principal motivo para a existência de um periódico científico foi a necessidade de intercâmbio rápido de informações de maneira formal e ampla.

As revistas científicas são uma manifestação da “comunicação formal, organizadas em fascículos ou números, publicados segundo uma periodicidade definida, por tempo indeterminado, com tiragem e disseminação amplas, contendo predominantemente artigos científicos” (LARA, 2006. p. 405). Para sua publicação, os artigos precisam passar pelo crivo dos seus pares (*peer review*), onde se verificam o mérito e o método científico (GONÇALVES; RAMOS; CASTRO, 2006). Se o periódico for categorizado como de divulgação científica, ele é destinado “ao público leigo ou iniciante [e] é de caráter amplo, abrange em geral várias áreas temáticas e é escrito em linguagem mais acessível” (LARA, 2006, p. 406)

Desde a sua origem, a finalidade das revistas científicas vai além da divulgação das pesquisas, pois elas também têm como função a comunicação formal dos resultados da pesquisa original; a preservação do conhecimento

registrado; estabelecimento da propriedade intelectual; e manutenção do padrão da qualidade na ciência (MUELLER, 2000, p. 75). Pela importância das revistas para a comunicação científica formal, elas têm despertado interesse para estudo sobre várias abordagens, como quanto à produção, análise e avaliação, além dos estudos sobre a evolução do pensamento científico e método das pesquisas (BUFREM, 2006).

Apesar de aceitos em formato impresso, com o desenvolvimento tecnológico, começaram a perceber-se como restrições a demora na publicação, os custos de publicação e aquisição, assim como a disponibilidade restrita pelo suporte físico dos exemplares na biblioteca. O periódico eletrônico é uma alternativa.

Do inúmero de informações que circulam na rede, existe um tipo de informação que está relacionada à produção acadêmica. Especificando ainda mais, uma parte da produção acadêmica se refere à comunicação científica e é dentro desta categoria que se encontram as revistas eletrônicas relacionadas a instituições de ensino e pesquisa. As revistas eletrônicas tem a particularidade de disponibilizar suas informações durante um tempo ilimitado abrangendo um público maior. Mais informações disponíveis resulta em mais material para pesquisa e daí, maior produção dos pesquisadores.

Como vantagens das revistas eletrônicas, Mueller (2000, p. 81) diz que o meio eletrônico é uma alternativa que satisfaz na medida que “oferece mais rapidez na comunicação e flexibilidade de acesso, tem largo alcance e baixo custo relativo, disponibilidade imediata, é capaz de diminuir a capacidade de manutenção de coleções, barateando os custos.” A autora acrescenta que as publicações eletrônicas são aquelas que se acessam por meio de recursos eletrônicos.

Os mesmos cuidados que os periódicos impressos são mantidos nos eletrônicos visando a qualidade dos mesmos, quais sejam: a periodicidade e pontualidade das edições; a duração temporal da revista que indica a tradição e o êxito da publicação; a adequação às normas; o trabalho editorial; a difusão e indexação em bases que propiciam visibilidade; grau de endogenia; indicadores bibliométricos (GONÇALVES; RAMOS; CASTRO, 2006). Com esta preocupação a CAPES possui uma política nesse sentido que incluem a recomendação de práticas e normas internacionais, além de qualificar mediante estratos as publicações nacionais.

O Sistema Eletrônico de Editoração de Revistas (SEER)¹⁰ é a ferramenta recomendada pela CAPES para o processo editorial, permitindo uma melhoria na qualidade dos periódicos e uma maior rapidez no fluxo das informações. O SEER visa instrumentar para a elaboração e gestão de uma revista eletrônica.

O estudo sobre o uso do SEER de Ferreira e Caregnato (2008, p. 179) constatou que este instrumento

facilita o processo editorial, rico em detalhes específicos em cada uma de suas etapas. Poupar tempo, minimizar custos e ter o controle do envio de artigos para os avaliadores, entre os diversos recursos oferecidos pelo SEER/OJS, contribui para que o editor possa dedicar-se a outras etapas do processo editorial.

Como resultado do cumprimento das exigências que asseguram a qualidade, a visibilidade é uma delas. Para Lara (2006) é a “capacidade de exposição de uma fonte de informação de influenciar seu público alvo e de ser acessada para responder a uma demanda.” Ela está relacionada com a qualidade e credibilidade do periódico e depende da indexação em índices nacionais e internacionais. Porque para que as informações possam se suceder e que o fluxo da informação não seja interrompido, o conteúdo das revistas precisará estar acessível ao maior número de interessados.

O entendimento do processo de publicação pode ser ilustrado pelo fluxo editorial que, geralmente, é composto pelas etapas de “recebimento de trabalhos para publicação, pré-seleção dos trabalhos, seleção de revisores, encaminhamento e acompanhamento do processo da revisão por pares, contato com o(s) autor(es) sobre comentários dos revisores, aprovação ou rejeição para publicação, revisão de texto, revisão gráfica e publicação “ (GONÇALVES; RAMOS; CASTRO, 2006, p. 172). O Corpo Editorial de uma revista é composto pelo “grupo de pesquisadores e membros da comunidade acadêmica que auxiliam o editor na gestão do fluxo editorial do periódico científico”, já o editor é “a pessoa ou instituição responsável pela gestão e definição da linha editorial de periódicos científicos e pela aprovação final dos trabalhos a serem publicados” (LARA, 2006, p. 396). Assim, ao editor cabe garantir a integridade científica do periódico; a manutenção do padrão de qualidade

¹⁰ SEER. Disponível em: <<http://seer.ibict.br/>>

e periodicidade; o grau de endogenia; adequação às normas; a indexação da revista e a estrutura dos bancos de dados, entre outros.

Dziekaniak (2010, p. 44) alerta para as transformações da ciência da Informação devido a

[...] duas revoluções: a **documentária**, processo pelo qual passou e passa a comunidade científica mundial, através da expressividade crescente da publicação de documentos, principalmente através dos periódicos e, principalmente a revolução **tecnológica**, focalizada no âmbito das tecnologias da informação e comunicação. [grifo nosso]

Lindoso (2011) postou um comentário no *blog* “*E Book Reader: o livro na era digital*”, falando que há dez anos houveram iniciativas no Brasil de tirar maior proveito dos metadados dos produtos editoriais e que depois não houve continuidade. Ele diz:

Em resumo, no começo da década passada já se esboçava uma visão de futuro sobre a importância da coleta de informação, e da padronização do tipo de informações. O esforço foi jogado fora e interrompido. [...] Mas o fato é que o mercado editorial brasileiro, nessa altura dos acontecimentos, tem que correr atrás do prejuízo.

Essa afirmação justifica estudos como esta pesquisa no sentido de não somente propiciar resultados positivos às demandas dos usuários através da qualidade das informações obtidas pelos metadados, mas também para o gerenciamento das publicações e dos negócios a elas relacionados. As publicações eletrônicas/objetos no ambiente digital são identificadas por um conjunto de elementos que as caracterizam e localizam. A normalização desses conjuntos, chamados formatos, propiciam as condições necessárias a sua recuperação e intercâmbio via rede (ROSETTO; NOGUEIRA; [s.d.], p. 5).

Assim, as revistas científicas de acesso livre são aquelas em que os usuários podem acessar, baixar, copiar, imprimir e distribuir seu conteúdo. No entanto, isso somente ocorrerá com a interoperabilidade dos repositórios onde as informações e conteúdos das revistas estejam armazenados. Quanto aos metadados que são efetivamente os responsáveis pela interoperabilidade, Rosetto e Nogueira ([sd], p. 6) salientam que possuem

diferentes níveis de especificidade, estrutura e complexidade, e seu propósito primário é: descrever, identificar, e definir um recurso eletrônico com o objetivo de modelar e filtrar o acesso, termos e condições para o uso, autenticação e avaliação, preservação, e interoperabilidade.

Deste modo, os editores precisam estar atentos aos benefícios das tecnologias de comunicação e informação e ao que os metadados das publicações podem trazer para o gerenciamento dos seus negócios, ou seja, precisa estabelecer políticas de metadados e, principalmente segui-las.

4.3 METADADOS

Os metadados são “dados sobre dados” como normalmente é encontrado na literatura. Vários autores identificam esta definição como pouco profunda, sem muitas informações e aberta a interpretações diversas. Rocha (2004, p. 113) complementa o conceito afirmando que

Quando se trata do mundo digital, chama-se de recursos o objeto descrito por metadados, pois este pode ser tanto um simples dado, quanto um documento, uma página da Web, ou até mesmo uma pessoa, uma coleção, um sistema, um equipamento, ou uma organização.

Portanto, um recurso pode ser qualquer objeto acessível através do seu endereço eletrônico e passível de ser descrito por um metadado. Ainda Ikematu (2001, p. 1) levantou algumas dessas interpretações, apresentadas a seguir:

- Metadados são dados que descrevem atributos de um recurso. Ele suporta um número de funções: localização, descoberta, documentação, avaliação, seleção, etc.
- Metadados fornecem o contexto para entender os dados através do tempo.
- Metadados é dado associado com objetos que ajuda seus usuários potenciais a ter vantagem completa do conhecimento da sua existência ou características.
- Metadados é o instrumental para transformar dados brutos em conhecimento.

Os conceitos resultam dos diferentes níveis de entendimento dos dados, onde percebe-se uma evolução dos dados como representações de um objeto para dados

como instrumentos para propiciar conhecimento. Ikematu (2001) propõe a análise dos metadados em função das necessidades informacionais e dos recursos a gerenciar. Ele salienta que os metadados sempre têm sido relacionados com o sistema de processamento da informação e que, ao longo do tempo, tem mudado seu escopo em decorrência da evolução das Tecnologias da Informação (TI).

Mediante os metadados são descritos os documentos, tanto em sua forma física quanto ao seu conteúdo, à temática.

[os] metadados não se restringem àqueles usados para descrever recursos para permitir sua descoberta e localização [e uso]. Por exemplo, metadados também descrevem os formatos dos recursos, a fim de permitir que computadores identifiquem quais aplicativos podem ser usados para manipular estes recursos. Aplicativos, ao analisarem metadados, verificam se o formato do recurso é compatível, as condições de uso do recurso, restrições de acesso etc. (ROCHA, 2004, p. 113)

De forma análoga a um catálogo de biblioteca que fornece informação para a localização de um livro, um metadado é utilizado para localizar uma informação, um registro (ROSETTO; NOGUEIRA, [s.d.], p. 4). O valor dos metadados está relacionado com a informação que se acessa a partir deles, eles não podem ser avaliados apenas como um recurso técnico. Para isso eles precisam ser gerenciados, com políticas que definam estratégias de metadados e seus procedimentos para ingresso, manutenção e uso. “A catalogação dos dados propiciará a maior utilização deles por usuários com múltiplos interesses. Sem uma documentação eficiente dos dados é dificultada aos usuários a localização de dados necessários para suas aplicações” (IKEMATU, 2001, p. 4). O autor ainda acrescenta que a falta de uma documentação dos dados leva, com o tempo, a duplicar esforços na coleta e manutenção dos registros. Isto pode resultar em inconsistências, perdas e, principalmente, a um aumento nos custos pela falta de uso ou uso inadequado da informação.

Os metadados por si só já trazem benefícios para a informação que circula na *Web*, Senso e Piñero (2003) levantaram uma série de características que ilustram isto. São elas:

- a) incremento da acessibilidade pela correta descrição dos recursos;

- b) redução do tráfego de informação na rede pelo trânsito dos metadados e não dos recursos, o objeto em si;
- c) expansão do uso da informação pela difusão de várias cópias de um único objeto;
- d) difusão de um mesmo recurso por meio de versões com diferentes graus de detalhamento em função do público alvo desejado;
- e) divulgação de aspectos legais que estipulem claramente as condições de uso e reuso do documento; e
- f) preservação do recurso ou objeto original.

Senso e Piñero (2003) trazem as palavras de Duval que diz que os metadados são parte da infraestrutura da informação necessária para auxiliar a criar ordem no caos da *Web*, proporcionando descrição, classificação e organização. (CHEN; CHANG¹¹, 1998 *apud* SENSO; PIÑERO, 2003, p. 105) Para os pesquisadores, principalmente, são um instrumento de suma importância por auxiliar na localização de fontes de informação (ROSETTO; NOGUEIRA; [s.d.], p. 6).

A informação organizada traz benefícios para o desenvolvimento das diversas áreas do conhecimento, pois a informação pode ser compartilhada infinitamente, e o seu valor cresce com o uso e, também como é conhecido, representa poder, prestígio. Na comunicação científica e, de forma geral, uma informação comparada e combinada com outras tem o seu valor aumentado.

Quando percebida a importância do acesso à informação, o foco passa a ser a qualidade dos metadados que a representam. Um ciclo virtuoso só ocorrerá se ela for recuperada e com qualidade, pois uma entrada de dados ineficaz implicará uma saída de dados também ineficaz, portanto, fornecerá informação incorreta ou nula. As falhas decorrem, fundamentalmente, das discrepâncias observadas na classificação dos recursos e nas linguagens empregadas na representação.

Em 2001 Ikematu (2001, p. 6) apontava como problemas ao trabalhar com metadados.

- a) Metadados tomam uma variedade de formas, especializadas ou gerais;
- b) Novos conjuntos de metadados irão ser criados à medida que a informação da rede fique mais madura;

¹¹ CHEN, H. H.; CHANG, Y. S. The role of metadata in national Taiwan university digital library / museum project. **Journal of Library and Information Science**, v. 23, n. 2, p. 51-65, 1998

- c) Diferentes comunidades irão propor, projetar e ser responsáveis por diferentes tipos de metadados;
- d) Existem muitos usuários de metadados;
- e) Adoção de diferentes vocabulários de metadados significa aumento de buscas usando vocabulários de metadados que não são familiares.

Com o passar do tempo, esses aspectos tornaram-se mais evidentes e várias ações foram levadas a efeito de superar as dificuldades, onde comunidades especializadas em diversas áreas do conhecimento passaram a criar padrões de Metadados.

Na Biblioteconomia, o padrão MARC (www.loc.gov/marc) é um exemplo de esquema de metadados. Estes metadados indicam propriedades do documento, como seu autor, sua data de publicação, seu título, seu assunto, e têm como finalidade permitir a descoberta e a localização destes documentos (ROCHA, 2004, p. 114).

Sené (2004) observa ainda que a catalogação descritiva é um tipo de metadado e que também podem ser vistos como tais os catálogos de bibliotecas, os serviços de indexação e resumo, e as diferentes ferramentas para a busca em arquivos que possuam metadados. Junto a estes metadados percebidos ou não como tais, surgiram propostas de criação de padrões de novos metadados.

Isso ficou evidente na *International Conference on Dublin Core and Metadata Applications* de 2001 onde estavam presentes iniciativas para criação de esquemas padronizados de metadados para diferentes áreas como ambiental, governo, educação, biblioteconomia, agricultura, matemática, etc. Sendo assim, pode-se estar frente a outro problema, que é metadados diferentes para os mesmos propósitos. “Esforços para evitar este surgimento caótico dos esquemas de metadados são os Registros de Metadados, os Perfis de Aplicação e a iniciativa Dublin Core” (ROCHA, 2004, p. 115).

4.4 DUBLIN CORE METADATA INICIATIVE (DCMI)

Os padrões de metadados servem para a representação dos documentos. O DCMI é um padrão de metadados que possibilita a catalogação de documentos eletrônicos na *Web*. Segundo Rocha (2004, p. 119)

para evitar a sobreposição de esquemas, os registros de esquemas apresentam-se como componentes de disseminação de esquemas já definidos, e Dublin Core apresenta-se como língua mãe para metadados, que pode ser estendida, através da criação de Perfis de Aplicação, para contemplar necessidades de comunidades específicas.

Sendo assim, a iniciativa *Dublin Core Metadata Initiative* (DCMI) tem sido amplamente adotada e o seu lema é “tornando mais fácil encontrar a informação” (DCMI, *on-line*), e a sua intenção é a de coexistir com outras normas de metadados com semânticas diferentes.

O DCMI é uma organização dedicada a promover e difundir a adoção da norma de interoperabilidade dos padrões de metadados e, ainda, o desenvolvimento de vocabulários especializados dos metadados para descrever recursos que permitam o uso de sistemas mais inteligentes (ALVES, 2009, p. 50).

A norma ISO 15.836 aborda os 15 elementos básicos do *Dublin Core Metadata Initiative*, são eles: Título, Criador, Assunto Descrição, Publicador, Colaborador, Data, Tipo, Formato, Identificador, Fonte, Idioma, Relação, Cobertura, Direitos Autorais. Esses 15 elementos, ainda podem ser ampliados pelo uso de elementos de refinamento, que compõem os qualificadores *Dublin Core*. A adoção dos qualificadores permite representar a informação com maior precisão. Na página oficial da DCMI na *internet*¹² os elementos são definidos, fornecidos exemplos e expostas as Melhores Práticas para o seu uso.

O documento *Dublin Core Metadata Element Set*¹³ que apresenta os 15 elementos do DCMI salienta que cada elemento refere-se ao recurso que está sendo descrito, onde recurso é um serviço ou recurso de informação, embora este conceito possa ser ampliado. Também observa que o nome do elemento não pretende defini-lo, é um código único que serve apenas para identificá-lo. Ele é sempre utilizado em língua inglesa, independente da utilizada no seu contexto para permitir a correta interpretação pelos sistemas e programas que interpretam metadados. No documento, cada elemento está descrito quanto à sua forma conceitual e representação. A Definição refere-se ao conceito semântico e Tipo de Dados e Comentário dizem respeito à representação dos dados, à sintaxe.

¹² *Dublin Core Metadata Initiative* <<http://dublincore.org/>>

¹³ *Dublin Core Metadata Element Set* <<http://dublincore.org/documents/2010/10/11/dces/>>

O refinamento dos elementos consta nas orientações sobre o uso no documento *Using Dublin Core – Dublin Core Qualifiers*¹⁴, é o resultado do processo de uso e busca de melhores práticas e não pretende ser exaustivo. A qualificação pode ser pelo Refinamento do Elemento ou pelo Esquema de Codificação. No documento, salienta que o refinamento procura especificar o elemento a fim de reduzir seu escopo. O sistema está projetado para que, se o usuário não reconhecer o elemento refinado, possa ignorar a qualificação e utilizar o elemento como se esta não existisse, seguindo o princípio *Dumb-Down*. Ex. *date.available* e *date.created* são refinamento do elemento *date*. O Esquema de Codificação auxilia a interpretar o valor do elemento e segue instruções de vocabulários controlados, notações formais e regras de análise conhecidas. Além do refinamento para os 15 elementos, também estão relacionados a Procedência, Audiência e Detentor dos Direitos.

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, 2009, p. 50).

As Melhores Práticas consistem em uma relação de instruções que são elaboradas a fim de orientar na implantação e uso do padrão DCMI. Para o caso dos 15 elementos do DCMI elas foram elaboradas em termos gerais e específicos. A seguir, passa-se a apresentar as definições, observações sobre sintaxe, exemplos e Melhores Práticas, com tradução da autora. Inicia-se pelos aspectos gerais.

Cada um dos 15 elementos utilizados precisa ter seus valores conforme a definição e instruções (demais recomendações) desse elemento no documento que define os Termos de Metadados do *Dublin Core Metadata Initiative*¹⁵.

Se algum dos elementos requer um refinamento, este deve ser utilizado em detrimento do termo mais genérico.

Quando necessário, optar pelo atributo *xsi:tipo* para especificar o esquema de codificação a fim de expressar o valor do registro com maior precisão.

¹⁴ *Using Dublin Core – Dublin Core Qualifiers*

<<http://dublincore.org/documents/2005/11/07/usageguide/qualifiers.shtml>>

¹⁵ DCMI Metadata Terms <<http://dublincore.org/documents/2010/10/11/dcmi-terms/>>

Quando necessário, utilizar o atributo *xml: lang* conforme valores do *Open Language Archives Community (OLAC) language*¹⁶ para identificar o idioma em que está escrito o conteúdo do elemento.

No caso de um elemento possuir mais de um valor, a recomendação é o uso de refinamentos do elemento onde serão apresentados todos os valores de forma qualificada, em detrimento de usar mais de um valor para um mesmo elemento.

Prossegue-se com a apresentação dos 15 elementos, os exemplos foram retirados dos metadados desta pesquisa.

4.4.1 Título (*title*)

Definição: nome dado ao recurso.

Comentário: é o nome pelo qual o recurso é formalmente reconhecido.

Refinamento: título:alternativo (*title:alternative*). Pode tratar-se de um título substituto, alternativo, abreviado ou traduzido; só utilizar quando este estiver expresso junto ao título do documento.

Melhores práticas recomendam:

- que cada registro deve possuir um elemento título;
- no caso em que o recurso não tenha um título formal, o catalogador deve proporcionar um título descritivo e colocá-lo entre colchetes;
- que exista um único elemento identificado como Título, que será o título original, com exceção dos títulos paralelos em obras de edição bilingue, todos os demais títulos deverão ser especificados no refinamento do elemento como título alternativo.

Ex: *Mapas imaginários sobre Porto Alegre: a cidade midiática.*

4.4.2 Criador (*creator*)

Definição: Entidade responsável em primeira instância pela existência do recurso.

Comentário: Em geral, tratar-se de uma pessoa, uma organização ou um serviço. Normalmente o nome de um Criador é utilizado para indicar a entidade.

¹⁶ OLAC <<http://www.language-archives.org/>>

Melhores práticas recomendam:

- usar Colaborador no lugar de Criador, com exceção no caso em que há uma nítida e importante participação de uma pessoa ou organização e não exista um termo adequado de refinamento no OLAC *Role Vocabulary*¹⁷ para utilizar com Colaborador.

Ex: Morigi, Valdir José

4.4.3 Colaborador (*contributor*)

Definição: entidade responsável por contribuir com o conteúdo do recurso.

Comentário: pode tratar-se de uma pessoa, uma organização ou um serviço.

Em geral, o nome do Colaborador pode ser utilizado para identificar a entidade.

Melhores práticas recomendam:

- Identificar o colaborador por um nome que permita o ordenamento alfabético;
- preferir colaborador a criador;
- utilizar um valor do OLAC *Role Vocabulary*¹⁸ para expressar o papel do Colaborador. Nos valores relacionados constam autor, ilustrador, consultor, entrevistador, editor e outros. O editor é definido como aquele que revisa, compila, corrige e/ou testa o recurso.

Ex: com uso de refinamento patrocinador (*contributor:sponsor*) Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

4.4.4 Assunto (*subject*)

Definição: tópico do conteúdo do recurso.

Comentário: tipicamente, um assunto deverá ser expresso por palavras-chave, frases, ou códigos de classificação que descrevam o conteúdo do recurso.

Esquema de Codificação: Cabeçalho de Assunto da Biblioteca do Congresso Norte Americano (*Library Congress Subject Headings – LCSH*); Cabeçalho de Assuntos Médicos (*Medical Subject Headings – MeSH*); Classificação Decimal de

¹⁷ OLAC Role Vocabulary <<http://www.language-archives.org/REC/role-20060406.html>>

¹⁸ OLAC Role Vocabulary <<http://www.language-archives.org/REC/role-20060406.html>>

Dewey (*Dewey Decimal Classification – DDC*); Classificação da Biblioteca do Congresso Norte Americano (*Library Congress Classification – LCC*); Classificação Decimal Universal (*Universal Decimal Classification – UDC*).

Melhores práticas recomendam:

- utilizar pelo menos um Assunto, Cobertura ou Descrição, ou um dos seus refinamentos, com o intuito de fornecer ao usuário uma idéia prospectiva sobre o conteúdo do recurso, além daquela oriunda apenas do título; o uso de todos esses elementos é incentivado;
- utilizar vocabulário controlado ou sistemas de classificação formais;
- identificar precisamente, de preferência baseado no código OLAC *Language*¹⁹, o idioma abordado, quando se trata de um assunto sobre linguagem humana;
- identificar os subcampos utilizando o esquema OLAC *Linguistic Field Vocabulary*²⁰ quando o assunto for referente ao campo da linguística;
- utilizar o elemento Cobertura para descrever a localização temporal e espacial do recurso.

Ex: Biblioteca Escolar

4.4.5 Descrição (*description*)

Definição: descrição do conteúdo do recurso.

Comentário: deve incluir, embora não restrito a isto, um resumo, um sumário, uma indicação de representação gráfica ou um texto onde relate pormenorizadamente o recurso.

Refinamento: descrição: lista de conteúdo, sumário (*description:tableofcontents*) que apresenta as subunidades do recurso em forma de lista e não como um texto descritivo; descrição:resumo (*description:abstract*) é o próprio resumo do documento, quando não consta no documento, descrever o conteúdo do documento com mesmo vocabulário a fim de evidenciar a profundidade da abordagem do assunto.

Melhores práticas recomendam:

¹⁹ OLAC Language Extensión <<http://www.language-archives.org/REC/language-20080222.html>>

²⁰ OLAC Linguistic Field Vocabulary <<http://www.language-archives.org/REC/field-20021127.html>>

- utilizar pelo menos uma Descrição, Cobertura, Assunto, ou um dos seus refinamentos, para permitir elaborar uma idéia sobre o conteúdo do recurso, indo além daquela oferecida apenas pelo título. O uso de todos esses elementos é estimulado.

Ex: *Editorial da Edição 2009, vol.15, n.2*. Este texto foi elaborado pelo indexador já que o recurso não possui resumo.

4.4.6 Cobertura (*coverage*)

Definição: identificação espacial ou temporal do recurso.

Comentário: A cobertura espacial é referente ao recurso diretamente ou à abrangência na qual ele é relevante. A identificação pode ser através do nome de um lugar ou especificação do local por suas coordenadas geográficas. A identificação temporal pode ser a denominação de uma época, data ou período. A abrangência pode ser identificada pelo nome da unidade administrativa ou geográfica do local onde o recurso se aplica.

Refinamento: espacial (*coverage:spacial*) e temporal (*coverage:temporal*)

Esquema de Codificação: segue instruções de pontos identificados pelas coordenadas geográficas pelo DCMI *Point*²¹, o DCMI *Box*²² para espaços limitados, regiões e ISO 3.166 e *Getty Thesaurus of Geographic Names TGN*²³, ara denominação de países e suas subdivisões. Para o refinamento temporal, a sintaxe segue o DCMI *Period Encoding Scheme*²⁴ onde é indicado o início e o fim do período como no elemento data (*date*).

Melhores práticas recomendam:

- que cada registro possua pelo menos um metadado de Cobertura, Descrição ou Assunto, ou um dos seus refinamentos, de forma a propiciar uma idéia melhor sobre o recurso do que apenas o potencial informativo do título. A utilização de todos aqueles metadados é aconselhada;

²¹ DCMI Point Encoding Scheme <<http://dublincore.org/documents/2006/04/10/dcmi-point/>>

²² DCMI Box Encoding Scheme <<http://dublincore.org/documents/2006/04/10/dcmi-box/>>

²³ TGN The Getty Research Institute

<<http://www.getty.edu/research/tools/vocabularies/tgn/index.html>>

²⁴ DCMI Period Encoding Scheme <<http://dublincore.org/documents/2006/04/10/dcmi-period/>>

- para dados sobre cobertura espacial, utilizar valores codificados que proporcionem maior precisão do recurso;
- utilizar vocabulário controlado como o tesauro *Getty Thesaurus of Geographic Names* [TGN];
- preferir nomes de lugares e períodos de tempo em vez de dados numéricos como coordenadas geográficas ou intervalos de datas.

Ex: *Século XX*

4.4.7 Fonte (*source*)

Definição: recurso relacionado do qual o presente recurso deriva.

Comentário: O recurso em questão pode derivar do recurso relacionado (fonte) tanto seja no todo ou em partes.

Melhores práticas recomendam:

- identificar o recurso relacionado seguindo um sistema de identificação formal;
- identificar pela especificação dos termos do DCMI *dcterms:URI*²⁵ para valores que possuam endereço digital fixo, um *Uniform Resource Locator* (URL);
- usar valores não literais, conforme definido em DCMI *Abstract Model*²⁶; que representem uma entidade física, digital ou conceitual;
- identificar tal como utilizado no elemento Relação.

Ex: Para o segundo número do volume 16 da Revista Em Questão de 2010:
<http://seer.ufrgs.br/EmQuestao/issue/view/1022>

4.4.8 Identificador (*identifier*)

Definição: referência que evite ambiguidades na identificação do recurso em um dado contexto.

Refinamento: identificador:citação bibliográfica (*identifier:bibliographiccitation*) se refere à forma como o recurso poderá ser citado; o detalhamento consta nas instruções para a codificação de citações bibliográficas²⁷.

²⁵ DCMI Metadata Terms <<http://dublincore.org/documents/dcmi-terms/>>

²⁶ DCMI Abstract Model <<http://dublincore.org/documents/abstract-model/>>

Melhores práticas recomendam:

- que se identifique o recurso de forma a ser coerente com um sistema de identificação formal;
- registrar conforme está na especificação dos termos do DCMI *dcterms:URI*²⁸ para valores que possuam endereço digital fixo, um *Uniform Resource Locator* (URL); outras opções são o *Digital Object Identifier* (DOI) e *International Standard Book Number* (ISBN).

Ex: <http://seer.ufrgs.br/EmQuestao/article/view/100>. No refinamento de citação bibliográfica: Em Questão. Porto Alegre, v. 15, n. 2, p. 349-357, jul./dez. 2004

4.4.9 Relação (*relation*)

Definição: recurso relacionado.

Comentário: identificar os demais recursos relacionados por meio de um sistema formal de identificação.

Refinamento: a relação pode ser por: é versão de (*relation:isversionof*) e tem versão (*relation:hasversion*) quando há mudanças substativas no conteúdo ou no formato; é substituída por (*relation:isreplacedby*) e substitui a (*relation:replaces*) são informações importantes para direcionar o usuário ao item vigente; é requerido por (*relation:isrequiresby*) e requer (*relation:requires*) pode referir-se a elementos físicos ou lógicos, pode ser um tipo de documento e um *software*, ou um *software* e um equipamento; é parte de (*relation:ispartof*) ou tem parte em (*relation:haspart*) evidencia relações hierárquicas tanto do tipo um para um como um para vários; é referenciado por (*relation:isreferencedby*) e referencia (*relation:references*) direciona para outro recurso que pode auxiliar ao usuário, como a sátira ao recurso descrito ou uma crítica ao artigo descrito; é novo formato de (*relation:isformatof*) e novo formato em (*relation:hasformat*) quando o recurso descrito se apresenta em novo formato ou existe em outro formato sem variar o conteúdo intelectual; conforme com (*relation:conformsto*) para informar que o recurso segue algum tipo de norma, como as de acessibilidade.

Melhores práticas recomendam:

²⁷ Guidelines for Encoding Bibliographic Citation Information in DC Metadata
<<http://dublincore.org/documents/dc-citation-guidelines/>>

²⁸ DCMI Metadata Terms <<http://dublincore.org/documents/dcmi-terms/>>

- identificar o recurso relacionado por meio do seu identificador OAI. Quando o recurso relacionado também for contido em um arquivo disponível. A relação que inicia com “oai” representará, normalmente, um provedor de serviços através de um direcionador (*link*) que retorna os metadados do mesmo;
- identificar pela especificação dos termos do DCMI *dcterms:URI*²⁹ para valores que possuam endereço digital fixo, um *Uniform Resource Locator* (URL);
- identificar o recurso relacionado que não for catalogado em um sistema de arquivos, por uma norma única de identificação;
- usar valores não literais, conforme definido em DCMI *Abstract Model*³⁰.

Ex: <http://seer.ufrgs.br/EmQuestao/article/view/10387>

4.4.10 Formato (*format*)

Definição: manifestação física ou digital do recurso.

Comentário: formato do arquivo, meio físico ou dimensões do recurso, tais como tamanho e duração.

Refinamento: extensão (*format:extent*) que denota o tamanho ou duração do recurso; meio (*format:medium*) apresenta a manifestação do recurso, seu meio material.

Melhores práticas recomendam:

- utilizar vocabulário controlado como o listado na *Internet Media Types* [MIME]³¹ para recursos digitais como mensagem, aplicativo, áudio, modelo, vídeo, texto, entre outros.

Ex: gif; jpeg. No refinamento extensão: 21 minutos; no refinamento meio: óleo sobre tela.

4.4.11 Tipo (*type*)

Definição: natureza ou gênero do conteúdo do recurso.

²⁹ DCMI Metadata Terms <<http://dublincore.org/documents/dcmi-terms/>>

³⁰ DCMI Abstract Model <<http://dublincore.org/documents/abstract-model/>>

³¹ MIME Media Types <<http://www.iana.org/assignments/media-types/index.html>>

Comentário: Podem descrever categorias genéricas, funções, gêneros, ou níveis de agregação para o conteúdo do registro.

Melhores práticas recomendam:

- que todo recurso tenha, pelo menos, uma descrição do tipo;
- utilizar vocabulário controlado como o encontrado no DCMI *Type Vocabulary* [DCMI-TYPE]³²;
- utilizar o esquema OLAC *Linguistic Data Type Vocabulary*³³ para identificar dados linguísticos;
- para descrever manifestação física ou digital do recurso, como formato do arquivo, meio físico ou dimensões do recurso, utilizar o elemento *Formato*.

Ex: Texto, imagem, som, evento, coleção, objeto físico, serviço, *software*.

4.4.12 Data (*date*)

Definição: pontual ou período de tempo associado com um evento no ciclo de vida do recurso.

Comentário: expressa uma informação temporal a qualquer nível de precisão. Normalmente associada à criação ou disponibilidade do recurso.

Refinamento: datas de criação (*date:created*), de validação (*date:valid*), de disponibilidade (*date:available*), de emissão (*date:issued*), de modificação (*date:modified*), de aceitação (*date:accepted*), do copyright (*date:copyrighted*), da submissão (*date:submitted*).

Esquema de Codificação: segue as instruções para períodos, indicando início e fim, como em DCMI *Period Encoding Scheme*³⁴

Melhores práticas recomendam:

- utilizar o esquema codificado como consta em W3CDTF³⁵ da ISO 8.601 [W3CDTF], para a data completa, ano, mês, dia, a sintaxe é: YYYY-MM-DD;
- utilizar pelo menos um registro de data ou dos seus refinamentos;

³² DCMI Type Vocabulary <<http://dublincore.org/documents/dcmi-type-vocabulary/>>

³³ OLAC Linguistic Data Type Vocabulary <<http://www.language-archives.org/REC/type-20060406.html>>

³⁴ DCMI Periodic Encoding Scheme <<http://dublincore.org/documents/dcmi-period/>>

³⁵ W3C <<http://www.w3.org/TR/NOTE-datetime>>

- colocar a informação entre colchetes quando for sem conformidade com o esquema de codificação ou for atribuída pelo catalogador;
- Utilizar o esquema codificado, para qualquer instância.

Ex: 2011-31-10

4.4.13 Idioma (*language*)

Definição: idioma do conteúdo intelectual do recurso.

Melhores práticas recomendam:

- utilizar vocabulário controlado como o RFC 4646; para identificação precisa, conforme o esquema *olac:language*³⁶. O mesmo utiliza os códigos de três letras da norma ISO 639-3:2007 e salienta que no momento da colheita dos metadados, pode-se converter o código de duas letras (ISO 639-1:2002) para três letras;
- opcionalmente acrescentar duas letras para identificar o país seguindo a norma ISO 3166;
- que todos os registros possuam pelo menos um valor para o elemento idioma; nos registros sem conteúdo de linguagem, com código para o elemento idioma “zxx”, identificar como “Sem conteúdo de linguagem.”

Ex: “por” para português; “spa” para espanhol, “eng” para inglês.

4.4.14 Editor (*Publisher*)

Definição: entidade responsável por tornar acessível o recurso.

Comentário: pode tratar-se de uma pessoa, uma organização ou um serviço.

O nome editor pode ser utilizado para indicar a entidade.

Melhores práticas recomendam:

- identificar o editor por um nome que possibilite o ordenamento alfabético.

Ex: *Universidade Federal do Rio Grande do Sul.*

³⁶ OLAC Language Extension <<http://www.language-archives.org/REC/language-20080222.html>>

4.4.15 Direitos (*rights*)

Definição: informação sobre os detentores dos direitos sobre os recursos.

Comentário: Normalmente informações sobre a gestão de direitos do recurso ou instruções onde encontrará essa informação. O elemento refere-se, normalmente, a direitos autorais associados ao recurso, incluindo propriedade intelectual.

Refinamento: direito de acesso (*rights:accessrights*) que informa sobre o grau de privacidade do elemento, público, restrito ou privado, relacionado ao grau de segurança e outras regulamentações. A restrição pode ser por critérios baseados na classe do documento ou no tipo do usuário. Licença (*rights:licence*) é o refinamento onde informa-se ao usuário sobre o uso que ele pode fazer daquele recurso, é um documento oficial.

Melhores práticas recomendam:

- no refinamento licença identificar um endereço digital fixo, um *Uniform Resource Locator* (URL) pela especificação dos termos do DCMI *dcterms:URI*³⁷;

Ex: com refinamento direito de acesso: “somente para assinantes”.

Para se ter uma noção da importância dessa padronização de metadados, recorreremos ao estudo de Sené (2004) que fez um levantamento da repercussão dos trabalhos do DCMI pela perspectiva do número de aprovações de normas internacionais logo após o congresso do DCMI em 2001. A autora lista:

- ISO 15.836, baseada totalmente na norma americana Z39-85:2001 Elementos do Dublin Core (*Information and documentation -- The Dublin Core metadata element set*);
- ISO/TR 19.033 Documentação técnica do produto – Metadados para a documentação de construção (*Technical product documentation -- Metadata for construction documentation*);
- ISO 19.115 Informação geográfica – Metadados (*Geographic information – Metadata*);

³⁷ DCMI Metadata Terms <<http://dublincore.org/documents/dcmi-terms/>>

- ISO 8.459-5 Informação e Documentação – Diretório de elementos do dado bibliográfico – Parte 5: Elementos de dados para o intercambio da catalogação e dos metadados (*Information and documentation -- Bibliographic data element directory for use in data exchange and enquiry*);
- ISO/IEC 11.179-3 Tecnologia da Informação – Registros de metadados (MDR) – Parte 3: Registro do metamodelo e atributos básicos (*Information technology -- Metadata registries (MDR)-- Part 3: Registry metamodel and basic attributes*);
- ISO/IEC 13.818-1 Tecnologias da Informação – Codificação genérica de imagens em movimento e a informação de áudio associada, Sistemas (*Information technology -- Generic coding of moving pictures and associated audio information: Systems*);
- ISO/IEC/TR 20.943-1 Tecnologias da Informação – Procedimentos para atingir a consistência de conteúdo do registro de metadado – Parte 1: elementos de dado (*Information technology -- Procedures for achieving metadata registry content consistency-- Part 1: Data elements*).

Para finalizar, Senso e Piñero (2003, p. 100) destacam um questionamento que, segundo os autores é comum quando se trata de metadados, “quem será o responsável de introduzir os documentos [na base]?”. Eles ressaltam que o questionamento não é sob o ponto de vista técnico e sim pela conscientização da necessidade da adoção dos padrões. Pois, hoje ninguém mais pergunta quem são os responsáveis pelas descrições bibliográficas dos acervos nas bibliotecas pois, na verdade, o que os bibliotecários fazem é introduzir metadados. E esse profissional deve estar ciente das suas responsabilidades e consequência de suas ações nos processos de busca e recuperação da informação.

4.5 QUALIDADE DOS METADADOS

Os trabalhos de Vanti (2007) abordam a comunicação científica na *internet*. A autora afirma que é necessário que se façam estudos de medição e validação da produção intelectual, identificando autores e instituições, de maneira semelhante aos

que se tem realizado com o material impresso. Decorre disso, a importância da informação disponibilizada nos metadados passíveis de colheita para tais fins.

Na bibliografia encontram-se diversas definições de qualidade de dados, mas todas tendem a coincidir em que um dado será considerado de qualidade quando atende ao propósito do seu usuário, apropriado para o uso específico de quem o requer. Portanto, assim, há uma percepção multidimensional da qualidade e ela é relativizada conforme o seu contexto (GUERRA-GARCIA et al., 2009, p. 126). Porém, quando uma mesma informação é acessada por usuários com demandas diferentes e em contextos diferentes, ocorre a dificuldade e contemplar a contento todos eles.

Ikematu (2001) afirma que muitas vezes é dada maior ênfase ao sistema onde os dados estão armazenados, do que a qualidade dos próprios dados e que as falhas derivam por encarar os metadados apenas como recurso técnico, não informacional. Mas há de reconhecer que um dos principais problemas de implementação é conciliar ambas, o que pode exigir uma equipe multidisciplinar.

Segundo Bruce e Hillmann (2004) há uma dificuldade em definir qualidade de metadados, embora o seu conceito esteja presente em diversos trabalhos. Isso significa, segundo os autores, que é possível reconhecer metadados de qualidade sem, de fato, poder definir o que isso significa, ou encontrar suas características essenciais. Os autores afirmam que existe um consenso em que qualidade é quantificável e mensurável. As primeiras comunidades que se preocuparam com a qualidade dos metadados estavam relacionadas a bibliotecas digitais e impressos eletrônicos.

Ainda os mesmos autores salientam que até em iniciativas que visam a interoperabilidade, como a *Open Archives Initiative Protocol for Metadata Harvesting* (OAI PMH) a qualidade dos metadados não é abordada em profundidade. Some-se a isso, a falta de coordenação de grupos isolados que abordam a questão a partir de conhecimentos específicos e não de forma abrangente. Eles observam que os projetos de bases de dados ou páginas na *internet*, por exemplo, raramente possuem um orçamento específico para garantir a qualidade dos seus metadados, por parecer que o investimento sempre é maior que o benefício esperado ou pela necessidade imediatista com outros aspectos. Isso revela que, na maioria das vezes, há um desconhecimento ou uma falta de cuidado com os possíveis

cruzamentos ou trabalhos posteriores que possam vir a ser feitos a partir dos metadados disponibilizados.

No texto, também mostram experiências onde há colaboração entre bibliotecas, os resultados de dados bibliográficos, de autoria e vocabulários controlados obtendo melhores resultados que esforços solitários. Da mesma forma, os autores propõem imprimir qualidade aos metadados através de iniciativas conjuntas envolvendo diversos profissionais, para encontrar soluções aplicáveis aos mais diversos usos. O trabalho em comunidades para desenvolver metadados de uso geral e com possibilidade de especificações particulares, requer uma série de documentação para expor tanto a semântica quanto a sintaxe dos metadados visando a interoperabilidade.

Bruce e Hillmann (2004) basearam-se em estudos canadenses para garantia da qualidade, *Quality Assurance Framework (QAF)* da *Statistics Canada (STC)*, e selecionaram sete características de qualidade: completude (*completeness*), exatidão (*accuracy*), origem (*provenance*), conformidade com as expectativas (*conformance to expectations*), consistência lógica e coerência (*logical consistency and coherence*), oportunidade (*timeliness*) e acessibilidade (*accessibility*).

A norma ISO/IEC 25.012 aborda a qualidade dos dados de forma multidimensional a fim de que possa ser medida. As dimensões da qualidade estão relacionadas com características da mesma que permitam identificá-la, defini-la e medi-la. Elas estão classificadas em duas categorias, as relativas à qualidade inerente dos dados, onde possa atender-se uma demanda informacional a partir dos dados em si, ou dependente do sistema, onde a qualidade do dado é melhorada ou dependente do sistema computacional em que opera. No entanto, também há casos em que ambas as categorias estão presente.

Então, as 15 dimensões arroladas na norma conforme as categorias definidas anteriormente são:

inerentes

- exatidão, (*accuracy*)

O grau no qual o dado tem atributos que corretamente representam o valor correto do atributo intencionado de um conceito ou evento em um contexto específico de uso.

- completude, (*completeness*)

O grau ao qual o dado do sujeito associado com uma entidade tem valores para todos os atributos esperados e instancias de entidades relacionadas em um contexto específico de uso.

- consistência (*consistency*)

O grau no qual o dado tem os atributos que são livres de contradição e são coerentes com outros dados em um contexto específico de uso.

- credibilidade, confiabilidade (*credibility*)

O grau no qual o dado tem atributos que são considerados como verdadeiros e fidedignos pelos usuários em um contexto específico de uso.

- atualidade (*currentness*)

O grau no qual o dado tem os atributos que são do período correto em um contexto específico de uso.

inerentes e dependentes do sistema

- acessibilidade (*accessibility*)

O grau no qual o dado pode ser acessado em um contexto específico de uso, em particular por quem necessita o suporte de tecnologia ou uma configuração especial decorrente de alguma necessidade especial ou incapacidade.

- conformidade (*compliance*)

O grau no qual o dado tem atributos que se aderem a normas, convenções ou regulamentações vigentes e regras similares relacionadas com a qualidade de dados em um contexto específico de uso.

- confidencialidade, segurança, privacidade (*confidentiality*)

O grau no qual o dado tem os atributos que garantam que este é somente acessível e interpretável por usuários autorizados em um contexto específico de uso.

- eficiência (*efficiency*)

O grau no qual o dado tem os atributos que podem ser procesados e proporciona os níveis esperados de funcionamento (desempenho) utilizando as quantidades e os tipos de recursos apropriados em um contexto específico de uso.

- precisão (*precision*)

O grau no qual o dado tem os atributos que são exatos ou que proporcionam a discriminação em um contexto específico de uso.

- rastreabilidade (*traceability*)

O grau no qual o dado tem atributos que proporcionam um rasto de auditoria de acesso aos dados e de qualquer alteração ocorrida com os mesmos em um contexto específico de uso.

- entendibilidade, facilidade de entendimento (*understandability*)

O grau no qual o dado tem atributos que lhe permitem ser lidos e interpretados por usuários, e é expresso em linguagens apropriadas, símbolos e unidades em um contexto específico de uso.

dependentes do sistema são:

- disponibilidade (*availability*)

O grau no qual o dado tem atributos que lhe permitem ser recuperados por usuários autorizados e/ou aplicações em um contexto específico de uso.

- portabilidade (*portability*)

O grau no qual o dado tem os atributos que lhe permitem ser instalados, substituídos ou removidos de um sistema a outro conservando a qualidade existente em um contexto específico de uso.

- recuperabilidade (*recoverability*)

O grau no qual o dado tem atributos que lhe permitem manter e conservar um nível especificado de operações e qualidade, mesmo em caso de falhas (do sistema), em um contexto específico de uso.

Porém Wang e Strong (1996, p. 16) afirmam que analisar casos práticos por todas as dimensões conjuntamente é difícil. Por isso, criaram quatro categorias onde foram agrupadas diversas dimensões da qualidade e, dessa forma, estabeleceu-se uma hierarquia na análise da qualidade dos dados. No Quadro 1 é representada esta categorização, conforme adaptação de VERBO e outros (2009 p. 362).

Quadro 1 - Modelo Genérico de Qualidade de Dados

CATEGORIA DA QUALIDADE DE DADOS	CONCEITOS MENSURÁVEIS	DESCRIÇÃO
Intrínseca	Precisão (<i>Accuracy</i>) Objetividade (<i>Objectivity</i>) Credibilidade (<i>Believability</i>) Reputação (<i>Reputation</i>)	Qualidade que os dados têm por si próprios
Acessibilidade	Acessibilidade (<i>Accessibility</i>) Segurança de Acesso (<i>Access Security</i>)	Propiciam significado sobre a facilidade de acesso aos dados
Contextual	Relevância (<i>Relevancy</i>) Valor Agregado (<i>Value-Added</i>) Oportunidade (<i>Timeliness</i>) Compleitude (<i>Completeness</i>) Quantidade de dados (<i>Amount of Data</i>)	Tratam do uso dos dados num contexto
Representacional	Interpretabilidade (<i>Interpretability</i>) Facilidade de Compreensão (<i>Easy of Understanding</i>) Representação concisa (<i>Concise Representation</i>) Representação Consistente (<i>Consistent Representation</i>)	Características da representação dos dados que os tornam úteis

Fonte: STRONG; LEE; WANG (1997 *apud* VERBO et al., 2009 p. 362)

Assim, determinadas as dimensões que refletem a qualidade dos dados, é importante definir o nível de aceitabilidade das inconformidades para efetuar a validação dos dados. Strong, Lee e Wang (1997, p. 104, [tradução nossa]) definem como *problema de qualidade de dados* toda dificuldade encontrada em todas ou na maioria das dimensões da qualidade que resultam em inadequação total ou fortemente inapropriada para o uso. Guerra-Garcia (2009, p. 3) afirma que as decisões devem ser tomadas sempre considerando os usuários e recomenda o uso

de técnicas e ferramentas clássicas da Gestão da Qualidade para a localização das falhas, assim como as causas da ocorrência.

Nesse sentido, é importante abordar o estudo de Wand e Wang (1996) que focam na dificuldade em representar o mundo real através de dados num sistema de informações. A dificuldade é decorrente da percepção daquilo que se quer representar, que é subjetiva; da própria representação e ainda da interpretação que o sistema ou usuário faz dos dados recebidos. No estudo identificam três tipos de não qualidade nos dados, quais sejam: representação incompleta, representação ambígua, e carência de sentido. A partir deles, a informação estará incompleta, ambigua, sem sentido e/ou incorreta.

Outro trabalho que aborda a falta de qualidade é o de Dushay e Hillmann, (2003) onde analisam os metadados com o intuito de uso e reuso através dos indicadores de dados faltando, dados incorretos, dados confusos e dados insuficientes. Os metadados com:

- a) dados faltando - são aqueles em que certos registros não estão presentes nos metadados colhidos;
- b) dados incorretos – quando os registros não correspondem ao padrão;
- c) dados confusos – onde há diversos valores para um único elemento de metadado; e
- d) dados insuficientes – no caso em que não há utilização de vocabulário controlado e o registro não retorna todas as informações necessárias.

Além do exposto, considera-se útil a compreensão dos conceitos de qualidade do projeto, qualidade de conformidade e qualidade do serviço. No caso dos repositórios, conta-se com diversos profissionais que são responsáveis por etapas específicas do processo de informação e que tomarão decisões em conjunto ou não. Assim, passa-se a um breve comentário sobre cada um.

Qualidade de projeto - Para alcançar um melhor resultado, antes de mais nada, é importante definir o objetivo do repositório, o vocabulário controlado, os procedimentos de catalogação e as ferramentas (interfases) de inserção dos metadados (GUY; POWELL; DAY, 2004).

Primeiro ter em mente o que será disponibilizado ao usuário e quem é ele. Podem ser buscas por título, autor, palavras-chave, tipo de documento, nome da

publicação ou evento, ano, etc. Também pode-se permitir ver atualizações, últimos documentos adicionados, uso de filtros, etc.

Segundo levantar os requisitos funcionais necessários a fim de disponibilizar os metadados para colheita por terceiros, mediante o uso de um protocolo aberto, como o OAI-PMH. Definir o que será disponibilizado.

Terceiro, selecionar os metadados, os esquemas de codificação e o vocabulário controlado para a codificação. O vocabulário controlado pode ser baseado em outro existente ou se haverá necessidade de criar um novo, vale salientar, que o uso de *thesaurus* ou vocabulários controlados consagrados poderão permitir melhores resultados na interoperabilidade.

Finalmente, é fundamental o estabelecimento de padrões de codificação. A falta ou desrespeito aos padrões dificultará a recuperação dos documentos, observar especialmente os metadados referentes a datas e nomes pessoais e institucionais. Segundo Guy, Powell e Day (2004), o uso de exemplos nos Padrões de Rotinas, poderão facilitar a inserção dos dados na base e reduzir o risco de erros de sintaxe e vocabulário.

Qualidade de Conformidade - Dado isso, é necessário conferir a conformidade aos padrões. O procedimento consiste em verificar se os dados inseridos estão corretos, que pode ser feito mediante a conferência semanal de dados escolhidos aleatoriamente. Existem dispositivos automáticos desenvolvidos nesse sentido (GUY; POWELL; DAY, 2004), inclusive alguns que enviam aos autores metadados predefinidos para que estes façam a conferência, o mesmo procedimento é usado enviando a especialistas. Após as conferências, são feitas as eventuais correções de forma automática ou manual.

Qualidade do Serviço - Porém, a qualidade dos metadados só será alcançada se o usuário recuperar os documentos desejados e obtiver outras informações sobre os mesmos, inclusive.

O autor Guerra-Garcia (2009) lista alguns motivos para a ocorrência da falta de qualidade dos dados, como o uso de múltiplas fontes para uma mesma informação, o uso de subjetividade que poderá incluir prejuízos ao inserir um dado, erros sucessivos que levam à perda da informação, facilidade de acesso à informação que pode levar a disponibilizar dados reservados ou privados e que entram em conflito com os requisitos de segurança, privacidade e confidencialidade. Também, em certos repositórios são os autores que inserem os metadados.

5 METODOLOGIA

A presente sessão tem o propósito de esclarecer os procedimentos utilizados na pesquisa a fim de alcançar os objetivos propostos.

As pesquisas, de maneira geral, iniciam pela escolha do tema, onde se define o aspecto de interesse de um assunto que se deseja desenvolver, o que será pesquisado, que pode ser algo original ou não. O tema escolhido foi a qualidade de dados e, a partir daí, foram realizadas leituras sobre o mesmo, onde se avançou na literatura sobre metadados, acesso livre e comunicação científica. Nas pesquisas bibliográficas parte-se de um “material já elaborado, constituído principalmente de livros e artigos científicos” (GIL, 1991, p. 48) com o propósito de encontrar informações em outros estudos com temas semelhantes e que poderiam auxiliar o estudo atual. A justificativa do tema está baseada na importância do mesmo no contexto da divulgação de resultados de descobertas e de estudos científicos. O resultado das primeiras impressões sobre o tema permitiu a formulação do problema e a definição do sujeito e objetivos da pesquisa. Dado isto, passou-se a delinear a metodologia e organizar os procedimentos da pesquisa.

Pela finalidade pretendida, que é conhecer a qualidade dos metadados de um conjunto particular, a pesquisa pode ser classificada como do tipo aplicada. Desde o ponto de vista dos objetivos, é um estudo exploratório, já que é uma primeira aproximação com o tema e se pretende recuperar informações disponíveis, como são os metadados da Revista Em Questão. Como amostra foi selecionado o próprio universo de dados disponíveis na página da Revista na *internet*, portanto esta é a fonte dos dados. Sendo assim, é uma pesquisa de campo já que os metadados foram coletados diretamente onde são produzidos através de um levantamento, que é o procedimento recomendado para as pesquisas exploratórias.

A página da publicação na *internet* é hospedada e gerenciada pelo Sistema Eletrônico de Editoração de Revistas (SEER). O instrumento de coleta foi o ambiente FABRICO, desenvolvido pelo Prof. Dr. Rafael Port da Rocha, da Faculdade de Biblioteconomia e Comunicação (FABICO/UFRGS).

O programa colhe os dados como inseridos no sistema pela Revista, apenas para os campos Assunto e Criador faz interferência. No campo Assunto, se houver ponto e vírgula (;) ou ponto (.) o programa separa as expressões assumindo que são assuntos diferentes embora tenham sido digitados num único campo. Da mesma

forma, no campo Criador, o programa assume tratar-se de autores diferentes quando houver um ponto e vírgula (;).

A colheita dos dados ocorreu em abril de 2011. A seguir, partiu-se para o tratamento dos dados, que é a etapa em que os mesmos são organizados e validados. Para isto, utilizaram-se os recursos do programa FABRICO em que são evidenciados os registros de cada um dos elementos do *Dublin Core Metadata Initiative* e utilizados tal qual se apresentaram. Os dados foram validados no sentido de estarem presentes os metadados de todas as edições do período de 2003 a 2010, totalizando 17 edições.

As análises foram fundamentadas no levantamento bibliográfico de forma a identificar semelhanças ou relações entre a teoria e a prática. No caso, a análise dos dados ocorreu de forma qualitativa. O procedimento consistiu na comparação dos dados encontrados com as recomendações do movimento DCMI. Assim, verificou-se se foram seguidas as recomendações quanto à seleção dos elementos, se houve uso dos elementos expandidos ou novos elementos e quanto a sua representação e codificação.

Nas Melhores Práticas do DCMI, consta que os valores dos metadados devem apresentar-se conforme as definições e instruções de cada elemento. A seguir, foi feita uma correspondência com as dimensões da qualidade definidas pela norma ISO/IEC 25.012, porém agrupadas conforme Wang e Strong (1996). E, por fim, os aspectos dos metadados quanto à não conformidade como nos estudos realizados por Wand e Wang (1996) e Dushay e Hillmann, (2003). Nesta etapa utilizaram-se cálculos estatísticos básicos, como frequência e média sobre aspectos específicos que são propiciados pelo programa FABRICO.

Como instrumento de pesquisa foi elaborado um quadro para auxiliar a análise a fim de garantir que todos os aspectos fossem observados. Ela foi inspirada nos modelos de Folha de Verificação (*Check List*), ferramenta da qualidade para coleta e organização de dados nos momentos de inspeção de processos e produtos ou serviços. Este tipo de ferramenta se caracteriza por ter um formato simples e fácil de usar.

Quadro 2 – Folha de Verificação para Análise dos Metadados

Qualidade	Atendimento ao DCMI			
	sim	não	parcial	não se aplica
uso de refinamento				
valor <i>conforme definição</i>				
intrínseca				
dados insuficientes				
dados confusos				
dados incorretos				
dados faltando				
<i>conforme instrução</i>				
representacional				
dados insuficientes				
dados confusos				
dados incorretos				
dados faltando				
ambiente				
acessibilidade				
contextual				

Os resultados obtidos nas análises são apresentados de maneira a evidenciar a observação e caracterizar os fenômenos. Então, foram elaboradas tabelas com esse propósito. Após a discussão dos resultados, estes foram sintetizados de maneira a compor as conclusões da pesquisa e se redigiu a monografia.

Ao finalizar a exposição da metodologia, cabe destacar as limitações do estudo, que podem ser apresentadas quanto:

- ao assunto: qualidade dos metadados de documentos publicados;
- ao espaço: metadados da Revista Em Questão;
- ao tempo: de 2003 a 2010;
- à equipe de trabalho: a aluna autora;
- à amostra: o universo dos dados disponíveis; e
- ao instrumento: ambiente FABRICO.

6 RESULTADOS

Os resultados são apresentados por elemento do *Dublin Core Metadata Initiative* (DCMI). Para cada um deles, são mostrados os relatórios com os levantamentos estatísticos calculados pelo FABRICO e os valores dos metadados colhidos. A seguir é apresentada a análise desses registros quanto a sua adequação às definições, instruções e recomendações de melhores práticas do DCMI, à não conformidade e comparando com o referencial teórico.

Para isso, são utilizados os documentos do DCMI, onde deixam explícitos os aspectos de definição, relacionado com a semântica dos metadados e quanto às instruções, que é a estrutura dos metadados e correspondem aos cuidados com a sintaxe. A não conformidade refere-se às categorias de qualidade e tipo de não conformidade.

Na análise, quando considerado apropriado, será orientado o uso de qualificadores, embora entenda-se que só o fato de padronizar a entrada com os critérios de um qualificador já traga benefícios, sem efetivamente qualificar o elemento. Dessa forma, a ação poderá ser imediata, dependendo somente de uma decisão gerencial e dispensando ações técnicas.

No relatório estatístico do FABRICO, os resultados informam o Total de Instâncias, que se refere ao número de recursos descritos através do elemento; o Total de Valores corresponde à quantidade de vezes que o metadado foi utilizado, ou seja, ao somatório das frequências; o Total de Valores indica o total dos universos de valores utilizados, não considerando repetições, trata da ocorrência dos termos, ou seja, as palavras distintas encontradas; e a Média de Valores por Instância é o quociente do Total de Valores pelo Total de Instâncias.

No levantamento, foram identificados 225 recursos diferentes representados pelos seus metadados. Eles correspondem aos artigos, resumos de dissertações e teses, editorial e expediente das 17 edições de 2003 a 2010.

6.1 TÍTULO (*TITLE*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 225

Total de Valores: 454

Total de Valores Diferentes: 446

Média de Valores por Instância: (454/225): 2.017777777778

Os títulos que ocorreram com uma frequência igual ou superior a dois são discriminados na Tabela 1, conforme se apresentam nos metadados.

Tabela 1 – Metadados do Elemento Título

Id.	TÍTULO	FREQUÊNCIA
1	Apresentação	5
2	Expediente	3
3	Editorial	2
4	Resumos das dissertações e teses defendidas no PPGCOM/UFRGS	2
5 - 446	Títulos com uma frequência	442
TOTAL		454

Fonte: FABRICO

O valor esperado para o Total de Instâncias era 225, o que de fato ocorreu. Também era esperado que não houvesse repetições para os Títulos de artigos, no entanto, foram encontrados títulos com mais de uma ocorrência e são mostrados na Tabela 1. Esses títulos correspondem às seções que se repetem ao longo das edições e que, realmente, poderiam apresentar registro com o mesmo valor, por exemplo, Editorial e Expediente.

Os metadados colhidos trouxeram 446 Valores Diferentes para Título, onde o mesmo item está representado pelo seu título no original e em outro idioma. Um exemplo é: *Publicidade: compromisso com o social*, em português; *Publicidad: compromiso con lo social*, em espanhol; e *Publicity: a social commitment*, em inglês. As recomendações do DCMI são para que exista apenas um elemento identificado como título. Nos metadados essas versões do título aparecem no *dc:title*: por isso, um mesmo recurso possui mais de um elemento título.

As recomendações também dizem que o título traduzido seja registrado como Título Equivalente somente se ele estiver junto ao título no original do recurso

descrito, para os demais, utilizar Título Alternativo como refinamento do elemento Título. Na Revista, os títulos e resumos em outro idioma encontram-se no final do texto, junto às Referências. No entanto, as versões para os dois idiomas não constam em todos os documentos.

Como o Total de Valores para o elemento Título foi 454, a Revista Em Questão possui uma média de dois títulos por recurso descrito. Isso se deve ao fato que a grande maioria dos recursos da Revista possui o título no original e pelo menos uma tradução. Como exceção, observamos 35 títulos referentes ao Editorial, Expediente, Resumos e outros semelhantes que não necessitariam de tradução.

Nas diferentes versões dos títulos, foram identificados caracteres alheios ao contexto, como no caso de *Reflexões sobre a televisão e a "falta de cultura" no Brasil*; *Reflexiones sobre la Televisión y la ?falta de cultura? en Brasil*; e *Reflexions on television and ?lack of culture? in Brazil*. Este tipo de erro pode ser resultante da migração de dados de um sistema a outro, resultante da utilização de um programa de tradução. Embora com esses caracteres, para a maioria das pessoas, será possível ler o título, o que não é garantido para um motor de busca.

O Editorial está presente em todas as edições a partir de 2005, o que perfaz 13 edições. Ele foi registrado duas vezes como Editorial; uma com o Editorial seguido pelo ano e edição, como em Editorial 2009/2; cinco como Apresentação, nas edições de 2007 até a primeira de 2008; e em cinco a palavra Apresentação foi seguida pelo ano e edição, por exemplo, Apresentação 2007/01. O somatório das diferentes variantes refere-se a 13 itens, como esperado.

O Expediente está presente em todas as 17 edições. Todos os itens possuem seus expedientes registrados no elemento Título. No entanto, somente em três apenas apresenta este termo, nos demais, Expediente está registrado seguido de ano e edição, como em Expediente 2006/01.

Também constam como título de seção *Resumos das dissertações e teses defendidas no PPGCOM/UFRGS (2003/1 e 2004/1)* e *Resumos das dissertações e teses defendidas pelo corpo docente da FABICO/UFRGS em 2003/1*, ambas expressões foram categorizados iguais a Expediente ou Edição. Portanto, mantém-se coerência.

No que se refere à qualidade dos dados na grafia ou sintaxe, não se verificaram problemas de digitação, incorreções ou omissão do título, apenas os já referidos nas traduções.

Como resumo da análise, pode-se dizer que a este elemento caberia um refinamento, utilizando a versão em outra língua como título alternativo. O elemento é utilizado conforme sua definição já que aqui é colocado o título do artigo publicado ou da seção a que se refere. Neste elemento não se encontraram problemas quanto às instruções de sintaxe, pois trata-se de copiar o título conforme se apresenta no artigo ou seção.

6.2 CRIADOR (*CREATOR*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 225

Total de Valores: 453

Total de Valores Diferentes: 336

Média de valores por Instância (453/225): 2.01333333333

Os valores atribuídos através do elemento Criador que ocorreram com frequências maior ou igual a três são discriminados na Tabela 2, conforme se apresentam nos metadados.

Tabela 2 – Metadados do Elemento Criador

Id.	CRIADOR	FREQUÊNCIA
1	Em Questão, Comissão Editorial	32
2	UFRGS	15
3	Morigi, Valdir Jose	6
4	PUCRS	6
5	Universidade Federal de Minas Gerais	5
6	Universidade Federal do Rio Grande do Sul	5
7	Caregnato, Sônia Elisa	4
8	Universidade de Brasília	4
9	Baldissera, Rudimar	3
10	Bonotto, Martha E. K. Kling	3
11	Golin, Cida	3
12	Unesp	3

13	Universidade Federal de Pernambuco	3
14	Universidade Federal de Santa Maria	3
15	Universidade Metodista de São Paulo	3
16	Vanz, Samile Andréa de Souza	3
17 - 336	Criadores com uma ou duas ocorrências	352
TOTAL		453

Fonte: FABRICO

Os dados fornecidos pelo programa indicam que 225 itens da Revista foram descritos através do elemento Criador. Foram encontrados 453 valores para este elemento, que leva a uma média de ocorrência de dois valores para cada item descrito. Destes 453 valores, 336 são nomes de criadores diferentes e 117 nomes de autores repetidos, indicando autores que apresentaram mais de uma contribuição à Revista. Sendo assim, segue as recomendações do DCMI de colocar pelo menos um registro para cada recurso.

Este elemento se destina a uma pessoa ou instituição que é responsável pela existência do item descrito, segue algumas considerações desde o ponto de vista da adequação à definição do elemento.

Para o Editorial e Expediente consta: *Em Questão, Comissão Editorial*, o que foi considerado correto.

O campo foi preenchido com o nome de pessoas e de instituições. Pelo DCMI é a entidade que é responsável pela existência do recurso, então, se a instituição de fato for, caberia seu uso. Nas recomendações constam a preferência pelo uso de Colaborador quando não é tão nítida ou evidente a participação da pessoa ou instituição, ou até, que não exista um termo adequado no refinamento de Colaborador. Neste ponto cabe destacar que o FABRICO interpreta como um segundo criador quando encontra no campo um sinal de ponto e vírgula (;), sendo assim, é possível que algumas instituições estejam colocadas junto ao nome do autor e interpretadas como um co-autor. Exemplo:

Quadro 3 – Elemento Título e Criador do artigo *A Mediação das cartas dos leitores na mídia: mapas imaginários sobre Porto Alegre*

ELEMENTO	METADADO
dc:title	A Mediação das cartas dos leitores na mídia: mapas imaginários sobre Porto Alegre
dc:creator	Morigi, Valdir Jose
dc:creator	UFRGS
dc:creator	Rocha, Carla Pires Vieira da
dc:creator	Borges, Márcia de Castro
dc:creator	Bacharel e Licenciada em História pela UFRGS
dc:creator	Mestre em Multimeios - UNICAMP.

Fonte: FABRICO

Pelo Quadro 3 já podem ser percebidas incorreções nos metadados quanto à falta de correspondências com a definição, onde resulta em dados confusos por criar a dúvida da atuação da UFRGS como criador; dado errados, já que *Bacharel e Licenciada em História pela UFRGS* não poderia ser um autor. As recomendações orientam para deixar este registro sem valor caso não seja extremamente evidente a identificação da autoria.

Ao analisar o elemento pela sua conformidade à instrução, encontraram-se erros na representação, sintaxe, do item e falta de padronização nos nomes próprios de entidades e pessoas. Em *Univrsidade de Aveiro* falta uma letra. Nos casos a seguir há falta de padronização.

Quadro 4 – Exemplos de falta de padronização no elemento Criador

ELEMENTO	METADADO
dc:creator	Schwaab, Reges
dc:creator	Schwaab, Reges Toni
dc:creator	Rozados, Helen Beatriz Frota
dc:creator	Rozados, Helen Frota
dc:creator	USP
dc:creator	Universidade de São Paulo, Escola de Comunicações e Artes

ELEMENTO	METADADO
dc:creator	Universidade de São Paulo, Escola de Comunicações e Artes, Departamento de Biblioteconomia e Documentação
dc:creator	UFRGS
dc:creator	Universidade Federal do Rio Grande do Sul
dc:creator	Universidade Federal Fluminense
dc:creator	Universidade Federal de Santa Catarina – UFSC

Fonte: FABRICO

No Quadro 4 verificam-se exemplos de representação do mesmo criador de forma diferente, com o nome dos criadores usando expressões completas ou abreviadas, uso de siglas antes ou depois da denominação por extenso.

Em suma, para este elemento, considerou-se que seria útil a adoção de refinamento, já que a orientação é para o uso de Colaborador qualificado segundo o DCMI *Role* para melhor entendimento. Os dados encontrados apresentam desconformidade quanto à definição do elemento, imprecisos e até incorretos. Assim, ao considerar a definição de Criador, que é aquele responsável em primeira instância pela existência do elemento, há incorreções no julgamento de quem é o criador. Entende-se que decisões poderiam ser tomadas no sentido de definir em que fontes buscar o valor correto. A sintaxe também apresenta incorreções que comprometem o entendimento da informação, pois a dúvida é se as instituições são, de fato, autores ou se o intuito foi identificar a instituição à qual está vinculado o autor. Julga-se que este elemento é muito importante para a recuperação de documentos e que merece um cuidado maior.

6.3 COLABORADOR (*CONTRIBUTOR*)

Os dados oferecidos pelo programa FABRICO são:

Total de Itens Descritos pelo Elemento: 28

Total de Valores do Elemento: 28

Total de Valores Diferentes: 18

Média de valores por Elemento (28/28): 1

Os valores para Colaborador são discriminados na Tabela 3, conforme se apresentam nos metadados.

Tabela 3 – Metadados do Elemento Colaborador

Id.	COLABORADOR	FREQUENCIA
1	CNPq	6
2	Capes	4
3	Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq	2
4	Fapemig	2
5	-	1
6	CAPES/PICDT	1
7	CAPES/PRODOC	1
8	CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico	1
9	CNPq e FAPERJ	1
10	Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)	1
11	Conselho Nacional de Desenvolvimento Científico e Tecnológico ? CNPq ? Brasil	1
12	Facultad de Comunicación de la Universidad de Piura (Perú) y UNIVALI (Brasil)	1
13	FAPEMA	1
14	Flacso-Argentina; Universidad Nacional de Quilmes (Argentina)	1
15	Intercom	1
16	PPGCOM / UFRGS	1
17	UNIJUÍ; CNPq; Fapers; PPGAS/UFSC	1
18	Universidade de Caxias do Sul	1
TOTAL		28

Fonte: FABRICO

Para identificar o Colaborador, os valores correspondem, principalmente, a entidades que aportam recursos financeiros para as pesquisas. Na intenção de identificar duas entidades, foram utilizados o conectivo (y), hífen (-) ou ponto e

vírgula (;). Porém, dessa forma, o dado é interpretado como uma nova entidade e não como duas que, provavelmente, já constam em outros registros. Foram encontradas até quatro instituições para um único valor de um item, ou seja, representadas apenas como uma: *UNIJUÍ; CNPq; Fapers; PPGAS/UFSC*.

O mesmo ocorre quando uma instituição é representada de duas ou mais formas diferentes, nesse caso, será contabilizada como mais de uma instituição. Exemplo disto é o uso do nome da instituição por extenso e/ou abreviado. Quando utilizados em conjunto, a abreviação é colocada antes ou após a expressão e entre parênteses ou não, o CNPq foi registrado de cinco formas diferentes:

CNPq

Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

Conselho Nacional de Desenvolvimento Científico e Tecnológico ? CNPq ?

Brasil

As entidades encontradas foram: cinco agências de fomento, dois programas específicos ligados a uma instituição, quatro universidades, duas faculdades, dois programas de pós-graduação, uma sociedade sem fins lucrativos e uma fundação assistencial previdenciária. Considerando cada uma como uma entidade, pode-se dizer que houve colaboradores diferentes, quais sejam:

agências de fomento:

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq);

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES);

Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG);

Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ);

Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA);

programas específicos:

Programa Institucional de Capacitação Docente e Técnica da CAPES (PICDT/CAPES)

Programa de Apoio a Projetos Institucionais com a Participação de Recém-Doutores da CAPES (PRODOC/CAPES)

universidades:

Universidade do Vale do Itajaí (UNIVALI)

Universidad Nacional de Quilmes (UNQ/Argentina)

Universidade Regional do Noroeste do Estado do Rio Grande do Sul (UNIJUÍ)

Universidade de Caxias do Sul (UCS)

faculdades

Facultad de Comunicación de la Universidad de Piura (UDEP/Perú)

Facultad Latinoamericana de Ciencias Sociales (FLACSO/Argentina)

Programas de Pós Graduação

Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul (PPGCOM/UFRGS)

Programa de Pós-Graduação em Antropologia Social da Universidade Federal de Santa Catarina (PPGAS/UFSC)

instituição sem fins lucrativos:

Sociedade Brasileira de Estudos Interdisciplinares da Comunicação (INTERCOM)

outro

Fundação Assistencial e Previdenciária da Extensão Rural no Rio Grande do Sul (FAPERS)

A última instituição, classificada como outro, consta no campo colaborador apenas pela sua sigla. Sendo assim, pode tratar-se de um erro, já que é provável que a intenção fosse escrever FAPERGS em vez de FAPERS. A FAPERGS é a Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul, e há outros registros com as correlatas de outros estados. Como a instituição não consta no documento, trata-se de uma suposição.

Nos casos que referenciam instituições de fomento, corresponderia utilizar o refinamento indicando o papel do Colaborador, extraído do OLAC *Role Vocabulary*. Neste caso ficaria: *contributor:sponsor* que identifica a participação de quem contribui financeiramente para a existência do recurso. Inclusive, nas melhores práticas, recomendam o uso deste elemento em detrimento de criador, utilizando a versão com refinamento.

Também se constatou um caso em que a intenção era não adjudicar alguém como colaborador e foi colocado um hífen (-) para isso. Pelo fato de haver um dado no campo, é considerado um registro preenchido com a denominação do colaborador "-". No entanto, o campo vazio já denota que este não tem um valor para seu uso.

Evidencia-se a falta de padronização, tanto para os dados existentes quanto para identificar campos vazios. A norma ABNT NBR 14724:2011 Informação e documentação – Trabalhos acadêmicos – Apresentação diz que, na primeira ocorrência de uma sigla, deve-se escrever a expressão por extenso seguida pela sigla entre parênteses. (ABNT, 2011, p. 11). Como no caso em estudo é para utilização em um único registro, pode-se optar pelo uso da sigla e a expressão ser disponibilizada num glossário. Também poderia ter-se o cuidado de permitir a recuperação do documento tanto pelo uso por extenso quanto pela sigla.

Em resumo, quanto ao elemento Colaborador, pode-se dizer que há dúvidas se houve uma compreensão do significado do mesmo. A escolha do termo para representação do item não é totalmente coerente com a definição dada pelo DCMI, o que resulta em falta de qualidade intrínseca dos dados, representada por dados incorretos e confusos. Ao colocar o nome de uma universidade ou programa como Colaborador não se sabe qual a sua participação na criação do item.

Neste elemento podem ser arrolados uma série de colaboradores identificados pelo refinamento do elemento, mas se a opção for usar somente uma espécie de colaborador, isto precisa ser expresso nas normas de procedimentos, assim como sua forma de expressão. A qualidade também é prejudicada pela falta de padronização evidenciada na variação da forma de representar o elemento: por extenso, sigla ou ambas.

6.4 ASSUNTO (*SUBJECT*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 190

Total de Valores: 970

Total de Valores Diferentes: 699

Média de valores por instância (970/190): '5.10526315789

Os valores para Assunto que ocorrem em uma frequência de quatro ou mais são discriminados na Tabela 4, conforme se apresentam nos metadados.

Tabela 4 – Metadados do Elemento Assunto

Id.	ASSUNTO	FREQUENCIA
1	comunicação	42
2	ciência da informação	19
3	jornalismo	18
4	biblioteconomia	10
5	internet	10
6	ciências sociais aplicadas	9
7	televisão	7
8	cinema	6
9	comunicação científica	6
10	análise do discurso	5
11	bibliometria	5
12	bibliotecário	5
13	cidadania	5
14	comunicação social	5
15	discurso	5
16	publicidade	5
17	tecnologia	5
18	cibercultura	4
19	cultura	4
20	educação	4
21	imagem	4
22	informação	4
23	mídia	4
24-699	Assuntos com uma, duas ou três ocorrências	779
TOTAL		970

Fonte: FABRICO

No elemento Assunto foram encontrados 190 recursos com valores. Do total de 255 itens, 65 não foram descritos através do elemento Assunto. Nestes 65, incluem-se os 35 dedicados a Editorial, Expediente e Resumos, o que é considerado aceitável, restando ainda 30 itens sem justificativa para o não uso do elemento

O levantamento encontrou que o elemento assunto foi usado 970 vezes dos quais 699 correspondem a valores diferentes, ou representam assuntos não repetidos. Os assuntos em destaque são Comunicação, Ciência da Informação, Jornalismo e Biblioteconomia o que, por tratar-se da revista da FABICO, já é um resultado esperado. Porém, ao somarem-se as suas frequências, chega-se a 89 recursos com esses assuntos, o que representa 10% do total, considerado baixo já que o escopo da revista abrange os campos da Informação e Comunicação, conforme descrito na sua página na *internet*³⁸.

Quanto aos valores diferentes, que representam 72% do total de valores, podem ser apontadas duas justificativas, ou há uma diversidade de assuntos abordados na revista ou inexistente um vocabulário controlado capaz de representar os assuntos iguais por termos iguais. A segunda opção parece ser a correta, pois há assuntos idênticos representados por termos diferentes, como por exemplo:

ciência da informação	-	ciências da informação
biblioteconomia	-	bibliotecología
análise de citação	-	análise de citações

De forma geral, os assuntos são representados pelas palavras-chave dos recursos, porém não exclusivamente, selecionaram-se dois exemplos, como segue:

Comunicação móvel: das potencialidades aos usos e aplicações.

Palavras-chave: Comunicação móvel. Internet. Localização geográfica. Comunidades virtuais.

Do talk show ao televisivo: mais espetáculo, menos informação

Palavras-chave: *Talk show*. Televisão. Espetáculo.

³⁸ <http://seer.ufrgs.br/EmQuestao/about/editorialPolicies#focusAndScope>

Quadro 5 – Exemplos de Representação de Assunto

ELEMENTO	METADADO
dc:title	Comunicação móvel: das potencialidades aos usos e aplicações
dc:subject	Cibercultura
	Comunicação
	comunicação móvel
dc:subject	comunicación móvil
dc:subject	comunidades virtuais
	comunidades virtuales
	geographical location
	Internet
	localização geográfica
	mobile communication
dc:subject	ubicación geográfica
	virtual communities
dc:title	Desde el talk show hasta la televisión: más espectáculo, menos información
dc:subject	talk show
dc:subject	televisão
dc:subject	espetáculo
dc:subject	6.09.05.00-0 comunicação visual
dc:subject	6.09.00.00-8 comunicação
dc:title	Bibliotecas escolares e web 2.0
dc:subject	biblioteca 2
dc:subject	biblioteca escolar
dc:subject	Biblioteconomia
dc:subject	brasil
dc:subject	ciências sociais
dc:subject	Portugal
dc:subject	web 2

Fonte: FABRICO

Pelo Quadro 5, percebe-se que no primeiro exemplo, foram incluídos os termos *Comunicação* e *Cibercultura*; no segundo acrescentou-se *Comunicação Visual* e *Comunicação*. Em ambos os casos, a inclusão poderá permitir uma revocação melhor dos recursos e é possível que este seja o motivo do acréscimo. No entanto, a incorporação de novos termos deve seguir regras gerais, aplicadas a todos os recursos e não somente em alguns, deixando ao critério do indexador.

O DCMI relaciona uma série de fontes com o intuito de padronizar a representação, tanto na escolha do vocabulário quanto na sintaxe, uma delas é o uso das palavras-chave, assim, nesse ponto, são seguidas as recomendações de boas práticas. Porém, o uso das palavras-chave escolhidas pelo autor nem sempre são representações adequadas para recuperação na *Web*, seja pela falta de um termo mais geral, como o caso de *Comunicação*, pelo uso de um termo regional, ou até ambíguo. Isto reduz a qualidade dos dados no sentido intrínseco e, principalmente, no contextual, no ambiente onde é disponibilizada a Revista.

Na representação sintática também foram identificados erros. Nos registros de assunto, encontraram-se como assuntos números seguidos por termos que, efetivamente seriam metadados de assunto.

Esses caracteres inesperados identificados nos assuntos, tais como *09.05.00-0 comunicação visual*. Como esse número é a representação do assunto Comunicação Visual na Classificação Decimal Universal (CDU), evidencia-se a inclusão de ambos os valores como um registro só. No entanto, podem ter sido originados em tradutores automáticos, já que encontram-se algumas palavras-chaves traduzidas, ou na migração de sistemas.

Assim, o elemento assunto precisa de atenção nos dois sentidos das recomendações do DCMI, quanto a sua definição e quanto a sua instrução. Dessa forma, os dados precisam ser melhor escolhidos, com normas bem definidas para o elemento, a fim de propiciar-lhe qualidade intrínseca, que é aquela relativa ao dado em si; contextual que é a qualidade considerando o usuários, o seu ambiente e possíveis estratégias de busca; e representacional, que refere-se à forma como o é registrado o valor do recurso.

6.5 DESCRIÇÃO (*DESCRIPTION*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 194

Total de Valores: 422

Total de Valores Diferentes: 421

Média de valores por instância (422/194): 2.17525773196

Os valores para Descrição que apresentaram duas ou mais frequências são discriminados na Tabela 5, conforme se apresentam nos metadados.

Tabela 5 – Metadados do Elemento Descrição

Id.	DESCRIÇÃO	FREQUENCIA
1	.	2
TOTAL		2

Fonte: FABRICO

Os resultados recebidos mostram 194 recursos registrados. O Total de Valores Diferentes foi 421 e o Total de Valores encontrados foi 422. Os itens com valores repetidos foram o Editorial (Apresentação) e Prefácio da Edição Especial, onde utilizou-se o ponto (.) para representar sua descrição devido à inexistência de resumo. Porém, outro Editorial, mesmo sem resumo, é descrito como *Editorial da Edição 2009, vol.15, n.2*.

Fora os dois casos referidos anteriormente, os demais recursos foram descritos segundo a orientação do DCMI, que diz que a representação poderá ocorrer por um resumo. O resumo já é um produto elaborado para a representação do documento, portanto julga-se a melhor opção para a descrição. Os demais elementos sem resumo podem ser representados por um texto padronizado elaborado pelo indexador, como o exemplo do Editorial.

Cada recurso possui um ou mais valores para o elemento Descrição. Trata-se do resumo em original e uma ou duas versões para outro idioma. Mas poderão ser considerados como um ou mais valores do elemento Descrição ao estar registrado em um ou mais campos do elemento.

Então, para este elemento entende-se que segue as recomendações do DCMI em quase a totalidade dos registros e não há motivo para refinamento.

Apenas necessita-se redigir um texto padrão para os recursos sem resumo e que descrevam o documento. Quanto à sintaxe, poder-se-ia registrar apenas um resumo por campo.

6.6 COBERTURA (*COVERAGE*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 24

Total de Valores: 61

Total de Valores Diferentes: 51

Média de valores por instância (61/24): 2.54166666667

Os valores para Cobertura são discriminados na Tabela 6, conforme se apresentam nos metadados.

Tabela 6 – Metadados do Elemento Cobertura

Id.	COBERTURA	FREQUÊNCIA
1	-	1
2	2006/2007	1
3	1930-1980	1
4	1995-2007	1
5	2000 - 2007	1
6	2003 a 2005	1
7	2005	1
8	2006	1
9	Anos 1968 a 1978	1
10	Anos 30-40	1
11	Anos 90 até a atualidade	1
12	Anúncios sobre saúde pública publicados na Revista do Globo	1
13	Argentina	1
14	Artigos publicados quinzenalmente na revista	1

Id.	COBERTURA	FREQUÊNCIA
15	Atualidade	1
16	Bibliográfica	2
17	Brasil	7
18	Brasil colônia; anos 1922; anos 1930; período JK; anos 60; anos 80; atual	1
19	Brasil e Rio de Janeiro	1
20	Brasil, Estados Unidos, Inglaterra, Índia	1
21	Brasil/Mundo/2005	1
22	Descritiva.	1
23	El proceso de reestructuración de telecomunicaciones em decada 1990	1
24	Época moderna e atual	1
25	Final do século XX e início do século XXI	1
26	Histórica	2
27	Inédita	1
28	Informação estatística e nas entrevistas à profundidade feita os trabalhadores das companhias telefônicas que operam em Argentina	1
29	Intencional	1
30	Interior Rio Grande do Sul; Brasil	1
31	Latino América	1
32	Não possui	1
33	Paraná; São Paulo; Rio de Janeiro	1
34	período entre os anos 2000 e 2007.	1
35	Pesquisa antropológica sobre recepção da mensagem televisiva; significado social da televisão	1
36	pesquisa qualitativa	1
37	publicações científicas.	1
38	Publicações nacionais e internacionais.	1
39	qualitativa	1

Id.	COBERTURA	FREQUÊNCIA
40	Resultados parciais da tese em andamento	1
41	revisão bibliográfica	1
42	Revisão teórica e de literatura sobre bibliometria	1
43	Rio de Janeiro	1
44	Rio Grande do Sul	2
45	São Paulo; Rio de Janeiro; Porto Alegre	1
46	Século XX	2
47	Século XXI	1
48	Serra Gaúcha	1
49	siglo XX	1
50	últimas décadas	1
51	universal	1
TOTAL		61

Fonte: FABRICO

Este elemento se destina à identificação espacial ou temporal do recurso. Os resultados obtidos mostraram que ele foi utilizado em apenas 24 recursos, o que representa 9,3% do total. Devido à diversidade dos valores encontrados, preferiu-se apresentar na Tabela 6 o universo dos resultados.

Considera-se que mais itens poderiam ter valores para este elemento. Destaca-se como exemplos *Impactos Socio-culturais de Telecentros Comunitários: o caso do Telecentro Chico Mendes*, onde cabe a identificação espacial; e *Reconfiguração do mercado editorial brasileiro de livros didáticos no início do século XXI: história das principais editoras e suas práticas comerciais*, poderia contar com identificação de cobertura temporal. Desta forma, evidenciam-se incorreções por dados faltando.

As recomendações também dão preferência ao uso de denominação de período do que a intervalo de datas. Nos dados verifica-se como período *Século XXI*, considerado como corretamente identificado, porém há expressões como *Época moderna e atual*; *últimas décadas*. Nestas últimas, os dados estão incorretos e confusos. Outras não conformidades encontradas foram dados ambíguos, como

nos registros: *Atualidade e Universal*; e dados incorretos, como *Revisão teórica e de literatura sobre bibliometria*; *Resultados parciais da tese em andamento*; que não atendem à definição do elemento Cobertura. Para designação espacial, não foram encontrados erros terminológicos significativos.

No entanto, quanto aos erros de sintaxe, encontraram-se vários que, inclusive, dificultam a recuperação do documento por este elemento. Inicialmente, destaca-se o uso de hífen (-) para indicar a pretensão de deixar o campo vazio, embora o sistema considere como campo com valor. Somente para indicar intervalo de datas, que não é o recomendado, encontraram-se diversas expressões diferentes, com uso de hífen, espaços, barras inclinadas, palavras; alguns exemplos: *1995-2007*; *2000 – 2007*; *2003 a 2005*; *2006/2007*; *Anos 1968 a 1978*; *Anos 30-40*.

Pelo apresentado, o elemento não segue as orientações do DCMI quanto à definição e instrução, assim como também não apresenta dados de qualidade, pois encontraram-se dados faltando, incorretos, confusos e ambíguos.

6.7 FONTE (SOURCE)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 225

Total de Valores: 225

Total de Valores Diferentes: 17

Média de valores por instância (225/225): 1

Os valores para Fonte são discriminados na Tabela 7, conforme se apresentam nos metadados.

Tabela 7 – Metadados do Elemento Fonte

Id.	FONTE	FREQUÊNCIA
1	Em Questão; v. 10, n. 1 (2004)	14
2	Em Questão; v. 10, n. 2 (2004)	16
3	Em Questão; v. 11, n. 1 (2005)	12
4	Em Questão; v. 11, n. 2 (2005)	13
5	Em Questão; v. 12, n. 1 (2006)	11

Id.	FONTE	FREQUÊNCIA
6	Em Questão; v. 12, n. 2 (2006)	12
7	Em Questão; v. 13, n. 1 (2007)	12
8	Em Questão; v. 13, n. 2 (2007)	10
9	Em Questão; v. 14, n. 1 (2008)	10
10	Em Questão; v. 14, n. 2 (2008)	15
11	Em Questão; v. 15, n. 1 (2009)	14
12	Em Questão; v. 15, n. 2 (2009)	16
13	Em Questão; v. 16 (2010): Em Questão - Edição Especial	10
14	Em Questão; v. 16, n. 1 (2010)	15
15	Em Questão; v. 16, n. 2 (2010)	20
16	Em Questão; v. 9, n. 1 (2003)	16
17	Em Questão; v. 9, n. 2 (2003)	9
TOTAL		225

Fonte: FABRICO

Todos os 225 recursos possuem uma única fonte, que é a revista estudada nesta pesquisa. O metadado Fonte identifica o volume e número da edição a que o recurso pertence, de onde resulta o ano de publicação. A frequência para cada um dos valores representa o número de recursos que cada edição disponibiliza e que varia de nove a vinte no período levantado.

As melhores práticas indicam o uso de uma identificação formal da fonte, um URI ou URL, dar preferência a valores não literais e que o critério seja tal qual para o elemento Relação. Pode-se dizer que o valor do registro não é literal, embora também não seja uma citação, que é um tipo de identificação formal. Para o elemento Relação constatou-se o uso do endereço na *Web* que leva ao documento, o que poderia ocorrer também no caso do elemento Fonte.

Assim, para este elemento a adequação foi considerada insatisfatória quanto à escolha dos valores para a representação dos recursos, apresentando apenas coerência interna no elemento.

6.8 IDENTIFICADOR (*IDENTIFIER*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 225

Total de Valores: 225

Total de Valores Diferentes: 225

Média de valores por instância (225/225): 1

Os valores de 12 registros de Identificador são discriminados na Tabela 8 para servirem de exemplo, conforme se apresentam nos metadados.

Tabela 8 – Metadados do Elemento Identificador

Id.	IDENTIFICADOR	FREQUÊNCIA
1	http://seer.ufrgs.br/EmQuestao/article/view/1	1
2	http://seer.ufrgs.br/EmQuestao/article/view/100	1
3	http://seer.ufrgs.br/EmQuestao/article/view/10060	1
4	http://seer.ufrgs.br/EmQuestao/article/view/101	1
5	http://seer.ufrgs.br/EmQuestao/article/view/10387	1
6	http://seer.ufrgs.br/EmQuestao/article/view/10400	1
7	http://seer.ufrgs.br/EmQuestao/article/view/10475	1
8	http://seer.ufrgs.br/EmQuestao/article/view/95	1
9	http://seer.ufrgs.br/EmQuestao/article/view/96	1
10	http://seer.ufrgs.br/EmQuestao/article/view/9976	1
11	http://seer.ufrgs.br/EmQuestao/article/view/9984	1
12	http://seer.ufrgs.br/EmQuestao/article/view/9995	1
TOTAL		225

Fonte: FABRICO

Todos os valores encontrados são diferentes, como era para ser. A cada recurso é dado como valor do elemento Identificador o seu endereço de URI, que fornece o local único e fixo do recurso na *Web*. O que diferencia cada documento da revista é o último número na expressão do URI, que varia de um a cinco caracteres, por exemplo, em <http://seer.ufrgs.br/EmQuestao/article/view/101> é o 101.

As orientações do DCMI são para que o Identificador não deixe dúvidas nem suscite ambiguidades sobre a identificação do recurso. A orientação é seguida ao utilizar-se o URI. Porém, há a possibilidade de utilizar outros recursos, como a referência do documento. Cabe destacar que todos os endereços testados levaram ao documento identificado.

Pelo exposto, este elemento é considerado que corresponde às definições e instruções do DCMI, assim como aos critérios de qualidade.

6.9 RELAÇÃO (*RELATION*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 18

Total de Valores: 65

Total de Valores Diferentes: 65

Média de valores por instância (65/18): 3.611111111111

Os valores de 10 registros de Relação são discriminados na Tabela 9 para servirem de exemplo, conforme se apresentam nos metadados.

Tabela 9 – Metadados do Elemento Relação

Id.	RELAÇÃO	FREQUÊNCIA
1	http://seer.ufrgs.br/EmQuestao/article/download/10815/35528	1
2	http://seer.ufrgs.br/EmQuestao/article/download/10815/35529	1
3	http://seer.ufrgs.br/EmQuestao/article/download/10815/35530	1
4	http://seer.ufrgs.br/EmQuestao/article/download/10815/35531	1
5	http://seer.ufrgs.br/EmQuestao/article/download/10815/35532	1
6	http://seer.ufrgs.br/EmQuestao/article/download/10815/35533	1
7	http://seer.ufrgs.br/EmQuestao/article/download/10815/41956	1
8	http://seer.ufrgs.br/EmQuestao/article/download/16073/56799	1
9	http://seer.ufrgs.br/EmQuestao/article/download/12975/44149	1
10	http://seer.ufrgs.br/EmQuestao/article/download/12975/44150	1
11 - 65	Demais valores	55
TOTAL		65

Fonte: FABRICO

Os resultados apresentaram 18 recursos com registro, ou seja, 8% do total. Se a intenção é colocar outros recursos com assuntos semelhantes ou relacionar com os demais recursos da mesma edição, em ambos os casos, o número de recursos com registros deveria ser maior ou a totalidade deles.

No entanto, o intuito do elemento, pelo DCMI, é relacioná-los com outras versões do mesmo recurso, com um software que permita ler o documento, como o caso do *Adobe Reader* para visualizar arquivos pdf, e outras relações similares.

Na Tabela 9 são apresentados alguns valores, onde os documentos são relacionados por um endereço na *Web*. No entanto, nenhum daqueles que foram verificados retornaram algum recurso publicado na Revista Em Questão, pois estes endereços não são válidos.

Os endereços dos documentos relacionados apresentam caracteres em comum. Toma-se, por exemplo, o seguinte valor de Relação:

<http://seer.ufrgs.br/EmQuestao/article/download/16073/56799>

Esse registro encontra-se no elemento Relação de <http://seer.ufrgs.br/EmQuestao/article/view/16073>. Percebe-se que os caracteres numéricos desta URI são iguais aos primeiros números daquele registro. Na Tabela 9 há sete recursos relacionados a <http://seer.ufrgs.br/EmQuestao/article/view/10815>. Por esta característica, entende-se o motivo de não existirem valores iguais, nem válidos como endereço particular de um documento, pois as relações são identificadas pela origem e destino.

Pelo exposto, os metadados de Relação não estão de acordo com as orientações do DCMI, seja pela definição ou sintaxe, nem com os critérios de qualidade. Este elemento precisa ser reformulado.

6.10 FORMATO (*FORMAT*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 225

Total de Valores: 264

Total de Valores Diferentes: 4

Média de valores por instância (264/225): 1.173333333333

Os valores para Formato são discriminados na Tabela 10, conforme se apresentam nos metadados.

Tabela 10 – Metadados do Elemento Formato

Id.	FORMATO	FREQUÊNCIA
1	application/pdf	171
2	application/force-download	54
3	text/html	38
4	audio/mpeg	1
TOTAL		264

Fonte: FABRICO

Todos os recursos têm um valor para o elemento Formato, pois refere-se à manifestação física ou digital do recurso. Os valores diferentes são quatro e há recursos que apresentam mais de um valor. Este é o caso dos arquivos que são disponibilizados em pdf abrindo na mesma tela, *application/pdf*, ou que baixam automaticamente, *application/force-download*; e os que possuem as opções de baixar automaticamente o pdf ou acesso em HTML, *text/html*. Só que foi identificado um caso em que ao escolher a opção *Baixar este arquivo pdf*, o arquivo recebido era em formato txt, ilegível. Isto ocorreu para o artigo *Base de Dados como Elemento Facilitador para o Tratamento dos Dados Terminológicos* no volume 10, número 2 de 2004.

O *áudio/mpeg* é uma entrevista ao Prof. Dr. Antonio Hohlfeldt disponibilizada em arquivo mp3 no volume 14, número 1 de 2008.

A designação do formato segue as especificações do DCMI, pois constam no MIME *Media Types* onde também estão os padrões *Request for Comments* (RFC) para aplicação, imagem, exemplos, texto, som, entre outros (IANA, 2007).

Para este elemento considera-se que atende a definição e instrução do DCMI, porém não está livre de não conformidades com relação ao ambiente ao oferecer um arquivo em formato pdf e entregar em formato txt impróprio para uso.

6.11 TIPO (*TYPE*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 141

Total de Valores: 155

Total de Valores Diferentes: 17

Média de valores por instância (155/141): 1.09929078014

Os valores para Tipo são discriminados na Tabela 11, conforme se apresentam nos metadados.

Tabela 11 – Metadados do Elemento Tipo

Id.	TIPO	FREQUÊNCIA
1	Avaliado por Pares	139
2	-	1
3	Abordagem dialética	1
4	Análise das publicações do Jornal Zero Hora sobre o turismo da Serra Gaúcha e suas influências	1
5	Análise de conteúdo	1
6	Análise de conteúdo; entrevistas	1
7	Análise de discurso	1
8	Análise Semiótica	1
9	Análise teórica, histórica e epistemológica	1
10	descritivo-interpretativo; dialética; fenomenologia	1
11	Informação estatística e entrevistas à profundidade	1
12	Interpretação	1
13	Método qualitativo; etnografia de audiência; estudos de recepção; interpretações mensagem televisão	1
14	Pesquisa de fontes empíricas	1
15	Qualitativo	1
16	Revisão de literatura	1

Id.	TIPO	FREQUÊNCIA
17	trabalho de campo com observação participante e entrevistas	1
TOTAL		155

Fonte: FABRICO

Encontrou-se apenas 141 itens descritos com este elemento, correspondendo um único valor para cada recurso. No entanto, todos os recursos possuem um tipo. Os valores diferentes constam na Tabela 11 e mostram uma diversidade de termos para identificar o tipo do recurso, o que reflete os diversos critérios de interpretação do elemento.

Os valores são relativos à natureza de pesquisa, *qualitativa*; ao local de realização e meios utilizados, *pesquisa de campo com observação participante e entrevistas*; ou até uma descrição do conteúdo, *Análise das publicações do Jornal Zero Hora sobre o turismo da Serra Gaúcha e suas influências*.

Segundo o DCMI este elemento informa sobre a natureza ou gênero do conteúdo do recurso e que deve dar-se preferência ao uso de um termo que conte no DCMI *Type Vocabulary*. Segundo este vocabulário controlado, o termo que corresponde a todos os recursos é Texto, com exceção de uma entrevista que é Som. No vocabulário para Tipo, define-se texto como um recurso composto prioritariamente por palavras para serem lidas e traz como exemplos: livros, cartas, dissertações, poemas, periódicos, artigos, arquivos de listas de correio (DCMI, 2010).

Assim, considera-se que neste elemento não são seguidas as recomendações do DCMI, principalmente quanto a sua definição. Os valores estão fora dos padrões de qualidade intrínseca, contextual e representacional, o que poderá causar confusão na interpretação por parte dos usuários.

6.12 DATA (DATE)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 225

Total de Valores: 225

Total de Valores Diferentes: 12

Média de valores por instância (225/225): 1

Os valores para Data são discriminados na Tabela 12, conforme se apresentam nos metadados.

Tabela 12 – Metadados do Elemento Data

Id.	DATA	FREQUÊNCIA
1	2006-12-09	80
2	2006-12-10	11
3	2007-03-02	12
4	2007-10-05	12
5	2008-01-24	10
6	2008-09-29	10
7	2009-03-13	15
8	2009-09-02	14
9	2010-03-10	16
10	2010-07-21	15
11	2010-11-09	10
12	2010-12-20	20
TOTAL		225

Fonte: FABRICO

O elemento Data está presente em todos os 225 recursos disponíveis. Os registros estão codificados ano-mês-dia (AAAA-MM-DD) conforme as recomendações do *Dublin Core* para este campo. Pelos valores encontrados, pode-se inferir que trata-se da data de inserção dos dados na base, resultado de um esforço inicial em 2006 inserindo 91 registros com dados de anos anteriores ou a data de migração do sistema, e depois apenas mantendo a base atualizada.

As recomendações indicam o uso de colchetes caso a data for atribuída pelo catalogador. Como o campo se refere a datas do ciclo de vida do documento, e há outras datas que são importantes para o usuário, como a data da publicação do documento, seria interessante o uso do refinamento do elemento.

Portanto, considera-se que o elemento atende as recomendações quanto à definição e sintaxe, no entanto a falta de informação que a data é de inclusão na base pode causar problemas de interpretação. Isto pode ser classificado como falhas na qualidade contextual, ao não considerar o que os usuários precisam, não são relevantes; e representacional pela escolha do valor do registro.

6.13 IDIOMA (*LANGUAGE*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 213

Total de Valores: 213

Total de Valores Diferentes: 3

Média de valores por instância (213/213): 1

Os valores para Idioma são discriminados na Tabela 13, conforme se apresentam nos metadados.

Tabela 13 – Metadados do Elemento Idioma

Id.	IDIOMA	FREQUENCIA
1	<u>pt</u>	206
2	<u>es</u>	5
3	<u>en</u>	2
TOTAL		213

Fonte: FABRICO

O metadado idioma tem valores preenchidos para 213 itens. A cada um deles corresponde apenas um valor, ou seja, o documento está escrito em português, inglês ou espanhol. No entanto, este elemento deveria estar presente em todos os recursos, mas na maioria dos Editoriais e Expedientes não está.

O metadado está correto desde o ponto de vista conceitual, pois as recomendações do DCMI é que expresse o idioma do conteúdo intelectual do documento. Quanto à sintaxe, as recomendações são no sentido de usar a norma ISO 639-3:2007. No documento OLAC *Language Extension* de 2008, a instrução é o

uso desse código de três letras para a representação do idioma (OLAC, 2008). Neste caso, os metadados da Em Questão seriam por, spa e eng, mas o mesmo documento informa que os programas que colhem os metadados podem fazer a interpretação do código de duas letras e transformá-los em três letras.

Assim, o elemento Idioma é considerado correto para os itens descritos, mas há 12 itens sem especificação, portanto, há dados faltando.

6.14 EDITOR (*PUBLISHER*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 225

Total de Valores: 225

Total de Valores Diferentes: 1

Média de valores por instância (225/225): 1

Os valores para Editor são discriminados na Tabela 14 para servirem de exemplo, conforme se apresentam nos metadados.

Tabela 14 – Metadados do Elemento Editor

Id.	EDITOR	FREQUENCIA
1	Universidade Federal do Rio Grande do Sul	225
TOTAL		225

Fonte: FABRICO

Neste elemento, os registros são todos idênticos, independente do tipo do item que representam. Todos eles identificam a Universidade Federal do Rio Grande do Sul.

O elemento Editor é quem disponibiliza o recurso, quem propicia seu acesso. A revista estudada é editada pela Faculdade de Biblioteconomia e Comunicação dessa universidade. Neste caso foi selecionado como Editor a instituição maior relacionada ao recurso.

As orientações das melhores práticas somente dizem que o valor deve ser passível de ordenação alfabética, como é o caso. Sendo assim, considera-se correta

a escolha do valor do elemento, além de atender aos requisitos para ser considerado um elemento com dados de qualidade.

6.15 DIREITOS AUTORAIS (*RIGHTS*)

Os dados oferecidos pelo programa FABRICO são:

Total de Instâncias: 225

Total de Valores: 225

Total de Valores Diferentes: 1

Média de valores por instância (225/225): 1

Os valores para Direitos Autoriais são discriminados na Tabela 15, conforme se apresentam nos metadados.

Tabela 15 – Metadados do Elemento Direitos Autorais

Id.	DIREITOS AUTORAIS	FREQUENCIA
1	Copyright for articles published in this journal is retained by the authors, with first publication rights granted to the journal. By virtue of their appearance in this open access journal, articles are free to use, with proper attribution, in educational and other non-commercial settings.	225
TOTAL		225

Fonte: FABRICO

Os direitos autorais de todo o material veiculado na revista está sob o mesmo regime, portanto, em todos os itens consta a mesma orientação para o uso da informação disponibilizada. Nos 225 registros a mensagem está em inglês e diz que os direitos intelectuais são dos autores, que cedem para a revista o seu ineditismo e que os artigos podem ser utilizados para fins educacionais e não comerciais.

Sendo assim, o Total de Valores Diferentes é um e, conseqüentemente, a Média de Valores por Instância também é.

Desde o ponto de vista das recomendações do DCMI o texto é adequado, trazem as informações necessárias para o uso do documento e salientam que a

propriedade intelectual é do autor. Não há necessidade de refinamento e, quanto à grafia, apresenta-se correta.

Em suma, este metadado é considerado adequado, padronizado e conforme o esperado.

7 CONCLUSÕES

Neste estudo foi abordada a qualidade dos metadados da Revista Em Questão, permitindo assim investigar e analisar sua adequação às recomendações do *Dublin Core Metadata Initiative* (DCMI). A análise foi subsidiada pelas teorias de qualidade de dados, principalmente pela abordagem de suas categorias e pela não conformidade. As leituras sobre verificação da qualidade de dados apontam os trabalhos pioneiros de Wang e Strong (1996) que foram utilizados e adaptados para as mais diversas aplicações. Tratou-se ainda do padrão de metadados do DCMI que disponibiliza material abundante na sua página na *internet*, porém de forma não linear.

Assim, neste sentido este trabalho se torna válido por apresentar as informações necessárias para tratar sobre metadados numa sequência lógica que auxilia na compreensão do assunto e facilita no planejamento para adotar este sistema de metadados.

O ambiente FABRICO se mostrou apropriado para o propósito, tanto para a coleta como para o exame dos metadados, visto que ele permite selecionar cada um deles por elemento, por edição da Revista e ainda fornece relatórios com dados estatísticos. O instrumento para análise baseado na Folha de Verificação de inspeção facilitou o procedimento e garantiu que todos os aspectos quanto às categorias de qualidade fossem considerados.

Os resultados apontaram deficiências na política de metadados da Revista, uma vez que não há padronização na adoção do DCMI, nem quanto à definição do elemento nem no que se refere à instrução de uso. Também foram constatados metadados com registros faltando, ambíguos e com incoerências no tratamento ao longo do tempo. As principais deficiências quanto à definição foram no uso do elemento Criador, Colaborador e Relação. No que tange à instrução, os principais não conformidades foram quanto ao elemento Cobertura, onde o período estava registrado de forma variada.

Recomenda-se o estabelecimento de uma política de metadados que esteja alinhada com a estratégia da Revista no sentido de tornar seu conteúdo visível, acessível e pronto para uso. Pois as revistas são componentes importantes no fluxo de comunicação científica e, neste caso, o de uma instituição consolidada como é a UFRGS espera-se que seja tida como de referência.

Ressalta-se também a importância de instituir o treinamento dos envolvidos para a melhoria da qualidade dos metadados e, por consequência, a da própria Revista. Este tipo de ação pode ser pensada de forma isolada ou conjunta com outras revistas hospedadas no SEER. Deste modo, abre-se espaço para reforçar o trabalho em colaboração institucional.

Seria interessante uma continuação da pesquisa na melhoria da qualidade no sentido da prevenção. Uma opção seria investigar o SEER para conhecer o quanto ele contribui de forma positiva ou negativa na qualidade dos metadados.

A experiência desta pesquisa permite valorizar o quanto pode-se aprender quando são identificadas as não conformidades, pois elas direcionam ao aperfeiçoamento e demonstram que sempre há espaço para melhorias. Acredita-se que é através de trabalhos deste tipo, desenvolvidos periodicamente, que permitirão elevar a qualidade dos metadados da Revista.

REFERÊNCIAS

ALVARENGA, Lídia. Representação do Conhecimento na Perspectiva da Ciência da Informação em Tempo e Espaço Digitais. **Encontros Bibli**, Florianópolis, v. 8, n. 15, 1º semestre, p. 18-40, 2003. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/viewFile/97/5233>>. Acesso em: 10 dez. 2010.

ALVES, Jaqueline Costa. **Organização e Representação da Informação na Biblioteca Digital de Teses e Dissertações da Universidade Federal do Estado de Santa Catarina (UDESC)**: uma análise focada em metadados sob a luz do padrão MTB-BR. Dissertação. (Mestrado em Ciência da Informação) Universidade Federal de Santa Catarina (UFSC). Orientadora: Ligia Maria Arruda Café. Florianópolis, 2009. 161 f. Disponível em: <<http://pgcin.paginas.ufsc.br/files/2010/10/ALVES-Jaqueline-Costa.pdf>>. Acesso em: 28 ago. 2011

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14.724:2011** Informação e documentação – Trabalhos acadêmicos – Apresentação. Rio de Janeiro, 2011, 11 p.

BRUCE, Thomas R.; HILLMANN, Diane I. The Continuum or Metadata Quality: defining, expressing, exploiting. In.: HILLMANN, D.; WESTBROOKS, E. **Metadata in Practice**. Chicago: ALA, 2004. 16 p. Disponível em: <<http://ecommons.library.cornell.edu/handle/1813/7895>>. Acesso em: 26 jun. 2011.

BUFREM, Leilah Santiago. Revistas Científicas: saberes no campo da ciência da informação. In.: POBLACIÓN, Dinah Aguiar; WITTER, Geraldina Porto; SILVA, José Fernando Modesto da. **Comunicação & Produção Científica**: contexto, indicadores e avaliação. São Paulo: Aditora, 2006. p. 191-214

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO (CNPq). Página oficial. Disponível em: <<http://www.cnpq.br/>>. Acesso em: 09 ago. 2011.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). Página oficial. Disponível em: <<http://www.capes.gov.br/>>. Acesso em: 27 jun. 2011

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Tabela de Áreas de Conhecimento**. [documento *on-line*, sem paginação], 2009. Disponível em: <<http://www.capes.gov.br/avaliacao/tabela-de-areas-de-conhecimento>> Acesso em: 27 jun. 2011

DUBLIN CORE METADATA INICIATIVE (DCMI). **DCMI Type Vocabulary**. [documento *on-line*, sem paginação], 2010. Disponível em: <<http://dublincore.org/documents/dcmi-type-vocabulary/>> Acesso em: 23 ago. 2011

DUSHAY, Naomi; HILLMANN, Diane. Analyzing Metadata for Effective Use and Re-Use. In.: DC 2003. In: **Proceedings of the international DCMI metadata conference and workshop**. Seattle: DCMI, 2003. Disponível em: <<http://ecommons.cornell.edu/handle/1813/7896>>. Acesso em: 12 dez. 2010. 10 p.

DZIEKANIAK, Gisele Vasconcelos. A Organização da Informação e a Comunicação Científica: implicações para os profissionais e usuários da informação. **Em Questão**. Porto Alegre, v. 16, n. 1, p. 45-59, 2010. Disponível em: <<http://seer.ufrgs.br/EmQuestao/article/view/11434>>. Acesso em: 18 jul. 2011.

FERREIRA, Ana Gabriela Clipes; CAREGNATO, Sônia Elisa. A editoração Eletrônica de revistas científicas brasileiras: o uso de SEER/OJS. **TransInformação**, Campinas, v. 20, n. 2, p. 171-180, maio/ago., 2008. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=227>>. Acesso em: 27 out. 2011

GIL, Antonio Carlos. **Como Elaborar Projetos de Pesquisa**. São Paulo: Atlas, 1991.

GONÇALVES, Andréa; RAMOS, Lucia Maria S. V. Costa; CASTRO, Regina C. Figueiredo. In.: POBLACIÓN, Dinah Aguiar; WITTER, Geraldina Porto; SILVA, José Fernando Modesto da. **Comunicação & Produção Científica**: contexto, indicadores e avaliação. São Paulo: Aditora, 2006. p. 163-190

GUERRA-GARCIA, César A. Introducción de Dimensiones de Calidad de Datos (DQ) como requisitos añadidos en el proceso de desarrollo de un producto de software. In: **Calidad y Medición de Sistemas de Información**. La Mancha: Escuela Superior de Informática, 2009. 17 p. Disponível em: <<http://alarcos.inf-cr.uclm.es/doc/cmsi/trabajos/Cesar%20Guerra.pdf>>. Acesso em: 25 ago. 2011

GUERRA-GARCIA, César; CABALLERO, Ismael; GUZMÁN, Ignacio García-Rodríguez de; PIATTINI, Mário. Modelado de Requisitos de Calidad de Datos en el Proceso de Desarrollo de Portales Web. In.: **Actas de los Talleres de las Jornadas de Ing. del Software y BBDD**, Ciudad Real, v. 3, n. 2, p. 124-133. 2009. Disponível em: <<http://www.sistedes.es/TJISBD/Vol-3/No-2/articles/DSDM-09-guerra-mrc.pdf>>. Acesso em: 29 ago. 2011.

GUY, Marieke; POWELL, Andy; DAY, Michael. Improving the Quality of Metadata in Eprint Archives. **Ariadne**, Bath, n. 38, [documento *on-line*, sem paginação], jan. 2004. Disponível em: Disponível em: <<http://www.ariadne.ac.uk/issue38/guy/>>. Acesso em: 09 dez. 2010.

IKEMATU, Ricardo Shoit. Gestão de Metadatos: sua evolução na Tecnologia da Informação. **DataGramZero** Revista de Ciência da Informação. Rio de Janeiro, v. 2; n. 6; p. 1-8, dez. 2001. Disponível em: <http://www.dgz.org.br/dez01/Art_02.htm>. Acesso em: 04 jun. 2011.

INTERNET ASSIGNED NUMBERS AUTHORITY (IANA). **MIME Media Types**, [documento *on-line*, sem paginação], 2007. Disponível em:

<<http://www.iana.org/assignments/media-types/index.html>>. Acesso em: 14 set. 2011.

KURAMOTO, Helio. Acesso Livre ou Acesso Aberto? Eis a questão. In.: **Blog do Kuramoto**. 2010. [documento *on-line*, sem paginação]. Disponível em: <<http://kuramoto.blog.br/2010/04/13/acesso-livre-ou-acesso-aberto-eis-a-questao/>>. Disponível em: <<http://kuramoto.blog.br/2010/04/13/acesso-livre-ou-acesso-aberto-eis-a-questao/>> Acesso em: 20 out. 2011.

LARA, Marilda Lopes Ginez. Termos e Conceitos da Área de Comunicação e Produção Científica. In: POBLACIÓN, Dinah Aguiar; WITTER, Geraldina Porto; SILVA, José Fernando Modesto da. **Comunicação & Produção Científica: contexto, indicadores e avaliação**. São Paulo: Aditora, 2006. p. 387-414

LIMA, Gercina Ângela Borém. Interfaces entre a ciência da informação e a ciência cognitiva. **Ciência da Informação**, Brasília, v. 32, n. 1, p. 77–87, jan./abr. 2003. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652003000100008&lng=pt&nrm=iso>. Acesso em: 25 jun. 2010.

LINDOSO, Felipe. **Metadata**, ou quando os editores brasileiros comem moscas, cobras e lagartos. [documento *on-line*, sem paginação, 2011]. Disponível em: <<http://ebookpress.wordpress.com/2011/07/27/metadata-ou-quando-os-editores-brasileiros-comem-moscas-cobras-e-lagartos/>>. Acesso em: 15 out. 2011.

MARCONDES, Carlos Henrique; SAYÃO, Luís Fernando. Integração e Interoperabilidade no Acesso a Recursos Informacionais Eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira. **Ciência da Informação**. Brasília, v. 30, n. 3, p. 24-33, set./dez., 2001. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652001000300004> Acesso em: 25 set. 2011.

MEADOWS, A. J. **A Comunicação Científica**. Brasília: Briquet de Lemos, 1999.

MENDONÇA, Maria de Fátima Almeida. **O Moderno Profissional da Informação e seu Perfil Face aos Novos Tempos**. [documento *on-line*, sem paginação, 2006]. Disponível em: <<http://www.biblioteca.sebrae.com.br/>>. Acesso em: 25 jun. 2010.

MUELLER, Suzana Pinheiro Machado. O Periódico Científico. In.: CAMPELLO, Bernadete Santos; CENDÓN, Beatriz Valadares; KREMER, Jeannette Merguerite (Org). **Fontes de Informação para Pesquisadores e Profissionais**. Belo Horizonte: UFMG, 2000. p. 73-95.

OLAC, **Open Language Archives Community**. [documento *on-line*, sem paginação], 2008. Disponível em <<http://www.language-archives.org/REC/language-20080222.html>>. Acesso em: 13 ago. 2011.

PACKER, Abel; MENEGHINI, Rogério. Visibilidade da Produção Científica. In: POBLACIÓN, Dinah Aguiar; WITTER, Geraldina Porto; SILVA, José Fernando Modesto da. **Comunicação & Produção Científica: contexto, indicadores e avaliação**. São Paulo: Aditora, 2006. p. 235-259

PAVAN, Cleusa. **Práticas Sociais na Comunicação Científica**: a avaliação pelos pares nas revistas brasileiras de Ciência da Informação. 2008. 171 f. Dissertação (Mestrado em Comunicação e Informação) Universidade Federal do Rio Grande do Sul, (UFRGS). Porto Alegre, 2008. Disponível em: <<http://www.lume.ufrgs.br/handle/10183/13798>>. Acesso em: 29 set. 2011.

REITZ, Joan M. **On-line Dictionary for Library and Information Science (ODLIS)**. 2010. Disponível em: <<http://www.abc-clio.com/ODLIS/searchODLIS.aspx>>. Acesso em: 30 out. 2011.

ROCHA, Rafael Port da. Metadados, *Web Semântica*, Categorização Automática: combinando esforços humanos e computacionais para a descoberta e uso dos recursos da *web*. **Em Questão**, Porto Alegre, v. 10, n. 1, p. 109-121, jan./ jun. 2004. Disponível em: <<http://seer.ufrgs.br/EmQuestao/article/view/86/1127>>. Acesso em: 23 jun. 2011.

ROCHA, Rafael Port da. Ambiente FABRICO. Disponível em: <<http://www6.ufrgs.br/fabrico/f2/notes/tipos.php>>. Acesso em: 10 dez. 2010.

ROSETTO, Marcia; NOGUEIRA, Adriana Hypólito. **Aplicação de Elementos Metadados Dublin Core para descrição de Dados bibliográficos On-Line da Biblioteca Digital de Teses da USP**. 14 p. Disponível em: <<http://www.sibi.ufrj.br/snbu2002/oralpdf/82.a.pdf>>. Acesso em: 26 set. 2011.

SAYÃO, Luis Fernando; MARCONDES, Carlos Henrique. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. **TransInformação**, Campinas, v. 20, n. 2 p. 133-148, maio/ago., 2008. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=215>>. Acesso em: 26 out. 2011.

SEER. Página oficial do Sistema Eletrônico de Editoração de Revistas. Disponível em: <<http://seer.ibict.br/>>. Acesso em: 24 mar. 2011.

SENÉ, María Luisa. Los Metadatos y su Lugar en la Arena Internacional. In.: **CONGRESO INTERNACIONAL DE INFORMACIÓN (INFO)**, 5. abr. 2004, La Habana, 8 p. Disponível em: <<http://www.mendeley.com/research/los-metadatos-y-su-lugar-en-la-arena-internacional/>>. Acesso em: 22 out. 2011.

SENSO, José A.; PIÑERO, Antonio de la Rosa. El Concepto de Metadato. Algo más que descripción de recursos electrónicos. **Ciência da Informação**. Brasília, v. 32, n. 2, p. 95-106, maio/ago. 2003. Disponível em: <<http://www.scielo.br/pdf/ci/v32n2/17038.pdf>>. Acesso em: 6 ago. 2011.

STUMPF, Ida Regina C. Avaliação das Revistas de Comunicação pela Comunidade Acadêmica da área. **Em Questão**. Porto Alegre, v. 9, n. 1, p. 25-38, jan./jun., 2003. Disponível em: <<http://seer.ufrgs.br/EmQuestao/article/view/57>>. Acesso em: 14 jul. 2011.

STRONG, D. M.; LEE, Y. W.; WANG, R. Y. Data Quality in Context. **Communications of the ACM**, v. 40, n. 5, p. 103-110, 1997. Disponível em: <<http://www.fing.edu.uy/inco/cursos/caldatos/articulos/StrongLeeWangCACMMay97.pdf>>. Acesso em: 10 ago. 2011.

VALÉRIO, Palmira Mariconi. Da Comunicação Científica à Divulgação. **TransInformação**, Campinas, v. 20, n. 2, p. 159-169, maio/ago., 2008. Disponível em: <<http://ridi.ibict.br/bitstream/123456789/28/1/Transinformacao2008Pinheiro.PDF>>. Acesso em: 03 nov. 2011.

VERBO, E.; CABALLERO, I.; PÉREZ, R.; CALERO, C.; PIATTINI, M. MEPLAMECAL: a methodology based on ISO/IEC 15939 to elaborate data quality measurement plants. **IEEE Latin American Transactions**. v. 7, n. 3. jul. 2009. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5336635>. Acesso em: 22 set. 2011.

VANTI, Nadia. **Links Hipertextuais na Comunicação científica**: análise webométrica dos sítios acadêmicos latino-americanos em Ciências Sociais. Porto Alegre, 2007. 293 p. Tese de Doutorado em Comunicação e Informação. Universidade Federal do Rio Grande do Sul (UFRGS). Disponível em: <<http://www.lume.ufrgs.br/handle/10183/11175>> Acesso em: 13 ago. 2011.

WAND, Yair; WANG, Richard Y. Data Quality Dimensions in Ontological Foundations. **Communications of the ACM**. v. 39, n. 11, nov. 1996. Disponível em: <<http://www.mendeley.com/research/anchoring-data-quality-dimensions-in-ontological-foundations-2/>> Acesso em: 02 set. 2011.

WANG, Richard Y; STRONG, Diane M. Beyond Accuracy: what data quality means to data consumers. **Journal of Management Information Systems**, v. 12, n. 4, p. 5-33, spring, 1996. Disponível em: <[http://w3.cyu.edu.tw/ccwei/PAPER/ERP/data%20quality\(JMIS\).pdf](http://w3.cyu.edu.tw/ccwei/PAPER/ERP/data%20quality(JMIS).pdf)> Acesso em: 29 ago. 2011

WEITZEL, Simone. **Os repositórios de e-prints como nova forma de organização da produção científica**: o caso da área das Ciências da Comunicação no Brasil. Tese. (Doutorado em Ciências da Informação). Orientadora Prof. Dr. Sueli Mara dos Santos Pinto Ferreira. Universidade de São Paulo (USP). 2006. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/27/27151/tde-14052009-133509/pt-br.php>>. Acesso em: 25 set. 2011.

ZIMAN, John. **Conhecimento Público**. Belo Horizonte: Itatiaia/São Paulo: USP, 1979. 164 p. (Série O Homem e a Ciência, v.8)