

MURILO MILTON MACHADO

OPEN ARCHIVES: panorama dos repositórios

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina como requisito parcial à obtenção do grau de Mestre em Ciência da Informação.

Florianópolis, 2006

MURILO MILTON MACHADO

OPEN ARCHIVES: panorama dos repositórios

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina como requisito parcial à obtenção do grau de Mestre em Ciência da Informação.

Orientador: Prof. Dr. Angel Freddy Godoy Viera

Florianópolis, 2006

MURILO MILTON MACHADO

OPEN ARCHIVES: panorama dos repositórios

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina como requisito parcial à obtenção do grau de Mestre em Ciência da Informação.

Orientador: Prof. Dr. Angel Freddy Godoy Viera

Prof^a. Dra. Ursula Blattmann

Prof^a. Dra. Patrícia Zeni Marchiori

Florianópolis, 2006

Pode alguém andar errado de diversos modos, porém, apenas um caminho pode chegar à verdade; por isto, é fácil fracassarmos e difícil conquistar a vitória.

Aristóteles

RESUMO

MACHADO, Murilo Milton. **Open Archives**: panorama dos repositórios. 2006, 101 f. Dissertação (Mestrado em Ciência da Informação). Universidade Federal de Santa Catarina, Florianópolis, 2006.

A presente dissertação possui como objetivo principal identificar o panorama dos *Open Archives* (Arquivos Abertos) na comunidade científica, como uma nova alternativa de comunicação apoiada nas tecnologias oferecidas pela *Open Archives Initiative* – OAI (Iniciativa de Arquivos Abertos). Para isso, realizou-se um levantamento sobre informações que tratam do histórico, da motivação e das instituições que oferecem suporte a *Open Archives Initiative* - OAI. Classificaram-se os *Open Archives* conforme o tipo de arquivo seguindo um padrão disponível no site oficial da organização Eprints.Org. Identificou-se as comunidades que adotaram as tecnologias da *Open Archives Initiative* - OAI, as principais tecnologias envolvidas e os objetivos para criação dos *Open Archives* participantes de cada classe. Descreveu-se o fluxo de informação dos *Open Archives* coletando e analisando dados sobre a disponibilidade de acesso ao texto completo e possibilidade de interação dos usuários com os autores. Verificou-se a variação do número de participantes e tempo de existência dos *Open Archives*. Por intermédio de uma pesquisa bibliográfica e dados coletados diretamente na amostra evidenciou-se alguns fatores que levaram ao surgimento dos *Open Archives*. A pesquisa foi realizada em *Open Archives* nacionais e internacionais que utilizavam o protocolo OAI-PMH para disponibilizarem seus metadados. A metodologia utilizada para realização da pesquisa foi composta por um conjunto de etapas e procedimentos envolvendo o método de dedução científica, uma pesquisa descritiva, a definição da população e amostra, a obtenção dos dados, a aplicação dos instrumentos de coleta de dados, o processamento dos dados e a análise exploratória dos dados. Com base nos resultados apresentados na pesquisa, concluiu-se que a utilização dos *Open Archives* como nova alternativa de comunicação científica apresenta pontos bastante positivos no cenário nacional e internacional. Entre eles possibilita de maneira prática e eficiente a criação de espaços virtuais para o compartilhamento do conhecimento científico entre os pesquisadores; oferece oportunidades para instituições com baixos recursos em investimentos criarem seus *Open Archives* para divulgarem e disponibilizarem seus trabalhos; inseriu em âmbito internacional uma nova solução tecnológica criando sistemas de informação interoperáveis de fácil aplicabilidade em um universo disperso em dados e informações como a Internet; criou uma nova filosofia tecnológica em relação à comunicação e produção do conhecimento científico; possibilitou a disseminação da informação em âmbito internacional, aproximando pesquisadores de várias áreas da ciência dispersos geograficamente; o protocolo OAI – PMH está se tornando parte da infraestrutura da *Web* de forma tão importante como o protocolo HTTP é atualmente; e os *Open Archives* que fizeram parte da amostra atendem comunidades específicas e com alcance direcionado a comunidades locais.

Palavras-chave: Arquivos abertos. Comunicação científica. Comunidade científica. Metadados. Repositório digital.

ABSTRACT

MACHADO, Murilo Milton. **Open Archives**: panorama dos repositórios. 2006, 101 f. Dissertação (Mestrado em Ciência da Informação). Universidade Federal de Santa Catarina, Florianópolis, 2006.

Present dissertation has principal objective to investigation Identify the panorama of Open Archives in the science community, how a new alternative of communication sustain in the technologies offering by Open Archives Initiative – OAI. That’s just it, to realized a survey about information to treat of history, motivation and institutions that offering support the OAI. Classified the Open Archives accordant the type of archive. Make identical the community that adopt the technologies OAI, principals involve technologies and the objectives of Open Archives partnership class every. Make a description of information flow of the Open Archives collecting and analyzing data about available of access at complete text and the possibility interaction of the users’ whit authors. Make verify the variation of partnership number and existence time of the Open Archives. Across of a research bibliography and directing data collection in the sample to make evidence any factories that take at arise of Open Archives. Resource was realized at Open Archives internationals and nationals that use protocol OAI-PMH to availability your metadata. Methodology utilized to realization research was composition by a stage whole and proceeding that involve method science deduction, description research, population and sample definition, data obtainment, instrument collection of data application, data processing and data analysis exploratory. With base in the results presented research concluded that the utilization of Open Archives how a new alternative of science communication present enough positive item in the international and national scenery. Between they, allow way efficient and practice the creation of virtual space for participate among researcher; offer opportunity for institutions with recourse little to creation theirs own Open Archives to theirs works divulge and availability; introduced in international ambit a new solution technology to creating information systems interoperate of applicable easy; created a new technology philosophy for communication and production of science knowledge; allowed dissemination of information in international ambit, bring near researchers of the large quantity science area dispersing geographic; OAI-PMH protocol is making part Web framework as important as the HTTP protocol is actually; and the Open Archives that make part of the sample attend specifics communities and with reach conduced to locals communities.

Key Word: Open Archives. E-prints. Metadata. Science community. Science communication. Repository digital.

RESUMEN

MACHADO, Murilo Milton. **Open Archives**: panorama dos repositórios. 2006, 101 f. Dissertação (Mestrado em Ciência da Informação). Universidade Federal de Santa Catarina, Florianópolis, 2006.

La presente disertación tiene como objetivo principal identificar el panorama de los *Open Archives* (Ficheros Abiertos) en la comunidad científica, como una nueva alternativa de comunicación apoyada en las tecnologías ofrecida por la *Open Archives Initiative – OAI* (Iniciativa de Ficheros Abiertos). Para eso, realizó-se un levantamiento sobre informaciones que tratan del histórico, de la motivación y de las instituciones que ofrecen soporte para la *Open Archives Initiative – OAI*. Clasificarán-se los *Open Archives* conforme lo tipo de fichero segundo un padrón disponible el lo *site* oficial de la organización Eprints.Org. Identificó-se las comunidades que adoptaran las tecnologías de la *Open Archives Initiative – OAI*, las principales tecnologías y los objetivos para creación de los *Open Archives* participantes de cada clase. Describió-se el flujo de información de los *Open Archives* colectando y analizando datos sobre la disponibilidad de acceso lo texto completo y la posibilidad de interacción de los usuarios con los autores. Verifico-se la variación del número de participantes y el tiempo de existencia de los *Open Archives*. A través de una pesquisa bibliográfica y datos colectados directamente de la muestra evidencio-se algunos factores que llevaran a lo aparecimiento de los *Open Archives*. La pesquisa fue realizada en *Open Archives* nacionales e internacionales que utilizaban el protocolo OAI-PMH para disponibilidad de sus metadados. La metodología utilizada para la realización de la pesquisa fue compuesta por uno conjunto de etapas y procedimientos que envolvió el método de deducción científica, una pesquisa descriptiva, la definición de la población y la muestra, la obtención de los datos, la aplicación de los instrumentos de coleta de datos, o procesamiento de los datos y la análisis exploratoria de los datos. Con base en los resultados presentados en la pesquisa, concluí-se que la utilización de los *Open Archives* como una alternativa de comunicación científica presenta pontos muy positivos en el escenario nacional y internacional. Entre ellos, posibilita de manera práctica y eficiente la creación de espacios virtuales para el intercambio de lo conocimiento científico entre los pesquisadores; ofrece oportunidades para las instituciones con pocos recursos en investí mentó crearen sus *Open Archives* para divulgaren y disponibilizaren sus trabajos; inserid en ámbito internacional una nueva solución tecnológica creando sistemas de información interoperáveis de fácil aplicabilidad; creo una nueva filosofía tecnológica en relación a la comunicación y producción de lo conocimiento científico; posibilito la diseminación de la información en ámbito internacional, aproximando pesquisadores de varias áreas de la ciencia dispersos geográficamente; el protocolo OAI-PMH está se tornando parte de la infra-estructura de la *Web* de forma tan importante como el protocolo HTTP es actualmente; y los *Open Archives* que hicieran parte de la muestra atended comunidades específicas y con alcance diseccionado a comunidad local.

Palavras-chave: Fichero abiertos. Comunicación científica. Comunidad científica. Metadados. Repositório digital.

LISTA DE FIGURAS

Figura 01 – Elementos da comunicação humana.....	39
Figura 02 - Fluxo de informação procedente nos Provedores de Dados.....	45
Figura 03 - Esquema de interação dos Provedores de Serviços.....	46
Figura 04 - Esquema de funcionamento básico do protocolo OAI-PMH.....	56
Figura 05 – Vários Provedores de Serviço realizando a Colheita em vários Provedores de Dados utilizando OAI-PMH.....	59
Figura 06 – Agregador acomodado entre os Provedores de Serviço e os Provedores de Dados.....	60
Figura 07 – Operação de colheita de metadados complementada com pesquisa baseada em Z39.50.....	61
Figura 08 - Distribuição dos <i>Open Archives</i> conforme à região de origem e tipo de <i>Open Archives</i>	71
Figura 09 – Distribuição dos <i>Open Archives</i> conforme os termos que caracterizam seus objetivos.....	76
Figura 10 - Distribuição dos <i>Open Archives</i> conforme os softwares utilizados para implantação no Brasil.....	81

LISTA DE TABELAS

Tabela 01 – População dos <i>Open Archives</i> conforme as classes por tipo de arquivo.....	64
Tabela 02 – Amostra dos <i>Open Archives</i> indicando o número de elementos.....	65
Tabela 03 - <i>Open archives</i> brasileiros que fizeram parte da primeira amostra.....	66
Tabela 04 – Distribuição de Frequências das variáveis conforme o número de elementos que apresentaram resposta.....	69
Tabela 05 - Distribuição dos <i>Open Archives</i> conforme a sua variação entre o período de 06/2005 a 08/2005.....	70
Tabela 06 - Distribuição dos <i>Open Archives</i> conforme à região de origem e o tipo de <i>Open Archives</i>	71
Tabela 07 – Distribuição dos <i>Open Archives</i> conforme a área de atuação	72
Tabela 08 – Distribuição de Frequências dos <i>Open Archives</i> conforme a descrição do idioma da primeira amostra.....	73
Tabela 09 – Distribuição dos <i>Open Archives</i> conforme o idioma utilizado para comunicação entre os pares.....	74
Tabela 10 – Distribuição dos <i>Open Archives</i> conforme a tecnologia adotada para implantação.....	75
Tabela 11 – Distribuição dos <i>Open Archives</i> conforme o percentual de disponibilidade de acesso ao texto completo.....	77

Tabela 12 – Distribuição dos *Open Archives* conforme as possibilidades de interação dos leitores visitantes.....78

Tabela 13 - Distribuição dos *Open Archives* conforme o tempo de existência.....79

Tabela 14 - Distribuição dos *Open Archives* conforme as categorias que predominam no Brasil.....80

LISTA DE QUADROS

Quadro 01 – Resumo do divisor tecnológico das décadas de 1960, 1970, 1980 e 1990 baseado em Castells.....	35
Quadro 02 – Resumo histórico da evolução do OAI-PMH.....	55

LISTA DE SIGLAS

ARL – *Association of Research Libraries*

ARPA – Agência de Projetos de Pesquisa Avançada do Departamento de Defesa Norte-Americano

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CLIR – *Council on Library Information Resources*

CNI – *Coalition for Networked Information*

DCMI – *Dublin Core Metadata Initiative*

DLF – *Digital Library Federation*

FTP – *File Transport Protocol*

HTTP – *Hypertext Transport Protocol*

IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia

IMPA - Instituto Nacional de Matemática Pura e Aplicada

NSF – *National Science Foundation*

OA – *Open Archives*

OAI – *Open Archives Initiative*

OAI-PMH – *Open Archives Initiative – Protocol Metadata Harvesting*

SEER - Sistema Eletrônico de Editoração de Revistas

XML – *Extensible Markup Language*

UPS – *Universal Preprint Service*

WWW – *World Wide Web*

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivos	18
1.1.1	Objetivo geral	18
1.1.2	Objetivos específicos.....	18
1.2	Justificativa	18
2	COMUNIDADE, COMUNICAÇÃO e INFORMAÇÃO	20
2.1	Comunicação científica	22
2.1.1	História da comunicação científica	24
2.1.2	Canais de comunicação científica	26
2.1.3	Comunicação eletrônica	27
2.1.4	Documentos digitais	29
2.2	Comunicação científica na Sociedade da Informação	30
2.2.1	Evolução tecnológica	33
2.2.2	Tecnologias de informação e comunicação.....	33
3	OPEN ARCHIVES	37
3.1	Open Archives Initiative – OAI	40
3.1.1	Características principais da OAI	42
3.1.2	Arquitetura dos Open Archives.....	43
3.1.2.1	Provedores de dados.....	43
3.1.2.2	Provedores de serviços.....	45
3.1.3	Software para implantação dos Open Archives.....	47
3.2	Instituições que oferecem suporte a Open Archives Initiative – OAI	48
3.2.1	Iniciativas para criação de Open Archives no Brasil.....	50
3.3	Open Archives Initiatives – Protocol for Metadata Harvesting – OAI – PMH	52
3.3.1	Histórico do OAI – PMH.....	53
3.3.2	Tecnologia do OAI - PMH.....	55
3.3.3	Considerações sobre implementação do OAI – PMH.....	58
3.3.4	Configurações para uso do protocolo OAI – PMH.....	59
4	METODOLOGIA	62
4.1	Delimitação do problema	62
4.2	Tipo de pesquisa	62
4.3	Método	63
4.4	Definição da população e amostra	63

4.4.1	População	64
4.4.2	Amostra	65
4.5	Obtenção dos dados.....	66
4.6	Aplicação do formulário de pesquisa	67
4.7	Processamento dos dados.....	67
4.7.1	Análise exploratória dos dados.....	67
5	APRESENTAÇÃO DOS RESULTADOS DA PESQUISA.....	69
5.1	Características que apresentam a variação dos <i>Open Archives</i>	69
5.2	Características dos elementos referentes à região de origem dos <i>Open Archives</i>.....	71
5.3	Características das comunidades que adotaram as tecnologias da OAI.....	72
5.4	Características dos fatores declarados como motivadores para o surgimento dos <i>Open Archives</i>.....	75
5.5	Características dos elementos relacionados com o fluxo de informação dos <i>Open Archives</i>.....	77
5.6	Características dos elementos dos <i>Open Archives</i> conforme o tempo de existência.....	78
5.7	Características dos <i>Open Archives</i> existentes no Brasil.....	79
6	CONCLUSÕES	82
6.1	Conclusões.....	82
6.2	Sugestões.....	85
7	GLOSSÁRIO.....	86
	REFERÊNCIAS	90
	APÊNDICES.....	95
	APÊNDICE A – Formulário de pesquisa de dados primários	96
	APÊNDICE B – Formulário de pesquisa de dados secundários	99

1 INTRODUÇÃO

A Sociedade da Informação é resultado da constante busca de melhorias nos meios de acesso, distribuição e uso da informação. Uma sociedade formada por grandes redes de comunicação que agregam valor à informação, utilizando-a como insumo de produção e elemento de poder econômico e político. As tecnologias de informação e comunicação assumem papéis importantes nessa nova sociedade, oferecendo suporte para as práticas relacionadas com o compartilhamento da informação e do conhecimento. A comunidade científica, incentivada pela corrida bélica e seus altos investimentos em ciência e tecnologia, criam novos meios de comunicação, que ultrapassam as fronteiras acadêmicas e militares, contribuindo consideravelmente para formação da Sociedade da Informação.

A comunicação científica passa pelo processo evolutivo das tecnologias sempre na busca de melhorar as práticas de compartilhamento da informação e do conhecimento entre os pares, visando à aprovação dos membros da comunidade e a institucionalização das ciências. As comunidades científicas apropriam-se dos recursos tecnológicos de comunicação conforme seus objetivos e normas, adaptando-se as transformações dos suportes de escrita e dos canais de comunicação formais e informais.

Com o avanço das novas tecnologias de informação e comunicação, surgiram recursos alternativos para comunicação científica em formato eletrônico, tais como os documentos multimídia, a Internet, a *World Wide Web*, listas de discussão, fóruns eletrônicos, conferências em linha, bancos de *preprints*, *e-prints*, correio eletrônico, periódico eletrônico, entre outros, os quais otimizaram o processo de troca de informações entre os pares e a criação de espaços virtuais para o compartilhamento do conhecimento científico. Esses novos recursos possibilitaram o intercâmbio de idéias e experiências entre pesquisadores das várias áreas das ciências com maior agilidade.

As formas tradicionais de comunicação científica, como os periódicos em papel, passaram por processos de transição com intuito de acompanhar a demanda informacional e atender as necessidades de adaptação dos pesquisadores. Muito embora, ainda apresentam questionamentos por parte de alguns membros da comunidade científica. Marcondes e Sayão (2001) relatam que os paradigmas de comunicação científica, tendo por base o periódico científico em papel, com seu esquema de revisão por pares e o monopólio das

grandes editoras científicas, vêm sofrendo grande impacto com o surgimento da Internet e grande questionamento em relação à publicação científica por parte da comunidade científica de todo o mundo. Moreira (2005) afirma que as grandes editoras ainda dominam a cena, mas a facilidade de acesso-cópia-envio certamente as assustou e as tem obrigado a rever suas práticas de mercado.

O desenvolvimento dos mecanismos de publicação eletrônica para comunidade acadêmica, com o intuito de aumentar sua visibilidade, tornou-se uma questão essencial para o desenvolvimento e a ampliação das atividades de pesquisa científica no cenário internacional. O surgimento do primeiro arquivo eletrônico de *eprints* (textos digitais de artigos revisados, utilizado também como termo geral) ou *preprints* (textos digitais de artigos antes da arbitragem e publicação), o ArXiv, no *Los Alamos National Laboratory*, criado em 1991 pelo físico Paul Ginsparg, possibilitou à própria comunidade científica oferecer uma alternativa prática para a publicação de seus trabalhos. Esse arquivo eletrônico foi denominado de *Open Archives* (Arquivos Abertos) e deu origem a criação de vários outros.

Open Archives Fórum (2005) os chamados *Open Archives* são bibliotecas digitais desenvolvidas na *Web* por cientistas e para cientistas, constituindo-se em fóruns privilegiados para difusão de resultados e debates científicos, é um exemplo da gradativa mudança de paradigma na comunidade científica, seus recursos enquanto meio de comunicação científica estão mobilizando vários setores da comunidade científica e possibilitando a prática de atividades inviáveis pelos recursos tradicionais.

No site oficial do Open Archives Fórum (2004) está relatado que:

A grande aceitabilidade por alguns membros da comunidade científica no cenário mundial propiciou o surgimento de uma iniciativa que oferecesse suporte para esta nova prática. Esta iniciativa foi denominada de OAI – *Open Archives Initiative* (Iniciativa de Arquivos Abertos), formada com o intuito de oferecer soluções de interoperabilidade entre os *Open Archives*. Esta iniciativa foi criada com o objetivo principal de gerar banco de dados que ofereça um sistema de operação de fácil compreensão e aplicabilidade à comunidade científica, para servir de estímulo à produção e publicação científica e adicionar funções de compartilhamento e interação entre as academias.

As origens da *Open Archives Initiative* – OAI se posicionam no crescente interesse em alternativas para os paradigmas tradicionais de publicação acadêmica. Embora isto possa estar em desacordo com a natureza de quais mudanças serão necessárias, existe um

grande consenso da possibilidade de uma mudança radical ser inevitável. Existem vários fatores motivadores para esta mudança. Um crescimento no número de disciplinas acadêmicas, especialmente aquelas chamadas de “*hard sciences*” (Física, Ciência da Computação, Ciências Humanas), estão produzindo resultados em um crescimento muito rápido. Esta velocidade demanda mudanças nos mecanismos para registrar resultados com baixos intervalos de tempo que os sistemas de revistas científicas não proporcionam (LAGOSE; SOMPEL, 2001).

No Brasil, também há iniciativas de incentivo para utilização de *softwares* livres seguindo a filosofia *Open Archives*. Kuramoto (2005, b) afirma que o Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT realiza um trabalho de prospecção tecnológica com vistas a identificar tecnologias de informação aplicadas ao tratamento e disseminação da informação científica e tecnológica. Como resultado, diversas tecnologias foram identificadas com características técnicas compatíveis com os padrões e ideais do modelo chamado *Open Archives*. O Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT realizou uma experiência com esse modelo, implantando a Biblioteca Digital de Teses e Dissertações – BDTD, e obteve excelentes resultados, que demonstraram a total interoperabilidade com todas as bibliotecas que utilizavam este modelo.

As tecnologias oferecidas pela *Open Archives Initiative – OAI* estão presentes em vários países, indiferente do nível de desenvolvimento, cultura ou idioma. Apresenta formas de comunicação científica em suporte digital para o compartilhamento e disseminação do conhecimento científico em âmbito mundial.

Esta dissertação identifica o panorama sobre os *Open Archives* a nível nacional e internacional. Devido a grande quantidade de palavras criadas especialmente para identificar funções e ferramentas dos *Open Archives*, esta dissertação apresenta um glossário com termos técnicos e seus respectivos significados para auxiliar no entendimento dos assuntos tratados.

1.1 Objetivos

A seguir serão apresentados o objetivo geral e os objetivos específicos da dissertação

1.1.1 Objetivo geral

Identificar o panorama dos *Open Archives* na comunidade científica como uma nova alternativa de comunicação científica apoiada nas tecnologias oferecidas pela *Open Archives Initiative* - OAI.

1.1.2 Objetivos específicos

A – Levantar informações sobre o histórico, motivação e instituições que oferecem suporte a OAI.

B – Evidenciar os fatores que levaram ao surgimento dos *Open Archives*.

C – Identificar as principais tecnologias envolvidas na OAI.

D – Identificar as comunidades que adotaram as tecnologias da OAI.

E – Verificar a variação do número dos participantes da OAI.

F – Identificar os objetivos para criação dos *Open Archives*.

1.2 Justificativa

Esta dissertação abrange fatores relacionados com as propriedades e o comportamento da informação, tratando de estudos referentes às tecnologias utilizadas, os processos, o material humano envolvido e o fluxo de informação.

Partindo da idéia de que na sociedade em rede, dificilmente um cientista solitário, consiga realizar uma descoberta relevante, do início ao fim, sem interagir com outros colegas cientistas, tendo em vista que a ciência é um trabalho coletivo, interdisciplinar e de interesse universal. A justificativa da construção deste trabalho está presente no intuito de evidenciar uma forma de comunicação científica e mudança no suporte de escrita, que possa, por sua vez, desenvolver soluções para problemas referentes ao compartilhamento e a construção do conhecimento científico.

Estes problemas se caracterizam como fatores dificultadores e proibitivos para novas descobertas científicas. Marcondes e Sayão (2002) destacam alguns deles:

A lentidão do ciclo de edição das revistas científicas comparadas à rapidez da geração de novos conhecimentos, a renúncia sobre a obra imposta pelas revistas, que impede a ampla disseminação pelo autor de suas idéias, a perspectiva extremamente rígida e conservadora dos esquemas de *peer review* (revisão por pares) adotados pelas revistas, que não raro são impeditivos para o surgimento de idéias inovadoras, entre outros fatores.

Kuramoto (2005, a) trata do mesmo assunto afirmando que:

Os pesquisadores da comunidade científica ainda enfrentam hoje dificuldades na publicação do resultado de suas pesquisas. Muitas vezes, as revistas demoram de seis meses a um ano para publicar um artigo. E isso é demasiado longo em fase do desenvolvimento científico que ocorre no mundo todo. Às vezes, está com um resultado quente na mão, e a demora na publicação pode tornar o artigo totalmente defasado. Em outras situações, o pesquisador não tem sequer a oportunidade publicar o seu trabalho, seja pela demora das revistas, seja porque as revistas têm um conselho editorial bastante rígido.

O surgimento dos *Open Archives* retratam o avanço de antigas formas de comunicação - como por exemplo os periódicos em suporte de papel, os congressos e os fóruns que exige a presença física dos pesquisadores nos locais – as quais apresentam-se com um objetivo em comum, facilitar o intercâmbio de idéias e experiências entre os membros de uma comunidade científica e a sociedade em geral.

Em se tratando do avanço das formas de comunicação científica, Barroso (2004) afirma que:

A tendência mundial nos aponta para a disponibilização de conteúdos digitais por meio de ferramentas de acesso aberto. A adesão a essas tecnologias proporciona simultaneamente a sistematização dos processos editoriais, resultando em maior agilidade na edição da publicação, assim como na diminuição dos custos de

impressão, solução para um contexto restrito de recursos orçamentários. O IBICT coloca à disposição da comunidade a revista automatizada pelo novo Sistema Eletrônico de Editoração de Revistas – SEER. O SEER é uma adaptação, desenvolvida pelo IBICT, do *software* livre *Open Journal System*, criado pela Universidade de British Columbia, seguindo a filosofia dos *Open Archives*.

O número de registros de publicações em *Open Archives* institucionais (*Research Institutional or Departmental*), no período de julho de 2001 a março de 2004 obteve crescimento de 48%, ou seja, foram criados 60 *Open Archives* institucionais novos, totalizando 70.000 novos registros nas bases de dados, e seu crescimento está cada vez mais acelerado (EPRINTS.ORG, 2004).

A *Open Archives Initiative - OAI* está possibilitando a criação de um consenso de interoperabilidade entre as comunidades científicas, visando organizar e fomentar o uso dos *Open Archives* para que não se torne um novo caos informacional e fuja dos interesses da comunidade científica. Esta iniciativa veio viabilizar o processo de mudança e possibilitar o bom uso desse novo recurso tecnológico de comunicação, aumentando sua credibilidade e conseqüentemente sua aceitabilidade por parte da comunidade científica.

O uso dos *Open Archives* é uma prática nova na comunidade científica, tendo sua origem no início dos anos de 1990, levanta a questão de qual o impacto causará e quais as conseqüências de sua influência no âmbito científico como alternativa de comunicação, esta indagação é um questionamento comum no surgimento de um novo fenômeno. Moreira (2005) discorre que “a comunidade científica e suas novas ferramentas e possibilidades interativas ou de interação ressurgem como campo complexo de investigação e são necessários novos estudos com o fim de verificar onde e como se dão as novas confluências”.

Outro fator relevante para execução deste trabalho é a importância dos *Open Archives* para sociedade, principalmente nos países em desenvolvimento que dispõem de poucos recursos para investimentos em ciência e tecnologia.

2 COMUNIDADE, COMUNICAÇÃO E INFORMAÇÃO

Dentro de um contexto social as comunidades são formadas por pessoas com particularidades, interesses comuns e o compartilhamento de espaço e tempo. A

comunicação liga os membros dessas comunidades por intermédio da linguagem e o envio de mensagens. A informação, por sua vez, no sentido etimológico da palavra, dar forma, apresenta-se como elemento utilitário na construção de significados e símbolos.

Barichello (2001) afirma que a comunidade faz a mediação e é mediada por sua vez pelo material e o simbólico, o experimentado, o vivido e o imaginário, o espaço local e o local. As relações sociais se constroem, sobretudo, pela inserção de uma comunidade em um espaço e um tempo histórico e pelos vínculos que esta comunidade consegue gerar entre seus membros.

Maffesoli (2004) sustenta que a comunicação é cimento social, é a cola do mundo pós-moderno, está implícita na sociedade. A comunicação é uma forma de reencarnação do velho simbolismo, simbolismo arcaico, pelo qual percebemos que não podemos nos compreender individualmente, mas que só podemos existir e compreendermo-nos na relação com o outro.

A informação é o resultado da decodificação de um dado ou um conjunto de dados emitidos através de linguagens simbólicas que manifestam a interpretação do decodificador. Para Maffesoli (2004) informar significa ser formado por, trata-se da forma que forma, a forma formante. Quer dizer que na era da informação, não se pensa por si mesmo, mas se é pensado, inserido numa comunidade de destino. Como a comunicação, a informação também liga, une, junta.

Na Sociedade da Informação o que conta é a partilha cotidiana e segmentada de emoções e de pequenas emoções. As pessoas não querem somente informação na mídia, querem ver-se, ouvir-se, participar, contar o próprio cotidiano para si mesmo e para aqueles com quem convivem.

Maffesoli (2004, p. 24) discorre que “a informação e a comunicação, no sentido de partilha de sentimentos e emoções, só podem dirigir-se a tribos que comungam em torno de um totem. Os totens são formados pelo desejo de criação de um grupo a partir de um elemento comum”. Neste contexto, as relações de empatia são decisivas na estruturação do tecido social. O universalismo cede lugar às apropriações particulares do global.

As tecnologias de informação e comunicação aproximam as pessoas vencendo as dimensões territoriais e rompendo fronteiras. O espaço perde para o tempo em uma guerra na qual as novas técnicas possibilitam o envio de mensagens quase instantaneamente.

Bauman (2001) considera este fato, afirmando que o tempo adquire história uma vez que a velocidade do momento através do espaço se torna uma questão de engenho, da imaginação e da capacidade humana. O espaço é o lado sólido e impassível, pesado e inerte, capaz apenas de uma guerra defensiva, de trincheiras, é apenas um obstáculo nos avanços do tempo.

As novas tecnologias criaram um novo espaço chamado de ciberespaço, onde a velocidade do tempo é fator predominante na formação e aproximação das comunidades e de seus membros. Lévy (2000) expõe que o ciberespaço é hoje o sistema com o desenvolvimento mais rápido de toda a história das técnicas de comunicação. Encarna um dispositivo de comunicação qualitativamente original, que se deve distinguir das outras formas de comunicação de suporte técnico.

A imprensa, a edição, o rádio e a televisão funcionam segundo um esquema em estrela, ou “um para todos”. O correio e o telefone desenham um esquema em rede, ponto a ponto, “um para um”. O ciberespaço permite, ao mesmo tempo, a reciprocidade na comunicação e a partilha de um contexto, trata-se de comunicação conforme um dispositivo “todos para todos”.

Na Sociedade da Informação existe um mundo paralelo, um novo mundo, um mundo virtual ligado ao real – o ciberespaço – constituído por um emaranhado de redes de computadores. Composto por indivíduos e comunidades estruturadas que se formam e se transformam em um processo interativo. A comunicação e a informação criam novos mundos dentro do ciberespaço, aproximando, unindo, ligando pessoas e comunidades, possibilitando a apropriação de particularidades e compartilhando o conhecimento existente no planeta.

2.1 Comunicação científica

A comunicação científica é parte fundamental da ciência e encontra-se sujeita às influências e exigências de um grupo de normas imposto por uma determinada comunidade científica. A prática de divulgar resultados, apresentar novas idéias, disponibilizar informações sobre pesquisas em andamento e obter retorno dos membros da comunidade

interessada, exige a criação de recursos capazes de suprir a demanda informacional produzida e recebida pelas comunidades científicas.

Kneller (1980, p. 272) quando trata da verdade científica afirma que o cientista busca a verdade a cerca da ordem da natureza. Ao fazê-lo, aceita uma rígida disciplina. Obedece a certos padrões de conduta ou “normas” que, de fato, estabelecem o que ele deve fazer para chegar à verdade. Como um meio para um fim que é almejado por seu valor intrínseco, essas normas são valores morais, respeitá-las constitui a integridade científica.

A conceituação de comunicação científica, segundo Rodrigues (2001) consiste no tipo de comunicação que abrange as etapas de produção, divulgação e utilização da informação científica, a partir da concepção da idéia a ser investigada pelo pesquisador, até o momento de divulgar seus resultados para serem avaliados pelos pares e incorporados ao acervo do conhecimento científico. Para Meadows (1999) a comunicação se situa no próprio coração da ciência. É para ela tão vital quanto à própria pesquisa, pois a esta não pode reivindicar com legitimidade este nome enquanto não houver sido analisada e aceita pelos pares.

O conjunto de normas impostas por uma comunidade científica para sua comunicação, demonstra qual o campo de atuação disciplinar que participa, quais são suas contribuições e atribuições no campo científico e afirma sua identidade enquanto ciência. Na visão de Kneller (1980, p. 182) “a comunidade científica é uma associação de pessoas que não estão vinculadas entre si por leis ou cadeias de comando, mas pela comunicação de informações – através de revistas especializadas, conferências, discussões informais e outros canais”. Khun (1990) conceitua comunidade científica como um grupo composto por cientistas que praticam uma mesma especialidade científica e cujos membros tiveram a mesma educação e iniciação profissionais, tiveram contato com a mesma literatura técnica e absorveram dela os mesmos ensinamentos.

As comunidades científicas, por sua vez, encontram-se sujeitas a certos padrões de comportamento que determinam um ideal tanto para a atividade de pesquisa quanto para comunicação dos resultados das pesquisas para seus membros. Aragão e Oliveira (1992) relatam que a psicologia da comunicação científica enfatiza a interação entre o cientista e seu ambiente do qual faz parte, como elemento principal, outro cientista.

Kneller (1980, p. 182) trata da comunicação científica afirmando que:

A comunicação é coordenada por instituições tais como as sociedades de especialistas e colégios invisíveis. Por meio dessas instituições e canais, e do mecanismo de recompensa que opera através deles, a comunidade científica procura atingir certos objetivos que contribuem para o propósito geral de ampliar o conhecimento da natureza. Esses objetivos consistem em manter os padrões de pesquisa, harmonizar os interesses do cientista individual e da empresa científica, promover a competição e cooperação, e estimular a inovação.

Meadows (1999) afirma que cada área temática contém um grupo relativamente pequeno de pessoas que dominam suas áreas preferidas. Elas são visíveis não só para quem pertence à mesma especialidade, mas também para pessoas de fora. Essa imagem condiz com a maneira como a maioria dos próprios pesquisadores vê seus pares.

Entre as informações colocadas nesta seção, pode-se afirmar que o tipo de comunicação utilizada por uma comunidade científica, depende de seus objetivos e especificações como ciência e de seus recursos de acesso às informações disponibilizadas. A transferência de dados e informações aparece na comunidade científica como fator de otimização no intercâmbio de idéias, experiências e conhecimento. Diferentes comunidades utilizam formas distintas de comunicação, as escolhas são feitas através de suas particularidades, as quais determinam os melhores meios de divulgar suas produções e compartilhar seus conhecimentos.

2.1.1 História da comunicação científica

O ato de compartilhar idéias, experiências, informações sobre pesquisas em andamento e de seus resultados e esperar contribuições e aprovação de seus pares com o intuito de construir um conhecimento coletivo deu origem à prática da comunicação científica. No início utilizavam-se meios de comunicação e suporte de escrita (manuscritos, pergaminhos, cartas entre outros) que exigiam muito tempo para serem produzidos e disponibilizados, dificultando o acesso às informações e conseqüentemente aumentando o espaço de tempo entre os pesquisadores e o material produzido.

Meadows (1999) afirma que “as atividades mais remotas que tiveram impacto na comunicação moderna foram inquestionavelmente as dos Gregos Antigos. A pesquisa era e é até hoje comunicada de várias formas, sendo que as duas mais importantes são a escrita e a fala, os gregos se valiam de ambas”.

No que se refere da tradição da pesquisa comunicada em forma escrita, são ainda as obras dos gregos, sendo de Aristóteles, as que mais tiveram a contribuir. Seus debates

foram registrados em manuscritos e copiados repetidas vezes, influenciando primeiro os Árabes e depois a Europa Ocidental.

A invenção da imprensa, no século XV, por Gutenberg, acelerou a disponibilidade de acesso ao conhecimento registrado. Muito embora, a maior parte dos livros não era sobre ciência, alimentou-se o desejo da transmissão dos resultados das pesquisas científicas de maneira mais consistente, com maior alcance geográfico entre os pesquisadores e maior rapidez na produção de exemplares. Meadows (1999) relata que a capacidade de multiplicar os exemplares de um livro representou um passo importante rumo a uma difusão melhor e mais rápida das pesquisas.

Schwarcz (2002) comenta o fato da distribuição do conhecimento afirmando que:

Os acervos medievais ficaram quase restritos aos mosteiros, onde se cultivavam formas de leitura individual e coletiva. Devido seu regime de internato, os mosteiros possuíam e produziam muitas obras, qualificando-se como verdadeiras editoras. As universidades e as escolas com um novo modelo escolástico ganharam espaço e conseguiram substituir o antigo formato monástico da escrita. As bibliotecas ocuparam um outro lugar fora dos mosteiros e o livro tomou nova forma e passou para mão dos comerciantes e dos príncipes. Entre os séculos XV e XVIII ocorreram mudanças de conceito e as bibliotecas passam a ser abertas para o público.

Com a publicação impressa os pesquisadores tiveram acesso a diversas obras distribuídas nas principais bibliotecas e em vários lugares da Europa. Os correios oficiais percorriam caminhos regulares, em viagens oficiais para o Estado, levando junto, correspondências particulares. Esta prática se estendeu passando a assumir um caráter mais formal, chegando aos sistemas postais atuais.

Esses serviços estimularam a difusão de notícias e com eles surgiram os jornais. No começo eram apenas folhetos que noticiavam um único acontecimento. Com a invenção da tipografia e a criação de vários sistemas de difusão de notícias que interligavam os centros comerciais, construiu-se um novo meio de disseminação de informação. A distribuição esporádica tornou-se mais regular e originou o ancestral do jornal moderno, que também serviu de modelo para o surgimento da revista científica.

A mudança da forma manuscrita para impressa aconteceu de maneira gradativa. Os noticiários científicos, quando se destinavam a um público pequeno, continuavam a ser produzidos de forma manuscrita e esta prática permaneceu durante o século XVII até o Século XVIII. Para os pesquisadores, era mais viável a circulação das idéias através de

cartas manuscritas para serem analisadas, testadas e depois respondidas. No caso das idéias possuírem grandes dimensões científicas e se destinassem a um público maior, eram impressas e enviadas. Desta forma, surgiram na segunda metade do século XVII, as primeiras revistas científicas.

Com o advento das novas tecnologias de informação e comunicação, as pesquisas comunicadas tornaram-se mais acessíveis, as revistas científicas se apropriaram de recursos mais abrangentes de comunicação e o compartilhamento de idéias entre pesquisadores se tornou mais rápido e eficiente.

2.1.2 Canais de comunicação científica

Na comunicação científica existem canais formais e informais que são utilizados para possibilitar o intercâmbio e o compartilhamento de informações entre uma ou várias comunidades científicas. A escolha de qual tipo de canal de comunicação atende melhor as necessidades de uma comunidade para publicação de suas atividades depende das pretensões, objetivos e especificações determinados pela comunidade interessada.

Os canais formais de comunicação científica compreendem os meios de comunicação escrita, com destaque para os livros, periódicos, obras de referência, relatórios técnicos e revisões de literatura.

Ramos (1994) afirma que nos sistemas de comunicação formal, o meio de comunicação mais característico é o artigo de periódico, o qual pública o resultado das pesquisas desenvolvidas pelos cientistas. Os artigos apresentam características significativas que retratam a forma de trabalho de determinada comunidade científica.

Meadows (1999) discorre que as comunicações formais, como livros e periódicos, têm uma resistência duradoura e dependem basicamente da visão, ou seja, da leitura.

Os canais informais dizem respeito aos contatos interpessoais, telefonemas, cartas, visitas interinstitucionais, reuniões científicas como congressos, conferências e pequenos encontros de grupos locais.

Meadows (1999) afirma que as comunicações informais são, por definição, efêmeras e isso é verdadeiro no caso da informação transmitida pela fala. No entanto, Macedo (1999) afirma que tendo em vista, muito embora, a satisfação das necessidades de informação dos

próprios cientistas, os canais informais representam um papel da maior importância, pela troca de idéias e pelos *feedbacks* imediatos com os pares.

Rodrigues (2001) considera que a tradicional subdivisão da comunicação científica informal e formal, não implica necessariamente isolamento de ambas as modalidades. Estes segmentos não são estanques, suas relações formam uma espécie de rede na qual fluem cientistas e os produtos de seus trabalhos em constante interação. No entanto, apresentam objetivos diferenciados, no que tange à operacionalização das pesquisas, sendo ambos indispensáveis à comunicação da produção científica, mas usadas em momentos e cronologias diferenciadas. Para Moreira (2005) a comunicação formal/informal realizada por meio de recursos das tecnologias de comunicação e informação cresce principalmente como meio de suprir falhas relacionadas à atualização.

A construção do conhecimento científico está visivelmente relacionada com o uso dos diferentes canais de comunicação. O objetivo da ciência é agregar recursos que possam facilitar o intercâmbio e compartilhamento de informação entre os cientistas. Por conseguinte, as novas tecnologias de informação e comunicação estão aproximando os canais formais e informais e colocando-os ao alcance das comunidades científicas, mudando os sistemas tradicionais de acesso e criando novos meios de comunicação. Esta evolução já é um fato inevitável e possui grande aceitabilidade no âmbito científico.

2.1.3 Comunicação eletrônica

Com a grande demanda de informações e conhecimentos produzidos pela comunidade científica, tornou-se necessário à criação de sistemas capazes de gerenciar de forma eficiente todo esse fluxo informacional. O advento das novas tecnologias de informação e comunicação viabilizou a organização e a disseminação da avalanche informacional disponibilizadas pelas várias áreas das ciências.

Gomés (2004) trata a questão das grandes massas documentais afirmando que:

Do início à metade do século XX, as inovações da fotografia, do rádio, do cinema reformularam os espaços do texto e da memória, renovando os saberes em torno dos sistemas sociais de comunicação e alargando-as com novas zonas de indagação e reflexão. Entre estas questões, algumas alertavam a cerca da importância do controle e gestão de massas documentais cada vez mais volumosas e diversificadas. A passagem das tecnologias culturais e de reprodução (imprensa,

fotografia, reprografia, cinema e rádio) às denominadas tecnologias intelectuais (aumentam e modificam capacidades cognitivas) de suporte digital modifica novamente as formas de inscrição, armazenagem e transmissão da produção social discursiva, modificando-se ao mesmo tempo as possibilidades e demandas de intervenção institucional técnica e profissional no tratamento, circulação e gestão dessa produção.

O sistema tradicional de registro do conhecimento, ou seja, a impressão em suporte de papel passou por intensas modificações chegando aos documentos digitais. Essa nova forma de registro teve origem com o surgimento dos computadores, que por sua vez foram aperfeiçoados e produzidos de maneira que fossem acessíveis às pessoas em grande número. Tecnologias como Internet e posteriormente a *Web* possibilitaram a construção de redes de computadores e a transferência de dados, informações, gráficos, sons e imagens em tempo real.

A alternativa de utilizar recursos eletrônicos de registro e transferência de conhecimento rompeu alguns paradigmas e mudou certos hábitos enraizados pela comunidade científica no uso das formas tradicionais. Meadows (1999) afirma que assim como os veículos de comunicação de massa possuem características diferentes, também a comunicação eletrônica e a impressa podem produzir percepções distintas. A diferença importante está na flexibilidade do processamento eletrônico. Por conseguinte, definições e hábitos que se desenvolveram no ambiente da palavra impressa talvez não se apliquem a um outro ambiente dominado pela comunicação eletrônica. Para Moreira (2005) “certamente a gratuidade de acesso a diversos periódicos e documentos eletrônicos já influi no modo como se organiza a comunicação científica”.

Souza (2000) trata o mesmo assunto afirmando que a análise das mudanças existentes quanto ao desenvolvimento do pensamento científico é fundamental para compreensão das instituições científicas, considerando que o primeiro objetivo dessas instituições é a produção do novo conhecimento. Uma das primeiras questões sociológicas que devem ser levantadas é se a comunidade científica e a variação no padrão de comunicação entre cientistas afetam atualmente o desenvolvimento do conhecimento.

Os meios de comunicação eletrônicos causaram grandes mudanças nas comunidades científicas. Seus recursos tecnológicos possibilitaram o acesso e disponibilidade de dados e informações entre pesquisadores de todas as partes do mundo.

A condição de estar conectado a uma rede de computadores e poder interagir com seus pares em tempo real, proporcionou uma aceleração considerável no processo de compartilhamento e construção do conhecimento científico. Novas formas de expor trabalhos à comunidade científica surgiram com a mesma rapidez dos avanços tecnológicos.

Os periódicos científicos são um bom exemplo de adequação aos novos recursos de comunicação eletrônica. Meadows (1999) considera que uma das principais virtudes do processamento eletrônico deriva da capacidade de fazer buscas rapidamente no meio de imenso volume de informações. O texto de um artigo de periódico não é tão grande que não possa ser lido na tela, e o acesso rápido e a possibilidade de buscas podem ser um recurso útil para os leitores. Desta forma, os meios eletrônicos de comunicação ocupam seu espaço no ambiente científico, melhorando a dinâmica da construção e do compartilhamento do conhecimento.

Outras formas de comunicação entre os pesquisadores estão se adaptando aos recursos eletrônicos em um processo cada vez mais rápido. A migração do suporte de papel para o digital demonstra ser relevante aos interesses de diversas comunidades científicas. A utilização de novas tecnologias de informação apresenta-se como plausíveis soluções para questões de tempo e espaço e modulam novos paradigmas na comunicação científica na sociedade da informação.

2.1.4 Documentos digitais

A transição do documento impresso em suporte de papel para o documento digital está sendo bastante acelerada no âmbito da comunicação científica. A transferência de dados de forma rápida em grandes distâncias geográficas por intermédio de redes de computadores, apresenta-se como uma exigência para os padrões atuais das comunidades científicas, e esta prática só é possível com a elaboração dos documentos digitais.

Outros fatores motivadores que aceleram o processo de transição dos documentos impressos são a facilidade de manipulação e editoração apresentado pelos documentos digitais, os quais atendem o conjunto de normas impostas por algumas comunidades científicas que precisam compartilhar seus conhecimentos de forma rápida e precisa.

Os documentos digitais são dados e informações de origens variadas registradas em suporte digital, o acesso ao seu conteúdo é possível somente através de recursos tecnológicos. O armazenamento e a transferência de seus conteúdos são realizados de forma prática e rápida, facilitando a manipulação de grandes volumes de dados e informação.

Na comunidade científica existe uma considerável aceitabilidade por parte dos pesquisadores na transição do documento impresso para o digital. Marcondes e Sayão (2002) comentam que o maior retorno que a comunidade acadêmica almeja, publicando o resultado de suas pesquisas, é que estes possam servir de bases para outras pesquisas, sendo citados por outros trabalhos.

A melhoria contínua de recursos facilitadores para divulgação de trabalhos científicos é fundamental para o desenvolvimento da ciência. O documento digital é mais um elemento que visa somar um conjunto de alternativas para este fim. Desta forma, analisa-se o documento digital como um meio de otimizar as atividades realizadas pelos pesquisadores em benefício da ciência. Muito embora, existam casos em que ocorram alguns problemas como a necessidade de utilizar a tecnologia apropriada para acessar a informação registrada, a necessidade de migrar os dados conforme a evolução das tecnologias e a compatibilidade dos programas para acessar os documentos digitais.

2.2 Comunicação científica na Sociedade da Informação

A comunicação científica na Sociedade da Informação se apresenta como elemento formador e transformador dentro de um processo evolutivo, no qual a sociedade se apropria de práticas utilizadas por cientistas e seus meios de comunicação para melhorar suas atividades cotidianas e para a criação de outras. A Internet e as tecnologias que envolvem seu uso são exemplos evidentes da influência direta dos meios de comunicação científica na construção da Sociedade da Informação. A utilização da Internet no ambiente social formou um novo paradigma fundado no acesso e distribuição da informação, criando espaços para interação de diferentes comunidades e possibilitando o surgimento de novas relações sociais e novas comunidades.

Kneller (1980, p. 284) trata da interação da ciência com o mundo em geral afirmando que “no século atual, a ciência tornou-se um objeto de vital interesse público, porque pela primeira vez na história o conhecimento científico se tornou a base da tecnologia e, por conseguinte, um parceiro na vasta gama de benefícios e malefícios pelos quais a tecnologia é responsável”.

A informação, por sua vez, está presente nas mais variadas formas de relacionamento humano. A história das sociedades relata a evolução de sistemas e mecanismos que otimizaram a prática da coleta, armazenamento e distribuição da informação para construção do conhecimento no planeta. Com a evolução tecnológica tornou-se possível o acesso à informação de maneira rápida e globalizada. A partir do acesso a conhecimentos de sociedades distintas surgiram novas ciências, ideologias, religiões e outras estruturas organizacionais.

Durante muito tempo, a informação limitou-se a pequenos grupos de determinadas regiões, estando ilhada e condenada, muitas vezes, ao esquecimento em função das grandes dificuldades de registro e distribuição. O avanço gradativo das tecnologias de informação e comunicação proporcionou a mudança desta realidade. A busca por novas informações para resolver problemas, aumentar conhecimentos e melhorar forças militares fez com que o homem aumentasse seu contato com grandes volumes de informação.

Rodriguez e Ferrante (1995) afirmam que nos últimos trinta anos do século XX, foram produzidos mais informações que nos últimos cinco mil anos anteriores, uma edição do The New York Times contém mais informações do que um cidadão médio do século XVI poderia obter em toda sua vida e que o volume de informações dobra a cada cinco anos.

A sociedade da informação é fruto de evoluções tecnológicas que influenciaram as práticas sociais, fundadas no acesso, transformação, disseminação e uso da informação. Castells (2002) comenta a importância da revolução tecnológica na transformação e formação de uma nova sociedade, relatando que:

No fim do segundo milênio da Era Cristã, vários acontecimentos de importância histórica transformaram o cenário social da vida humana. Uma revolução tecnológica concentrada nas tecnologias da informação começou a remodelar a base material em ritmo acelerado. Economias por todo o mundo passaram a manter interdependência global, apresentando uma nova forma de relação entre a economia, o estado e a sociedade em um sistema de geometria variável.

O contato com outras culturas proporcionando a construção de novos conhecimentos, a abertura de novos nichos de mercado transformando a economia global e as buscas constantes de melhorias na indústria bélica e espacial são características peculiares na formação da sociedade da informação. A corrida por meios que proporcionassem um diferencial competitivo entre as nações consideradas como grandes potências também impulsionou a criação de um novo paradigma social, baseado no acesso e na distribuição rápida e eficiente da informação.

Kneller (1980, p. 13) trata da relação da ciência dentro do movimento histórico da civilização afirmando que:

A ciência é intrinsecamente histórica. Não só o conhecimento científico mas também as técnicas pelas quais ele é produzido, as tradições de pesquisa que o produzem e as instituições que as apóiam, tudo isso muda em resposta a desenvolvimentos nelas e no mundo social e cultural a que pertencem. Se quisermos entender o que a ciência realmente é, devemos considerá-la em primeiro lugar e acima de tudo como uma sucessão de movimentos dentro do movimento histórico mais amplo da própria civilização.

A tecnologia como um dos pilares da sociedade da informação provocou transformações consideráveis nas relações científicas, otimizou a comunicação entre os membros de determinadas comunidades, aproximou comunidades de interesses comuns e diferenciados, possibilitou o surgimento de novas ciências e facilitou o compartilhamento do conhecimento científico.

A Sociedade da Informação apresenta-se como uma comunidade global conectada num emaranhado de redes de comunicação, a qual agrega valor à informação e os elementos que possibilitam sua produção, transformação, distribuição e uso.

A interatividade compõe o processo dessa construção social e científica, proporcionando ao indivíduo a condição de agente ativo. O ambiente virtual passa a ser um espaço de criação de significados, símbolos e novas oportunidades. A cultura dá ênfase ao tangível sede lugar ao intangível como nova forma de produção econômica, cultural, social e política. O conhecimento científico é considerado insumo para criação de novas alternativas de comunicação, tanto no âmbito científico como social, político e econômico.

No processo de transformação social pode-se perceber a importância do conhecimento científico no contexto global. A comunicação científica na sociedade da

informação é resultado da constante busca de melhorias na distribuição, no uso e acesso à informação, que, por sua vez, ao sair dos meios acadêmicos, influencia as práticas sociais e rompe paradigmas enraizados nas diferentes culturas espalhadas pelo mundo.

2.2.1 Evolução tecnológica

A evolução tecnológica é proporcional ao poder criativo do homem e está presente em suas constantes conquistas, tanto em âmbito científico, militar ou social. A tecnologia é o próprio conhecimento científico aplicado, fruto da coleta, análise e uso da informação. No decorrer da história, constata-se que a evolução tecnológica esteve presente nos grandes acontecimentos, tendo participação direta ou influenciando as principais mudanças sociais.

As tecnologias na sua essência foram elaboradas para facilitar e melhorar as atividades realizadas pelo homem. Muito embora, o que impulsionou os grandes avanços tecnológicos teve objetivos militares. A corrida armamentista tornou-se muitas vezes o grande trampolim para descobertas de intensa proporção tecnológica. A busca das nações de identificarem-se como grandes potências exigiu investimentos acentuados em pesquisas científicas cujo, o intuito era ser revertido em tecnologia bélica e em comunicação. Os resultados dessas pesquisas proporcionaram muitas descobertas que hoje em dia são encontradas nas residências, empresas e instituições de ensino.

A evolução tecnológica demonstra ser infinita como a imaginação de quem a cria. A própria tecnologia abre portas para novas descobertas. A maior contribuição tecnológica para sociedade está concentrada no campo da comunicação. No final dos últimos trinta anos do século XX, torna-se clara sua importância e o seu poder de transformação. Com o advento de poderosas tecnologias de informação e comunicação surge outro mundo conhecido como ciberespaço, no qual, surgem novas comunidades, culturas, atividades sociais e oportunidades.

2.2.2 Tecnologias de informação e comunicação

As tecnologias de informação e comunicação aparecem com destaque no cenário mundial a partir dos anos de 1970. O processo interativo entre tecnologia e sociedade passa

a ganhar ênfase à medida que sua aplicabilidade alcança os setores civis e comerciais. A acessibilidade e custos cada vez menores com maior qualidade provocam grandes transformações no contexto da comunicação científica.

Castells (2002) apresenta um conjunto de acontecimentos nos anos de 1970 para demonstrar o momento que praticamente deu origem à inclusão das tecnologias de informação e comunicação de maneira incisiva. O autor conta que O primeiro computador eletrônico industrial apareceu em 1969 e o computador digital foi desenvolvido em meados dos anos de 1970, sendo distribuído no comércio em 1977.

O microprocessador, o principal dispositivo de difusão da microeletrônica, foi inventado em 1971 e começou a ser difundido em meados dos anos de 1970. A fibra ótica foi produzida na Palo Alto, matriz de muitas tecnologias de *software* para PCs dos anos de 1990, foi desenvolvida nos laboratórios PARC em Palo Alto, em 1973. O microcomputador foi inventado em 1975, e o primeiro produto comercial de sucesso, o Apple II, foi introduzido em abril de 1977, por volta da mesma época em que a Microsoft começava a produzir sistemas operacionais para microcomputadores em escala industrial pela primeira vez pela Corning Glass, no início da década de 1970. Por fim, foi em 1969 que a ARPA (Agência de Projetos de pesquisa Avançada do Departamento de Defesa Norte-Americana) instalou uma nova e revolucionária rede eletrônica de comunicação que se desenvolveu durante os anos 1970 e veio a se tornar a Internet.

A Internet, por sua vez, revolucionou os meios de comunicação, seu objetivo inicial foi militar, mas com o decorrer do tempo tornou-se difícil separar a pesquisa voltada para fins militares das comunicações científicas e das conversas pessoais.

Em 1983, houve a divisão entre ARPANET, dedicada a fins científicos, e a MILINET, orientada diretamente às aplicações militares. A rede das redes que se formou durante a década de 1980 chamava-se ARPA-INTERNET e depois passou a ser chamada de INTERNET.

A ARPANET encerrou suas atividades em 1990, depois de se tornar obsoleta após vinte anos de serviços prestados. Pressões comerciais e o crescimento de redes privadas e redes cooperativas sem fins lucrativos renunciaram a privatização total da Internet. Em 1995, a tecnologia de transmissão em gigabits alcançou a capacidade de transmitir o equivalente à Biblioteca do Congresso dos EUA em um minuto.

No entanto, foram necessários avanços tecnológicos para a Internet se transformar na atual teia mundial de comunicação. O salto tecnológico que permitiu a difusão da Internet na sociedade em geral, foi a criação de um protocolo chamado *World Wide Web* - WWW em 1990, que organizava o teor dos sítios da Internet por informação, e não por localização, oferecendo aos usuários um sistema fácil de pesquisa para procurar as informações desejadas.

O quadro 01 apresenta um resumo do divisor tecnológico dos anos de 1969 a 1995, destacando algumas tecnologias, instituições e empresas responsáveis pela criação e suas respectivas datas. Este resumo é baseado nas informações apresentadas por Castells no livro “A Sociedade em Rede”.

Quadro 01 – Resumo do divisor tecnológico das décadas de 1960, 1970, 1980 e 1990 baseado em Castells.

TECNOLOGIA	INSTITUIÇÃO / EMPRESA	DATA
Primeiro computador eletrônico industrial	-----	1969
Rede eletrônica de comunicação	ARPA (Agencia de Projetos de pesquisa Avançada do Departamento de Defesa Norte-Americana)	1969
Fibra ótica	Corning Glass	Início – 1970
Microprocessador	-----	1971
Microcomputador	-----	1975
Palo Alto (matriz de muitas tecnologias de software para PCs dos anos 90)	Laboratórios PARC	1973
Sistemas operacionais para microcomputadores.	Microsoft	1975 – 1977
computador digital	-----	1975(desenvolvimento) – 1977(distribuição)
Apple II / microcomputador	-----	1977
ARPANET (dedicada a fins científicos) – MILINET (orientada diretamente às aplicações militares)	ARPA	1983
ARPA-INTERNET	ARPA	Década de 80
Protocolo World Wide Web	Tim Berners-Lee	Fim dos anos 80
ARPANET	ARPA	Término das Atividades em 1990
INTERNET (privatização)	Redes privadas e redes cooperativas	Início dos anos 90
INTERNET (aumento da velocidade de transmissão)	-----	1995

Outras tecnologias surgiram para atender os interesses dos usuários da Internet e suprir algumas carências existentes. Novas formas de comunicação científica e suportes de escrita também foram criados a partir das alternativas oferecidas pela internet, possibilitando a construção de espaços virtuais para o compartilhamento e criação do conhecimento científico. Os *Open Archives* se apresentam como um resultado neste processo de evolução tecnológica, direcionada à comunicação científica e possibilitando a utilização do documento digital para suporte de escrita.

O próximo capítulo apresenta informações sobre os *Open Archives*, tratando de assuntos referentes à tecnologia, história, forma de representação, softwares, iniciativas a nível nacional e internacional, principais características da *Open Archives Initiatives – OAI* e instituições que oferecem suporte. Sobre o protocolo OAI-PMH é apresentado informações relacionadas com a sua evolução, as tecnologias envolvidas e algumas considerações sobre implementação e configurações.

3 OPEN ARCHIVES

O surgimento da Internet otimizou a comunicação entre os pesquisadores das várias áreas das ciências. Por intermédio de uma mídia de fácil acesso, rápida disponibilização de dados e de grande alcance geográfico tornou-se possível o intercâmbio de idéias entre os pares de forma mais eficiente. Esses novos recursos tornaram visível o contraste de tempo de produção e distribuição das publicações impressas e as publicações digitais. Proporcionaram também a quebra de antigos paradigmas sobre os sistemas tradicionais de divulgação de publicações e tornaram possível a elaboração de novos sistemas de comunicação científica.

Os *Open Archives* são um bom exemplo da quebra de antigos paradigmas e da criação de novas alternativas para comunicação científica. Os *Open Archives* são repositórios de trabalhos científicos em forma digital, o primeiro deles, o arXiv.org, surgiu em 1991, por iniciativa de Paul Ginsparg, do *Los Alamos National Laboratory*, e desempenhou papel importante junto à comunidade científica de Matemática, Física, Astronomia e Ciência da Computação.

Os *Open Archives* são arquivos digitais que incluem não só versões eletrônicas preliminares de documentos científicos (*pre-prints*), mas também artigos aceitos para publicação pelo processo de revisão tradicional por pares (*e-prints*). Além disso, aceitam também anotações e atualizações feitas pelos autores ou comentários feitos pelos pares. Na realidade, os arquivos abertos são Bibliotecas Digitais desenvolvidas na *Web* por cientistas e para cientistas, constituindo-se em fóruns privilegiados para difusão de resultados e debate científico, ficando entre a comunicação formal e a informal (OPEN ARCHIVES FORUM, 2005).

Bailer (2005) realizou um levantamento bibliográfico exaustivo sobre os *Open Archives*, sua pesquisa obteve o resultado de mais 1300 referências em relação ao assunto.

Sena (2000) afirma que nos Estados Unidos e em alguns países europeus, existe um número crescente de coleções de *e-prints*. O fluxo de trabalhos científicos que está sendo direcionado para esses arquivos já os consagrou como modelo alternativo de comunicação para algumas áreas de pesquisa. Os pares tornam-se seus próprios editores. A publicação dos trabalhos é de inteira responsabilidade do autor que faz a forma automatizada, seja por

correio eletrônico, seja por transferência de arquivos via ftp ou por envio através de um formulário disponível na *Web*. A revisão dos pares e o quesito de ineditismo não são barreiras para publicação.

A comunicação informal por pares, com poucas restrições para disponibilidade de informações e retorno rápido é uma prática desenvolvida antes mesmo do advento da Internet. Macedo (1999) discorre sobre os colégios invisíveis identificados nos anos 1960, constituindo-se em comunidades informais de cientistas que se comunicando e reunindo-se periodicamente para troca de experiências e informações, passaram a desempenhar papel fundamental para publicação formal e disseminação dos avanços no campo do conhecimento científico. Contrastados com os canais mais formais, apresentam vantagens na atualização das informações, oportunidades de *feedbacks* e do compartilhamento de estágios formativos de suas pesquisas. Kneller (1980, p. 183) conceitua colégio invisível afirmando que “um colégio invisível é um grupo ou escola de cerca de dez a uma centena de cientistas trabalhando numa tradição de pesquisa. Os seus membros mantêm-se em contato assíduo, usualmente verbal, e evitam os canais mais lentos de comunicação formal.” O surgimento da Internet possibilitou a prática deste modelo ser realizada em ambiente virtual, passando a ser chamada de colégio virtual. Moreira (2005) discorre sobre os colégios virtuais relatando que “a comunidade mantém-se a par dos respectivos trabalhos e troca informações via correio eletrônico ou listas de discussão, preferencialmente, embora utilize outros meios. O colégio virtual atua como rede de comunicação e intercâmbio, como fórum de educação e de socialização dos novos cientistas”.

A comunicação humana pressupõe e exige quatro elementos: o emissor, uma pessoa ou um grupo de pessoas com objetivo ou uma razão para empenhar-se em comunicação; a mensagem, tradução de idéias, objetivos e intenções num código, num conjunto sistemático de símbolos; o receptor que capta a mensagem; e o retorno do receptor que concretiza a comunicação em uma prática de acuso de recebimento e entendimento. A figura 01 apresenta os elementos da comunicação humana.

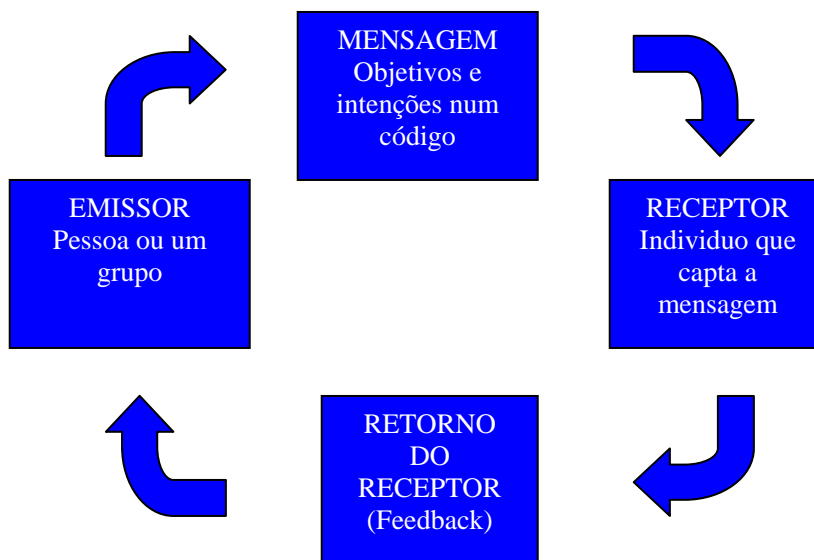


Figura 01 – Elementos da comunicação humana

No caso da comunicação científica o processo é o mesmo, a comunidade científica ou o cientista se comunica para atender seus objetivos, suas necessidades e seus interesses conforme o conjunto de normas impostas pelo campo que está inserido, envia mensagens para participar suas idéias e espera obter retorno de seus pares. Os *Open Archives* como meio de comunicação científica possibilitam a execução dessa prática. Seguindo algumas normas para possibilitar a interoperabilidade dos metadados, o autor disponibiliza seus trabalhos a comunidade em que está inserido, tendo autonomia para o auto-arquivamento e editoração do seu trabalho, assumindo, desta forma, o papel de emissor. A mensagem, no caso o trabalho do autor (emissor), é enviado e disponibilizado na base de dados do *Open Archives*. O receptor, membro da comunidade, acessa o trabalho (mensagem) por intermédio dos recursos disponíveis pelo *Open Archive*, analisa as informações e apresenta suas considerações sobre o trabalho, possibilitando desta forma *Feedback* imediato com seus pares.

A utilização dos *Open Archives* está caracterizado por um processo de democratização da informação no contexto acadêmico e científico. A disponibilização de *softwares* livres por parte da *Open Archives Initiative - OAI* com intuito de possibilitar o uso, o acesso, a distribuição e o compartilhamento do conhecimento científico apresenta

uma política voltada ao nivelamento de condições para publicação acadêmica no cenário global.

Kuramoto (2005, a) trata da questão da democratização da informação utilizando os *Open Archives* afirmando que:

Os *Open Archives* podem ser uma efetiva ação de inclusão. À medida que é facilitado o acesso à informação, com os repositórios livres, qualquer um pode acessar as informações que estão nestes repositórios. Tanto aqueles que têm acesso às revistas estrangeiras, quanto àqueles que não têm. Os *Open Archives* são um forte instrumento de inclusão, e não de exclusão. Quanto maior o acesso público à informação, maior será a possibilidade de ampliarmos a comunidade de usuários, até mesmo de levar esta informação para as comunidades que não têm este acesso.

A elaboração de formas mais eficientes para comunicação científica é imprescindível para o desenvolvimento da ciência. Os *Open Archives* são uma realidade e estão em processo de adaptação para atender as necessidades dos pesquisadores. Para isso, a comunidade científica está criando um conjunto de iniciativas com o intuito de tornar esta prática acessível e viável para todos.

3.1 Open Archives Initiative – OAI

No ano de 1999, Paul Ginsparg, Rick Luce e Herbert Van Sompel organizaram uma grande discussão sobre os serviços de auto-arquivamento por parte dos pesquisadores para a transição de seus trabalhos acadêmicos. Este encontro foi chamado de Convenção de Santa Fé, onde se apresentou uma nova filosofia para a publicação científica. A *Open Archives Initiative – OAI* nasceu neste evento com o objetivo de desenvolver e promover soluções de interoperabilidade que facilitassem uma disseminação eficiente do conteúdo dos *Open Archives* (OPEN ARCHIVES FORUM, 2004).

A política de decisões sobre a *Open Archives Initiative* é realizada por um Comitê Administrativo. A infra-estrutura de interoperabilidade foi desenvolvida por um comitê técnico, o qual continua tratando das experiências desenvolvidas pela iniciativa. Herbert Van de Sompel e Carl Lagoze são responsáveis pela coordenação das atividades da *Open Archives Initiative - OAI*, os quais estão concentrados na *Cornell University*.

A *Open Archives Initiative – OAI* possui na sua raiz um esforço para intensificar o acesso aos arquivos de *e-prints* como um meio de aumentar a disponibilidade da comunicação acadêmica. O auxílio continuado deste trabalho traça uma base do programa de *Open Archives* proposto pela iniciativa. O fundamento de estrutura tecnológica e os padrões que são desenvolvidos para auxiliar este trabalho são conseqüentemente, independentes dos tipos de conteúdos oferecidos e do mecanismo econômico que os envolve. A *Open Archives Initiative – OAI* visa ter ampla relevância na abertura para acesso a grandes quantidades de materiais digitais. Como um resultado de seu trabalho, a *Open Archives Initiative – OAI* está movimentando uma organização e um nítido esforço em transição, e está também comprometida a explorar e habilitar este novo e grande campo de aplicações, no caso, o universo que envolve os *Open Archives* (OPEN ARCHIVES INITIATIVE, 2005).

O objetivo principal desta iniciativa é gerar um banco de dados que ofereça um sistema operacional de fácil compreensão e aplicabilidade à comunidade científica, para servir de estímulo a produção e publicação científica e acionar funções de compartilhamento e interação entre as academias (OPEN ARCHIVES FORUM, 2004).

Lagoze e Sompel (2003) afirmam que as raízes da *Open Archives Initiative – OAI* encontram-se dentro de uma visão que busca estimular o crescimento dos repositórios de *e-prints*. O propósito que conduziu a criação das tecnologias foi de confederar e agregar valor a informação nos repositórios individuais via serviços através dos próprios repositórios.

Weitzel (2005, p. 162) discorre sobre *Open Archives Initiative - OAI* afirmando que:

A *Open Archives Initiative* ou Iniciativa dos Arquivos Abertos (OAI), movimento instituído por um grupo de pesquisadores, constitui iniciativa pioneira na reflexão sobre a revisão no processo de comunicação científica, tomando como referência a integração de soluções tecnológicas e a sua implantação. A OAI apóia-se em ferramentas e procedimentos fundamentais, quais sejam: auto-arquivamento pelos autores; metadados padronizados para descrição; acesso livre à produção científica e o protocolo OAI. Este viabiliza a interoperabilidade entre os diferentes arquivos e permite, dentre outras funções, a recuperação e o compartilhamento das informações.

Em se tratando de propriedade intelectual a *Open Archives Initiative – OAI* não define ou prescreve algum esquema de gerenciamento de direitos. Publicações de acesso restrito ou gerenciamento de propriedade intelectual por metadado exposto são de responsabilidade do Provedor de Dados ou de Serviço que adota o protocolo OAI-PMH. A

Open Archives Initiative – OAI espera que os repositórios adeptos ao protocolo permitam mais acessos aos seus metadados e menos restrições.

3.1.1 Características principais da *Open Archives Initiative – OAI*

A *Open Archives Initiative – OAI* possui quatro características principais que especificam sua usabilidade e disponibilidade, as quais se dividem categoricamente em auto-arquivamento, interatividade, interdisciplinaridade e interoperabilidade.

A primeira característica é chamada de auto-arquivamento pela possibilidade do autor ter autonomia para editar e arquivar seus textos. Esta liberdade permite em alguns casos ratificações e atualizações da obra, valorizando o conteúdo informacional constantemente.

A segunda, a interatividade é uma característica que diz respeito à interação entre a comunidade acadêmica, possibilitando a troca de informações, críticas, comentários e sugestões do texto disponível neste sistema. Esta forma de comunicação garante legitimidade e aponta o prestígio e aceitação das informações contidas no texto. Isto acontece através de convergências e confrontos de idéias, discussões, críticas e consensos entre a comunidade participante.

A terceira característica trata da interdisciplinaridade do sistema por envolver um leque de obras de diversas áreas do conhecimento em um repositório digital.

A quarta característica é a interoperabilidade, a qual se destaca por ser a mais peculiar de todas, pois objetiva configurar todos os arquivos, padronizando-os globalmente, para que o acesso a eles seja ilimitado, sem restrições de plataformas.

As características principais da OAI apontam pontos positivos. Muito embora, a utilização do modelo de comunicação *Open Archives* apresentar questões em aberto quanto ao seu futuro promissor no ambiente científico.

Entre as questões referentes ao uso do modelo *Open Archives* como alternativa de comunicação científica destaca-se a absorção dos custos de criação e as dificuldades de aceitação como item de produtividade.

A absorção dos custos de criação dos *Open Archives* desperta a questão de como as instituições interessadas poderão custear os gastos com material tecnológico, humano e

infra-estrutura. Além da manutenção, divulgação e muitas vezes a incerteza da continuidade das iniciativas que oferecem os *softwares* livres para implementação dos *Open Archives*.

Em relação aos direitos autorais, a Open Archives Initiative (2005) discorre que "o gerenciamento da propriedade intelectual dos conteúdos dos documentos digitais são de responsabilidade do Provedor de Dados ou de Serviço que adota o protocolo OAI-PMH." Este fato, levanta indagações sobre a confiabilidade nos conteúdos dos trabalhos disponibilizados nos *Open Archives*.

Apesar de suas características serem positivas a OAI é uma iniciativa que foi criada em 1999 e por ser considerada nova no ambiente científico, sua existência e suas atividades levantam interrogações acerca de sua utilização enquanto alternativa de comunicação e suporte de escrita para disseminação e publicação da produção científica nacional e internacional. A próxima seção trata da arquitetura dos *Open Archives* e suas formas de representação.

3.1.2 Arquitetura dos Open Archives

Na arquitetura dos Open Archives existem duas maneiras de representação, os Provedores de dados (Data Provider) e os Provedores de Serviço (Service Provider).

3.1.2.1 Provedores de dados

Os Provedores de Dados (*Data Provider*) são os *Open Archives* que organizam os documentos digitais por meio de seu protocolo, coletando dados e expondo os metadados dos documentos. (OPEN ARCHIVES FORUM, 2005).

Os Provedores de Dados controlam o depósito e editoram os recursos em um repositório, expondo-os para colheita do metadado conforme os recursos disponíveis. Eles são os criadores e mantenedores dos metadados e dos recursos dos repositórios. (OPEN ARCHIVES INITIATIVE, 2005).

Em se tratando de vantagens a OAI não promete algum benefício imediato na adoção do OAI-PMH. As motivações para adoção do protocolo dependem do tipo de participação da comunidade interessada. (OPEN ARCHIVES INITIATIVE, 2005).

Um Provedor de Dados adota o OAI-PMH por várias razões incluindo: membros em uma comunidade que coletivamente tem dedicado o uso do protocolo como um meio de aumentar a exposição de seus trabalhos, interesse no funcionamento oferecido por Provedores de Serviço específicos ou interesse geral na noção de exposição de metadados como um meio de compartilhar conhecimento. (OPEN ARCHIVES INITIATIVE, 2005).

O Provedor de Dados pode estar informado de outros repositórios em conformidade com o OAI-PMH. Um exemplo é o Provedor de Dados membro de um grupo ou comunidade de Provedores de Dados associados. No *friends container* (*Containers* que fazem parte de uma lista disponível na OAI) o identificador de resposta pode ser usado por repositórios para lista de repositórios associados. Estes provedores podem de maneira automática usar os coletores para descobrir outros provedores de seu interesse. O uso espalhado de seu *container* proveria um mecanismo descentralizado pelos quais os coletores podem descobrir outros repositórios. (OPEN ARCHIVES INITIATIVE, 2005).

Os Provedores de Dados que estão em conformidade com o OAI-PMH podem escolher se registrar diretamente na página de registros no endereço <http://www.openarchives.org/data/registry/provider.html>. O registro na base de dados serve como uma lista de acesso público dos repositórios em conformidade com o OAI-PMH, tornando possível para os Provedores de Serviço descobrirem sites em que os metadados podem ser coletados. O registro na base de dados contém todas as informações disponíveis direto do identificador de solicitação OAI incluindo: nome do repositório – um nome humanamente legível para o repositório, como por exemplo “*The University of Spa E-print System*”; *baseURL* – para fazer o protocolo de solicitação ao repositório; *protocolVersion* – a versão do OAI-PMH usada pelo repositório; *adminEmail* – o endereço eletrônico do administrador do repositório; e informação adicional do protocolo e o pacote de descrições de um repositório específico. Os Provedores de Dados que estão registrados na OAI concordam em tornar as informações públicas, não existem provisões para restringir acesso ao registro da base de dados. (OPEN ARCHIVES INITIATIVE, 2005).

A figura 02 apresenta o fluxo de informação procedente nos Provedores de Dados. No exemplo, os autores realizam o depósito e a editoração de seus *eprints*, os Provedores de Dados fazem o controle dos recursos e expõem os metadados para colheita.

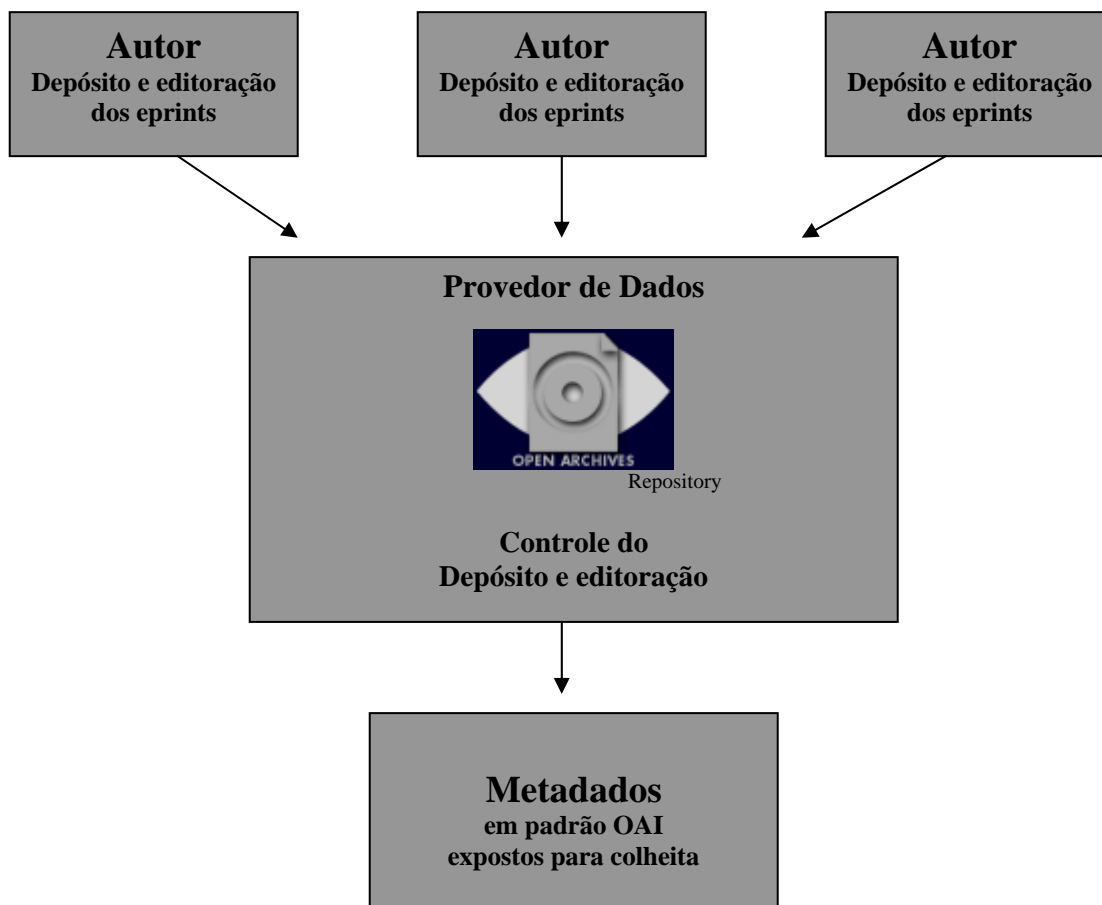


Figura 02 - Fluxo de informação procedente nos Provedores de Dados.

O fluxo de informação procedente nos Provedores de Dados mostra que a simplicidade existente no processo de auto-arquivamento e possibilidade de editoração dos *e-prints* por parte dos autores otimiza as atividades referentes a exposição de seus trabalhos.

3.1.2.2 Provedores de serviços

Os Provedores de Serviços funcionam como uma biblioteca virtual, que busca as informações de forma aleatória em vários depósitos informacionais dos *Open Archives* ao mesmo tempo (OPEN ARCHIVES FORUM, 2005).

Como nos Provedores de Dados a OAI não pode prometer algum benefício imediato aos Provedores de Serviço que adotam o OAI-PMH. As motivações para adoção dependem do tipo de participação e dos objetivos da comunidade interessada. Os Provedores de

Serviço devem adotar o OAI-PMH para seus interesses e competência em combinar e agregar valor ao metadado disponível em um campo de Provedores de Dados (OPEN ARCHIVES INITIATIVE, 2005).

Os Provedores de Serviço fazem a colheita dos metadados nos Provedores de dados, usando o metadado colhido com o propósito de promover um ou mais serviços através de todos os dados. Os tipos de serviços que podem ser oferecidos incluem uma interface de pesquisa, sistemas de revisão por pares, etc. Destaca-se que um “Provedor de Serviço” pode ter ambas as funções, tanto oferecer dados para colheita como serviços para usuário final. A mudança chave na Arquitetura dos Provedores de Serviço foi alteração nas interfaces de usuário final de suporte exclusivo humano para o suporte de ambos os sistemas, no caso das interfaces de usuário final e interfaces mecânicas para colheita. (OPEN ARCHIVESFORUM, 2005).

A figura 03 apresenta um esquema de interação dos Provedores de Serviço. Neste exemplo, O Provedor de Serviço solicita ao Provedor de Dados seus registros e obtém resposta em padrão OAI.

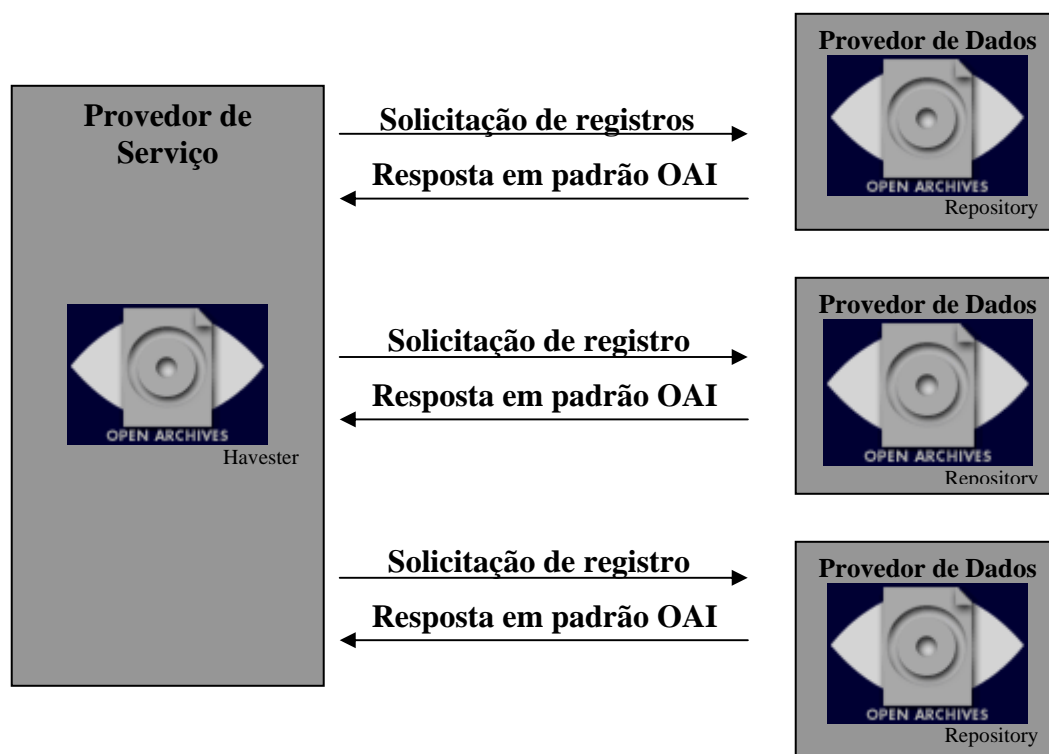


Figura 03 - Esquema de interação dos Provedores de Serviços.

No esquema de interação apresentado na figura 03 os Provedores de Serviço além de terem acesso aos metadados dos Provedores de Dados e de outros Provedores de Serviço têm a possibilidade de agregar valor às informações coletadas.

A seguinte seção trata de softwares utilizados para implantação dos *Open Archives* no Brasil e no mundo. Estes programas são distribuídos livremente e compatíveis com o *Open Archives Initiative – Protocol for Metadata Harvesting - OAI-PMH*.

3.1.3 Software para implantação dos *Open Archives*

Os *softwares* para implantação dos *Open Archives* são desenvolvidos por instituições que na sua maioria os distribuem livremente, incentivando outras instituições a adotarem os sistemas para disseminação e publicação de suas produções. No cenário internacional existente vários *softwares*, nesta seção dá-se destaque ao *GNU EPrints*, *DSpace*, *CDSware*, *OPUS*, *Fedora*, *Open Journal Systems* e o *SEER*.

O *GNU EPrints* é um *software* genérico para arquivos desenvolvido pela universidade de *Southampton*. Ele é destinado para criar arquivos *Web* altamente configuráveis. O objetivo principal é ser configurado como um *Open Archives* para trabalhos de pesquisa e o padrão de configuração reflete isto, mas ele pode ser facilmente usado por outros documentos como imagens, dados de pesquisa, arquivos de áudio, documentos que podem ser armazenados digitalmente (EPRINTS.ORG, 2006).

O *DSpace* é um sistema de repositório digital que captura, armazena, indexa, preserva e redistribui materiais de pesquisa em formato digital. O *DSpace* é desenvolvido conjuntamente pela *MIT Libraries* e a *Hewlett – Packard* (HP) (DSPACE.ORG, 2006).

O sistema integrado de biblioteca digital *CDSware* é um conjunto de aplicações as quais fornecem estrutura e ferramentas para construir e gerenciar um servidor autônomo de biblioteca digital. Ele é um *software* livre e licenciado pela *GNU General Public Licence* (GPL). A tecnologia oferecida pelo *software* cobre todos os aspectos de gerenciamento da biblioteca digital. O *CDSware* é compatível com o *OAI-PMH* e usa o *MARC 21* como padrão bibliográfico (CDSWARE SOFTWARE CONSORTIUM, 2006).

O sistema *OPUS* foi disponibilizado para todos os membros da universidade de *Stuttgard* em 1998. Os resultados dos projetos são destinados para serem distribuídos e o

sistema estabelecido em outras universidades ou instituições para que mais documentos digitais sejam disponíveis à comunidade científica (OPUS-DATABASE, 2006).

O *Fedora* é um sistema de repositório digital desenvolvido pela *Cornell University Information Science* em conjunto com a *University of Virginia Library*. O projeto *Fedora* é desenvolvido com o objetivo de promover *softwares* de repositórios *open source*. O software é disponibilizado seguindo os termos da *Educational Community License – ECL* (FEDORA.INF, 2006).

O *ETD-db* foi desenvolvido na *Virginia Tech* como um projeto de junção entre a *Graduate School at Virginia Tech, Digital Library and Archives* e a *National Digital Library of Theses and Dissertations*. O *software* é uma série de páginas *Web* e roteiros que interage com um *MySQL database*. Estes roteiros proporcionam uma interface padrão para usuários *Web* e pesquisadores, acadêmicos e bibliotecas acadêmicas utilizar e gerenciar os arquivos e metadados de coleções eletrônicas de teses e dissertações (ETD-DB, 2006).

O *Open Journal Systems - OJS* é um sistema de gerenciamento de revista eletrônica que tem sido desenvolvido pelo *Public Knowledge Project* da Universidade de *British Columbia* para expandir e melhorar o acesso às pesquisas (PUBLIC KNOWLEDGE PROJECT, 2006).

O Sistema Eletrônico de Editoração de Revistas – SEER é um *software* desenvolvido para construção e gestão de uma publicação periódica eletrônica. Esta ferramenta completa ações essenciais à automação das atividades de editoração de periódicos científicos. O SEER foi traduzido e customizado pelo IBICT baseado no *software Open Journal Systems*. (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2006).

Na próxima seção são apresentadas informações sobre instituições ligadas a *Open Archives Initiative – OAI* que oferecem suporte para suas atividades.

3.2 Instituições que oferecem suporte a *Open Archives Initiative – OAI*

A *Open Archives Initiative – OAI* é formada por um comitê diretor que é composto por conceituadas universidades e outras entidades com interesse na disponibilização de informações digitais e por um comitê técnico presidido por Carl Lagoze e Herbert Van de

Sompel. A iniciativa é patrocinada por *Digital Library Federation, Coalition for Networked Information* e *Natural Science Foundation* (OPEN ARCHIVES INITIATIVE, 2005).

A *Digital Library Federation - DLF* está ultimamente subordinado ao *Council on Library Information Resources - CLIR*. A *DLF* é uma organização com uma abrangência limitada e com recursos limitados. Além disso, como uma “federação”, a *DLF* é um conselho planejado para ter uma organização central limitada. Deste modo, a *DLF* precisa de uma estrutura com cinco componentes essenciais: o escritório do diretor executivo, um comitê administrativo, grupos consultivos, iniciativa *DLF* e o Fórum *Digital Library Federation - DLF* (DIGITAL LIBRARY FEDERATION, 2005).

O escritório do diretor executivo consiste em um pequeno corpo de trabalho e as várias estruturas de assistência que a *Council on Library Information Resources - CLIR* proporciona. Ele é responsável pelos objetivos e direção da Federação, gerenciando o programa de atividades e finanças da *DLF*, auxiliando os membros da *DLF*, comunicando sobre a *DLF* e suas atividades para toda comunidade, estabelecendo alianças externas e recrutando novos parceiros; o comitê administrativo é o corpo administrativo da *DLF*, ele se encontra três vezes por ano e proporciona supervisão para o programa de atividades e iniciativas da *DLF*; os grupos consultivos podem ser conveniados ao diretor executivo para tratar sobre programas desenvolvidos em áreas chaves, eles compreendem em grupos de especialistas de instituições da *DLF* e ajudam a articular e priorizar necessidades para pesquisa, desenvolvimento e compartilhar informação entre essas áreas; a iniciativa- *DLF* discute a pesquisa específica, o desenvolvimento, ou o compartilhamento de informação necessário que é identificado pelos grupos consultivos ou pelo amplo número de membros; e os Fóruns *Digital Library Federation* são realizados periodicamente e incluem um número de patrocinadores de biblioteca digitais de cada instituição membro da *DLF*, esses fóruns servem como lugares de encontro, lugares para divulgação e congressos. (DIGITAL LIBRARY FEDERATION, 2005).

A *Coalition for Networked Information - CNI* é uma organização dedicada a auxiliar no compromisso de transformação das redes de tecnologia de informação para o avanço da comunicação acadêmica e enriquecimento da produção intelectual. A *CNI* representa cerca de 200 instituições de ensino superior, editoras, redes e telecomunicações, tecnologia de informação, bibliotecas e organizações de bibliotecas.

A *CNI* é patrocinada pela *Association of Research Libraries - ARL* e *EDUCAUSE*, e é gerenciado por um comitê administrativo dirigido por Richard P. West da *California State University*. Possui um diretor executivo, Clifford A. Lynch, que coordena um corpo de trabalho no escritório central em *Washington*. (COALITION for NETWORKED INFORMATION, 2005).

A *National Science Foundation - NSF* é uma agência federal independente criada pelo congresso americano em 1950, para promover o progresso da ciência, avanço da saúde e segurança nacional. Com um orçamento anual \$ 5.5 bilhões, é fonte de apoio para aproximadamente 20% de todo suporte básico para pesquisas conduzidas pelas faculdades e universidades americanas. Em muitos campos semelhantes ao da matemática, ciência da computação e ciências sociais a *NSF* é o maior recurso de apoio federal dos Estados Unidos (NATIONAL SCIENCE FOUNDATION, 2005).

Como descrito no seu plano estratégico, a *NSF* é a única agência federal americana cuja missão é oferecer suporte a todos os campos da ciência fundamental e engenharia, exceto para ciência médica. A *National Science Foundation - NSF* está incumbida de conduzir as principais descobertas nas áreas da astronomia para geologia e para zoologia nos Estados Unidos. Em função de financiar pesquisas nas áreas acadêmicas tradicionais, a agência também oferece suporte a idéias de alto risco, novas colaborações e numerosos projetos que podem parecer a princípio ficção científica, mas para a agência pode levar as grandes descobertas no futuro. Em todos os casos, a *National Science Foundation - NSF* assegura que a pesquisa é totalmente integrada com educação e todos os dias são realizados trabalhos revolucionários que irão preparar cientistas e engenheiros de primeira linha para o futuro. (NATIONAL SCIENCE FOUNDATION, 2005).

No Brasil existem algumas iniciativas para incentivar a adoção do modelo *Open Archives* para produção e disseminação da produção científica. Estas iniciativas são apresentadas na seção seguinte.

3.2.1 Iniciativas para criação de Open Archives no Brasil

No cenário internacional, muitos países, indiferente da situação econômica e cultural, estão se apropriando das tecnologias oferecidas pela *Open Archives Initiative* –

OAI e de outras iniciativas a nível internacional para criarem seus *Open Archives*. No Brasil, existem algumas iniciativas que incentivam instituições a usarem tecnologias baseadas na filosofia *Open Archives* para produção e disseminação da publicação científica nacional.

Entre os países que utilizam as tecnologias *Open Archives*, o Brasil se encontra no topo, apresentando presença bastante considerável entre os países desenvolvidos e em desenvolvimento. O Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT está apresentando algumas iniciativas para incentivar a adoção dos *Open Archives* no Brasil e as ferramentas eletrônicas que possibilitarão a multiplicação dos repositórios nacionais de informação científica.

Kuramoto (2005, a) coordenador de Projetos Especiais do IBICT apresenta algumas razões para incentivar instituições no Brasil a implantarem o modelo *Open Archives*:

Uma das razões principais é que faz parte da missão do IBICT o registro e a disseminação da produção científica brasileira. Assim, devemos, por exemplo, registrar as teses e dissertações, registrar e disseminar os artigos dos nossos pesquisadores nas várias áreas do conhecimento e registrar outros materiais e documentos, como relatórios técnicos, capítulos de livro etc [...] Outra razão é que o acesso à informação é muito caro. Nosso país gasta cerca de 20 milhões de dólares para manter o Portal de Periódicos da Capes. Os periódicos estrangeiros são muito caros, o acesso é muito caro, a manutenção é muito cara para o país [...] A idéia é que, por meio da criação destes repositórios de acesso livre, seja possível facilitar ao pesquisador publicar e ter acesso à informação, desenvolvendo, em melhores condições, a sua pesquisa [...] E, se no mundo todo este movimento está se fortalecendo, com a criação de novos repositórios e com a facilidade ao acesso à informação, o país pode se beneficiar deste processo, pois, além de criar seus repositórios, pode se integrar e interoperar os nossos repositórios com os repositórios estrangeiros. Isso daria acesso, a um custo muito menor, às informações que estão tanto no país quanto fora do país.

Os *Open Archives* apresentam vários tipos de conteúdos conforme os objetivos das instituições e das comunidades que os criam. Existem *Open Archives* de revistas, teses e dissertações, alguns que possibilitam o acesso a artigos científicos, a relatórios técnicos e a anais de congressos. O Brasil possui material com informações de grande relevância para pesquisadores de todo mundo, muitas dessas informações estão sem a possibilidade de serem acessadas por pesquisadores dispersos nacionalmente e globalmente. O Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT está desenvolvendo algumas iniciativas para promover a criação dos *Open Archives*.

Kuramoto (2005, a) relata que:

Uma das iniciativas que o IBICT vem desenvolvendo é a Biblioteca de Teses e Dissertações – BDTD, que já está em curso e hoje tem cerca 15 mil trabalhos depositados, com texto integral [...] Outra iniciativa é disseminar e transferir, de forma massiva, o *software* de editoração eletrônica de revistas, ou melhor, o Sistema Eletrônico de Editoração de Revistas – SEER [...] Outra iniciativa que é o *software* DSPACE, que é instrumento para criação de repositórios, que aceita tanto artigos, teses e dissertações, quanto material didático, e tem seus utilitários, os *Open Sources*.

Experiências pioneiras como repositório de *pre-prints* do Instituto Nacional de Matemática Pura e Aplicada – IMPA, relatado em artigo por Chataignier e Silva (2001) e o Sistema de Arquivos Abertos da Sociedade Brasileira de Genética tendo o apoio do Instituto Brasileiro de Informação em Ciência e Tecnologia (SISTEMA DE ARQUIVOS ABERTOS DA SOCIEDADE BRASILEIRA DE GENÉTICA, 2006), são exemplos das primeiras instituições que adotarão o modelo *Open Archives* para auxiliar suas atividades de produção científica.

No contexto nacional e internacional as iniciativas de incentivo assumem papel importante na criação dos *Open Archives* no Brasil, pois divulgam o modelo, disponibilizam as tecnologias e oferecem as instituições interessadas segurança na adoção do sistema.

A próxima seção apresenta informações sobre a tecnologia do protocolo OAI-PMH e outras tecnologias utilizadas para seu desenvolvimento.

3.3 Open Archives Initiative – Protocol for Metadata Harvesting - OAI – PMH

O *Open Archives Initiative – Protocol for Metadata Harvesting - OAI-PMH* é um protocolo que possibilita aos participantes da iniciativa OAI compartilhar seus metadados. Muitas promessas são vistas quando se trata do uso do protocolo dentro da abordagem dos *Open Archives*. O suporte de um novo modelo para comunicação acadêmica é o maior benefício em potencial que o protocolo oferece para publicação. (OPEN ARCHIVES INITIATIVE, 2005).

Na visão de Arellano, Ferreira e Caregnato (2005, p.196) o OAI-PMH, “trata-se de um protocolo que estimula a geração de acesso aberto, disseminando a produção científica para acessos locais e globais, alinhando-se, portanto, aos movimentos dos arquivos abertos e do livre acesso à informação e ao conhecimento em ciências e humanidades”.

Para a *Open Archives Initiative - OAI* a maior possibilidade dos *Open Archives* é fazer emergir recursos escondidos e apresentar solução de baixo custo de interoperabilidade. Apesar do OAI-PMH ser tecnicamente muito simples, a construção de serviços coerentes reúne solicitações do usuário com traços complexos. O OAI-PMH poderá se tornar parte da infra-estrutura da *Web* de forma tão importante como o protocolo HTTP é atualmente, desde que a combinação de suas respectivas simplicidades e o sucesso provado na facilidade dos seus implementos alcancem um contexto nivelador para o conhecimento espalhado pelas organizações de pesquisa, editoras, e outras organizações (OPEN ARCHIVES FORUM, 2004).

3.3.1 Histórico do OAI-PMH

Com o reconhecimento por parte da comunidade *e-print* em relação aos benefícios advindos da interoperabilidade alcançada por *Open Archives* distribuídos, fundou-se a *Open Archives Initiative - OAI* como um meio para acordar uma estrutura de interoperabilidade. No ano de 2000 a *Open Archives Initiative - OAI* elaborou um encontro em Santa Fé para estabelecer uma infra-estrutura técnica para possibilitar interoperabilidade. Em 2001 a *Open Archives Initiative - OAI* especificou um protocolo para colheita de metadados como uma base para provisão de serviços interoperáveis. O protocolo foi testado através de instrumentos durante 2001 e 2002. Contudo, cresceu o reconhecimento por parte da comunidade *e-print* em relação ao modelo técnico e de serviço oferecido pela *Open Archives Initiative - OAI*, mostrando potencial de grande alcance para comunidade *e-print*. Todavia, a *Open Archives Initiative - OAI* oferece potencial para novas opções de serviço, ambas como relação de conteúdo e modelo de serviço (OPEN ARCHIVES FORUM, 2005).

A convenção de Santa Fé foi à primeira encarnação do *Open Archives Initiative Protocol for Metadata Harvesting - OAI-PMH*. Desenhado sobre o protótipo do *Universal*

Preprint Service - UPS, o modelo *RePEc/SODA Service/Data provider*, o protocolo *Dienst* e o trabalho do grupo Santa Fé. O OAI-PMH 1.0 introduziu o conjunto de elementos *Dublin Core* como uma base para interoperabilidade dos metadados, desenhado a partir da Convenção de Santa Fé, encontros da *Digital Library Federation*, trabalhos na *Cornell University* e *feedback* da *alpha-testers*. O protocolo tornou-se então uma barreira muito pequena para especificação de interoperabilidade, baseado em um modelo de colheita de metadados, e baseado em HTTP, usando HTTP GET / POST para solicitação e XML para resposta. Destaca-se que o OAI-PMH não é um protocolo de pesquisa, mas sim um sistema baseado em um modelo de colheita de metadados. O OAI-PMH 1.1 foi uma revisão das especificações da versão 1.0 utilizando as mudanças para a adoção das especificações do esquema XML. Ambas as versões 1.0 e 1.1 foram de natureza experimental (OPEN ARCHIVES FORUM, 2005).

A versão inicial do OAI-PMH, 1.0, foi realizada em janeiro de 2001. A menor versão atualizada, 1.1, foi realizada em julho de 2001 conforme as mudanças nas especificações do esquema XML. A versão atual do OAI-PMH, 2.0, foi realizada em junho de 2002. Em adição as mudanças técnicas, a versão 2.0 do OAI-PMH marca o fim do período de experimentação inicial para o protocolo. Esse período inicial durou 18 meses e foi o meio de testar os aspectos funcionais do protocolo e a idéia geral da utilidade da colheita de metadados. O uso substancial do OAI-PMH durante esse período teve validade tanto para validar a colheita de metadados como uma ferramenta de informação para federação e promover a formação das mudanças técnicas e otimização na versão 2.0 (OPEN ARCHIVES INITIATIVE, 2005).

A versão 2.0 do OAI-PMH não é compatível com as versões anteriores. Os resultados no período de experimentação da versão 1.0 e 1.1 indicaram que um número de mudanças fundamentais foi necessário e o fato de manter as versões anteriores compatíveis teria resultado em uma liberação no que trata o compromisso de produção (OPEN ARCHIVES INITIATIVE, 2005).

A OAI está comprometida em considerar a compatibilidade de versões anteriores em futuras liberações. O objetivo da OAI foi sair da versão 1.1 do OAI-PMH no fim de 2002. Certamente Provedores de dados e de Serviço individuais podem continuar suportando versões anteriores, mas os Provedores de dados e de Serviço registrados

somente incluirão e aceitarão sites que estão na versão 2.0 OAI-PMH compliant (OPEN ARCHIVES INITIATIVE, 2005).

Com pouco tempo experimental, a versão 2.0 tornou-se um protocolo estável e a OAI tem se comprometido a fazer revisões subseqüentes do protocolo em relação à compatibilidade das versões anteriores. O quadro 02 apresenta um resumo histórico da evolução do OAI-PMH segundo informações obtidas no site do *Open Archives Forum*.

Quadro 02 – Resumo histórico da evolução do OAI-PMH.

Evolução Mudanças Estruturais	Convenção de Santa Fé	OAI-PMH v. 1.0/1.1	OAI-PMH v. 2.0
Natureza	Experimental	Experimental	Estável
Verbos	Dienst protocol	OAI-PMH	OAI-PMH
Solicitações	HTTP GET / POST	HTTP GET / POST	HTTP GET / POST
Respostas	XML	XML	XML
Transporte	HTTP	HTTP	HTTP
Metadado	OAMS	Dublin Core	Dublin Core
Temática	Eprints	Objetos relacionados aos documentos	Recursos
Modelo	Colheita de Metadado	Colheita de Metadado	Colheita de Metadado

As mudanças estruturais e a evolução do protocolo demonstraram a busca da *Open Archives Initiative – OAI* em alcançar um nível de interoperabilidade que caminhasse em direção da possibilidade dos repositórios exporem seus metadados de forma global.

3.3.2 Tecnologia do OAI-PMH

O OAI-PMH é um protocolo para coleta de metadados que define um mecanismo para a coleta dos registros que contêm os metadados dos repositórios. O OAI-PMH dá uma opção técnica simples para que os fornecedores dos dados tornem seus metadados

disponíveis aos Provedores de Serviços, baseado no HTTP (protocolo de transferência de *hypertext*) e no XML (*Extensible Markup Language*) (OPEN ARCHIVES FORUM, 2005).

A figura 04 mostra um esquema de funcionamento básico do protocolo OAI-PMH. No exemplo, o Provedor de serviço realiza uma solicitação baseado no protocolo HTTP e obtém os metadados codificados em XML. Em seguida, o Provedor de Serviço disponibiliza ao usuário um serviço de valor agregado.

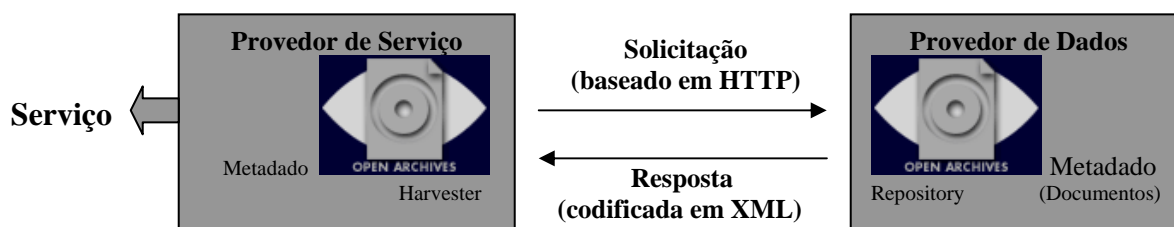


Figura 04 - Esquema de funcionamento básico do protocolo OAI-PMH

O OAI-PMH define um mecanismo para colheita de metadados em formato XML nos repositórios. O protocolo não proporciona um mecanismo para colheita de dados (conteúdo) que não é codificado em XML. O protocolo também não coordena o meio de associação entre o metadado e o conteúdo relacionado. Desde que muitos interessados possam desejar acesso ao conteúdo associado com o metadado colhido, os provedores de dados podem torná-lo apropriado para definir um link no metadado para o conteúdo. O próprio mandatário em formato *Dublin Core* proporciona o elemento identificador que pode ser usado para esse propósito (OPEN ARCHIVES INITIATIVE, 2005).

Conforme Howe (1993) o termo metadado foi apresentado por Jack Myers nos anos de 1960 para descrever conjuntos de dados. A primeira interpretação apresentada foi de “dados sobre os dados”, já que proporcionam a informação mínima necessária para identificar um recurso. No ponto de vista da Open Archives Initiative (2005) Metadado é a informação estruturada sobre recursos (incluindo recursos digitais e não digitais). O Metadado pode ser usado para auxiliar no suporte de uma ampla rede de operações sobre esses recursos. No contexto de serviços baseados em colheita de Metadados via OAI-PMH,

a operação mais comum é à descoberta e recuperação dos recursos. Os metadados podem ser: título; autor; instituição; resumo; etc.

Os metadados de muitas fontes podem ser recolhidos juntos em uma base de dados, e os serviços podem ser fornecidos baseados nesta colheita centralizada, ou “agregar” dados. A ligação entre este metadado e o índice relacionado não é definida pelo protocolo da OAI, pois o OAI-PMH não fornece uma busca através destes dados, ele simplesmente torna possível trazer junto os dados em apenas um lugar, ou seja, fornece os dados colhidos à base na qual foi realizada a consulta (OPEN ARCHIVES FORUM, 2004).

O metadado a ser colhido pode estar em qualquer formato previamente escolhido por uma comunidade (ou pelo conjunto discreto de fornecedores dos dados e de serviços), embora o núcleo de *Dublin Core* seja especificado para fornecer um nível básico de interoperabilidade.

O *Dublin Core* foi criado por uma organização que surgiu em Dublin, Ohio (EUA) em 1995 e que se dedica a promover a adoção de padrões de interoperabilidade em metadados, esta organização chama-se Iniciativa de Metadados *Dublin Core* (*Dublin Core Metadata Initiative* – DCMI)

O *Dublin Core* é um conjunto de 15 elementos: *title* (título); *Creator* (entidade responsável); *Subject* (palavra-chave); *Description* (resumo); *Publisher* (entidade responsável); *Contributor* (entidade responsável por contribuições); *Date* (data); *Type* (natureza ou gênero); *Format* (tipo de mídia); *Identifier* (Identificador); *Source* (referência); *Language* (Linguagem do conteúdo); *Relation* (referência a um recurso relacionado); *Coverage* (extensão ou escopo do conteúdo); e *Rights Management* (direitos autorais) (DUBLIN CORE, 2005).

A OAI realizou um mapeamento entre vários formatos de metadados em Provedores de Serviços com considerável peso, os quais fazem à colheita dos metadados e os utilizam para construção de serviços de alto nível. Embora exista trabalho de pesquisa referente ao crescimento de serviços em interfaces de pesquisa comum, sendo realizados através de formatos de metadados heterogêneos, ultimamente as soluções mais organizáveis requererem repositórios com mapa para um simples e comum formato de metadados. Os quinze elementos do *Dublin Core* nos anos anteriores evoluíram como um fator de padronização para simples metadados através de disciplinas e este fato destinou a escolha

de um conjunto de metadados comuns. A cooperação entre a OAI e a Dublin Core Metadata Initiative guiou de um esquema XML comum para o *Dublin Core* que está disponível no endereço <http://dublincore.org/schemas/xmls/simpledc2002312.xsd>. (OPEN ARCHIVES INITIATIVE, 2005).

No caso do repositório desejar expor metadados em formato diferente do *Dublin Core* o protocolo de colheita de metadados suporta a noção de múltiplos conjuntos de metadados, permitindo as comunidades exporem os metadados em formatos que são específicos para suas aplicações e domínios. A estrutura técnica de campos sem limitações sobre a natureza dos conjuntos paralelos e outros que os registros dos metadados estão estruturados como dados de XML tem um correspondente esquema XML para validação (OPEN ARCHIVES INITIATIVE, 2005).

3.3.3 Considerações sobre Implementação do OAI-PMH

O OAI-PMH tem sido designado com propósito de possuir fácil implantação. Por essa razão, a pergunta geral de configuração de um servidor *Web* para lidar com as solicitações do OAI-PMH e fazer a análise gramatical dos argumentos envolve menos que um dia de trabalho por alguém com experiência em configurar servidores web e escrita *CGS scripts* (OPEN ARCHIVES INITIATIVE, 2005).

Programar o protocolo, de qualquer modo, envolve mais que simples análise gramatical do protocolo de solicitação. Responder o protocolo de solicitação também envolve acessar ou extrair seu metadado. Dá-se um tempo estimado para esta dificuldade desde que a natureza da tarefa é inteiramente peculiar. No caso dos dados estarem bem organizados, já possuir metadados e ter estabelecido mecanismos para extrair ou derivar de metadados a OAI ostenta que esta tarefa não deve ser onerosa (OPEN ARCHIVES INITIATIVE, 2005).

No que trata a relação entre o OAI-PMH com outros protocolos a estrutura técnica da OAI é intencionalmente simples com o intuito de proporcionar pequenas barreiras para os participantes. Os protocolos como o Z39.50 tem uma funcionalidade mais completa, por exemplo, ele parte com uma sessão de gerenciamento e conjuntos de resultados permitindo a especificação de predicados que filtram os registros retornados. De qualquer modo, esta

funcionalidade resulta em aumento na dificuldade de implantação e custo. A estrutura técnica da OAI não é elaborada para substituir outras abordagens, mas proporcionar uma fácil implantação e um fácil desenvolvimento como alternativa para diferentes instituições e diferentes propósitos nos interessados nas soluções de interoperabilidades (OPEN ARCHIVES INITIATIVE, 2005).

3.3.4 Configurações para uso do Protocolo OAI-PMH

O *Open Archives Initiative – Protocol Metadata Harvesting* possibilita uma organização flexível. Por ser um protocolo simples e baseado no HTTP e XML permite rápida organização. As ferramentas são disponíveis pela OAI e os sistemas podem ser organizados em uma variedade de configurações. Os recursos de metadados e os textos completos são tipicamente disponíveis livremente, mas isto não se trata de uma obrigação. O OAI-PMH também pode ser usado entre grupos fechados, para disponibilizar somente partes de metadados e aplicações em divulgação. Nos diagramas abaixo é apresentado algumas características do OAI-PMH.

O diagrama apresenta vários Provedores de Serviço realizando a colheita através do OAI-PMH em vários Provedores de Dados.

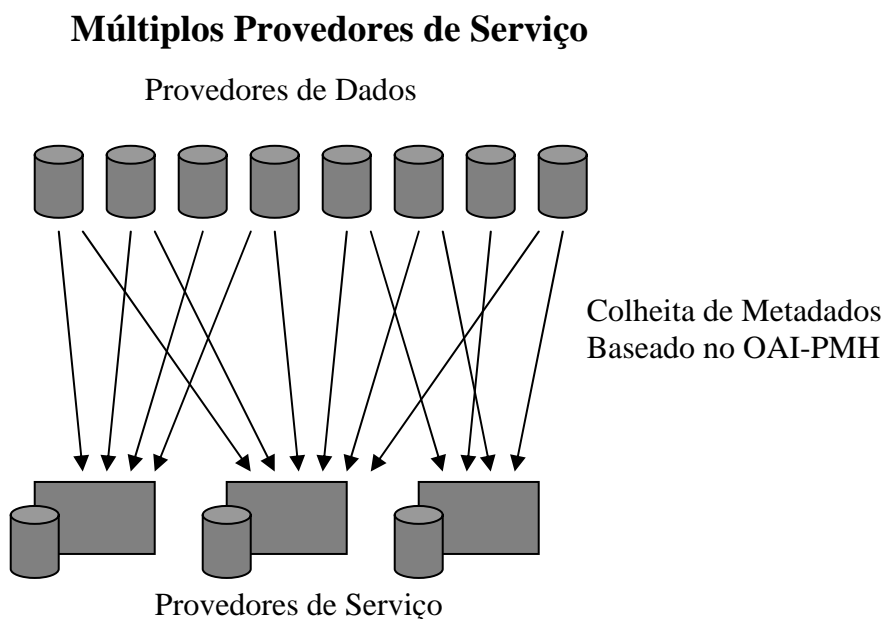


Figura 05 – vários Provedores de Serviço realizando a colheita em vários Provedores de Dados utilizando OAI-PMH (OPEN ARCHIVES FORUM, 2005).

O diagrama trata dos *Aggregators* (agregadores) que podem se acomodar entre os Provedores de Serviço e os Provedores de Dados. “O OAI *Aggregator* é um serviço que junta os registros dos metadados de vários Provedores de Dados e os torna disponíveis para se juntarem com outros usando para isso o OAI-PMH” (OPEN ARCHIVES FORUM, 2005).

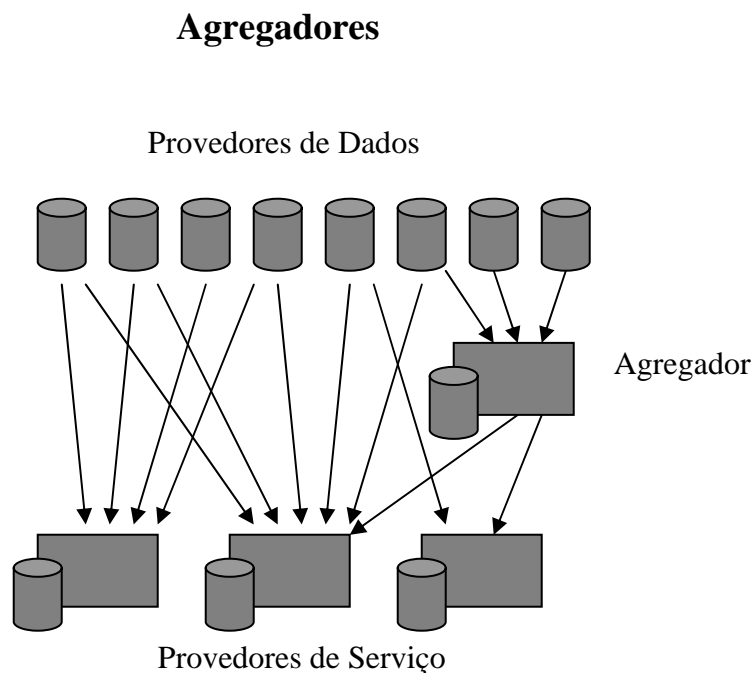


Figura 06 – Agregador acomodado entre os Provedores de Serviço e os Provedores de Dados (OPEN ARCHIVES FORUM, 2005)

O diagrama se refere à operação de colheita de metadados que pode ser complementada com pesquisa baseada em Z39.50. No exemplo, o Provedor de Serviço executa uma pesquisa baseada no Z39.50 e realiza uma colheita de metadados através do OAI-PMH nos Provedores de dados. O complemento possibilita o usuário obter resultados direcionados aos seus interesses de pesquisa, torna uma simples colheita de metadados em uma sessão de gerenciamento e conjuntos de resultados que permite a especificação de predicados que filtram os registros retornados.

Colheita de Metadados Complementada com Pesquisa

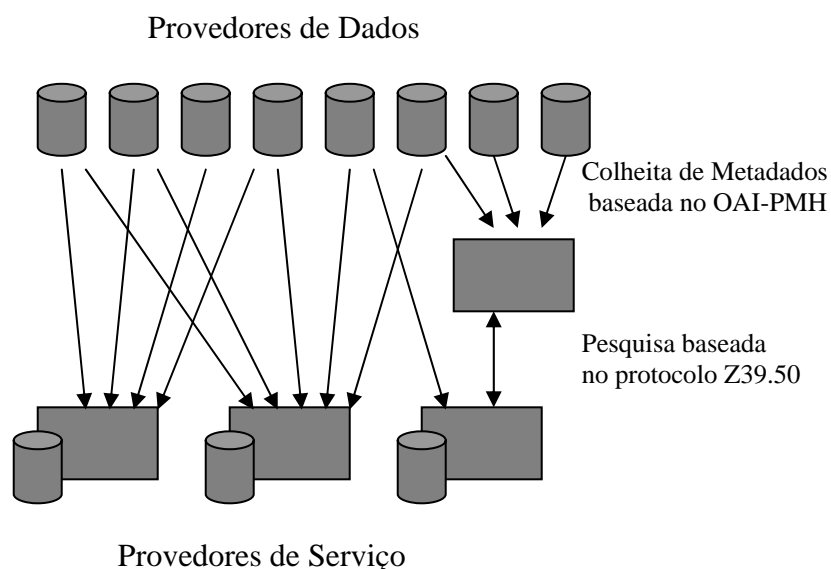


Figura 07 – Operação de colheita de metadados complementada com pesquisa baseada em Z39.50 (OPEN ARCHIVES FORUM, 2005).

As configurações do OAI-PMH apresentam algumas possibilidades e serviços disponíveis aos repositórios e seus usuários. A utilização desses recursos depende dos objetivos e pretensões das comunidades interessadas em criar seu *Open Archives*.

Na seguinte seção é apresentada a metodologia utilizada para realização da pesquisa, sendo descritos os passos e os instrumentos usados.

4 METODOLOGIA

A metodologia utilizada para realização da pesquisa foi composta por um conjunto de etapas e procedimentos envolvendo o método de dedução científica, uma pesquisa descritiva, a definição da população e amostra, a obtenção dos dados, a aplicação dos instrumentos de coleta de dados, o processamento dos dados e a análise exploratória dos dados.

4.1 Delimitação do Problema

Este trabalho foi delimitado quanto ao espaço e tempo do estudo, quanto aos termos e quanto aos seus participantes.

Tendo como problema da pesquisa identificar o panorama dos *Open Archives* na comunidade científica como uma nova alternativa de comunicação, o tempo que abrange o problema desta pesquisa refere-se ao período de 1999 quando é fundada a *Open Archives Initiative* e se estende até o ano de 2005.

Quanto ao espaço que envolve o problema à investigação restringe-se acerca das classes de *Open Archives* conforme o tipo de documento disponibilizado pelo Open Archives. Nesse caso, as classes são Pesquisa Institucional ou Departamental, Pesquisa Inter-Institucional, Teses e Dissertações, Revistas Eletrônicas, Base de Dados e Outros.

Já a delimitação quanto aos participantes indica quem são os elementos envolvidos no estudo. Nesse caso, trata-se dos *Open Archives* que utilizam as tecnologias disponíveis pela *Open Archives Initiative* no cenário nacional e internacional.

Em se tratando dos termos que envolvem o fenômeno, limita-se em tratar de aspectos referentes às principais tecnologias envolvidas, as comunidades, o fluxo de informação e a variação dos *Open Archives*.

4.2 Tipo de pesquisa

Para o desenvolvimento da pesquisa foi utilizada a pesquisa descritiva. Segundo Bervian e Cervo (1983, p.55) A pesquisa descritiva observa, registra, analisa e correlaciona

fatos ou fenômenos (variáveis) sem manipulá-los. Procura descobrir a precisão possível, a frequência com que um fenômeno ocorre, sua relação e conexão com outros, sua natureza e características.

A pesquisa bibliográfica foi realizada para explicar o problema a partir de referências teóricas publicadas em documentos. Fez parte da pesquisa descritiva com a finalidade de conhecer e analisar as contribuições científicas do passado existente sobre o problema estudado. Foi utilizada com o intuito de recolher informações e conhecimentos prévios acerca do problema para o qual se procura respostas (CERVO ; BERVIAN, 1983).

A pesquisa bibliográfica foi realizada baseada na bibliografia existente, tal como artigos, manuais e periódicos científicos (Revista Ciência da Informação, Informação e Sociedade, Gestão do Conhecimento entre outras), livros, teses e dissertações (nas áreas da Comunicação, Ciência da Informação, Computação entre outras), e também nas constantes das páginas oficiais das entidades (*Eprints.org*, *Open Archives Initiative*, *Dublin Core Metadata Initiative* entre outras).

4.3 Método

Para a realização da pesquisa foi utilizado o método de dedução científica, que segundo Bervian e Cervo (1983, p. 40) “a dedução é a argumentação que torna explícitas verdades particulares contidas em verdades universais. O ponto de partida é o antecedente que afirma uma verdade universal, e o ponto de chegada é o conseqüente que afirma uma verdade menos geral ou particular contida implicitamente no primeiro”.

4.4 Definição da população e amostra

A definição da população e a amostra foram realizadas em duas etapas a partir de dados coletados no site oficial da Organização *Eprints*, a qual disponibiliza as informações na página Institutional Archives Registry. A primeira etapa foi realizada na data de 09 de abril de 2005 com abrangência nacional e internacional. A segunda etapa teve abrangência somente nacional e foi realizada na data de 15 de agosto de 2005.

4.4.1 População

A população estudada foi dividida em classes conforme o tipo de *Open Archives*. As classes apresentadas na população são representadas por *Open Archives* de pesquisa institucional ou departamental (*Research Institutional or Departmental*), *Open Archives* de pesquisa inter-institucional (*Research Cross-Institution*), *Open Archives* de teses e dissertações (*e-Theses*), *Open Archives* de revistas eletrônicas e publicações (*e-Journal/Publication*), *Open Archives* de base de dados (*Database*) e outros (*Other*). A divisão das classes segue um padrão de classificação por tipo de *Open Archives* apresentado no site oficial da Eprints.org (EPRINTS.ORG, 2004). A tabela 01 apresenta o número de elementos da população dos *Open Archives*.

Tabela 01 – População dos *Open Archives* conforme as classes por tipo de arquivo

Classes	População
Pesquisa Institucional ou Departamental	194
Pesquisa Inter-Institucional	53
Teses e Dissertações	55
Revistas Eletrônicas e Publicações	39
Base de Dados	09
Outros	37
Total	387

Os *Open Archives* da classe Pesquisa Institucional ou Departamental tratam de repositórios de universidades ou faculdades, departamentos ou disciplinas oferecidas pelas instituições. Caracterizados por centralizar, preservar e possibilitar acesso ao capital intelectual da instituição (JOHNSON, 2002); Pesquisa Inter-Institucional tratam de repositórios que fornecem serviços através das instituições (GERARD ; LYNCH, 2005); Teses e Dissertações tratam de repositórios institucionais que disponibilizam teses e dissertações em formato digital (EPRINTS.ORG, 2005); Revistas Eletrônicas e Publicações são repositórios de revistas que utilizam as tecnologias *Open Archives* para disponibilizarem e divulgar suas coleções (EPRINTS.ORG, 2005); Base de Dados são repositórios utilizados para armazenar, preservar e possibilitar acesso a materiais digitais das instituições ou centros de pesquisa (EPRINTS.ORG, 2005); Outros é uma classe criada

para incluir repositórios que não possuem características semelhantes às demais classes (EPRINTS.ORG, 2005).

4.4.2 Amostra

Para a definição da amostra foi utilizada a técnica de amostragem estratificada que consiste em dividir a população em subgrupos, os quais são denominados de estratos. (BARBETTA, 1998). Nesta pesquisa os estratos foram definidos pelas classes dos *Open Archives* conforme o tipo de arquivo.

Estes estratos são internamente mais homogêneos do que a população toda, com respeito as variáveis em estudo. Sobre os estratos da população foram realizadas seleções aleatórias de forma independente. A amostra estratificada proporcional garantiu que os elementos da amostra fossem adequados às proporções existentes na população.

Para melhor entendimento a tabela 02 apresenta o número de elementos por classe indicando o volume da primeira amostra.

Tabela 02 – amostra dos *Open Archives* indicando o número de elementos.

Classes	População	Amostra
Pesquisa Institucional ou Departamental	194	24
Pesquisa Inter-Institucional	53	08
Teses e Dissertações	55	08
Revistas Eletrônicas e Publicações	39	06
Base de Dados	09	02
Outros	37	06
Total	387	54

Na definição da população e da primeira amostra realizada a partir de dados coletados no site oficial da Organização *Eprints* na data de 09 de abril de 2005, a qual disponibiliza as informações na página *Institutional Archives Registry*, fizeram parte da amostra 14 *Open Archives* brasileiros, dos quais 2 (dois) são da classe Pesquisa Institucional ou Departamental (*Research Institutional or Departmental*), 3 (três) da classe Pesquisa Inter-institucional (*Research Cross-Institution*), 3 (três) da classe Teses e

Dissertações (*E-Teses*) e 6 (seis) da classe Revistas Eletrônicas e Publicações (*E-Journal*). Na tabela 06 são apresentados os *Open archives* brasileiros que fizeram parte da primeira amostra.

Tabela – 03 *Open archives* brasileiros que fizeram parte da primeira amostra.

Tipo de Arquivo (classe)	Número de Elementos	Percentual
Revistas Eletrônicas e Publicações	6	42,65%
Pesquisa Inter-institucional	3	21,52%
Teses e Dissertações	3	21,52%
Pesquisa Institucional ou Departamental	2	14,28%
Total	14	100,00%

A definição da amostra com abrangência somente nacional realizada na data de 15 de agosto de 2005 foi composta por 29 *Open Archives* que são apresentados e analisados no capítulo que trata dos resultados da pesquisa.

4.5 Obtenção dos dados

Para a obtenção dos dados foram utilizados instrumentos de coleta de dados compostos por variáveis qualitativas e quantitativas conforme os objetivos da pesquisa.

Os dados primários que fizeram parte da pesquisa se constituirão em dados que não foram trabalhados por outros pesquisadores e se encontravam nas constantes das páginas dos *Open Archives* e das entidades envolvidas. Foi criado um instrumento de coleta de dados primários – Apêndice A – Formulário de pesquisa de dados primários – o qual se utilizou diretamente na amostra. O instrumento de coleta de dados primários foi usado para coletar dados em cada *Open Archives* conforme sua classificação, seguindo um conjunto de questões segundo os objetivos da pesquisa. A coleta de dados primários foi realizada através do acesso às páginas dos *Open Archives*.

Os dados secundários que fizeram parte da pesquisa foram os dados que já haviam sido trabalhados por outros pesquisadores e se apresentavam em artigos de periódicos científicos, livros, teses e dissertações e também nas constantes das páginas oficiais dos *Open Archives* e das entidades envolvidas. Para a obtenção dos dados foi criado um

instrumento de coleta de dados – Apêndice B – Formulário de pesquisa de dados secundários – para ser utilizado em fontes de dados secundárias, ou seja, obter informações que já haviam sido analisadas e disponibilizadas.

4.6 Aplicação do formulário de pesquisa

Os formulários de pesquisa foram aplicados conforme a definição da amostra. Os formulários tornaram-se instrumentos que ofereceram suporte as atividades de pesquisa, possibilitando a coleta precisa dos dados e a garantia da uniformidade na interpretação dos dados e dos critérios pelos quais foram coletados.

Para a validação dos formulários de pesquisa foram realizados testes antecipados em parte da amostra e algumas modificações necessárias para o ajustamento e melhoria na sua qualidade.

4.7 Processamento dos dados

Depois dos dados serem coletados, todos foram organizados para facilitar a realização da análise. Os dados foram categorizados e tabulados sendo armazenados em uma matriz, onde cada coluna referiu-se a uma variável e cada linha a uma resposta.

4.7.1 Análise exploratória dos dados

Após os dados serem adequadamente resumidos e apresentados em tabelas e gráficos, foram observados determinados aspectos relevantes a respeito da estrutura do fenômeno em estudo, entre os dados coletados foram explorados dados de variáveis qualitativas e quantitativas, visando cercar de forma mais precisa todos os aspectos referentes à manifestação e causa do fenômeno.

Para realização da análise exploratória dos dados foram utilizadas técnicas referentes à descrição e exploração dos dados categorizados, os quais se dividiram em classificação simples e representações gráficas.

A classificação simples possui como primeiro passo a análise isolada de cada variável, seguido da construção de uma distribuição de frequências. A distribuição de frequências compreende a organização dos dados de acordo com as ocorrências dos diferentes resultados observados em cada categoria (BARBETTA, 1998, p. 66).

No caso da presente pesquisa foi utilizada as representações gráficas e as tabelas, pois em geral, fornecem uma visualização mais sugestiva. Elas constituíram-se numa forma alternativa de apresentação de distribuições de frequências.

O próximo capítulo trata da apresentação dos resultados da pesquisa. Apresenta informações e análises a cerca dos *Open Archives* pertencentes a amostra a nível nacional e internacional.

5 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS DA PESQUISA

Neste capítulo são apresentados os resultados da pesquisa e análise dos dados coletados. As interpretações dos dados são embasadas nos conceitos e definições apresentados na pesquisa bibliográfica.

A pesquisa é composta por duas etapas as quais se dividem em âmbito nacional e internacional. A primeira etapa foi realizada em âmbito internacional, utilizando uma técnica de amostragem estratificada, a qual teve participações de *Open Archives* do Brasil. A segunda etapa utilizou apenas *Open Archives* do Brasil, totalizando uma amostragem de 29 elementos.

A tabela 04 apresenta a taxa de resposta das variáveis com seus respectivos percentuais de respostas.

Tabela 04 – Distribuição de Freqüências das variáveis conforme o número de elementos que apresentaram resposta.

Variável	Freqüência do n°. de elementos	Percentual
Tecnologia adotada para implantação	53	98,14%
Tempo de existência (Celestial)	53	98,14%
Termos que caracterizam objetivos	44	72,13%
Número total da amostra	54	100,00%

Foram utilizados 54 (cinquenta e quatro) formulários nos *Open Archives* na primeira amostra. Entre os *Open Archives* pesquisados alguns não apresentavam os dados necessários para preencherem todas as variáveis do formulário.

5.1 Características que apresentam a variação dos *Open Archives*

Nesta seção é apresentado elementos que caracterizam a variação dos *Open Archives* no cenário internacional no período de 06/2005 a 08/2005. Este período foi

escolhido para possibilitar maior compreensão no processo de análise dos dados e garantir maior representação nos resultados da pesquisa. Os dados foram coletados no décimo quinto dia de cada mês.

A tabela 05 apresenta a distribuição dos *Open Archives* conforme a sua variação entre o período de 06/2005 a 08/2005.

Tabela 05 - Distribuição dos *Open Archives* conforme a sua variação entre o período de 06/2005 a 08/2005.

Tipos de AO	Número de OA		
	06/2005	08/2005	Variação Total
Pesquisa Institucional ou Departamental	216	230	14
Pesquisa Inter-Institucional	55	60	05
Teses e Dissertações	54	55	01
Revistas Eletrônicas e Publicações	41	44	03
Base de Dados	10	11	01
Outros	50	59	09
Total	426	459	33

Em análise aos dados apresentados na tabela 05 pode-se constatar que entre as seis classes de *Open Archives* pesquisados, todos tiveram aumento no número de elementos. A diferença total alcançou o número de 33 elementos a mais, computando um percentual de 7,18%, entre os meses de junho, julho e agosto de 2005.

A variabilidade apresentada nos elementos pesquisados entre os meses de junho, julho e agosto de 2005 apontaram aumento na criação de *Open Archives*, mostrando que existe crescimento no grau de utilização do modelo de comunicação por parte da comunidade científica.

5.2 Características dos elementos referentes à região de origem dos *Open Archives*

Na presente seção é apresentado as características dos elementos referentes à região de origem dos *Open Archives*. A tabela 06 apresenta a distribuição de frequência de cada classe conforme a região de origem.

Tabela 06 - Distribuição dos *Open Archives* conforme à região de origem e tipo de *Open Archives*

Tipo de OA x Localidade	Base de dados	Revistas eletrônicas	Teses e dissertações	Outros	PI-I	PIouD	Frequência	Percentual
Europa	02		03	02	03	14	24	44,44%
América do Sul		06	03	01	03	02	15	27,77%
América do Norte			02	03	02	04	11	20,37%
Oceania						03	03	5,55%
Ásia						01	01	1,65%
Frequência	02	06	08	06	08	24	54	100,00%

No gráfico da figura 08 é apresentado o percentual de *Open Archives* conforme a região de origem e tipo de *Open Archives*.

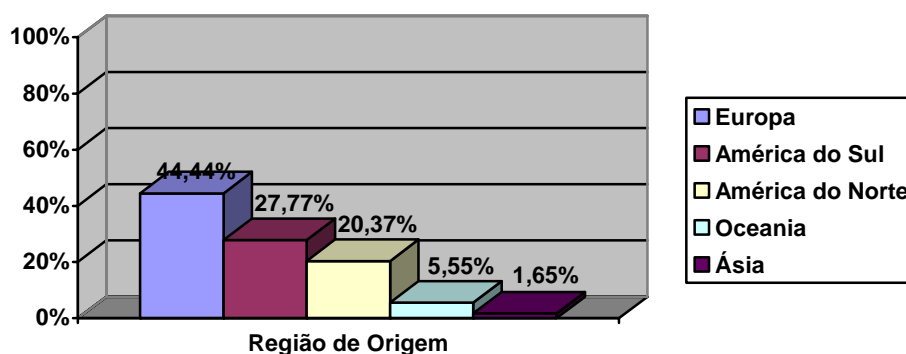


Figura 08 - Distribuição dos *Open Archives* conforme à região de origem e tipo de *Open Archives*

Em análise ao gráfico da figura 08 os dados apresentados demonstram que há maior presença de *Open Archives* no continente europeu, com percentual de 44,44%, mostrando grande aplicabilidade no uso da tecnologia *Open Archives*. Enquanto existe uma distribuição semelhante entre os elementos da amostra que estão localizados na América do Sul e América do Norte, possuindo ambas expressiva presença no cenário mundial. No caso do continente asiático e Oceania, ambos apresentaram índice bastante baixo, sendo representados apenas pelos países do Japão e Austrália.

5.3 Características das comunidades que adotaram as tecnologias da OAI

Nesta seção são apresentados elementos que caracterizam as comunidades que adotaram as tecnologias da *Open Archives Initiative*. A tabela 07 apresenta a distribuição dos *Open Archives* conforme sua área de atuação.

Tabela 07 – Distribuição dos *Open Archives* conforme a área de atuação

Área de atuação	Frequência	Porcentagem
Ensino Superior (vários cursos)	27	50,00%
Biblioteconomia e Ciência da Informação	5	9,25%
Ciências Humanas	3	5,55%
Matemática e Estatística	3	5,55%
Informática e Computação	3	5,55%
Governamental	2	3,78%
Pesquisa Aeroespacial (áreas afins)	2	3,78%
Ciência e Tecnologia	1	1,65%
Agronomia	1	1,65%
Geociência	1	1,65%
Física	1	1,65%
Engenharia eólica	1	1,65%
Ciências Empresariais	1	1,65%
Educação	1	1,65%
Genética	1	1,65%
Psicologia	1	1,65%
Total de resposta	54	100%

Nesta tabela pode-se constatar que a maior parte da área de atuação dos *Open Archives* estão direcionados para o Ensino Superior, alcançando um percentual de 50%, o qual abrange instituições que disponibilizam esta alternativa de comunicação para os cursos oferecidos no seu campo curricular. Os *Open Archives* que atuam em outras áreas apresenta percentual baixo, tendo destaque para Biblioteconomia e Ciência da Informação, o qual apresentou percentual de 9,25%.

Estes dados corroboram os interesses da OAI quando afirma que “possui na sua raiz um esforço para intensificar o acesso aos *Open Archives* como um meio de aumentar a disponibilidade da comunicação acadêmica” (OPEN ARCHIVES INITIATIVE, 2005).

Os resultados apresentados apontam a utilização dos *Open Archives* como alternativa de comunicação científica e atingem os interesses das Instituições de Ensino Superior, possibilitando oferecer a sua comunidade uma alternativa viável para comunicação entre os pares e divulgação de seus trabalhos.

A tabela 08 mostra dados sobre a descrição do idioma encontrado nos *Open Archives* que fizeram parte da primeira amostra.

Tabela 08 – Distribuição de freqüências dos *Open Archives* conforme a descrição do idioma da primeira amostra

Idioma	Freqüência	Percentagem
Inglês	42	59,74%
Português	17	22,07%
Francês	5	6,49%
Alemão	4	5,19%
Espanhol	3	3,89%
Italiano	2	2,59%
Grego	2	2,59%
Esloveno	1	1,29%
Finlandês	1	1,29%
Total de respostas	77	100%

Os dados apresentados na tabela 08 mostram que o idioma inglês é o mais utilizado pelos *Open Archives* para comunicação e divulgação, apresentando percentual de 59,74%, sendo usado pelos membros das comunidades científicas como recurso lingüístico para o

acesso aos seus trabalhos. O idioma Português obteve incidência bastante expressiva, com percentual de 22,07%. O idioma espanhol que é bastante utilizado para comunicação no cenário internacional, nesta amostra apresentou presença pouco expressiva, com percentual de 3,89%.

A tabela 09 apresenta a distribuição dos *Open Archives* conforme o idioma utilizado para comunicação entre os pares.

Tabela 09 – Distribuição dos *Open Archives* conforme o idioma utilizado para comunicação entre os pares.

Idioma	Frequência	Porcentagem
Nativo	47	87,03%
Não nativo	7	12,96%
Total de respostas	54	100%

Na tabela 09 pode-se constatar que a maior parte dos *Open Archives* utilizam o idioma nativo para se comunicar entre seus pares e divulgar seus trabalhos. Com uma representação percentual de 87,03% contra apenas 12,96% da categoria idioma não nativo.

Estas ocorrências mostram que as comunidades utilizam o idioma nativo para alcançar um leque de abrangência direcionado aos membros que usam o idioma de sua nacionalidade, ou seja, o idioma do país de origem da comunidade criadora do *Open Archives*. A disseminação da informação possuiu um alcance voltado às comunidades locais, isto é bastante aparente na variável área de atuação quando apresenta grande ocorrência no item Ensino Superior, o qual abrange instituições que disponibilizam esta alternativa de comunicação para os cursos oferecidos no seu campo curricular.

A tabela 10 trata da distribuição dos *Open Archives* conforme a tecnologia adotada para implantação no cenário internacional. Os dados apresentados na tabela 10 foram coletados de fonte secundária a qual apresentava a categoria outros softwares (vários) para agrupar os *softwares* com baixo índice presença.

Tabela 10 – Distribuição dos *Open Archives* conforme a tecnologia adotada para implantação.

Tecnologia adotada para implantação	Frequência	Porcentagem
GNU Eprints	23	43,39%
Outros softwares (various)	14	26,41%
Open Journal System - OJS	7	13,20%
Dspace	6	11,32%
OPUS (Open Publications System)	1	1,88%
CDSWare	1	1,88%
Fedora	1	1,88%
Total de respostas	53	100,00%

Em análise a tabela 10 pode-se constatar que a tecnologia predominante para implantação dos *Open Archives* é o *GNU Eprints*, apresentando um percentual de 43,39%. Os dados apresentados contabilizam as duas versões do *software*, tecnologia disponibilizada pela organização *Eprint.org*. A possível explicação deste *software* possuir maior aceitabilidade por parte das comunidades interessadas é o fato de ter sido lançado com maior antecedência que os demais, apresentando estabilidade na evolução e desenvolvimento da tecnologia oferecida. Destaca-se também o *software Open Journal Systems* que apresentou índice bastante expressivo na pesquisa seguido do *software Dspace*, os quais juntos totalizaram o percentual de 24,52% de toda amostra.

5.4 Características dos fatores declarados como motivadores para o surgimento dos *Open Archives*.

Nesta seção são apresentados os fatores que foram declarados pelas comunidades nas constantes das páginas dos *Open Archives* como motivadores para o seu surgimento. Na figura 09 é apresentado a distribuição dos *Open Archives* conforme os termos que caracterizam seus objetivos.

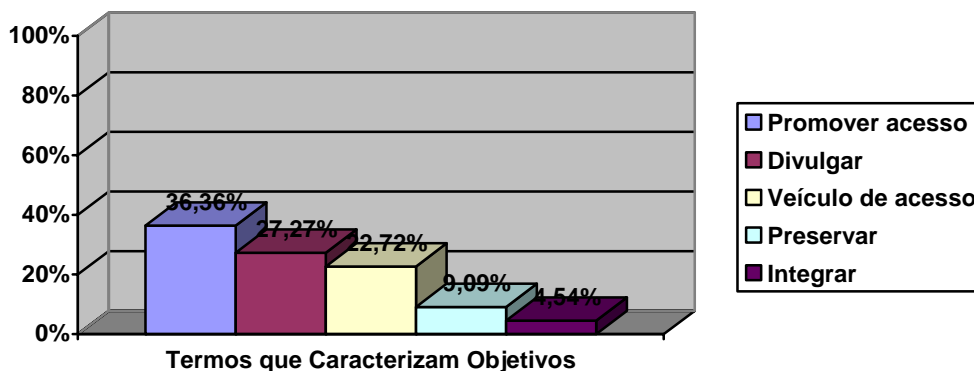


Figura 09 – Distribuição dos *Open Archives* conforme os termos que caracterizam seus objetivos.

Entre os 54 *Open Archives* pesquisados obteve-se resposta de 81,48%, ou seja, foram coletados dados somente de *Open Archives* que disponibilizavam informações com termos que caracterizavam seus objetivos.

Para chegar aos termos citados foi realizada uma aproximação dos termos equivalentes, com o intuito de eliminar respostas iguais entre termos semelhantes.

No gráfico da figura 09 se pode observar maior presença do termo PROMOVER ACESSO, com percentual de 33,36%, seguido do termo DIVULGAR, com percentual de 27,27%, o termo VEÍCULO DE ACESSO apresentou percentual de 22,72%. No caso dos termos, PRESERVAR e INTEGRAR apresentaram juntos 13,63% do total dos dados coletados na amostra.

Em relação aos termos encontrados nas páginas oficiais dos *Open Archives* usados para expressar seus objetivos, as ocorrências apontam para o termo PROMOVER ACESSO como a criação dos *Open Archives* em função de possibilitar aos seus membros e outros pesquisadores interessados o acesso a seus trabalhos. No caso do termo DIVULGAR está relacionado com a criação dos *Open Archives* como um meio utilizado para apresentar e informar a existência de trabalhos realizados pelos membros das comunidades envolvidas com os *Open Archives*. Em relação ao termo VEÍCULO DE ACESSO refere-se à criação dos *Open Archives* como a construção de um espaço virtual que possa servir de acesso aos trabalhos realizados pelos membros das comunidades. O termo PRESERVAR apresenta-se como meio de garantir a guarda e manutenção dos trabalhos e o termo INTEGRAR

descreve a criação dos *Open Archives* no sentido de aproximar os membros das comunidades.

5.5 Características dos elementos relacionados com o fluxo de informação dos *Open Archives*

A presente seção apresenta as características dos elementos relacionados com o fluxo de informação dos *Open Archives*, tratando da disponibilidade de acesso livre ao texto completo e das possibilidades de interação. A tabela 11 apresenta a distribuição dos *Open Archives* conforme o percentual de disponibilidade de acesso ao documento.

Tabela 11 – Distribuição dos *Open Archives* conforme o percentual de disponibilidade de acesso ao texto completo.

Percentual de acesso ao texto completo	Frequência	Percentual
0% a 30%	02	3,70%
31% a 50%	01	1,85%
51% a 70%	05	9,25%
71% a 100%	46	85,18%
Total de respostas	54	100,00%

Na tabela 11 observa-se maior número de *Open Archives* que disponibilizam acesso ao documento entre 71% a 100%, apresentando percentual de 86,88%, contra apenas 3,27% dos *Open Archives* que disponibilizam de 0% a 30%. Em relação à disponibilidade de acesso livre ao texto completo os resultados mostram que a maior parte dos *Open Archives* visam disponibilizar seus trabalhos para alcançar o maior número de interessados possíveis e permitir o acesso a novas idéias e descobertas inéditas. A ocorrência muito baixa dos *Open Archives* que disponibilizam apenas de 0% a 30% reforçam a idéia de que os objetivos dos *Open Archives Initiative* estão sendo alcançados, a qual incentiva o acesso aberto aos documentos digitais.

A tabela 12 apresenta a distribuição dos *Open Archives* conforme as possibilidades de interação dos leitores visitantes com os autores que disponibilizam seus trabalhos.

Tabela 12 – Distribuição dos *Open Archives* conforme as possibilidades de interação dos leitores visitantes.

Possibilidades de interação (Leitor visitante)	Frequência	Percentual
Apenas acesso	50	92,59%
Comentários	04	7,4%
Total de respostas	54	100,00%

Na tabela 12 constata-se que a categoria leitor visitante apresenta percentual de 92,59% no item “apenas acesso” e 7,4% no item comentários. As ocorrências demonstram que a possibilidade de interação dos *Open Archives* na categoria leitor visitante é muito pequena, já que a maior parte oferece apenas acesso aos trabalhos e somente 7,4% possibilitam realizar comentários. Estes resultados fogem de uma das características principais da OAI, a interatividade, a qual diz respeito à interação entre a comunidade acadêmica, possibilitando a troca de informações, críticas, comentários e sugestões do texto disponível neste tipo sistema.

5.6 Características dos elementos dos *Open Archives* conforme o tempo de existência

A tabela 13 apresenta a distribuição dos *Open Archives* conforme seu tempo de existência. Os dados apresentados são dos *Open Archives* registrados no *Institutional Archives Registry* e seus metadados são sucessivamente coletados pelo *software Celestial*.

Tabela 13 - Distribuição dos *Open Archives* conforme o tempo de existência

Tempo de existência (Celestial)	Frequência	Percentual
Até 1 ano	24	45,28%
Até 2 anos	12	22,64%
Até 3 anos	11	20,75%
Até 4 anos	3	5,66%
Até 5 anos	3	5,66%
Total de respostas	53	100,00%

Os resultados demonstram que a maior concentração de *Open Archives* possuem entre 1 a 3 anos de existência, totalizando percentual de 88,67% do total de elementos da pesquisa. Estes dados mostram que a utilização dos *Open Archives* como meio de comunicação obteve bastante crescimento nos últimos três anos.

5.7 Características dos *Open Archives* existentes no Brasil

Para realização da análise das características dos *Open Archives* no Brasil foi executada uma coleta de dados no dia 09 de agosto de 2005, no site oficial da Organização *Eprints*, a qual disponibiliza as informações na página *Institutional Archives Registry*. O total da população apresentou 29 elementos, os quais todos participaram da pesquisa.

A tabela 14 apresenta a distribuição dos *Open Archives* conforme as categorias que predominam no Brasil.

Tabela 14 - Distribuição dos *Open Archives* conforme as categorias que predominam no Brasil.

Categorias que predominam no Brasil	Frequência	Percentual
Outros	9	31,03%
Revistas Eletrônicas e Publicações	8	27,58%
Pesquisa Institucional ou Departamental	5	17,24%
Pesquisa Inter-Institucional	3	10,34%
Teses e Dissertações	3	10,34%
Base de Dados	1	3,44%
Total	29	100,00%

Com base nos dados apresentados constata-se que entre os *Open Archives* existentes no Brasil predomina a categoria Outros, a qual trata dos *Open Archives* que não se enquadram nas demais categorias, demonstrando grande diversidade, com percentual de 30,03%. Entre as demais, a categoria Revistas Eletrônicas e Publicações apresenta maior incidência, mostrando considerável contribuição do uso dos *Open Archives* para o desenvolvimento e criação de novas revistas eletrônicas, com percentual de 27,58% de toda amostra. Em seguida destaca-se a categoria Pesquisa Institucional ou Departamental apresentando percentual de 17,24%.

O gráfico da figura 10 trata das características dos *Open Archives* existentes no Brasil, apresentando a distribuição conforme os *softwares* utilizados para implantação. Os dados apresentados no gráfico foram coletados de fonte secundária a qual apresentava a categoria outros *softwares* (vários) para agrupar os *softwares* que apresentam baixo índice de presença.

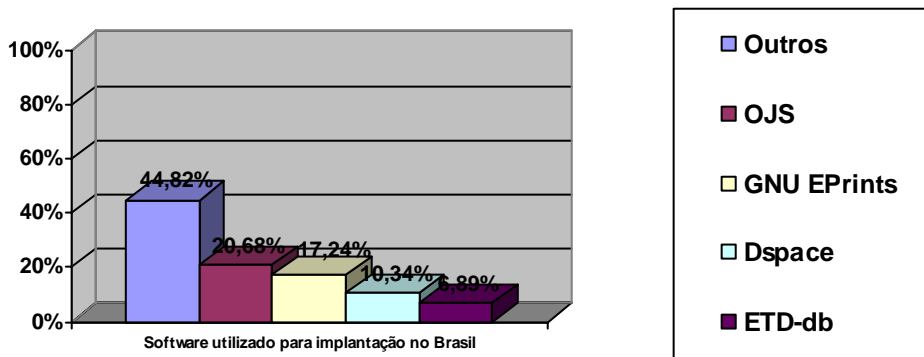


Figura 10 - Distribuição dos *Open Archives* conforme os *softwares* utilizados para implantação no Brasil

Em análise aos dados apresentados no gráfico constatou-se que entre os softwares utilizados para implantação dos *Open Archives* no Brasil existe grande variedade, pois o item *Outros softwares* (vários) apresentou maior índice de frequência, com percentual de 44,82%. Seguido do *software Open Journal Systems*, com percentual de 20,68%, o qual totalizou 75% dos *Open Archives* da categoria Revistas e Eletrônicas existentes no Brasil. O *software GNU EPrints* também apresentou um percentual bastante expressivo totalizando 17,24% de toda amostra.

Os dados coletados sobre os *softwares* para análise refletem somente o número de OAI-PMH *records* (registros do protocolo) que um *Open Archives* possui e apenas trata dos *Open Archives* que são registrados e sucessivamente coletados pelo *software Celestial*. “O *software Celestial* é um programa que realiza a colheita de metadados de repositórios OAI-compliant e os expõem novamente para outros serviços” (E-PRNTS.ORG, 2004).

No seguinte capítulo são apresentadas as conclusões obtidas a partir da análise e interpretação dos resultados da pesquisa e da literatura especializada sobre o tema tratado. Sugestões são apontadas para possíveis investigações e auxiliar outros pesquisadores executarem novos trabalhos de pesquisa referentes ao modelo de comunicação *Open Archives*.

6 CONCLUSÕES

Neste capítulo são apresentadas as conclusões obtidas a partir da análise e interpretação dos resultados e algumas sugestões para realização de novas pesquisas.

6.1 Conclusões

A presente pesquisa teve como principal objetivo identificar o panorama sobre os *Open Archives* na comunidade científica como uma nova alternativa de comunicação científica apoiada nas tecnologias oferecidas pela *Open Archives Initiative* (OAI).

Para isso, realizou-se um levantamento sobre informações que trataram do histórico, da motivação e das instituições que oferecem suporte a OAI. Utilizou-se uma classificação dos *Open Archives* segundo o tipo de repositório conforme seu conteúdo. Verificou-se a variação do número de participantes e o tempo de existência. Identificou-se as comunidades que adotaram as tecnologias da OAI, as principais tecnologias envolvidas e os objetivos dos *Open Archives* participantes de cada classe.

Descreveu-se o fluxo de informação dos *Open Archives* coletando e analisando dados sobre a disponibilidade de acesso ao texto completo e a possibilidade de interação dos usuários com os autores.

Por intermédio de uma pesquisa bibliográfica e dados coletados diretamente na amostra evidenciou-se alguns fatores que levaram ao surgimento dos *Open Archives*.

Em relação às instituições que oferecem suporte à *Open Archives Initiative* percebeu-se que todas estão ligadas ou representando outras instituições de diversos setores das ciências, porém com um objetivo em comum, aumentar o acesso e o desenvolvimento da produção científico/acadêmica.

Com relação à variação do número de participantes da *Open Archives Initiative* destaca-se o aumento na criação de *Open Archives*, demonstrando aceitabilidade desta nova alternativa de comunicação científico/acadêmico pela comunidade científica no cenário nacional e internacional, dando destaque no Brasil para a classe Revistas Eletrônicas e Publicações.

Com base nos resultados apresentados, conclui-se que a utilização dos *Open Archives* como nova alternativa de comunicação científica apresenta pontos bastantes positivos no cenário nacional e internacional. Entre eles possibilita de maneira prática e eficiente a criação de espaços virtuais para o compartilhamento do conhecimento científico entre os pesquisadores; oferece oportunidades para instituições com baixos recursos em investimentos criarem seus *Open Archives* para divulgarem e disponibilizarem seus trabalhos; inseriu em âmbito mundial uma nova solução tecnológica criando sistemas de informação interoperáveis de fácil aplicabilidade em um universo disperso em dados e informações como a Internet; criou uma nova filosofia tecnológica em relação à comunicação e produção do conhecimento científico; possibilitou a disseminação da informação em âmbito mundial, aproximando pesquisadores de várias áreas da ciência dispersos geograficamente; o protocolo OAI – PMH está se tornando parte da infraestrutura da *Web* de forma tão importante como o protocolo HTTP é atualmente; e os *Open Archives* que fizeram parte da amostra atendem comunidades específicas e com alcance direcionado a comunidades locais.

Entre os fatores declarados como motivadores para o surgimento dos *Open Archives*, no que trata dos termos que caracterizam seus objetivos, o termo integrar, o qual está relacionado com o processo de promover integração entre os membros da comunidade, apresentou índice de 4,54%, tendo uma representação pouco expressiva. Conclui-se que este fato se caracteriza pela diversidade dos *Open Archives*, apresentando tipos diferentes de arquivos, os quais são utilizados para divulgar e promover acesso à produção científico/acadêmica. Um grande exemplo são os *Open Archives* de Teses e Dissertações que disponibilizam pesquisas finalizadas (*e-prints*), impossibilitando a mudança e atualização do documento. Diferente dos *Open Archives* que disponibilizam documentos *pre-prints*, os quais possibilitam aos pares realizar contribuições para o desenvolvimento da pesquisa.

A apropriação do uso das tecnologias dos *Open Archives* por diversas comunidades com objetivos diferentes demonstrou que sua aplicabilidade está voltada para divulgação e acesso à informação científica. A interatividade entre autor e leitor apresentou índice de representação bastante baixo. Conclui-se que o fato está relacionado com os tipos de arquivos existentes no universo *Open Archives* tanto no cenário nacional como

internacional, já que poucos apresentaram recursos que proporcionassem contato direto com os autores.

Em relação à disponibilidade de acesso aos trabalhos desenvolvidos pelas comunidades que utilizam as tecnologias dos *Open Archives*, conclui-se que os objetivos da *Open Archives Initiative – OAI* estão sendo alcançados no que trata do incentivo ao acesso aberto aos documentos digitais.

Quanto ao Brasil a utilização da tecnologia dos *Open Archives* está alcançando destaque na criação de novas revistas eletrônicas e este fato está aumentando de forma considerável a visibilidade e o acesso à produção científico/acadêmica. Revistas tradicionais sem a possibilidade de investir recursos em tecnologia para disponibilizar suas produções em suporte digital estão se beneficiando com as soluções oferecidas pela *Open Archives Initiative - OAI*. Os altos gastos com impressão gráfica e distribuição das revistas impressas se apresentam como fator considerável para criação de revistas eletrônicas, além da facilidade de acesso e a maior visibilidade proporcionada por estarem disponíveis em um ambiente sem barreiras de tempo e espaço como a Internet e sendo visualizadas por diversos provedores nacionais e internacionais.

Por fim, as iniciativas de incentivo desenvolvidas pelo Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT estão auxiliando as instituições a adotarem o novo modelo *Open Archives*, divulgando, disponibilizando as tecnologias e oferecendo treinamento as instituições interessadas.

6.2 Sugestões

Em função de particularidades e limitações, esta dissertação deixa algumas questões em aberto. Consequentemente, este estudo pode ser continuado em diversas frentes, de forma preencher e enriquecer os espaços existentes a respeito dos *Open Archives* na comunidade científica como nova alternativa de comunicação.

Alguns aspectos deste trabalho merecem maior investigação e consequentemente são apresentados como sugestões para futuros estudos:

Utilizar os resultados obtidos para realizar uma investigação direcionada a uma classe de *Open Archives*, já que cada classe apresentou particularidades diferentes.

Executar uma pesquisa em um determinado *Open Archives* com objetivo de realizar um estudo sobre os membros participantes da comunidade que criou, gerencia e utiliza seus recursos.

Realizar um estudo sobre as políticas de incentivo ao uso das tecnologias *Open Archives* tanto no cenário nacional como internacional.

Evidenciar a razão de alguns *Open Archives* estarem no Portal da Capes e outros não utilizarem o portal para disponibilizarem seus conteúdos.

Por ultimo, estudar a produção de *Open Archives* de Teses e Dissertações realizando parâmetros com a produção das revistas eletrônicas.

7 GLOSSÁRIO

Aggregator – o OAI *Aggregator* (agregador) é um serviço que junta os registros dos metadados de vários Provedores de Dados e os torna disponíveis para se juntarem com outros usando para isso o OAI-PMH (OPEN ARCHIVES FORUM, 2005).

Archive – o termo “*archive*” no nome *Open Archives Initiative* reflete as origens da própria Iniciativa. Na comunidade *E-prints* o termo *archive* é aceito como um sinônimo para repositório de publicações acadêmicas. A OAI utiliza o termo *archive* em um amplo consenso como um repositório para armazenamento de informação (OPEN ARCHIVES INITIATIVE, 2005).

Containers - Os *Containers* são espaços nos OAI-PMH *responses* onde o XML ao concordar com algum esquema externo pode ser substituído. Os *Containers* são proporcionados para possibilitar maior extensibilidade e na otimização de comunidades específicas. O guia on-line de implementação da OAI lista os *Containers* opcionais existentes e promove *links* para os esquemas (OPEN ARCHIVES FORUM, 2005).

Data Provider – um Provedor de Dados mantém um ou mais repositórios (servidor *Web*) que estejam em conformidade com o OAI-PMH como um meio de expor seus metadados (OPEN ARCHIVES INITIATIVE, 2005).

Dublin Core – é um formato de metadado definido a partir de uma base de consenso internacional. O conjunto de elementos de metadados da *Dublin Core* define quinze elementos para simples descrição e descoberta de recursos, todos dos quais são recomendados e nove são elementos chaves (mandatários). O *Dublin Core* tem se estendido com a promoção de elementos opcionais, qualificadores de elementos e vocabulários de termos (OPEN ARCHIVES FORUM, 2005).

Eprints – são textos digitais de artigos de pesquisa que passam pelo processo de *peer-review* (revisão por pares). Antes da arbitragem e publicação, o documento é chamado de

“*pre-print*”. Depois da arbitragem e finalmente aceito é chamado de “*postprint*”. O termo “*eprints*” também é usado de forma geral (EPRINTS.ORG, 2005).

Eprint Archive – um *Eprint Archive* é uma coleção de documentos digitais. Um OAI-compliant *Eprint Archives* possui em comum o mesmo metadado, fazendo seus conteúdos interoperáveis com outro. Seu metadado pode ser coletado globalmente dentro de um arquivo virtual navegável por algum usuário (exatamente como um índice comercial ou uma base de dados de resumos é navegável, mas com acesso ao texto completo) (EPRINTS.ORG, 2005).

Harvesting – no contexto da OAI, “*harvesting*” refere –se especialmente a colheita reunida de metadados de um número de repositórios distribuídos em um único banco de dados (OPEN ARCHIVES FORUM, 2005)

Interoperabilidade – é a habilidade de sistemas, serviços e organizações trabalharem juntos em direções comuns ou com objetivos diferentes. Na área técnica é suportado por padrões abertos para comunicação entre sistemas e para descrição de recursos e coleções, entre outros (OPEN ARCHIVES FORUM, 2005).

Metadata – informação estruturada sobre recursos (incluindo recursos digitais e não digitais). O Metadado pode ser usado para auxiliar no suporte de uma ampla rede de operações sobre esses recursos. No contexto de serviços baseados em colheita de Metadados via OAI-PMH, a operação mais comum é à descoberta e recuperação dos recursos. Os metadados podem ser: título; autor; instituição; resumo; etc (OPEN ARCHIVES INITIATIVE, 2005).

OAI-compliant – é uma forma de utilizar os *tags* dos metadados da *Open Archives Initiative - OAI*. Um documento pode ser *OAI-compliant* e um *Eprint Archive* pode ser *OAI-compliant*. Todos os documentos *OAI-compliant* em arquivos *OAI-compliant* são interoperáveis. Esta forma de distribuição de documentos pode ser considerada como se eles estivessem todos em um único lugar e em um único formato (EPRINTS.ORG, 2005).

Open – o termo “*open*” está relacionado com a perspectiva de arquitetura dos repositórios, não significa acesso livre ou ilimitado para repositórios de dados que esteja em conformidade com o OAI-PMH (OPEN ARCHIVES INITIATIVE, 2005).

Open Archives – é uma interface de computador para acessar uma coleção de dados, que esteja em conformidade com as especificações ditadas pelo Protocolo para Colheita de Metadados da Iniciativa *Open Archives* (GARCIA, 2003).

Open Archive Initiative (OAI) – a OAI é uma iniciativa que desenvolve e promove padrões de interoperabilidade objetivando facilitar a disseminação eficiente de documentos digitais (OPEN ARCHIVES INITIATIVE, 2005).

Protocol – um “protocolo” é um campo de funções definidas para comunicação entre sistemas. *File Transfer Protocol - FTP* e *Hypertext Transport Protocol - HTTP* são exemplos de outros protocolos usados para comunicação entre sistemas através da Internet (OPEN ARCHIVES FORUM, 2005).

Resource – um *Resource* (recurso) é alguma coisa que possui identidade. Exemplos semelhantes incluem documento eletrônico, uma imagem, um serviço e uma coleção de outros recursos. Nem todos os recursos são recuperáveis na rede. As pessoas, corporações e documentos em uma biblioteca podem ser considerados recursos (OPEN ARCHIVES FORUM, 2005).

Self-archive – o “*Self-archive*” ou Auto-arquivamento é o depósito de um documento digital em um *website* de acessibilidade pública, preferencialmente um *Eprint Archive* submetido a OAI (*OAI-compliant*). Depositando em uma simples interface web onde o depositante inclui cópia/conteúdo em forma de “metadado” (data, autor-nome, título, jornal-nome, etc.) e então anexando o documento em texto completo (EPRINTS.ORG, 2005).

Service Provider – um Provedor de serviços realiza uma solicitação através do OAI-PMH para o Provedor de Dados e usa o metadado como uma base para construir serviços de valor agregado (*Value-added service*) (OPEN ARCHIVES FORUM, 2005).

Value-added service – um serviço que é baseado sobre a colheita de metadado e agrega valor para usuários por meio dos quais podem incluir normalização e enriquecer a colheita de metadado. Os tipos de serviços que podem ser oferecidos incluem serviços de pesquisa, links de citações, *overlay journals* (cobertura de periódicos), serviços de revisão por pares, entre outros (OPEN ARCHIVES FORUM, 2005).

Verbos - instrução em forma de código ou caractere anexada no início da mensagem (OPEN ARCHIVES FORUM, 2005).

XML – é a abreviação para *Extensible Markup Language*. O XML é uma linguagem para criação de outras linguagens. Ele define um meio de descrição de dados e pode ser validado diante de um *Document Type Definition - DTD* ou esquema de configuração dos elementos da linguagem criada (OPEN ARCHIVES FORUM, 2005).

REFERÊNCIAS:

ARAGÃO, Esmeralda Maria de; OLIVEIRA, Margarida Pinto. Padrões de comunicação científica na Universidade Federal da Bahia. **Ci. Inf.**, Brasília, v. 21, n. 3, p. 201-215, set./dez. 1992.

ARELLANO, Miguel Angel Márdero; CAREGNATO, Sônia Elisa; FERREIRA, Sueli Mara Soares Pinto. Editoração eletrônica de revistas científicas com suporte do protocolo OAI. In: FERREIRA, Sueli Mara Soares Pinto; TARGINO, Maria das Graças (Org.). **Preparação de revistas científicas: teoria e prática**. São Paulo: Reichmann e Autores, 2005.

BAILER JR., Charles W. **Open Access Bibliography: liberating scholarly, literature with e-prints, and open access journal**. Association of Research Libraries, 2005. Disponível em: <<http://www.escholarlypup.com/oab/oab>>. Acesso em: 25 maio 2006.

BARBETTA, Pedro Alberto. **Estatística aplicada às Ciências Sociais**. 2. ed. Florianópolis: Ed. da UFSC, 1998.

BARICHELLO, Eugenia M. da R. Comunicação e sociabilidades. In: SILVEIRA, Ada Cristina Machada da (Org.). **Comunicação & Sociabilidades**. Santa Maria: FACOS, 2001.

BARROSO, Paulo Blanco. Revista Ciência da Informação 100% eletrônica. **Ci. Inf.**, Brasília, v. 33, n. 1, 2004.

BAUMAN, Zygmund. **Modernidade líquida**. Rio de Janeiro: Zahar Ed., 1999.

CASTELLS, Manuel. **A sociedade em rede**. 6. ed. totalmente rev. e ampl. São Paulo: Paz e Terra, 2002. v.1

CDSWARE SOFTWARE CONSORTIUM. CDSware. Disponível em: <<http://cdsware.cern.ch/cdsware/overview.html>>. Acesso em: 22 maio 2006.

CERVO, Amado Luiz; BERVIAN, Pedro Alcino. **Metodologia científica: para estudantes universitários**. 3. ed. São Paulo: McGraw-Hill do Brasil, 1983.

CHATAIGNIER, Maria Cecília Pragana; SILVA, Margareth Prevot da. Biblioteca digital: a experiência do Impa. **Ci. Inf.**, Brasília, v. 30, n. 3, p. 7-12, set./dez. 2001.

COALITION FOR NETWORKED INFORMATION. Disponível em: <http://www.cni.org/>. Acesso em: 25 abril 2005.

DIGITAL LIBRARY FEDERATION. Disponível em: <<http://www.diglib.org/>>. Acesso em: 26 abril 2005.

DSPACE.ORG. Disponível em: <<http://dspace.org/introduction/index.html>>. Acesso em: 23 maio 2006.

DUBLIN CORE. Disponível em: <<http://www.dublincore.org/>>. Acesso em: 26 maio 2004.

DUBLIN CORE. Disponível em: <<http://www.dublincore.org/>>. Acesso em: 15 abril 2005.

EPRINTS. ORG. Disponível em: <<http://www.software.eprints.org/>>. Acesso em: 23 maio 2004.

EPRINTS. ORG. Disponível em: <<http://www.software.eprints.org/>>. Acesso em: 20 jun. 2005.

EPRINTS. ORG. Disponível em: <<http://www.software.eprints.org/>>. Acesso em: 25 maio 2006.

ETD-DB. Digital library or archives: formerly the scholarly communications project. Disponível em: <<http://scholar.lib.vt.edu/ETD-db/>>. Acesso em: 25 maio 2006.

FEDORA.INFO. Disponível em: <<http://www.fedora.info/about/history.shtml>>. Acesso em: 22 maio 2006.

GARCIA, Patrícia de Andrade Bueno. **Provedores de dados de baixo custo**: publicação digital ao alcance de todos. 2003, 128 f. Dissertação (Mestrado em Informática) Universidade Federal do Paraná, Curitiba, 2003.

GERARD, Westrienen Van; LYNCH, Clifford A. Academic institutional repositories: deployment status in 13 nations as of mid 2005. **D-Lib Magazine**, v. 11, n. 9. set. 2005. Disponível em: <<http://www.D-Lib.org/>>. Acesso em: 20 maio 2006.

GOMÉS, Maria Nélida González. Novas fronteiras tecnológicas das ações de informação: questões e abordagens. **Ci. Inf.**, Brasília, v. 33, n. 1, p. 55-67, jan./abril. 2004.

HOWE, Denis. Free on-line dictionary of computing. Disponível em: <http://wonbat.doc.ic.ac.uk/foldoc/>. Acesso em: 03 abril 2006.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Disponível em: <<http://www.ibict.br/secao.php?cat=SEER/>>. Acesso em: 23 maio 2006.

JOHNSON, Richard R. Institutional repositories: partnering with faculty to enhance scholarly communication. **D – Lib Magazine**, v. 8, n. 11, nov. 2002. Disponível em: <<http://www.D-Lib.org/>>. Acesso em: 20 maio 2006.

KNELLER, George Frederick. **A ciência como atividade humana**. Rio de Janeiro: Zahar; São Paulo: Ed. Univ. S. Paulo, 1980. 310 p.

- KUHN, T. S. **A estrutura das revoluções científicas**. 3. ed. São Paulo: Perspectiva, 1990.
- KURAMOTO, Hélio. IBICT estimula adoção dos Open Archives no Brasil. Instituto Brasileiro de Informação em Ciência e Tecnologia. 2005. Disponível em: <<http://www.ibict.br/>>. Acesso em: 20 maio 2006.
- KURAMOTO, Hélio. Os Open Archives e as políticas para a informação científica. Instituto Brasileiro de Informação em Ciência e Tecnologia. 2005. Disponível em: <<http://www.ibict.br/>>. Acesso em: 20 maio 2006.
- LAGOZE, Carl; SOMPEL, Herbert Van de. The Making of the Open Archives Initiative Protocol for Matadata Harvesting. 2003. Disponível em: <<http://public.lanl.gov/herbertv/>>. Acesso em: 20 maio 2006.
- LAGOZE, Carl; SOMPEL, Herbert Van de. The Open Archives Initiative: Building a low-barrier interoperability framework. 2001. Disponível em: <<http://public.lanl.gov/herbertv/>>. Acesso em: 20 maio 2006.
- LÉVY, Pierre. A revolução contemporânea em matéria de comunicação. In: MARTINS, Francisco Menezes; SILVA, Jurenir Machado da (Org.). **Para navegar no século XXI: tecnologias do imaginário e cibercultura**. 2. ed. Porto Alegre: EDIPUCRS, 2000.
- MACEDO, Tonia Marta Barbosa. Redes informais nas organizações: a co-gestão do conhecimento. **Ci. Inf.**, Brasília, v. 28, n. 1, p. 94-100, jan./abr. 1999.
- MAFFESOLI, Michel. A comunicação sem fim: teoria pós-moderna da comunicação. In: MARTINS, Francisco Menezes; SILVA, Jurenir Machado da (Org.). **A genealogia do virtual: comunicação, cultura e tecnologias do imaginário**. Porto Alegre: Sulina, 2004.
- MARCONDES, Carlos Henrique; SAYÃO, Luis Fernando. Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca digital brasileira. **Ci. Inf.**, Brasília, v. 30, n. 3, p. 71-78, set./dez. 2001.
- MARCONDES, Carlos Henrique; SAYÃO, Luis Fernando. Documentos digitais e novas formas de cooperação entre sistemas de informação em C&T. **Ci. Inf.**, Brasília, v. 31, n. 3, p. 71-78, set./dez. 2002.
- MATTELART, Armand. **História da sociedade da informação**. São Paulo: Loyola, 2002.
- MEADOWS, A. J. **A comunicação científica**. Brasília: Briquet de Lemos, 1999.
- MOREIRA, Walter. Os colégios virtuais e a nova configuração da comunicação científica. **Ci. Inf.**, Brasília, v. 34, n. 1, p. 57-63, jan/abr. 2005.
- NATIONAL SCIENCE FOUNDATION. Disponível em: <<http://www.nsf.gov/>>. Acesso em: 26 abril 2005.

OPEN ARCHIVES FORUM. Disponível em: <<http://www.oaforum.org/>>. Acesso em 26 maio 2004.

OPEN ARCHIVES FORUM. Disponível em: <<http://www.oaforum.org/>>. Acesso em 15 abril 2005.

OPEN ARCHIVES INITIATIVE. Disponível em: <<http://www.openarchives.org/>>. Acesso em 25 maio 2004.

OPEN ARCHIVES INITIATIVE. Disponível em: <<http://www.openarchives.org/>>. Acesso em 15 abril 2005.

OPUS-DATABASE. Publication server of the University of Stuttgart. Disponível em: <<http://elib.unistuttgart.de/opus/doku/about.php?la=en>>. Acesso em: 25 maio 2006.

PUBLIC KNOWLEDGE PROJECT. PKP@SFU. Disponível em: <<http://pkp.sfu.ca/ojs/>>. Acesso em: 25 maio 2006.

RAMOS, M. G., Modelos de comunicação e divulgação científicas: uma revisão de perspectivas. **Ci. Inf.**, Brasília, v. 23, n. 3, p. 340-348, set./dez. 1994.

RODRIGUES, Janete de Páscoa. **A produção técnico-científica dos docentes do departamento de educação física da Universidade Federal do Piauí: um estudo exploratório sob o enfoque da Ciência da Informação.** 2001, 159 f. Dissertação (Mestrado em Ciência da Informação) Universidade Federal do Piauí, Piauí, 2001.

RODRIGUEZ, M.V.R. (Martius V.R.); FERRANTE, Agustin J. **A tecnologia de informação e mudança organizacional.** Rio de Janeiro: Infobook, 1995. 391p.

SCHWARCZ, Lilia Moritz. Uma nova biblioteca: uma nova Lisboa. In: **A longa viagem da biblioteca dos reis: do terremoto de Lisboa à independência do Brasil.** São Paulo: Companhia das Letras, 2002. cap. 4.

SENA, Nathália Knipp. Open archives: caminho alternativo para a comunicação científica. **Ci. Inf.**, Brasília, v. 29, n. 3, p. 71-78, set./dez. 2000.

SISTEMA DE ARQUIVOS ABERTOS DA SOCIEDADE BRASILEIRA DE GENÉTICA. Disponível em: <<http://mangueira.ibict.br/>>. Acesso em: 23 maio 2006.

SOUZA, Maria da Paixão Neres de. Efeitos das tecnologias da informação na comunidade de pesquisadores da Embrapa. **Ci. Inf.**, Brasília, v. 32, n. 1, set./dez. 2000.

YEARLEY, Steven. A sociologia do conhecimento e a sociologia da comunidade científica. In: GONÇALVES, Maria Eduarda (Coord.). **Comunidade científica e poder.** Lisboa: Edições 70, 1993.

WEITZEL, Simone da Rocha. E-Prints: modelo da comunicação científica em transição. In: FERREIRA, Sueli Mara Soares Pinto; TARGINO, Maria das Graças (Org.). **Preparação de revistas científicas: teoria e prática**. São Paulo: Reichmann e Autores, 2005.

APÊNDICES

APÊNDICE A
Formulário de pesquisa de dados primários

FORMULÁRIO DE PESQUISA – DADOS PRIMÁRIOS

1 Dados sobre o *Open Archives*

Nome:

Comunidade Científica:

País de origem:

Endereço eletrônico:

Responsáveis:

Idioma:

Área de atuação:

Ano de criação:

Tipo de *Open Archives* conforme sua classificação:

Research Institutional or Departmental ()

Research Cross-Institution ()

E-Theses ()

Demonstration ()

E-Journal/Publication ()

Database ()

Outros ()

Volume de registros:

Ano: ___/___/___ Número de registros: _____

2 Motivos para Criação:

3 Objetivos do *Open Archives*:

4 Tecnologia adotada para implantação:

5 Fluxo de informação do *Open Archives*:

- Disponibilidade de acesso ao documento:

Categorias de usuários

Leitor Visitante ()
Leitor Cadastrado ()
Autores ()

Possibilidades de interação

Leitor visitante ()
Troca de informações ()
Comentários ()
Apenas acesso ()
Nenhum ()

Leitor cadastrado ()
Troca de informações ()
Comentários ()
Apenas acesso ()
Nenhum ()

Autores ()
Troca de informações ()
Comentários ()
Apenas acesso ()
Nenhum ()

APÊNDICE B

Formulário de pesquisa de dados secundários

FORMULÁRIO DE PESQUISA – DADOS SECUNDÁRIOS

1 Fonte / Referência:

2 Instituições que oferecem suporte a OAI

Nome: _____

História: _____

Motivação: _____

Objetivos: _____

3 Fatores que levaram ao surgimento dos Open Archives

1 _____

2 _____

3 _____

4 _____

5 _____

4 Principais tecnologias envolvidas na OAI

1 _____

2 _____

3 _____

4 _____

5 _____