



**UNIVERSIDADE DE BRASÍLIA**  
**FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, CONTABILIDADE E**  
**CIÊNCIA DA INFORMAÇÃO E DOCUMENTAÇÃO – FACE**  
**DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO E DOCUMENTAÇÃO – CID**  
**Programa de Pós-Graduação em Ciência da Informação**

**ROBSON LOPES DE ALMEIDA**

## **DISSEMINAÇÃO DE CONTEÚDOS NA WEB:**

**A TECNOLOGIA RSS COMO PROPOSTA PARA A COMUNICAÇÃO CIENTÍFICA**

BRASÍLIA

2008

ROBSON LOPES DE ALMEIDA

## DISSEMINAÇÃO DE CONTEÚDOS NA WEB:

A TECNOLOGIA RSS COMO PROPOSTA PARA A COMUNICAÇÃO CIENTÍFICA

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação do Departamento de Ciência da Informação e Documentação da Universidade de Brasília como requisito parcial para obtenção do título de Mestre.  
Área de Concentração: Transferência da Informação  
Linha de Pesquisa: Comunicação da Informação

Orientadora: Dra. Elmira Luzia Melo Soares Simeão

BRASÍLIA

2008

A447d Almeida, Robson Lopes de.

Disseminação de conteúdos na Web: a tecnologia RSS como proposta para a comunicação científica / Robson Lopes de Almeida. -- Brasília, 2008.  
192 f.

Dissertação (Mestrado) - Universidade de Brasília, Departamento de Ciência da Informação e Documentação.

Orientadora: Dra. Elmira Luzia Melo Soares Simeão.

1. Periódicos eletrônicos. 2. Periódicos científicos. 3. Comunicação científica. 4. Disseminação seletiva de informação. 5. Sindicação de conteúdo. 6. RSS. I Título.

CDU025.5.001.8:004.738.5(043.3)

Autorizo a reprodução desta dissertação por processo de fotocopiadoras e outros meios eletrônicos para fins exclusivamente acadêmicos e científicos.

Brasília, 29 de fevereiro de 2008

Robson Lopes de Almeida



## FOLHA DE APROVAÇÃO

**Título:** “Disseminação Seletiva de Conteúdos na Web: a tecnologia RSS como proposta para a Comunicação Científica”

**Autor:** Robson Lopes de Almeida

**Área de concentração:** Transferência da Informação

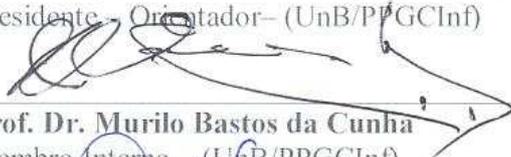
**Linha de pesquisa:** Comunicação da Informação

Dissertação submetida à Comissão Examinadora designada pelo Colegiado do Programa de Pós-graduação em Ciência da Informação do Departamento de Ciência da Informação e Documentação da Universidade de Brasília como requisito parcial para obtenção do título de **Mestre** em Ciência da Informação.

Dissertação aprovada em: 29 de fevereiro de 2008.

Aprovado por:

  
\_\_\_\_\_  
**Profa. Dra. Elmira Luzia Melo Soares Simeão**  
Presidente - Orientador - (UnB/PPGCInf)

  
\_\_\_\_\_  
**Prof. Dr. Murilo Bastos da Cunha**  
Membro Interno - (UnB/PPGCInf)

  
\_\_\_\_\_  
**Profa. Dra. Sueli Mara Soares Pinto Ferreira**  
Membro Externo - USP

  
\_\_\_\_\_  
**Profa. Dra. Sely Maria de Souza Costa**  
Suplente - (UnB/PPGCInf)

***PARA JOÃO GABRIEL E ANDREA,  
RAZÕES DA MINHA VIDA***

## AGRADECIMENTOS

À CAPES, pelo apoio financeiro.

À Profª Elmira Simeão, pelo acompanhamento e orientação nos momentos necessários, que ajudaram na definição do foco desse trabalho.

Aos professores do Programa de Pós-Graduação em Ciência e Informação (PPGCINF/CID), em especial às professoras Sely Costa e Marisa Brascher, pelo apoio em diversos momentos, mas principalmente pelo carinho.

Às meninas da secretaria do PPGCINF/CID, Jucilene e Martha, sempre tão prestativas e atenciosas.

Aos pesquisadores da ANCIB e a todos os coordenadores de seus Grupos de Trabalho, por terem colaborado de maneira fundamental para a realização dessa pesquisa em sua totalidade.

Aos colegas do Laboratório de Metodologias de Tratamento e Disseminação da Informação do IBICT, em especial a coordenadora, Bianca Amaro, por acreditarem e incentivarem o uso do RSS no contexto acadêmico-científico.

Aos colegas de mestrado, por compartilharem seus conhecimentos em momentos tão ricos nas disciplinas que cursei. Em especial à Alessandra, pela torcida, e Milton, pelo apoio técnico.

Aos amigos de Brasília, do Rio e os compadres de POA (usuários de RSS), que aceitaram minha ausência e reclusão nos últimos meses e nem por isso deixaram de torcer por mim.

À minha irmã Renata, minha amiga Denise e à minha cunhada, Lynette, pelo valioso apoio, que foi além do carinho e amizade, mas me ajudaram a fazer desse trabalho algo palpável.

À querida família, mãe, avó, sobrinhos, sobrinhas, sogro, sogra e até os cunhados pelo incentivo e torcida, mesmo à distância.

E, *last but not least*, minha gratidão mais do que especial ao meu filho João Gabriel e minha esposa Andrea, que estiveram incondicionalmente ao meu lado, vibrando com as pequenas vitórias e compreendendo os momentos mais difíceis, minhas ausências e minhas angústias...

Muito obrigado!!!

*“Assim como uma lâmpada se apaga devido ao excesso de óleo,  
também a mente se extingue por excesso de conhecimento”  
(Montaigne, séc. XIV)*

ALMEIDA, Robson Lopes de. **Disseminação de conteúdos na Web: a tecnologia RSS como proposta para a comunicação científica**. 2008. 192 f. Dissertação (Mestrado em Ciência da Informação). Departamento de Ciência da Informação e Documentação, Universidade de Brasília, Brasília. 2008.

## RESUMO

Os efeitos da sobrecarga de informação verificada após o surgimento da Web, nos anos 90, vêm sendo motivo de preocupação de pesquisadores quanto à capacidade dos usuários na absorção e aproveitamento de conteúdos relevantes e úteis. A comunicação científica, tanto no seu aspecto informal quanto no formal, representado principalmente pelos periódicos científicos, enfrenta o mesmo problema, pois também sofre a influência dos avanços promovidos pelas tecnologias de informação e comunicação. Notadamente, verifica-se um crescimento nos títulos de periódicos científicos eletrônicos, conseqüência da agilidade no processo de publicação em meio eletrônico, que induz ao aumento da produção bibliográfica. Desse modo, se faz cada vez mais necessária a utilização de filtros, permitindo uma disseminação seletiva de conteúdos para usuários especializados, conforme suas reais necessidades. É nesse contexto que se insere a presente pesquisa, focada no estudo do potencial de uma tecnologia emergente, o RSS (*Rich Site Summary*, *RDF Site Summary* ou, ainda, *Really Simple Syndication*), voltada para a agregação e distribuição de conteúdos digitais. Como um estudo exploratório, o objetivo principal foi perceber as possíveis aplicações deste recurso no âmbito da comunicação científica, especialmente quanto ao aspecto da disseminação. Adicionalmente, comenta-se as vantagens e barreiras da tecnologia, além dos modos pelos quais os formatos RSS vêm sendo empregados no meio acadêmico. Além da análise da literatura pertinente, foi realizado um estudo de comportamento informacional junto a um grupo de pesquisadores da área da Ciência da Informação envolvidos com o processo de comunicação científica a fim de testar a viabilidade na adoção da tecnologia proposta. Complementando a pesquisa, descreve-se a criação de um protótipo de um serviço de agregação de conteúdo temático para ilustrar as funcionalidades da tecnologia. Os resultados obtidos indicam um terreno fértil e propício para implantação do RSS no contexto estudado, sugerindo ampliação da visibilidade da informação científica, além de integração com outros recursos tecnológicos, tais como bases de dados, bibliotecas e repositórios digitais.

**Palavras-chave:** periódico científico eletrônico; periódico científico; disseminação seletiva de informação, DSI, comunicação científica; sindicância de conteúdo; RSS.

ALMEIDA, Robson Lopes de. **Dissemination of contents in the Web: RSS technology as a proposal for scientific communication**. 2008. 192 f. Dissertation (Master's in Information Science). Department of Information Science and Documentation, University of Brasilia, Brasilia. 2008.

## ABSTRACT

The effect of the information overload verified after the advent of the Web, in the 90s, has become a concern for researchers as to the capability of the users in the process of acquisition and utilization of relevant and useful contents. Scientific communication, in its informal as well as in its formal aspect, represented mainly by scientific periodicals, faces the same problem, since it also suffers with the influence of the advancements caused by the information and communication technologies. Distinctively, a growth in the scientific electronic periodical headings is verified, as a consequence of the agility in the electronic publication process, what leads to the increase of the bibliographical production. This way, the use of filters is becoming more and more required, what allows a selective dissemination of contents for specialized users, according to their real needs. Inserted in this context, the present research focusses on the study of the potential of an emergent technology: the RSS (Rich Site Summary, RDF Site Summary or even Really Simple Syndication), intended to the aggregation and distribution of digital contents. As an inquisitive study, the main objective was to perceive the possible applications of this resource in the scope of scientific communication, especially as to the aspect of dissemination. Additionally, a comment about the advantages and barriers of the technology is made, as well as the ways by which formats RSS have been employed in the academic field. Besides the analysis of pertinent literature, a study of informational behavior of a group of Information Science researchers involved with the process of scientific communication in order to test the viability of the proposal technology. Complementing the research, the creation of a prototype of an aggregation of thematic contents service is described as to illustrate the functionalities of the technology. The obtained results indicate a fertile and propitious field for implantation of the RSS in the studied context, suggesting an increase of the visibility of scientific information, besides the integration with other technological resources, such as databases, digital libraries and repositories.

**Key-words:** electronic scientific journal; scientific journal; selective dissemination of information, SDI; scientific communication; content syndication; RSS.

## LISTA DE FIGURAS

Figura 2.1	Modelo genérico de um sistema de filtragem de informação .....	33
Figura 2.2	Busca de informação e busca em sistemas de informação.....	40
Figura 3.1	Modelo de comunicação científica, segundo Garvey e Griffith (1979).....	44
Figura 3.2	Modelo híbrido de comunicação científica, segundo Costa (1999) .....	49
Figura 5.1	Esquema de funcionamento de um sistema baseado em RSS.....	92
Figura 5.2	Crescimento do número de <i>feeds</i> RSS entre 09/2001 a 12/2007 .....	93
Figura 5.3	História do desenvolvimento dos feeds RSS e Atom .....	98
Figura 5.4	Estrutura básica de um <i>feed</i> RSS (versão 0.92) .....	103
Figura 5.5	Modelo representativo da especificação RSS 1.0 (HAMMERSLEY, 2003) ...	104
Figura 5.6	Modelo representativo da especificação RSS 2.0 (HAMMERSLEY, 2003) ...	106
Figura 5.7	Página inicial da CAPES, visualizada pelo navegador Mozilla Firefox .....	110
Figura 5.8	Exibição do <i>feed</i> RSS da CAPES pelo navegador Mozilla Firefox .....	110
Figura 5.9	Tela do FeedReader, exemplo de aplicativo agregador de conteúdo .....	111
Figura 5.10	Tela do Google Reader, leitor de RSS baseado na Web .....	113
Figura 6.1	Exemplo de <i>feed</i> RSS utilizando vocabulários de metadados DC e PRISM .	124
Figura 6.2	Documentos que tratam de “RSS”, segundo a base LISTA/EBSCO.....	133
Figura 6.3	Criação dinâmica de <i>feed</i> RSS correspondente a um resultado de busca....	134
Figura 6.4	Tela do <i>Bibliorandum</i> : <i>feeds</i> RSS da produção científica recente.....	137
Figura 7.1	Atributos funcionais dos artigos científicos.....	157
Figura 7.2	Possíveis aplicações de RSS em periódicos científicos eletrônicos.....	159
Figura 7.3	Tela do FeedForAll, ferramenta de autoria para criação de <i>feeds</i> RSS .....	161
Figura 7.4	Tela inicial do agregador temático em Ciência da Informação .....	163
Figura 7.5	Resultado da busca pelo termo “gestão” em fontes simultâneas .....	165
Figura 7.6	Serviço <i>pipe</i> criado para filtragem dos <i>feeds</i> agregados.....	166
Figura 7.7	Filtragem de conteúdo a partir de <i>feeds</i> agregados.....	167

## LISTA DE QUADROS

Quadro 2.1 Diferenças fundamentais entre os sistemas de Recuperação de Informação (RI) e Filtragem de Informação (FI) .....	32
Quadro 3.2 Características dos canais de comunicação formal, informal e eletrônica (TARGINO, 2000) .....	48
Quadro 3.3 Evolução dos conceitos dos periódicos científicos eletrônicos.....	53
Quadro 3.4 Características dos formatos de comunicação intensiva e extensiva .....	60
Quadro 3.5 Ranking de <i>blogs</i> científicos no mundo.....	70
Quadro 4.1 Relacionamento entre objetivos específicos, capítulos e método .....	77
Quadro 4.2 Periódicos científicos eletrônicos em Ciência da Informação disponíveis na plataforma SEER .....	87
Quadro 5.1 Principais formatos de sindicância de conteúdo .....	99
Quadro 5.2 Descrição dos sub-elementos requeridos do RSS 1.0 .....	105
Quadro 5.3 Descrição dos sub-elementos opcionais no RSS 1.0.....	105
Quadro 5.4 Descrição dos sub-elementos requeridos do RSS 2.0 .....	107
Quadro 5.5 Descrição dos sub-elementos opcionais do RSS 2.0.....	108
Quadro 5.6 Diretórios de indexação e divulgação de feeds RSS.....	114
Quadro 6.1 Editoras científicas que utilizam elementos de metadados dos vocabulários DC e PRISM na criação de feeds RSS .....	125
Quadro 6.2 Metadados DC e PRISM utilizados juntamente com o elemento <i>&lt;channel&gt;</i> de feeds RSS de editoras científicas .....	126
Quadro 6.3 Metadados DC e PRISM utilizados juntamente com o elemento <i>&lt;item&gt;</i> de feeds RSS de editoras científicas .....	127
Quadro 6.4 Editoras científicas que não utilizam elementos de metadados na descrição de suas publicações .....	129
Quadro 6.5 Utilização de <i>feeds</i> RSS por bibliotecas universitárias .....	130
Quadro 6.6 Exemplos de bases de dados online que disponibilizam <i>feeds</i> RSS .....	135
Quadro 7.2 <i>Feeds</i> RSS das fontes de informação reunidas pelo agregador temático em Ciência da Informação.....	163

## LISTA DE TABELAS

Tabela 4.1 Evolução da literatura sobre RSS segundo a base de dados LISA .....	76
Tabela 7.1 Total das respostas .....	140
Tabela 7.2 Faixa etária .....	140
Tabela 7.3 Titulação acadêmica .....	141
Tabela 7.4 Estados de origem .....	141
Tabela 7.5 Perfil de atuação .....	142
Tabela 7.6 Vínculo aos GTs da ANCIB.....	142
Tabela 7.7 Freqüência de utilização das revistas em Ciência da Informação .....	144
Tabela 7.8 Importância da notificação de conteúdo .....	145
Tabela 7.9 Uso de serviços de alerta.....	145
Tabela 7.10 Tipo de serviço de alerta.....	146
Tabela 7.11 Classificação do serviço de alerta .....	146
Tabela 7.12 Hábito de leitura .....	147
Tabela 7.13 Utilização de ferramentas de busca .....	149
Tabela 7.14 Necessidade de buscas simultâneas .....	149
Tabela 7.15 Conhecimento da tecnologia RSS.....	151
Tabela 7.16 Relação entre faixa etária e uso de RSS.....	152
Tabela 7.17 Hábito de assinatura de <i>feeds</i> RSS.....	153
Tabela 7.18 Assinatura de <i>feeds</i> RSS em periódicos científicos eletrônicos.....	153
Tabela 7.19 Barreiras identificadas na utilização de <i>feeds</i> RSS .....	155

## LISTA DE ABREVIATURAS E SIGLAS

AAAS	AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE
ANCIB	ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
BIREME	CENTRO LATINO-AMERICANO E DO CARIBE DE INFORMAÇÃO EM CIÊNCIAS DA SAÚDE
BLEND	BIRMINGHAM AND LOUGHBOROUGH ELECTRONIC NETWORK DEVELOPMENT
CAPES	COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR
CRAYON	CREATE YOUR OWN NEWSPAPER
COGPRINTS	COGNITIVE SCIENCES EPRINT ARCHIVE
CORE	CHEMISTRY ONLINE RETRIEVAL EXPERIMENT
DCMI	DUBLIN CORE METADATA INICIATIVE
DOAJ	DIRECTORY OF OPEN ACCESS JOURNALS
DSI	DISSEMINAÇÃO SELETIVA DE INFORMAÇÃO
EIES	ELECTRONIC INFORMATION EXCHANGE SYSTEM
ELIB	ELECTRONIC LIBRARIES PROGRAMME
E-LIS	E-PRINTS IN LIBRARY AND INFORMATION SCIENCE
ENANCIB	ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO
FAPESP	FUNDAÇÃO DE AMPARO A PESQUISA DO ESTADO DE SÃO PAULO
HTML	HYPERTEXT MARKUP LANGUAGE
IBICT	INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA
IETF	INTERNET ENGINEERING TASK FORCE
JSTOR	JOURNAL STORAGE
LISA	LIBRARY AND INFORMATION SCIENCE ABSTRACTS
LIST	LIBRARY/INFORMATION SCIENCES & TECHNOLOGY ABSTRACTS
MARC	MACHINE READABLE CATALOGING
MCF	META CONTENT FRAMEWORK

NCSA	NATIONAL CENTER FOR SUPERCOMPUTING APLICATIONS
NCSTRL	NETWORKED COMPUTER SCIENCE TECHNICAL REFERENCE LIBRARY
NIB	NÚCLEO DE INFORMÁTICA BIOMÉDICA
NSF	NATIONAL SCIENCE FOUNDATION
OAI	OPEN ARCHIVES INICIATIVE
OAI-PMH	OPEN ARCHIVES INICIATIVE – PROTOCOL OF METADATA HARVESTING
OCLC	ON-LINE COMPUTER LIBRARY CENTER
OJS	OPEN JOURNAL SYSTEM
OPML	OUTLINE PROCESSOR MARKUP LANGUAGE
PRISM	PUBLISHING REQUIREMENTS FOR INDUSTRY STANDARDS METADATA
RDF	RESOURCE DESCRIPTION FRAMEWORK
PKP	PUBLIC KNOWLEDGE PROJECT
RSS	RICH SITE SUMMARY; RDF SITE SUMMARY; REALLY SIMPLE SYNDICATION
SCIELO	SCIENTIFIC ELECTRONIC LIBRARY ONLINE
SDI	SELECTIVE DISSEMINATION OF INFORMATION
SEER	SERVIÇO DE EDITORAÇÃO ELETRÔNICA DE REVISTAS
SGML	STANDARD GENERALIZED MARKUP LANGUAGE
TIC	TECNOLOGIAS DE INFORMAÇÃO E COMUNICAÇÃO
TULIP	THE UNIVERSITY LICENSING PROGRAM
URL	UNIFORM RESOURCE LOCATOR
WEB	WORLD WIDE WEB
W3C	WORLD WIDE WEB CONSORTIUM
XML	EXTENSIBLE MARKUP LANGUAGE

# SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>17</b>
1.1 APRESENTAÇÃO.....	17
1.2 PROBLEMA DE PESQUISA .....	20
1.3 OBJETIVO GERAL.....	21
1.4 OBJETIVOS ESPECÍFICOS .....	22
1.5 ORGANIZAÇÃO DA DISSERTAÇÃO.....	22
<b>2. CONCEITOS FUNDAMENTAIS .....</b>	<b>25</b>
2.1 SOBRECARGA DA INFORMAÇÃO.....	25
2.2 SISTEMAS DE FILTRAGEM DE INFORMAÇÃO .....	28
2.3 DISSEMINAÇÃO SELETIVA DE INFORMAÇÃO .....	34
2.4 COMPORTAMENTO INFORMACIONAL EM PERIÓDICOS CIENTÍFICOS ELETRÔNICOS .....	38
<b>3. A COMUNICAÇÃO CIENTÍFICA.....</b>	<b>43</b>
3.1 COMUNICAÇÃO CIENTÍFICA E AS TECNOLOGIAS DE INFORMAÇÃO .....	46
3.2 O PERIÓDICO CIENTÍFICO ELETRÔNICO .....	52
3.2.1 EVOLUÇÃO DOS PERIÓDICOS CIENTÍFICOS ELETRÔNICOS .....	55
3.3 COMUNICAÇÃO EXTENSIVA .....	60
3.3.1 INDICADORES DA COMUNICAÇÃO EXTENSIVA .....	60
3.3.1.1 INTERATIVIDADE .....	61
3.3.1.2 HIPERTEXTUALIDADE .....	62
3.3.1.3 HIPERMIDIAÇÃO.....	63
3.3.2 COMUNICAÇÃO CIENTÍFICA E RSS.....	64
3.4 ACESSO ABERTO: UM NOVO MODELO PARA A COMUNICAÇÃO CIENTÍFICA .....	65
3.5 BLOGS CIENTÍFICOS .....	68
<b>4. PROCEDIMENTOS METODOLÓGICOS .....</b>	<b>74</b>
4.1 TIPO DE PESQUISA .....	74
4.2 ETAPAS DA PESQUISA.....	75
4.3 POPULAÇÃO ESTUDADA.....	79
4.4 INSTRUMENTOS DE COLETA .....	81
4.4.1 CARACTERÍSTICAS DO QUESTIONÁRIO ELETRÔNICO.....	82
4.4.2 CARACTERÍSTICAS DO QUESTIONÁRIO APLICADO AOS COORDENADORES DE GTS.....	82
4.5 PRÉ-TESTE.....	83
4.6 PROTÓTIPO DE AGREGADOR DE CONTEÚDOS EM CI .....	84
4.6.1 AMOSTRA .....	84
<b>5. TECNOLOGIA RSS.....</b>	<b>89</b>
5.1 APRESENTAÇÃO.....	89
5.2 O QUE É RSS?.....	90
5.3 BREVE HISTÓRIA DA SINDICAÇÃO DE CONTEÚDO NA WEB .....	94
5.4 ARQUITETURA DE UM FEED RSS .....	100
5.5 FUNCIONAMENTO BÁSICO DA TECNOLOGIA RSS.....	108
5.6 BARREIRAS PARA UTILIZAÇÃO DE RSS.....	115

<b>6. A TECNOLOGIA RSS APLICADA À COMUNICAÇÃO CIENTÍFICA .....</b>	<b>119</b>
6.1 CONJUNTOS DE METADADOS .....	120
6.1.1 VOCABULÁRIO DUBLIN CORE .....	121
6.1.2 VOCABULÁRIO PRISM.....	123
6.2 APLICAÇÕES MAIS COMUNS .....	129
6.2.1 BIBLIOTECAS E UNIDADES DE INFORMAÇÃO.....	130
6.2.2 BASES DE DADOS .....	132
6.2.3 OUTRAS APLICAÇÕES .....	135
<b>7. ANÁLISE DOS RESULTADOS .....</b>	<b>139</b>
7.1 CARACTERÍSTICAS DA POPULAÇÃO ESTUDADA .....	139
7.2 COMPORTAMENTO INFORMACIONAL .....	143
7.2.1 FREQUENCIA DE USO DE PERIÓDICOS EM CIÊNCIA DA INFORMAÇÃO.....	143
7.2.2 UTILIZAÇÃO DE SERVIÇOS DE ALERTA .....	145
7.2.3 PROFUNDIDADE DE LEITURA E RECUPERAÇÃO DE INFORMAÇÃO .....	147
7.3 CONHECIMENTO E USO DA TECNOLOGIA RSS .....	150
7.4 ATRIBUTOS DOS ARTIGOS DE PERIÓDICOS .....	156
7.5 CONTEÚDOS RSS EM PERIÓDICOS CIENTÍFICOS.....	158
7.6 CRIAÇÃO DE FEEDS DE PERIÓDICOS NACIONAIS .....	159
7.7 MONTAGEM DO PROTÓTIPO DO AGREGADOR TEMÁTICO.....	162
<b>8. CONCLUSÕES E RECOMENDAÇÕES .....</b>	<b>169</b>
8.1 SUGESTÕES PARA ESTUDOS FUTUROS .....	174
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>177</b>
<b>APÊNDICES.....</b>	<b>186</b>
APÊNDICE A - MODELO DO QUESTIONÁRIO ON-LINE DISPONIBILIZADO.....	186
APÊNDICE B - FORMULÁRIO APLICADO AOS COORDENADORES DE GTS DA ANCIB .....	190

# CAPÍTULO 1

## Introdução

---

## 1.1 Apresentação

Um novo capítulo da história das comunicações começou a ser escrito a partir do início dos anos 90 com o advento da *World Wide Web* – a parte gráfica e multimídia da Internet. Em pouco tempo, as páginas da Web tomaram conta da chamada “grande rede”, provocando uma verdadeira “explosão informacional”, de impacto ainda maior do que o mesmo fenômeno observado na década de 40, com o aumento de títulos de periódicos especializados após a Segunda Guerra Mundial. Esse movimento acabou impulsionando o desenvolvimento da própria Ciência da Informação, que teve sua origem a partir da revolução científica e tecnológica decorrente do pós-Guerra. (SARACEVIC, 1996, p. 42).

Observa-se que o volume de informação disponibilizado na Web nos últimos anos vem se tornando um motivo a mais de preocupação para a nossa sociedade. E a razão não está somente na dificuldade em localizar determinado conteúdo que atenda às reais necessidades do usuário, mas, principalmente, no efeito de sua sobrecarga. O rápido desenvolvimento das tecnologias de informação e comunicação (TICs) tem colaborado para a disseminação desordenada de novos conteúdos, os quais não encontram mais limites nos suportes digitais. Essa situação implica no aparecimento de vários problemas, como a falta de organização e estruturação das informações, dificultando a sua recuperação.

É bem verdade que a simplicidade das atuais ferramentas de publicação na rede tem facilitado tanto quem produz quanto quem consome informação, oferecendo mecanismos ágeis e de baixo custo para difusão de novas idéias. Por outro lado, a velocidade com que a informação digital circula na rede tem exigido maior atenção em relação à qualidade do conteúdo a ser publicado e mais discernimento quanto ao seu uso. A comunicação científica também foi afetada por essas transformações, haja vista a proliferação dos periódicos científicos eletrônicos.

Profissionais das mais diversas áreas do conhecimento, comprometidos com a pesquisa e desenvolvimento de tecnologias voltadas para a recuperação da informação, têm buscado soluções que possam lidar com essa situação.

Historicamente, porém, já nos anos 50 surgiam os primeiros mecanismos capazes de selecionar informação relevante para um determinado usuário, considerando seu perfil de interesse.

Desse modo, faz-se cada vez mais necessária a utilização de serviços de intermediação que atuem como filtros, permitindo a coleta e a distribuição seletiva de informação aos usuários de acordo com sua necessidade. É dentro dessa perspectiva e, considerando o caos informacional da Web, que pretende-se estudar o processo de divulgação da produção científica a partir da aplicação de uma tecnologia emergente conhecida genericamente como *Web Syndication* (sindicação de conteúdo Web), a qual vem sendo popularizada pelos formatos RSS<sup>1</sup>. Por meio desse recurso é possível oferecer aos usuários notificações automáticas sobre a atualização de conteúdos disponibilizados sob a plataforma Web.

Criado originalmente para servir como padrão na construção de sistemas de publicação de *headlines* (manchetes) das páginas da Internet, o RSS, em última instância, propõe-se a auxiliar os consumidores de informação a satisfazerem suas necessidades imediatas, mantendo-os permanentemente atualizados sem que se percam diante do vasto manancial de informações disponíveis na rede.

Atualmente, a tecnologia vem sendo utilizada pela maioria dos sítios de notícias e *weblogs*<sup>2</sup>, já que ambos possuem como características a alta frequência na atualização de seus conteúdos. Porém, vem ganhando cada vez mais popularidade em outras aplicações, variando desde o *marketing*, por meio da notificação de um determinado produto, à divulgação de alertas sobre previsão do tempo ou condições de estradas.

No caso dos periódicos científicos eletrônicos, foco do presente estudo, percebe-se que a utilização do RSS pode trazer benefícios para os pesquisadores, já que lhes permite manterem-se informados a respeito de temas de seu interesse por meio do recebimento de avisos sobre os artigos de uma determinada área, por

---

<sup>1</sup> O termo RSS é o acrônimo de *Rich Site Summary*, *RDF Site Summary* ou, ainda, *Really Simple Syndication*, dependendo da especificação utilizada. Os diferentes significados atribuídos serão explicados no capítulo 4 – Tecnologia RSS.

<sup>2</sup> Entende-se por *weblog* ou *blog* um tipo de registro publicado na Internet relativo a um determinado tema e organizado cronologicamente como um histórico ou diário.

exemplo. Além disso, a agregação dos conteúdos categorizados por tópicos em um único local – permitido pelos programas leitores de documentos RSS – facilita a pesquisa e recuperação de informações que se encontram distribuídas em diferentes instâncias (publicações, bases de dados, eventos, etc.). Por outro lado, editores e autores também ganham com o aumento da visibilidade da produção científica. Dessa forma, incrementa-se o processo da comunicação científica, tornando mais fácil a divulgação desse tipo de informação, independentemente da ação do usuário.

Pretende-se, com esse trabalho, estudar as características da tecnologia de sindicância<sup>3</sup> de conteúdos, em particular a aplicação dos formatos RSS na promoção de recursos que facilitem a divulgação da produção dos periódicos científicos eletrônicos. Adicionalmente, será abordado o funcionamento das aplicações já existentes, suas vantagens e as possíveis dificuldades encontradas para implantação da referida tecnologia.

A metodologia adotada previu um estudo realizado junto a um grupo de usuários de revistas científicas eletrônicas com o objetivo de buscar subsídios para questões relacionadas à viabilidade na adoção do RSS como meio de distribuição de conteúdo científico. Adicionalmente, também foi feito um levantamento dos requisitos mais comuns dos documentos publicados por estes periódicos (especialmente os artigos científicos).

Percebe-se a contribuição da presente pesquisa, principalmente no que concerne a contribuição de estudos científicos nacionais a respeito do tema, uma vez que é verificado um crescente aumento na literatura que trata da sindicância de conteúdo na Web e dos formatos RSS localizados principalmente no exterior. Estudiosos de diferentes áreas e disciplinas começam a lidar com esses assuntos, em tópicos que contextualizam o uso da blogosfera<sup>4</sup> e aplicações da chamada Web 2.0.

---

<sup>3</sup> O termo “sindicância” é uma tradução da palavra *syndication*, em Inglês, e pode ser entendido como a atividade de coleta e replicação de conteúdos dinâmicos em ambientes digitais.

<sup>4</sup> Segundo a Wikipédia, “blogosfera” é o termo coletivo que compreende todos os *blogs* como uma comunidade ou rede social.

## 1.2 Problema de Pesquisa

Com o advento das tecnologias de informação e comunicação (TICs), em particular a Internet que se destaca como seu principal expoente, verifica-se um aumento considerável no volume de informações produzidas e disponibilizadas facilmente e a baixo custo na Web. Em excesso, os conteúdos quando não são inúteis, acabam sendo prejudiciais à absorção de novas informações.

A comunicação científica também foi – e continua sendo – influenciada pela evolução das tecnologias. A crescente utilização da Internet, que se configura como uma poderosa fonte de informação para os cientistas, tem tornado mais ágil o processo de disseminação dos resultados de pesquisas, principalmente no que tange ao aspecto informal da comunicação. Porém, as transformações verificadas nos últimos anos em relação à transição dos periódicos impressos para os eletrônicos apontam para o reconhecimento da contribuição das tecnologias de informação também na comunicação científica formal.

Outro aspecto observado é o incentivo ao auto-depósito da produção científica, promovido pela iniciativa dos arquivos abertos (*open archives*) que, de acordo com especialistas, vem modificando os fluxos do modelo tradicional da comunicação científica.

Atualmente vivemos o paradoxo da quantidade extraordinária de informação disponível na Internet ser justamente um obstáculo para que consigamos acessá-la de forma eficiente. Assim, considera-se pertinente a afirmação de Silva Filho e Delgado (2003), segundo a qual “o verdadeiro valor produzido por um fornecedor de informação reside em focalizar, filtrar e comunicar o que é útil para o consumidor”. Isso implica que o usuário tenha uma percepção do conteúdo publicado anteriormente para que possa detectar as alterações nos conteúdos de uma determinada página da Web, por exemplo.

Nesse sentido, encontrar alternativas que possibilitem filtrar as informações que atendam ao perfil de um determinado usuário representa uma iniciativa muito bem vinda e objeto de interesse de estudos da Ciência da Informação.

Para tratar do problema em questão, identificamos na tecnologia conhecida como *Web syndication* a capacidade de agregar e distribuir conteúdos publicados na Web, os quais poderão ser acessados pelos usuários que manifestarem interesse. Essa tecnologia é implementada por meio de programas específicos ou serviços conhecidos como “agregadores”, construídos a partir dos formatos RSS, a serem detalhados mais adiante.

Acredita-se que a utilização das aplicações de agregação de conteúdos ajudará a minimizar o problema da sobrecarga de informação e da dificuldade do usuário em se manter em dia com relação aos assuntos de seu interesse, sobretudo no âmbito da comunicação científica. Assim sendo, no decorrer desta dissertação, procuraremos responder à seguinte pergunta de pesquisa:

*Quais as possibilidades de aplicação da tecnologia de syndicação de conteúdos Web como ferramenta capaz de incrementar o processo de disseminação da produção científica, particularmente quanto aos conteúdos registrados pelos periódicos científicos eletrônicos?*

A partir da formulação da pergunta enunciada anteriormente, consideramos os seguintes objetivos, a serem verificados com base nos resultados obtidos com essa pesquisa:

### **1.3 Objetivo Geral**

Identificar o potencial de utilização da tecnologia de syndicação de conteúdos Web, viabilizada principalmente pelo conjunto de formatos RSS, a fim de verificar as contribuições deste recurso para maximizar o processo de disseminação da informação científica.

## 1.4 Objetivos Específicos

1. descrever as características e funcionamento da tecnologia de sindicância de conteúdo, particularmente os formatos RSS;
2. identificar como os formatos RSS vêm sendo empregados no contexto da comunicação científica;
3. identificar o comportamento informacional dos usuários de periódicos eletrônicos da área da Ciência da Informação;
4. levantar as possibilidades de aplicação da tecnologia RSS como serviço de disseminação seletiva de informação científica.

## 1.5 Organização da Dissertação

A presente dissertação encontra-se organizada em 8 (oito) capítulos, os quais podem ser resumidos da seguinte maneira:

**Capítulo 1** – Introdução: apresenta o contexto no qual a pesquisa se insere, os problemas que motivaram o estudo, seus objetivos e alguns pressupostos teóricos para aproximação do leitor à pesquisa.

**Capítulo 2** – Conceitos Fundamentais: capítulo de revisão de literatura sobre os tópicos considerados relevantes como suporte teórico para o estudo e que possuem relação com o problema de pesquisa. São esses: sobrecarga da informação, sistemas de filtragem de informação, disseminação seletiva de informação e estudos sobre comportamento informacional dos usuários de periódicos científicos eletrônicos.

**Capítulo 3** – A Comunicação Científica: revisa os tópicos essenciais do tema, buscando referenciais teóricos para o desenvolvimento do estudo. Comenta as mudanças no processo de comunicação científica após a introdução das

tecnologias de informação, como a proliferação dos periódicos científicos eletrônicos. Resgata o conceito de Comunicação Extensiva e sua contribuição para o fenômeno estudado. Aborda as definições, características e funcionalidades dos periódicos científicos eletrônicos, culminando com um panorama sobre a filosofia do “acesso aberto” como novo modelo para a comunicação científica e a contribuição dos *blogs* científicos.

**Capítulo 4** – Procedimentos Metodológicos: caracteriza o tipo de pesquisa, o universo estudado e os critérios para a seleção da amostra. Detalha os procedimentos adotados para o levantamento de dados necessários ao embasamento da análise e conclusões finais.

**Capítulo 5** – Tecnologia RSS: aborda as principais características da tecnologia RSS. Resgata o histórico e a evolução dos formatos de sindicância de conteúdos Web. Descreve a arquitetura e principais elementos de um *feed* RSS. Informa sobre os principais recursos, modos de utilização e barreiras para utilização da tecnologia.

**Capítulo 6** – A Tecnologia RSS Aplicada à Comunicação Científica: trata dos vocabulários de metadados aplicados às especificações RSS e sua importância na representação da informação científica. Descreve e exemplifica usos da tecnologia RSS relacionados à disseminação de informação científica em aplicações de bibliotecas, bases de dados, repositórios digitais, entre outros.

**Capítulo 7** – Análise dos Resultados: apresenta os resultados obtidos durante a etapa de coleta de dados à luz do referencial teórico, da compilação de informações reunidas durante a fase de pesquisa bibliográfica e documental .

**Capítulo 8** – Conclusões e Recomendações: discutem-se os aspectos mais relevantes da pesquisa e apresenta sugestões para estudos futuros.

## **CAPÍTULO 2**

### **Conceitos Fundamentais**

---

## 2.1 Sobrecarga da Informação

Em artigo que trata da disponibilidade de informação para a aquisição e transmissão do conhecimento humano, Krzyzanowski e Taruhn (1998) citam o pesquisador Joseph Henry, do *Smithsonian Institute*, que, já no ano de 1851, teria manifestado seu receio com o aumento da produção bibliográfica, ao afirmar:

a humanidade tem seu progresso baseado em pesquisa, estudo e investigação, que geram saber, conhecimento ou, simplesmente, informação. Praticamente para cada item de interesse existe algum registro de saber pertinente. A não ser que essa massa de informações seja armazenada com ordem e que se especifiquem bem os meios em que nos irão expor os respectivos conteúdos, tanto a literatura como a ciência perecerão esmagadas sob seu próprio peso (HENRY, 1852<sup>5</sup> apud KRZYZANOWSKI; TARUHN, 1998, p. 193).

Quase um século mais tarde, o volume da literatura científica e técnica produzida após a Segunda Guerra Mundial chegaria a um patamar ainda mais preocupante. Nesse período, merece destaque a publicação do histórico artigo *We May Think*<sup>6</sup>, do físico e matemático Vannevar Bush. Em seu texto, Bush (1945) trata das possíveis barreiras na organização e o repasse à sociedade das informações em ciência e tecnologia mantidas em sigilo durante o período da Segunda Guerra.

Como solução para o problema da informação, o pesquisador propõe a criação de um dispositivo, o qual chamou de “Memex” (abreviação de *Memory Extension*), destinado ao armazenamento e recuperação de documentos mediante a associação de palavras. Em sua concepção, o equipamento deveria conseguir reproduzir os processos mentais humanos de associações de idéias por meio de uma espécie de “indexação associativa” capaz de manter as ligações entre os assuntos, os quais seriam apresentados sob forma de documentos escritos, sonoros ou visuais.

---

<sup>5</sup> Annual report of the board of regents of the Smithsonian Institution...during the year 1851. Washington: 1852

<sup>6</sup> Uma versão eletrônica desse documento encontra-se disponível na Internet em: <<http://www.theatlantic.com/doc/194507/bush>>

Pouco tempo depois, as preocupações decorrentes da chamada “explosão bibliográfica” foram levadas à plenária da Conferência sobre Informação Científica, da *Royal Society*, em 1948, dando origem um ano mais tarde, à criação do *Institute for Information Scientists*, que acolheu as idéias dos pesquisadores dessa nova área, fato que teria contribuindo para o surgimento dos periódicos de resumos. (RUSSO et al., 2001; BARRETO, 2002, p. 69).

Com os avanços tecnológicos vieram novas possibilidades para otimizar o acesso e a disseminação da informação, alterando o paradigma de informação bibliográfica baseada em suporte impresso. O acesso a novos recursos informacionais, por meio de rede computadores, tornou-se uma realidade.

É fato que a edição de publicações eletrônicas vem agilizando o acesso aos seus conteúdos, contribuindo para a produção de outros conhecimentos e, conseqüentemente, para o aumento exponencial do volume de informações, verificado notadamente na segunda metade do século XX. Neste período, o número de títulos de periódicos impressos passou de cerca de 10 mil para 1 milhão de títulos (KRZYZANOWSKI; TARUHN, 1998, p. 193).

Portanto, a sobrecarga de informação não é, nem de longe, uma preocupação recente. Porém, com a proliferação desordenada de conteúdos publicados na Internet, torna-se urgente a criação de mecanismos capazes de filtrar informações relevantes para um público que necessite dela, tal como já preconizavam os serviços de disseminação seletiva de informação, criados na década de 60.

No contexto da Web, tais serviços devem ser aperfeiçoados e/ou acrescidos de ferramentas capazes de minimizar os efeitos do excesso de informação para os usuários. Assim, podem atuar como um sistema de intermediação entre o amplo conteúdo disponibilizado, considerando a necessidade de informação imediata do usuário.

Um estudo realizado por Lyman e Varian (2003), da Universidade da Califórnia, em Berkley, fez um levantamento sobre a quantidade de informação produzida anualmente em todo o mundo. Segundo a pesquisa – iniciada no ano 2000 e atualizada em 2003 – cerca de 1,5 bilhões de gigabytes de informação são

registradas em todo o planeta nos mais diferentes suportes (impresso, magnético, ótico e filmes). Na época, esse número representava uma quantidade de 250 Mb por pessoa na Terra. Outro indicador revelado pela pesquisa aponta uma taxa de crescimento de 30% no volume de informação nova produzida.

Especificamente nas redes eletrônicas, o excesso de informação também pode ser evidenciado quando verificamos a profusão de milhares de potenciais novos sítios que surgem todos os dias a partir do pedido de registro de nomes de domínio. Somente no Brasil, onde o número de domínios ultrapassa 1,2 milhões<sup>7</sup>, a média de novos registros gira em torno de 800 por dia.

Uma das primeiras iniciativas para minimizar o efeito da sobrecarga de informação produzida pelas páginas da Web foram os diretórios, como o *Yahoo!* (o primeiro catálogo da Web) e o pioneiro nacional *Cadê?*, que serviam como referência para localização temática das páginas. A característica marcante desse tipo de serviço é que a inclusão das informações é feita segundo a classificação de profissionais especializados. Já os mecanismos de busca automáticos, tais como o *Altavista*, *Lycos* e *Google*, realizam uma espécie de indexação na Web via programas-robôs (agentes). Cabe a esses *softwares* percorrerem os textos de cada página publicada e delas extraem palavras-chave que são armazenadas em banco de dados, a partir dos quais o usuário realiza sua pesquisa.

Apesar da boa intenção desses serviços os resultados fornecidos inicialmente apresentavam baixa precisão devido ao processamento automático sem controle terminológico. Posteriormente, algumas páginas passaram a incluir metadados por meio do rótulo *<meta>* da linguagem de marcação HTML a fim de melhorar a qualidade na recuperação da informação pelos robôs dos mecanismos de busca. Essa solução, entretanto, não surte efeito para o conteúdo da chamada Web profunda (*deep Web*), formado pelas páginas dinâmicas geradas a partir de bancos de dados como resultado de uma requisição ou consulta específica realizada pelo usuário. A Web profunda é considerada até 2 mil vezes maior do que a Web superficial, que reúne o conjunto de documentos estáticos armazenados nos

---

<sup>7</sup> As estatísticas de registro de domínios no Brasil são disponibilizadas diariamente em: <http://www.registro.br/estatisticas.html>

servidores como arquivos com extensão .html ou .htm. Esta situação faz com que a informação registrada na Web necessite cada vez mais de metadados associados aos documentos eletrônicos que eles descrevem (MARCONDES, 2006, p. 98).

Para Carlson (2003), a sobrecarga da informação é tratada como um fenômeno que possui causas objetivas e subjetivas. Objetivamente, o pesquisador afirma que o crescimento exponencial da quantidade de informação nas últimas cinco décadas é um fato e, até o momento, não há indicadores que demonstrem que essa taxa não poderá aumentar ainda mais em um futuro próximo. Já o componente subjetivo do fenômeno está localizado na percepção de que essa sobrecarga advém de uma disponibilidade de informação maior do que a capacidade de sua assimilação.

Segundo Bax et al. (2004), a maior disponibilidade de recursos de informação demanda serviços de intermediação que permitam a coleta e filtragem automática das fontes e a distribuição seletiva desse conteúdo aos usuários de acordo com seu perfil de interesse. Este é o campo de pesquisa dos Sistemas de Recuperação e de Filtragem da Informação.

## 2.2 Sistemas de Filtragem de Informação

O termo “recuperação da informação” foi criado em 1951 pelo cientista americano Calvin Mooers e seu conceito engloba

os aspectos intelectuais da descrição da informação e sua especificação para a busca, bem como qualquer sistema, técnicas ou máquinas que são empregadas para realizar esta operação (MOOERS, 1951).

Dentre os problemas tratados por esse campo destacam-se a representação e organização da informação, a especificação da busca por informação e o da criação de mecanismos para a recuperação. Mais especificamente, os Sistemas de Recuperação da Informação (SRIs) devem responder: a) como descrever a informação? b) como especificar sua a busca? e c) que sistemas, técnicas ou máquinas devem ser empregados? (BAX et al., 2004).

A partir da década de 70, o tema começou a aparecer mais frequentemente na literatura com o incremento dos sistemas de recuperação *on-line*. Na obra *Information Retrieval On-line*, Lancaster e Fayen (1973, p. 1) definem um sistema de recuperação de informação deste tipo como aquele em que o usuário pode consultar uma base de dados de documentos (ou suas representações) por meio de um computador. Nesta época, os sistemas eram desenhados essencialmente para recuperação de dados bibliográficos via terminal de computador. Mais tarde, os sistemas de recuperação *on-line* passaram a ser utilizados pelas companhias aéreas para armazenar e recuperar informações sobre reservas de vôos.

Baeza-Yates e Ribeiro Neto (1999, p. 1) afirmam que os Sistemas de Recuperação de Informação lidam com representação, armazenamento, organização e acesso a documentos. A função da representação e da organização da informação é facilitar a recuperação por parte do usuário. Os autores frisam, no entanto, que existe uma diferença entre a recuperar dados e recuperar informação. O primeiro processo se limita a atender as condições onde haja coincidência entre os documentos e as palavras-chave de uma determinada consulta. Geralmente esse resultado não é suficiente para satisfazer às necessidades de informação do usuário. A recuperação de informação, portanto, consiste na capacidade de oferecer um grau de relevância que atenda às solicitações do usuário. Assim, a principal função de um SRI deveria ser “recuperar todos os documentos que são relevantes a uma solicitação do usuário com uma quantidade mínima de documentos não-relevantes” (BAEZA-YATES; RIBEIRO NETO, 1999, p. 1) ou, ainda, segundo Belkin (1992, p. 29), “conduzir os usuários aos documentos originais para melhor satisfazerem suas necessidades”.

Na vasta literatura sobre o tópico, identificamos três modelos clássicos para o processo de recuperação de informação: booleano, vetorial e probabilístico. O primeiro considera uma consulta como uma expressão cujos termos são interligados por conectores lógicos (AND, OR, NOT). No modelo vetorial cada documento é representado por um vetor de termos e cada termo possui um valor associado que indica o grau de importância deste no documento. Finalmente, a técnica probabilística faz com que os documentos sejam ordenados conforme a probabilidade de relevância de uma consulta.

Todos esses modelos descritos anteriormente apresentam estratégias de busca para localização de documentos relevantes. No entanto, se desejarmos um maior refinamento na recuperação dos conteúdos, necessitaremos de um processo adicional de filtragem.

O termo “filtragem da informação”, aliás, foi cunhado a partir dessa concepção. A expressão surgiu a partir de uma carta de Peter Denning dirigida ao presidente da *Association for Computing Machinery* (ACM) e publicada por *Communications of the ACM* em março de 1982. Ao falar sobre as novas transações da ACM em sistemas de informação para escritório, Denning tratou da necessidade da criação de um “filtro de conteúdo” para selecionar a informação que chegava por meio de correio eletrônico de modo a separar mensagens urgentes das rotineiras.

A pesquisa sobre tecnologias de filtragem de informação ganhou força nos anos seguintes à comunicação de Denning. Em 1989, a Agência de Projetos de Pesquisa Avançada de Defesa dos Estados Unidos (*United States Defense Advanced Research Projects Agency*, DARPA) patrocinou a primeira de uma série de Conferências de Entendimento de Mensagens (*Message Understanding Conferences*) sobre técnicas de extração de informação para suportar a seleção de mensagens.

Outra iniciativa foi capitaneada em novembro de 1991 pela Bellcore e pelo Grupo de Interesse Especial em Sistemas de Informação para Escritório da ACM que, juntas, patrocinaram um *Workshop* em “Filtragem de Informação de Alto desempenho”. Desde então, a literatura tem mostrado vários avanços nas pesquisas sobre filtragem de informação a partir do desenvolvimento de sistemas, envolvendo a interseção de duas áreas de pesquisa historicamente distintas: recuperação e filtragem de informação e agentes de *software*.

Atualmente, o termo “filtragem da informação” é utilizado genericamente para descrever uma variedade de processos envolvendo entrega da informação para pessoas que precisem dela (BELKIN, 1999, p. 29). Os filtros, portanto, atuam como intermediários entre as fontes de informação e seus usuários finais, sendo capazes de gerenciar a distribuição de informação. Um sistema de filtragem deve levar em

conta as tarefas de coleta, seleção e exibição da informação selecionada, podendo ser relevante ou não para o usuário.

Belkin e Croft (1992, p. 30) relacionam as seguintes características dos sistemas de filtragem de informação:

- é projetado para dados não estruturados ou semi-estruturados, podendo incluir outros tipos de dados como imagens, áudio e vídeo (multimídia);
- envolvem grandes quantidades de dados. Suas aplicações lidam com grandes volumes de textos disponíveis em fontes diversas;
- envolvem tipicamente fluxos de dados de entrada, ou senso transmitidos por fontes remotas ou enviados diretamente por meio de outras fontes (correio eletrônico);
- é baseada em descrições de preferências de informação de indivíduos ou grupos (perfis), que representam os interesses dos usuários.

Dependendo da maneira com que os documentos são selecionados para o usuário, os sistemas de filtragem de informação podem ser classificados como sistemas cognitivos (baseados em conteúdo), sistemas sociais (colaborativos) e econômicos (SHETH, 1994, p. 13).

Os sistemas cognitivos escolhem os documentos baseando-se em perfis de seus conteúdos, enquanto sistemas sociais selecionam documentos com base em recomendações e anotações de outros usuários. Já os sistemas econômicos selecionam os documentos a partir de mecanismos que consideram o custo-benefício para o usuário. A escolha da melhor abordagem vai depender da área da aplicação.

Na avaliação de Sheth (1994, p. 8-9), um sistema de filtragem de informação deve reunir as seguintes propriedades: especialização, adaptação e exploração. Para o autor, o principal desafio para um sistema de filtragem reside na sua capacidade de prover informações compatíveis com as necessidades específicas dos usuários de modo oportuno (especialização), podendo acomodar eventuais

mudanças nas necessidades desses usuários (adaptação) e, ainda, um sistema de filtragem ideal deveria ser capaz não só de trabalhar com as necessidades conhecidas, mas também explorar novos domínios com o objetivo de encontrar informação de potencial interesse para o usuário (exploração).

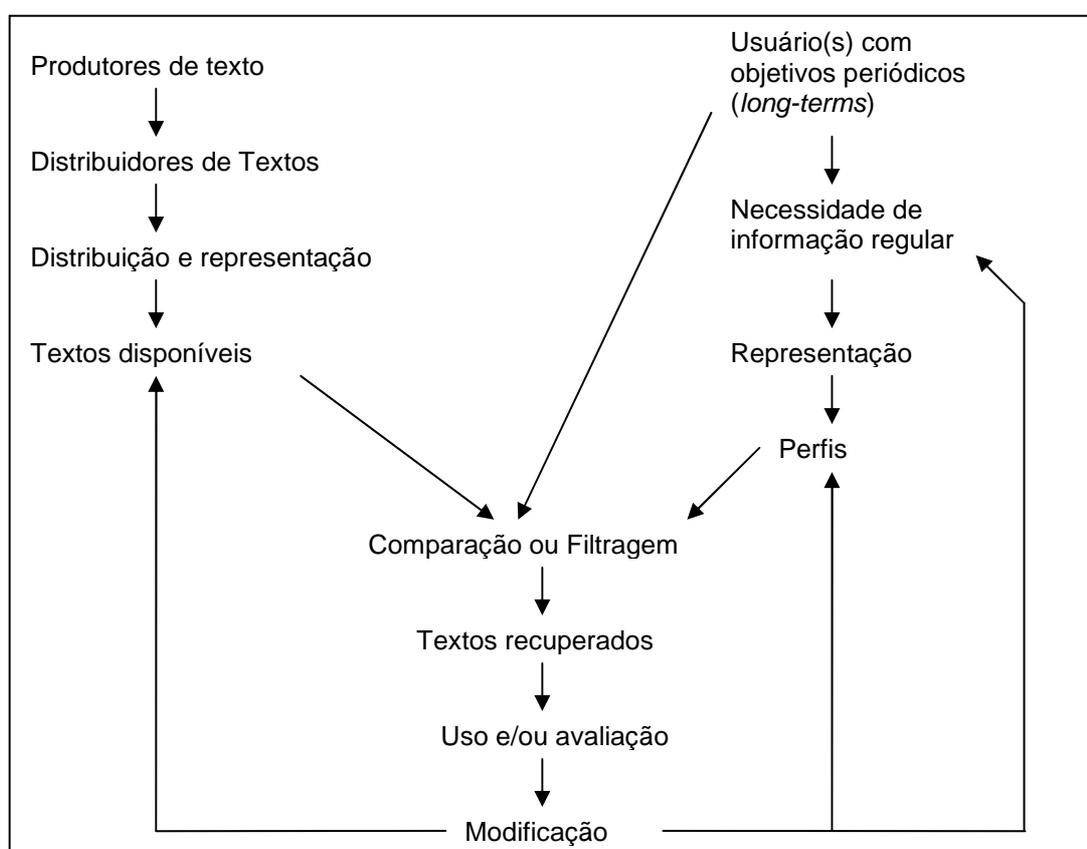
Apesar dos sistemas convencionais de recuperação e de filtragem possuírem relacionamentos estreitos quanto ao seu objetivo de selecionar informação relevante de acordo com a necessidade do usuário, os sistemas de filtragem geralmente dão uma ênfase maior na representação das necessidades do usuário, isto é, na definição do perfil do usuário, o que difere dos sistemas de recuperação de informação. Além disso, um problema que parece alcançar maior destaque em filtragem do que na recuperação é a representação de dados não textuais.

Belkin e Croft (1992) identificaram diferenças primárias entre ambos (Quadro 2.1). A partir desta comparação, concluem que a filtragem e recuperação são “dois lados da mesma moeda. Eles trabalham juntos para ajudar pessoas a obter informações necessárias para executarem suas tarefas” (BELKIN; CROFT, 1992, p. 37).

<b>Recuperação de Informação (RI)</b>	<b>Filtragem de Informação (FI)</b>
uso singular do sistema (pessoa com uma consulta por vez)	uso repetido do sistema (pessoas com interesses / objetivos periódicos)
consultas representam necessidade de informação nem sempre correta	perfis são especificações corretas do interesse de informação
consultas sobrevivem apenas até a construção da lista de documentos que a satisfazem	perfis são mantidos por um longo tempo e podem ser modificados
interesse na coleção e organização de textos	interesse na distribuição de textos para grupos ou indivíduos
seleção de textos de um banco de dados relativamente estático	seleção ou eliminação de textos de um fluxo de dados dinâmico ( <i>datastream</i> )
responde a interação do usuário com textos, num único episódio de busca de informação	mudanças nos “ <i>long-term</i> ” sobre uma série de episódios de busca de informação
encontram informações relevantes com todos os documentos classificados por relevância	removem informações irrelevantes decide se o documento é ou não relevante

**Quadro 2.1** – Diferenças fundamentais entre os sistemas de RI e FI, segundo Belkin e Croft (1992)

Ainda em relação à atividade de localização de informação pelo usuário, Saracevic et al. (1997) mostram que existem diversas abordagens relacionadas ao problema da especificação da busca por informação. Tais abordagens podem estar centradas no sistema, no usuário ou em ambos, como ocorre com os sistemas de disseminação seletiva de informação. As funções e entidades envolvidas no processo de filtragem de informação têm sido descritas por meio de modelos, como o de Belkin e Croft, ilustrado pela Figura 2.1.



**Figura 2.1** – Modelo genérico de um Sistema de Filtragem de Informação, segundo Belkin e Croft (1992 p. 34)

Pode-se dizer que uma das funções dos sistemas de filtragem de informação é a seleção de conteúdos que sejam realmente adequados à necessidade expressa do usuário ou de um determinado grupo. Essa é uma condição fundamental, a qual coincide com os objetivos dos chamados serviços de Disseminação Seletiva de Informação, a ser detalhado no próximo tópico.

### 2.3 Disseminação Seletiva de Informação

O conceito de Disseminação Seletiva de Informação ou *Selective Dissemination of Information* (SDI) foi criado na década de 50 por Hans Peter Luhn, da IBM Corporation, com a finalidade de aperfeiçoar os serviços de alerta oferecidos pelas bibliotecas, centros de documentação e sistemas especializados de informação documental. Em 1958, Luhn propôs a idéia de um Sistema Inteligente de Negócios (*Business Intelligence System*), incentivando os bibliotecários a criarem perfis individuais para os seus usuários, de forma a produzir listas de documentos que atendessem às suas necessidades.

Luhn (1961, p. 132) define o processo de Disseminação Seletiva de Informação (DSI) como “aquele serviço dentro de uma organização que se refere à canalização de novos itens de informação, vindos de quaisquer fontes, para aqueles pontos onde a probabilidade de utilização, em conexão com o interesse corrente [do usuário], seja alta”.

Como serviço de referência, o objetivo mais amplo de um DSI é reunir a literatura corrente e anunciá-la seletivamente para uma comunidade de usuários. Conceitualmente, um típico serviço de disseminação seletiva “é aquele que fornece ao usuário uma lista de referências bibliográficas em intervalos regulares, relacionada com sua área de interesse” (NOCETTI, 1980, p. 15).

Para Longo (1978, p.102-103), o processo de disseminação seletiva reúne uma série de fatores que contribui para que o usuário gaste menos tempo com o exame e a seleção da literatura corrente. Tais processos envolvem: 1) coleta da informação produzida; 2) indexação dessa informação; 3) divulgação da informação aos usuários; 4) acessibilidade da informação. Essas etapas dependem da eficiência das pessoas e serviços que estão envolvidos na execução dos mesmos (LONGO, 1978, p. 102).

Em artigo de revisão publicado no ARIST<sup>8</sup>, Housman (1973) comenta diversos modos de aplicação do conceito de DSI à época. Para o autor, um típico serviço de disseminação seletiva atuaria como a automação de uma função clássica da biblioteca tradicional, consistindo em informar o usuário sobre as novas aquisições, compatíveis com seus hábitos e interesses de consulta. Antes, existiam apenas os serviços manuais de alerta por meio do manuseio de títulos de periódicos e da compilação de resumos para as referências que fossem consideradas relevantes, distribuindo-as para os usuários. Esta modalidade era utilizada principalmente em bibliotecas especializadas, que geralmente apresentam um número reduzido de usuários.

Como um serviço voltado para a filtragem de informação, um dos aspectos mais importantes em relação à configuração de um DSI é a construção do “perfil de interesse” a partir de uma base de conhecimento específico sobre a necessidade do usuário, característica que o distingue dos demais serviços de alerta. Segundo Nocetti (1980, p. 24), um perfil pode ser entendido como “o conjunto de indicadores que caracterizam as necessidades informacionais dos usuários”. Outro ponto importante é que esse perfil deve ser prontamente modificado tão logo se perceba uma mudança na necessidade de informação do usuário.

Longo (1979, p. 85) afirma que

a melhor forma de se construir um perfil é através de uma entrevista pessoal com o usuário, na qual é feita uma narração por escrito de seu campo de atuação onde também são submetidas palavras-chave e referências que melhor definam seu interesse específico.

Sousa e Brigheti (1981, p. 28) atribuem às bibliotecas a responsabilidade de prover sua comunidade de serviços de recuperação e localização da informação compatíveis com a necessidade de sua clientela. Disseminar seletivamente a informação é uma das formas de disponibilizá-la ao usuário, uma vez que se trata de uma extensão dos serviços convencionais de referência.

Dentre as vantagens desse tipo de serviço, Longo (1978, p. 103) destaca a redução considerável do tempo gasto pelos usuários durante o exame e seleção de

---

<sup>8</sup> Annual Review of Information Science and Technology

literatura corrente e o maior uso da coleção das unidades de informação pelos usuários. Tais fatores induzem uma mudança no comportamento do profissional da informação, que passa a ter uma postura mais ativa em relação ao usuário.

Mesmo com tantas facilidades, esse tipo de sistema também apresenta alguns problemas, como a dificuldade na padronização da produção de base de dados, o elevado custo de operacionalização do serviço e o baixo percentual de usuários que retroalimentam o serviço (LONGO, 1978, p. 110-111). Na opinião de alguns autores o DSI seria considerado apenas um meio suplementar à disseminação da informação já que não é capaz de satisfazer todos os requisitos informacionais dos usuários (LONGO, 1979, p. 81).

Em alguns países os sistemas de DSI tiveram início na década de 60 e expandiram-se nos anos posteriores, tornando serviços correntes de disseminação nos grandes sistemas de recuperação de informações referenciais. No Brasil, a literatura registra algumas iniciativas de destaque, entre elas a do Sistema Orientado para Notificação Automática de Referências (Sonar), criado pela Comissão Nacional de Energia Nuclear/Centro de Informações Nucleares (CIN/CNEN), em 1972, e o serviço automatizado da Embrapa – Empresa Brasileira de Pesquisa Agropecuária, no final da década de 70 (BAX et al., 2004).

Com as mudanças introduzidas pela tecnologia, os serviços de disseminação seletiva foram gradativamente incorporados aos formatos eletrônicos. Com o surgimento das bases de dados em CD-ROM, a partir da década de 80, as pesquisas bibliográficas ganharam agilidade e aumentaram as chances de se obter resultados mais satisfatórios. Atualmente, os serviços de DSI continuam tendo um alto índice de aceitação entre os pesquisadores apesar das facilidades oferecidas pelas diversas opções de ferramentas de busca disponíveis gratuitamente na Internet. Isso porque o grau de relevância da informação recuperada por meio de um DSI é bastante superior às encontradas nas pesquisas bibliográficas realizadas pelo próprio usuário, sem contar com o auxílio de sistemas de filtragem ou a intermediação de um profissional habilitado.

A idéia geral de um DSI mostra-se atual nos serviços de personalização promovidos por catálogos *on-line* e bases de dados bibliográficas, que permitem o armazenamento dos parâmetros de uma busca para serem reutilizados posteriormente. É importante ressaltar que os mecanismos de personalização na Web geralmente envolvem o desenvolvimento de algoritmos complexos e exigem grande capacidade de processamento, por isso são encontrados mais freqüentemente nas aplicações comerciais. Entretanto, já existem relatos de pesquisas sobre serviços de personalização no contexto acadêmico, como Moura Speroni et al. (2006) sobre a Biblioteca Digital de Literatura Brasileira (BD-LB), que oferece um serviço de recuperação de informação e distribuição seletiva adaptáveis ao perfil dos usuários. Cardoso e Oliveira (2000), por sua vez, propõem a i-LIB, um protótipo de interface de biblioteca digital configurável de acordo com o perfil do usuário.

Apesar das possibilidades de criação de um serviço de DSI para a Web, identificamos algumas barreiras para a operacionalização de um modelo tradicional de disseminação seletiva. O primeiro problema reside na construção do perfil do usuário de modo adequado devido à dificuldade em selecionar fontes de informação relevantes. Essa tarefa é agravada pelo aumento desordenado da produção bibliográfica na Internet, além do tempo cada vez mais exíguo para aguardarmos pelo resultado de uma pesquisa mais sofisticada realizada por terceiros.

A tecnologia RSS, explorada no decorrer do presente trabalho, se aproxima do conceito de Disseminação Seletiva de Informação na medida em que propicia a os seus usuários a condição de se manterem constantemente informados de acordo com suas preferências e tópicos de interesse, com a diferença no fato de que é o próprio usuário quem define os filtros que irão selecionar o conteúdo desejado, além da periodicidade e o modo pelo qual irá receber as notificações (alerta sonoro, janela *pop-up*, etc.).

Assim sendo, a construção do “perfil” é diferenciada em relação aos serviços tradicionais de DSI, os quais geralmente são formados com o auxílio de um profissional de informação que atua como intermediário. O “ajuste fino” desejado para o recebimento de informações de modo mais dirigido possível pode ser obtido

por meio de recursos tecnológicos, tais como a geração automática de canais RSS a partir de resultados de buscas em bases de dados já integradas a essa tecnologia. Tais aplicações serão detalhadas no Capítulo 6 – A Tecnologia RSS Aplicada à Comunicação Científica.

## 2.4 Comportamento Informacional em Periódicos Científicos Eletrônicos

A fim de diagnosticar as oportunidades para a adoção da tecnologia proposta, percebeu-se a necessidade de realizar um breve apanhado da literatura pertinente a este tópico.

O comportamento de busca e uso da informação em si é considerado um tópico de pesquisa consagrado nas áreas de Biblioteconomia e Ciência da Informação. No artigo seminal *On user studies and information needs*, Wilson (1981) afirmava à época que, além da recuperação da informação, o campo de “estudo de usuários” era um dos mais pesquisados dentro da Ciência da Informação.

De fato, a literatura indica que os estudos sobre as necessidades de informação dos usuários podem ser verificados desde 1948, ano da Conferência sobre Informação Científica, promovida pela *Royal Society*. Embora não houvesse explicitamente o emprego dessa expressão, vários trabalhos da época foram produzidos com o objetivo de identificar o comportamento informacional de cientistas e tecnólogos. Muitos desses estudos estão mais relacionados à forma como as pessoas usam os sistemas e serviços do que os aspectos de seu comportamento genérico de busca, intercâmbio, uso e transferência da informação (WILSON, 1994).

Conceitualmente, o processo de busca de informação pode ser definido como uma sucessão de etapas pela qual o indivíduo passa desde que sente uma necessidade (GONZÁLEZ TERUEL, 2005). Já o uso é determinado pelos atos físicos e mentais envolvidos na incorporação da informação às estruturas de conhecimento do indivíduo (WILSON, 2000).

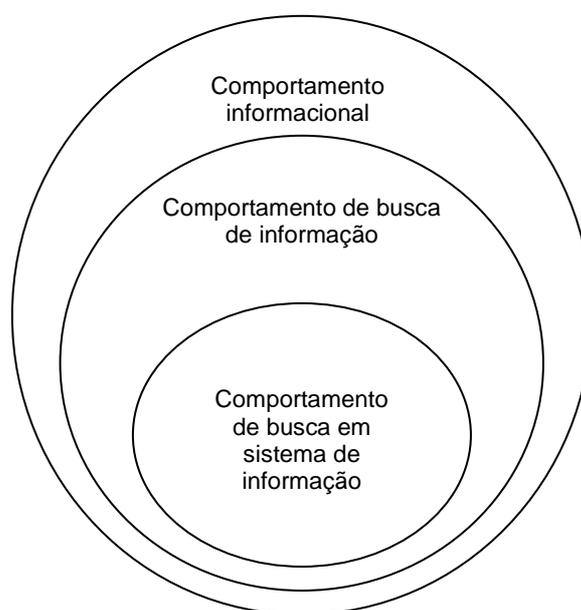
Para Le Coadic (2004, p. 38), “o conhecimento da necessidade de informação permite compreender por que as pessoas se envolvem num processo de busca da informação”. Mais que isso, é possível inferir como as pessoas se comportam no sentido de satisfazer suas necessidades informacionais. Em artigo de revisão, Bettiol (1990) conclui que a necessidade de informação varia de grau de intensidade de uma definição para outra. Sendo independente na sua essência e nas diversas situações que se apresenta a uma pessoa.

Diversos modelos de comportamento informacional foram elaborados levando em consideração os vários aspectos que podem influenciar – e até mesmo determinar – o comportamento informacional. O proposto por David Ellis, por exemplo, é centrado em aspectos cognitivos da busca da informação, compreendendo as características complexas da ação e interação nas quais as pessoas se engajam quando procuram por informações de qualquer tipo e para qualquer problema (CRESPO; CAREGNATO, 2003, p. 249). Já o modelo de Kuhlthau, denominado ISP (*Information Search Process*), foi baseado na teoria da construção pessoal de George Kelly, com foco no estudo das experiências vivenciadas pelos indivíduos, entre elas o desenvolvimento da afetividade, no processo de construção de sentido das informações absorvidas pelos usuários (CRESPO; CAREGNATO, 2003, p. 250).

Outros autores como Brenda Dervin, Nicholas Belkin e Robert Taylor também estudaram os processos de busca e uso da informação sob outras perspectivas. No entanto, todas elas convergem para a conclusão de que os pensamentos e sentimentos dos usuários são os locais adequados para a construção da informação, que ficará disponível tanto na vida do indivíduo quanto em seu ambiente de trabalho. É importante ressaltar que, após a década de 80, as pesquisas sobre o tema começam a focar o usuário como centro do processo de estudo, diferentemente dos trabalhos publicados antes desse período, quando o enfoque dado ao comportamento informacional era sobre o sistema.

A necessidade em definir os limites e aproximações das várias abordagens, representadas por modelos conceituais distintos, levou Wilson (1999) a propor um esquema para as áreas de pesquisa sobre comportamento informacional. Na visão

do autor, “comportamento informacional” (*information behavior*) pode ser compreendido como o campo mais geral de investigação, englobando o subcampo comportamento de busca de informação (*information-seeking behavior*), o qual, por sua vez, contém as atividades que contemplam o “comportamento de busca em sistemas de informação” (*information search behavior*), conforme representado na Figura 2.2.



Fonte: Wilson (1999, p. 263)

**Figura 2.2** – Busca de informação e busca em sistemas de informação

Wilson (1999, p. 249) define o termo comportamento informacional como “as atividades de busca, uso e transferência da informação, nas quais uma pessoa se engaja quando identifica as próprias necessidades de informação”. Assim, o comportamento informacional é fruto da necessidade do indivíduo em resolver um problema ou suprir um estado atípico de conhecimento, insuficiente ou inadequado. Para o autor, o “comportamento de busca de informação” refere-se à variedade de métodos que as pessoas utilizam para descobrir e ter acesso às fontes de informação. O subcampo “comportamento de busca em sistema de informação” está

---

relacionado com as interações entre o usuário e um sistema de informação baseado em computador (WILSON, 1999, p. 263).

Por estarmos investigando a viabilidade de uma nova ferramenta aplicada ao uso de periódicos científicos eletrônicos, procuramos localizar, na literatura, estudos que versassem a respeito do comportamento informacional dos usuários em relação a este tipo de veículo. Destacamos os trabalhos de Gomes (1999), cuja pesquisa analisou o processo de apropriação dos periódicos eletrônicos pelos acadêmicos vinculados aos cursos de pós-graduação *strictu sensu* no Brasil; Fachin (2002), que estudou os periódicos científicos eletrônicos visando à construção de um modelo de avaliação a partir de indicadores bibliográficos e telemáticos; Dias (2003) e sua tese sobre as dinâmicas de acesso e uso dos periódicos científicos brasileiros na área da Ciência da Informação a partir da análise dos arquivos de *logs* de acesso; Souza e Albuquerque (2005), que investigaram a utilização de periódicos científicos da mesma área pelos alunos do curso de Biblioteconomia da Universidade Federal da Paraíba (UFPB); Araújo et al. (2006), que coletaram dados entre os docentes dos programas de pós-graduação credenciados pela CAPES para analisar o impacto provocado sobre esses usuários pelo acesso e uso do periódico *Informação & Sociedade: estudos*, da UFPB, e Oliveira (2006), que estudou o comportamento de busca e uso de periódicos eletrônicos por docentes e alunos de pós graduação da área de geociências.

Mesmo com a importância desse tipo de estudo para a comunicação científica, percebe-se que o tema vem sendo pouco tratado pelos autores brasileiros. Diante deste fato, pretende-se que esta dissertação colabore, também, como estudo de avaliação das características do comportamento de informação de usuários de periódicos eletrônicos.

## **CAPÍTULO 3**

### **A Comunicação Científica**

---

Um dos propósitos das sociedades científicas, organizadas a partir do século XVII, é o de comunicar a ciência, compartilhando entre os membros dessas comunidades os resultados e métodos empregados nos experimentos realizados. Inicialmente, este tipo de comunicação era realizado por meio de correspondências pessoais enviadas pelos cientistas aos seus colegas com o intuito de relatar idéias e descobertas mais recentes. Naquela época o conhecimento técnico e científico também era comunicado nas atas ou memórias produzidas após as reuniões das sociedades, que, depois de impressas, serviam como fonte de consulta e referência para seus membros. Posteriormente, surgiram os primeiros periódicos científicos, considerados até hoje como um dos principais veículos utilizados tanto para divulgação do conhecimento como para comunicação entre os membros de uma comunidade científica.

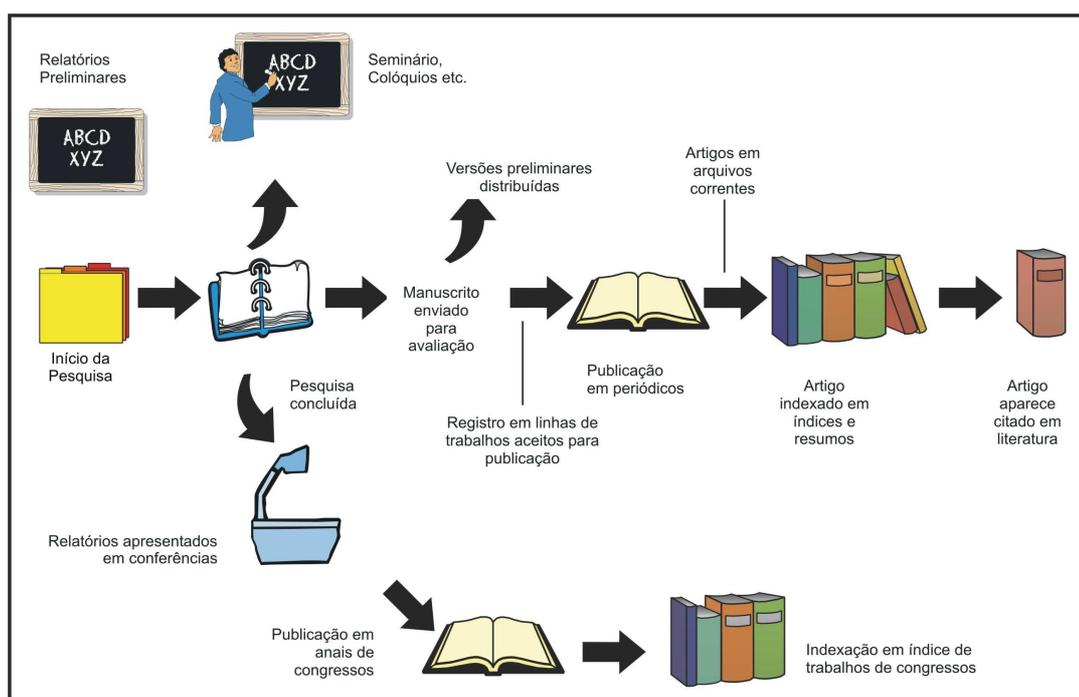
Três séculos se passaram e o modelo de desenvolvimento da ciência permanece praticamente inalterado, ou seja, totalmente dependente da comunicação científica. Afinal, a maturidade do conhecimento científico só pode se efetivar por meio da divulgação dos resultados das pesquisas e do processo de revisão pelos pares, considerado o selo de garantia de qualidade do conhecimento científico. O fluxo da comunicação científica só é concluído quando existe a disseminação desse conhecimento para que haja seu uso efetivo.

Para Ziman (1979, p. 116), tão importante quanto o embrião que deu origem a idéia da pesquisa é a forma como a literatura sobre um determinado tema é apresentada à comunidade científica, os primeiros resultados, e mesmo as críticas e as citações de outros autores. A idéia da comunicação científica, portanto, representa a essência do conhecimento científico, sendo “tão vital quanto a própria pesquisa” (MEADOWS, 1999, p. vii). É natural, pois, que a comunicação científica se configure como um tópico bastante explorado e discutido no âmbito da Ciência da Informação.

Segundo a clássica definição de Garvey (1979, p. ix), a comunicação científica:

inclui todo o espectro das atividades associadas à produção, disseminação e uso da informação, desde o momento em que o cientista concebe a idéia para sua pesquisa até que a informação sobre os resultados dessa pesquisa seja aceita como constituinte do conhecimento científico.

Após uma ampla pesquisa realizada sobre o processo de comunicação científica na área de Psicologia, Garvey e Griffith (1979) apresentaram um modelo que viria a se tornar referência para demais estudos sobre o assunto em diferentes áreas do conhecimento, devido à sua característica de adaptação com relação a fatores novos que venham influenciar no processo de comunicação científica. O modelo de Garvey e Griffith (Figura 3.1) representa os fluxos da comunicação científica, desde as fases iniciais da elaboração de um projeto de pesquisa até a utilização de seus resultados, após o reconhecimento do conhecimento científico.



Fonte: Hurd (2000, p. 1.280)

**Figura 3.1** – Modelo de comunicação científica, segundo Garvey e Griffith (1979)

Dentre as atividades associadas à disseminação de informações científicas – foco de interesse do presente estudo – destacam-se aquelas que conferem à projeção necessária para visibilidade, tanto do produto quanto dos produtores, no meio social no qual estão inseridos (MIRANDA; PEREIRA, 1996, p. 375). A função

da disseminação do conhecimento científico é fundamental para a sua preservação e para a evolução da ciência em si, já que os cientistas, uma vez informados, serão capazes de desenvolver novas pesquisas que corroboram ou refutam resultados de ensaios anteriores. Além disso, a confiabilidade das informações poderá ser aferida a partir da crítica e das citações de outros autores.

A comunicação do conhecimento científico obedece a padrões consensuais da comunidade científica, que prima por critérios de qualidade, confiabilidade e credibilidade do que é divulgado. O sistema de comunicação científica, por sua vez, geralmente é composto por canais informais e formais. Os canais informais são constituídos, basicamente, pelos contatos pessoais (conversas entre pesquisadores, telefonemas e divulgação dos primeiros resultados em eventos científicos). Já os canais formais são estabelecidos pela literatura científica primária (periódicos, relatórios técnicos, etc.), secundária (resumos, índices, etc.) e terciária (tratados, livros-texto, etc.).

Tradicionalmente, o processo de transferência da informação nos canais informais é mais ágil, seletivo e interativo. Já nos formais, o mesmo processo se dá de forma mais lenta, já que a informação precisa ser avaliada e comprovada antes de ser considerada legítima e, portanto passível de ser divulgada para se tornar pública. Porém, sua principal vantagem é a permanência da publicação, o que facilita sua recuperação e localização. É, pois, através dos canais formais, que a comunicação científica tradicional se estabelece, servindo de fonte para definição da produtividade e reconhecimento na comunidade científica.

É consenso que o periódico científico desempenha um papel de fundamental importância no processo de comunicação científica, constituindo-se uma das principais fontes para organização e veiculação de novos conhecimentos. Por esse motivo, o periódico é considerado por muitos como sendo o “arquivo oficial da comunidade científica”, já que a avaliação e crítica prévias feitas por editores e bancas de especialistas aos textos submetidos à publicação conferem ao periódico uma base mais sólida para a ciência (DIAS, 1999). Além disso, segundo Barbalho (2005, p. 126), “trata-se de um veículo de divulgação que reúne diversas autorias, é publicado em intervalos determinados de tempo e apresenta, de forma condensada,

os conhecimentos recém-gerados, dando-lhes visibilidade no meio acadêmico e científico”.

### 3.1 Comunicação Científica e as Tecnologias de Informação

A rápida evolução das tecnologias de informação e comunicação, em especial a Internet, tem contribuído para a modificação do processo de comunicação científica (MUELLER, 1994). A passagem do meio impresso para o meio eletrônico no sistema de comunicação entre pesquisadores foi analisada por Hurd, em 1996, tendo como base o modelo de Garvey e Griffith. Ao estudar a comunicação entre pesquisadores da área de ciências naturais, a autora propôs um modelo inteiramente baseado no meio eletrônico para o sistema de comunicação científica (COSTA, 2006, p. 167).

Hoje, parece ser consenso afirmar que existe uma categoria estabelecida para a comunicação científica que vai além dos canais formais e informais: a comunicação científica eletrônica, a qual é definida por Targino (2000) como “a transmissão de informações científicas através de meios eletrônicos”. A perspectiva da utilização da rede Internet como canal de comunicação, contudo, é objeto de atenção dos pesquisadores da Ciência da Informação há mais de uma década:

é certo que a Internet, sendo uma rede de redes com serviços de correio eletrônico e discussão temática em grupo, facilita a comunicação informal entre os pesquisadores nas diversas áreas da Ciência & Tecnologia. E quando se pensa que os canais informais são fundamentais para se trabalhar em níveis de maior valor agregado à informação (...) pode-se ter idéias do valor comunicacional da rede. Por outro lado, bases de dados e documentos publicados estão disponíveis através de serviços de compartilhamento ou de transferência de arquivos, caracterizando-se como canais formais de comunicação da informação (ARAÚJO; FREIRE, 1996, p. 53).

Essencialmente, a comunicação científica realizada por meios eletrônicos é considerada híbrida, já que possui características tanto de um canal informal quanto de um formal. Informalmente, a comunicação eletrônica agiliza o contato entre os pesquisadores, favorecendo a troca rápida de informações e o *feedback* imediato ao desenvolvimento de seus trabalhos. Os contatos “pessoais” têm sido cada vez mais intermediados pelo uso intensivo de correio eletrônico (e-mail), grupos de discussão

e, mais recentemente, com a utilização de novos dispositivos tecnológicos pelos pesquisadores, como as conversas em tempo real (*chats*) e os diários eletrônicos (*blogs*). A contribuição dos meios eletrônicos na comunicação formal, por outro lado, pode ser facilmente identificada pela praticidade e rapidez na divulgação do conhecimento produzido para um público bem mais amplo.

Ainda hoje a utilização da Internet como meio de comunicação científica ocorre com mais frequência nos estágios informais do processo de pesquisa, onde há envolvimento de discussões e troca de informações com os colegas. A própria estrutura de rede favorece o compartilhamento de informação e a interatividade entre a comunidade científica, como interligação de pessoas localizadas em diferentes regiões geográficas interagindo em tempo real. Nos meios formais, por sua vez, tem causado impacto relevante o número crescente de periódicos científicos publicados em formato eletrônico (OLIVEIRA; NORONHA, 2005).

Silva, Menezes e Bissani (2002, p. 3) consideram que a dificuldade em estabelecer os limites entre comunicação formal e informal no ambiente digital se deve à característica de retro alimentação praticamente instantânea entre usuários e produtores da informação que é proporcionada pelos meios eletrônicos, alterando o fluxo de transferência da informação. Esse fenômeno foi percebido por Barreto (1998, p. 125-126), quando estabeleceu alguns pontos que atestam a mudança na estrutura do fluxo da informação e do conhecimento provocada pela comunicação eletrônica:

- a) a interação do receptor com a informação: o receptor deixa a sua posição de distanciamento e passa a participar do fluxo da comunicação, como se estivesse posicionando em seu interior. Sua interação com a informação é direta, conversacional e sem intermediários;
- b) tempo de interação: o receptor conectado desenha a sua própria interação com o fluxo de informação em tempo real. Essa velocidade o coloca em uma nova dimensão para o julgamento de valor e da relevância da informação;
- c) estrutura da mensagem: o receptor pode elaborar a informação por meio de diversas linguagens (texto, imagem, som) para elaborar a informação; não

está mais preso a estrutura linear da informação, podendo criar seu próprio documento conforme sua decisão;

d) facilidade de ir e vir: a conexão em rede amplia a percepção da dimensão do espaço da comunicação, e o receptor pode acessar diferentes estoques de informação no momento que desejar.

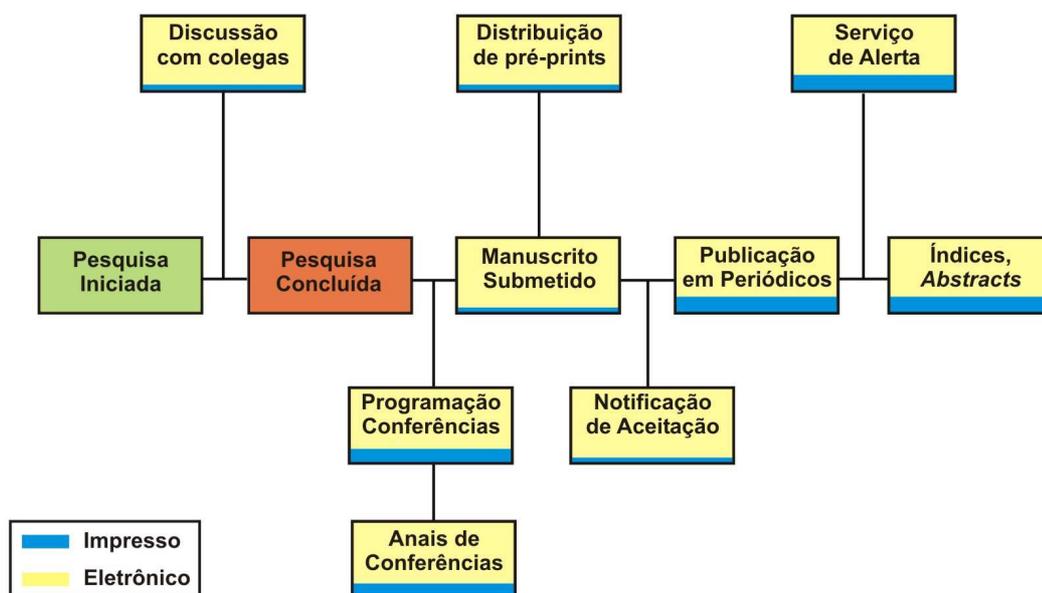
Targino (2000) relaciona as características básicas dos canais comunicação formal, informal e eletrônica. A autora, no entanto, faz uma ressalva de que a comunicação eletrônica possui uma inclinação para os aspectos dos sistemas informais de comunicação (Quadro 3.2).

<b>Canais formais</b>	<b>Canais informais</b>	<b>Canais eletrônicos</b>
público potencialmente grande	público restrito	público potencialmente grande
informação armazenada e recuperável	informação armazenada e não recuperável	armazenamento e recuperação complexos
informação relativamente antiga	informação recente	informação recente
direção do fluxo selecionada pelo usuário	direção do fluxo selecionado pelo produtor	direção do fluxo selecionada pelo usuário
redundância moderada	redundância, às vezes significativa	redundância, às vezes significativa
avaliação prévia	sem avaliação prévia	sem avaliação prévia, em geral
<i>feedback</i> irrisório para o autor	<i>feedback</i> significativo para o autor	<i>feedback</i> significativo para o autor

**Quadro 3.2** – Características dos canais de comunicação formal, informal e eletrônica (TARGINO, 2000)

Na avaliação de Oliveira (2006), o uso de tecnologias na comunicação formal parece estar ainda em desenvolvimento. Atualmente, convivemos com uma situação em que encontramos desde as publicações eletrônicas disponibilizadas integralmente em formato digital (como os anais de eventos científicos) às publicações em estágio híbrido, em que coexistem os títulos de revistas eletrônicas com o seu equivalente impresso.

Em estudo realizado com pesquisadores brasileiros, Costa, Silva e Costa (2001, p. 67) concluem que a “adoção da comunicação formal puramente eletrônica entre pesquisadores parece estar longe de se tornar ubíqua, como se pode observar em relação à comunicação informal”. Considerando as peculiaridades do meio eletrônico, Costa (1999) readaptou o modelo de Hurd (1996) – por sua vez foi baseado em Garvey e Griffith – e desenvolveu um modelo híbrido do processo de comunicação (Figura 3.2), no qual considera a coexistência dos meios impresso e eletrônico. No entanto, “os estágios informais do processo de comunicação são quase inteiramente baseados no meio eletrônico, enquanto os estágios formais ainda têm sido basicamente impressos, embora gradualmente evoluindo para o eletrônico” (COSTA; SILVA; COSTA, 2001, p. 65).



**Figura 3.2** – Modelo híbrido de comunicação científica, segundo Costa (1999)

Esse cenário, segundo os autores, representa uma visão complementar dos domínios dos canais de comunicação, mais do que um processo de substituição. “Como a penetração universal do meio eletrônico na comunicação tem sido amplamente reconhecida e o processo da tecnologia é gradual, parece necessário acompanhar de perto as mudanças na comunicação formal, também.” (COSTA; SILVA; COSTA, 2001, p. 65).

É comum encontramos na literatura estudos comparativos destacando as vantagens e desvantagens da utilização das redes eletrônicas como meio de comunicação científica. Nesse aspecto, McMurdo (1995, apud OLIVEIRA; NORONHA, 2005, p. 84), reconhece como vantagens a disseminação quase instantânea da informação, a possibilidade de fornecimento de cópias múltiplas, novas opções de leitura não linear, entre outros.

Apesar dos avanços, contudo, este autor identifica algumas desvantagens para esse novo tipo de comunicação, tal como a falta de etiqueta ou normas comportamentais e a exigência de recursos e instrumentos específicos, o que, na visão de McMurdo, pode ampliar as diferenças no fluxo informacional entre países ricos e pobres. Outras vozes também alertam que devem ser considerados aspectos relacionados à inconsistência das informações, a complexidade de armazenamento e controle bibliográfico, além da banalização da autoria.

Sabbatini (1999), por sua vez, relaciona as vantagens da publicação científica eletrônica em relação aos pontos de vista do editor e do usuário, consumidor final da informação. Quanto ao aspecto editorial, as publicações eletrônicas:

- a) permitem atingir uma audiência potencialmente grande, devido à disponibilidade universal da informação;
- b) oferecem mais acessibilidade, eliminando as limitações de acesso condicionadas pela comunicação científica baseada em papel;
- c) permitem novas formas de apresentação (áudio, vídeo, interação com o usuário final da informação), integração com outros *sites* e documentos da Web;
- d) diminuem os atrasos de publicação, e possibilitam a submissão eletrônica de manuscritos.

O mesmo autor ressalta como vantagens para o usuário:

- a) o rápido e baixo custo de acesso à informação, eliminando custos associados à reprodução e transporte, sujeitos nos tradicionais meios de comunicação científica;
- b) a facilidade de realização de cópia e impressão;
- c) informação mais atualizada e fácil de localizar, com a utilização de mecanismos de busca sofisticados;
- d) possibilidade de diálogo interativo com autores e editores.

Sabbatini (ibidem) alerta, também, para alguns empecilhos encontrados nas publicações eletrônicas científicas, especialmente aquelas que usam a Internet como suporte. A proteção ao direito autoral, por exemplo, encontra dificuldades na medida em que existe uma possibilidade de reprodução dos conteúdos praticamente ilimitada. Consequentemente, as questões relacionadas à legitimidade e qualidade da informação são prejudicadas. O próprio processo de avaliação por pares enquanto forma de certificação da qualidade científica é um ponto crucial na aceitação do modelo de publicação científica eletrônica (SABBATINI, op. cit.).

Ainda sobre o papel das tecnologias de informação na comunicação científica, citamos o estudo de Pinheiro e colaboradores (2003) realizado nos períodos de 1998-2000 e 2000-2002 sobre o impacto das redes eletrônicas para a comunicação científica de pesquisadores brasileiros. Segundo dados apontados pela pesquisa, o uso da Internet faz parte da rotina dos pesquisadores brasileiros, com a grande maioria (87%) utilizando diariamente a rede, além da quase totalidade (96,7%) reconhecer que as redes eletrônicas alteraram a comunicação entre pares e a produtividade científica, e que o ambiente virtual intensificou a comunicação já existente (PINHEIRO, 1996).

Em artigo que revisa estudos clássicos sobre os colégios invisíveis e o fluxo da informação científica, Mueller (1994) sugere, entre outros tópicos, a investigação das alternativas possíveis para substituição do periódico tradicional. Porém, já demonstrava preocupação com o problema da volatilidade do meio eletrônico quanto

à questão da autoria e garantia de prioridade da publicação e também em relação à preservação digital da informação. Sobre isso, a autora sustenta que:

as funções de comunicação do conhecimento são bem desempenhadas pelos periódicos eletrônicos. Mas as funções de reconhecimento da prioridade e autoridade e também a função de arquivo (preservação do registro do conhecimento) apresentam muitos problemas. Existe aqui um terreno promissor para pesquisa (MUELLER, 1994, p. 315).

A preocupação com a segurança e a preservação da informação ainda existe nos dias atuais, porém a situação melhorou bastante com o surgimento dos repositórios digitais (institucionais ou temáticos). Tais mecanismos funcionariam como fontes secundárias, destinados ao armazenamento e o acesso organizado aos documentos digitais primários, como os periódicos e demais produções científicas e/ou institucionais. A literatura sobre os repositórios digitais encontra-se geralmente associada às iniciativas de auto-depósito e à infra-estrutura fomentada pela iniciativa dos arquivos abertos, que serão mais detalhados no item 3.5 deste capítulo.

### 3.2 O Periódico Científico Eletrônico

Na avaliação de Gomes (1999, p. 8), há várias interpretações possíveis para o termo “publicação” quando este é seguido pelo adjetivo “eletrônico”. Para a autora, o conceito de publicação eletrônica designa qualquer etapa do processo de evolução na geração de produtos que, ou se utilizam dos meios computacionais para sua produção, ou se apresentam exclusivamente no formato eletrônico.

Para Kling e McKim (1999), o que define uma publicação eletrônica é o fato desta ser “primariamente distribuída através do meio eletrônico”. O formato de distribuição, portanto, é fator determinante de sua natureza, pois uma publicação eletrônica pode ser impressa, a *posteriori*, para leitura e circulação (KLING; MCKIM, 1999, p. 891). Já para Dias (2003, p. 7-8), uma publicação eletrônica pode ser qualquer material informativo gerado com auxílio de uma ferramenta eletrônica e disponibilizada para consulta através de CD-ROM, DVD, de um *website*, ou através de qualquer outro suporte eletrônico adequado para disponibilização da informação.

O periódico eletrônico (*electronic journal*), por sua vez, é um tipo específico de publicação eletrônica, que se distingue por suas características de controle de qualidade e periodicidade. E quando nos referimos aos periódicos científicos eletrônico, estamos tratando daqueles que, independente de seu formato, devem atender as mesmas normas e procedimentos quanto à apresentação dos elementos bibliográficos e telemáticos<sup>9</sup>, orientando também a produção editorial e gráfica do periódico (FACHIN, 2002, p. 32).

O conceito de periódico eletrônico ou termos correlatos, tais como periódico *on-line*, periódico digital, etc.), evoluiu no tempo, de acordo com a percepção dos autores em determinado momento. Em vários estudos que tratam deste tema, podemos encontrar um ou mais conceitos, dentre os quais, destacamos os apresentados no Quadro 3.3.

Ano	Autor	Conceito
1994	Universidade de Nebraska-Lincoln	publicação eletrônica com texto completo, que pode incluir imagens e que pretende ser publicada indefinidamente
1995	Lancaster	periódico criado para o meio eletrônico e disponível apenas neste suporte
1995	Harrison; Stephen	periódicos acadêmicos que são disponibilizados através da Internet e de suas tecnologias
1997	Mueller	periódicos aos quais se tem acesso mediante o uso de equipamentos eletrônicos
1997	Cunha	periódico eletrônico ou digital é a publicação editada em intervalos regulares e distribuída na forma eletrônica ou digital
1999	Bonbak	publicação cuja principal forma de envio aos assinantes é por meio de um arquivo de computador
2003	Cruz et al.	aquele que possui artigos em texto integral, disponibilizados via rede, com acesso <i>on-line</i> , e que pode ou não existir em versão impressa ou em qualquer outro tipo de suporte

Fonte: compilado de Oliveira (2006, p. 29-30)

**Quadro 3.3** – Evolução dos conceitos dos periódicos científicos eletrônicos

Ao recapitular propostas conceituais para os periódicos eletrônicos, Ornelas e Arroyo (2006) propõem a seguinte definição:

<sup>9</sup> Referentes às tecnologias de informação e comunicação

a revista eletrônica é aquela publicação arbitrada criada, produzida e editada em hipertexto com versão única digital difundida pela Internet, com propriedades editoriais que se apegam estritamente às normas de qualquer revista acadêmica ou científica.

Independente do conceito adotado, todos possuem em comum a aceitação do periódico científico eletrônico como sendo tanto a publicação disponibilizada apenas em meio eletrônico como aquela que contém versões em ambos os suportes (digital e impresso). “Ao usuário, em um primeiro momento, a distinção entre a existência do título apenas em formato eletrônico e àqueles que possuem também a versão impressa não é primordial, pois não influi no momento do acesso e obtenção da informação” (OLIVEIRA, 2005, p. 30).

Outro ponto de convergência entre os autores que estudam esse tópico é a noção clara de que o desenvolvimento das publicações eletrônicas e, conseqüentemente, os periódicos neste formato, está atrelado ao uso das tecnologias, verificadas nas últimas décadas (GOMES, 1999, p. 9). Esse desenvolvimento é dividido em quatro etapas, que podem coexistir simultaneamente, segundo Lancaster (1995, p. 518):

- a) uso de computadores para gerar a publicação impressa tradicional (processadores de texto, editoração eletrônica);
- b) distribuição do texto em formato eletrônico, com a versão eletrônica exatamente igual à versão impressa;
- c) publicação eletrônica com formato da versão impressa, mas agregando algumas funções, como a possibilidade de pesquisa, produção de metadados, serviços de alerta, etc;
- d) publicações elaboradas especificamente para o formato eletrônico, que exploram as possibilidades desse novo meio, com uso intensivo de hipertexto e hipermídia. Esta etapa pode ser subdividida em dois estágios:

- a. renovação de textos e gráficos já existentes mediante inovações eletrônicas;
- b. produção de publicações concebidas para explorarem todo o potencial eletrônico

Embora relevantes, não será aprofundado os tópicos relacionados às características de arquitetura dos periódicos eletrônicos e a discussão sobre as vantagens e desvantagens desse recurso. Os mesmos já foram tratados de modo exaustivo por outros autores, especialmente Biojone (2001) e Dias (2003), cujos trabalhos constituem ótimas referências.

### 3.2.1 Evolução dos Periódicos Científicos Eletrônicos

Lancaster (1995, p. 518) aponta N.E. Sondak e R.J. Schwartz como os idealizadores dos periódicos científicos em formato eletrônico, ainda na década de 70. A contribuição destes pesquisadores consistiu na concepção de um mecanismo de fornecimento de arquivos capazes de serem lidos por computadores nas bibliotecas e por assinantes individuais, por meio de microfichas.

Contudo, a primeira configuração de um periódico eletrônico dotado de padrões e procedimentos de revisão pelos pares foi o *Electronic Information Exchange System* (EIES), financiado pela *National Science Foundation* (NSF) e desenvolvido no *News Jersey Institute of Technology*, Estados Unidos, entre 1978 e 1980 (GOMES, 1999, p. 15). O objetivo principal do EIES era o de propor um modelo de publicação para os periódicos em meio eletrônico, onde se buscava satisfazer às necessidades e desejos de editores e usuários, além de respeitar os padrões já adotados pelos periódicos científicos disponíveis no formato impresso.

Desse modo, tentava-se tornar público o acesso à informação, disseminando-a de modo mais ágil (BIOJONE, 2001, p.52). Há, no entanto, uma controvérsia, já que outro periódico, o *Postmodern Culture*, também é citado como o primeiro dessa nova genealogia de produtos da comunicação científica. Porém, o mesmo teria sido

desenvolvido apenas em 1990, inicialmente no formato de correio eletrônico, depois em disquete e, em 1994, com uma versão na Internet (BRIQUET DE LEMOS, 2005).

Na década de 80, com o surgimento do suporte CD-ROM e das tecnologias de acesso via terminal remoto de computador, teve início um ambicioso projeto chamado BLEND – *Birmingham and Loughborough Electronic Network Development*, conduzido por um consórcio entre as universidades britânicas de Birmingham e Loughborough, com a tutela do departamento de pesquisa e desenvolvimento da *British Library*. O BLEND funcionou entre os anos de 1980 a 1984, preocupando-se com a automação da armazenagem dos artigos científicos, além da facilidade de acesso que esses poderiam ter ao estarem disponíveis em meio eletrônico. Nesses quatro anos, o projeto experimentou as possibilidades de comunicação por intermédio de algumas revistas criadas, editadas e publicadas. Assim nasceu a *Computer Human Factors*, acessada via terminal conectado a um computador central por meio de uma rede local. Este periódico testou pela primeira vez o sistema de revisão por pares em linha.

Outros projetos pioneiros foram o ADONIS, também de 1980, cujo foco era a difusão, em meio eletrônico, de periódicos em formato impresso, e o QUARTET, que tinha o propósito de investigar as implicações das tecnologias de informação para o processo da comunicação científica em meios acadêmicos. O fruto dessa iniciativa foi a criação do sistema HyperBIT-*Behavior and Information Technology*, um modelo para a publicação dos primeiros periódicos eletrônicos usando os recursos de hipertextos (McKNIGHT, 1993, p. 7).

Uma década mais tarde, em 1992, o projeto TULIP (*The University Licensing Program*), da editora científica Elsevier, lançou 43 publicações técnicas e científicas em rede para cerca de 15 instituições acadêmicas que, de sua parte, desenvolviam ou adaptavam sistemas para disponibilizar as revistas eletrônicas para seus usuários. Uma das metas do TULIP era examinar os aspectos econômicos, legais e técnicos que envolvem os periódicos eletrônicos (SILVA et al., 1996, p. 455).

Na mesma época, segundo Langschieid (1994, p. 90), a *American Association for the Advancement of Science* (AAAS) e a *On-line Computer Library Center*

(OCLC) lançaram o *The On-line Journal of Current Clinical Trials*, um periódico distribuído na Internet, baseado no formato SGML – *Standard Generalized Markup Language*, padrão de descrição de textos que possibilitou a publicação de gráficos e tabelas. Mais tarde, este formato deu origem à linguagem HTML, utilizada universalmente na publicação de páginas Web. O projeto, no entanto, acabou fracassando devido à falta de artigos de qualidade, exigidos pelos editores. Por ser uma mídia nova, portanto não consagrada pela comunidade científica, poucos autores enviaram seus textos para a revista que, não gozava de popularidade por ter um alcance bastante reduzido.

O desenvolvimento tecnológico dos periódicos eletrônicos entre os anos 80 e 90 prepararam o terreno para a grande revolução tecnológica que estava por vir: os periódicos disponibilizados na *World Wide Web*, os quais viabilizaram a adoção definitiva do formato eletrônico para a publicação científica. Juntamente com o fascínio pela publicação em um novo ambiente, veio a preocupação com a disponibilização e o armazenamento dos documentos eletrônicos publicados na Web. O início das publicações eletrônicas passa, então, a ser realizado com base nos projetos de experimentação e cooperação, destacando-se (SABBATINI, 1999; OLIVEIRA, 2006):

- *Chemistry On-line Retrieval Experiment (CORE)* – Funcionou entre os anos de 1991 a 1995. Reunia os textos e imagens (digitalizadas) dos periódicos mais relevantes da área de Química publicados pela *American Chemical Society*.
- *Red Sage Electronic Journal Project* – Criado a partir de um acordo firmado em 1993 entre os laboratórios AT&T Bell, a editora científica Springer Verlag e a Universidade da Califórnia, em São Francisco. O objetivo do projeto era identificar e estudar as questões técnicas, legais, econômicas e culturais envolvendo a criação e disponibilização de literatura biomédica. Funcionou até 1996.
- MUSE – Iniciado em 1993, por meio de uma parceria entre a editora universitária Johns Hopkins e a biblioteca Milton S. Eisenhower para a disponibilização de cerca de 40 títulos daquela editora. Atualmente

oferece acesso a textos integrais de mais de 380 periódicos nas áreas de Ciências Sociais e Humanidades publicados por 60 universidades.

- *Journal Storage* (JSTOR) – Concebido inicialmente como um projeto financiado pela Andrew W. Mellon Foundation, em 1994. Atualmente é uma organização sem fins lucrativos, com o propósito de manter um repositório confiável e acessível de periódicos científicos em diversas áreas do conhecimento.
- *SuperJournal Project* – Desenvolvido no âmbito do *Electronic Libraries Programme* (eLiB) a partir de 1995. Representa a colaboração entre editores, universidades e bibliotecas para o desenvolvimento de periódicos eletrônicos multimídia. Funcionou até 1998, disponibilizando acesso a 49 periódicos eletrônicos direcionados aos usuários de universidades inglesas.
- *HighWire Press* – Unidade da biblioteca da Universidade de Stanford, fundada no início de 1995. Desde então, disponibiliza cerca de 1.100 periódicos eletrônicos de diversas áreas.
- BioOne – Resultado de uma aliança de diferentes instituições envolvidas com o processo de comunicação científica (sociedades científicas, universidades, bibliotecas acadêmicas e editores comerciais). São mais de 70 periódicos relevantes na área de Ciências. Sua base de dados tornou-se pública em janeiro de 2001.

No Brasil, um dos primeiros projetos que tratou das publicações eletrônicas científicas na Internet foi o Grupo de Publicações Eletrônicas em Medicina (E-pub), criado em 1994 por iniciativa do Núcleo de Informática Biomédica (NIB), da Universidade Estadual de Campinas. O E-pub dedicou-se ao desenvolvimento de revistas eletrônicas com textos completos da área médica. Entre os títulos publicados estão os *Arquivos Brasileiros de Cardiologia* e a revista *Saúde e Vida On-line*, esta última dirigida ao público leigo.

Um outro projeto pioneiro direcionado ao desenvolvimento de periódicos científicos nacionais foi a Base de Dados Tropical (BDT), da Fundação Tropical de

Pesquisa e Tecnologia André Tosello, uma organização sem fins lucrativos voltada para a disseminação e divulgação de informação científica especializada na área biológica. Um de seus principais projetos é o serviço *Bioline Publications*, implementado em 1994, que permite o acesso à 26 revistas científicas da área de biociências.

Em 1997, teve início o projeto SciELO – *Scientific Electronic Library On-line*, resultado de uma parceria entre a Fapesp, Bireme e um grupo de editores científicos de diversas áreas do conhecimento que se uniram com a finalidade de desenvolver uma metodologia para a preparação, armazenamento, disseminação e avaliação de publicações científicas em formato eletrônico (PACKER et al., 1998, p. 109). “A base de dados em operação, além de fornecer um controle bibliográfico da literatura científica, produz indicadores bibliométricos para medir o uso e o fator de impacto dos periódicos disponibilizados” (BIOJONE, 2001, p. 64). Atualmente, a base do SciELO reúne cerca de 480 periódicos de sete países, além do Brasil: Argentina, Chile, Colômbia, Cuba, Espanha, Portugal e Venezuela.

De acordo com Meadows (1999, p.1), “a forma como as revistas apresentam a informação evoluiu gradualmente durante os três últimos séculos em resposta tanto às transformações tecnológicas quanto às exigências cambiantes da comunidade científica.” Com isso, verifica-se que os periódicos científicos passaram por grandes mudanças, desde sua forma tradicional, passando pelo hipertexto, chegando ao periódico eletrônico.

Este fenômeno foi evidenciado por Simeão (2003), ao estudar as alterações verificadas no formato dos periódicos científicos eletrônicos. Pesquisa realizada para sua tese de doutorado junto a 400 periódicos eletrônicos disponíveis no portal da CAPES (dos 2.412 títulos oferecidos em 2002) demonstrou a existência de mecanismos que indicam a prática de uma comunicação diferenciada, denominada “extensiva”, e própria do contexto da comunicação eletrônica em rede.

### 3.3 Comunicação Extensiva

O conceito da Comunicação Extensiva foi deduzido por Simeão (2003) a partir do pensamento de historiadores da cultura do livro, Roger Chartier e Guglielmo Cavallo, para explicar as transformações evidenciadas nas experiências e no tratamento da informação disponibilizada em meio digital. O conceito abrange as mudanças estruturais observadas nos documentos eletrônicos e é definido como:

o processo flexível, horizontal e instável que atualmente rege os princípios e ações das tecnologias de informação em rede, considerando todos os dispositivos de interação entre os sistemas de informação e usuários e a possibilidade do uso de dados multidimensionais. (SIMEÃO, 2006, p. 21)

Tomando por base esses princípios, Simeão (2006, p. 53) relaciona as principais características da Comunicação Extensiva, comparando-as com as verificadas nos formatos de comunicação tradicional utilizado nos suportes estáticos, como o papel, com pouca possibilidade de interação com o receptor (Quadro 3.4).

<b>Comunicação Intensiva</b>	<b>Comunicação Extensiva</b>
tradicionalismo	informalidade
normas rígidas, padrões fixos	regras flexíveis com padrões dinâmicos
restrições à leitura e edição centralizada	leitura expandida, sem limites, edição interativa
promove o reconhecimento	promove o inédito, o inesperado
referências idênticas, restritas à áreas específicas	referências diferentes e complementares
leitura lenta, íntima e linear	leitura rápida, superficial e comutativa
configuração vertical, informação em profundidade	configuração horizontal, inteligência múltipla

Fonte: SIMEÃO, 2006, p. 53

**Quadro 3.4** – Características dos formatos de Comunicação Intensiva e Extensiva

#### 3.3.1 Indicadores da Comunicação Extensiva

O estudo das particularidades da Comunicação Extensiva nos traz elementos para melhor compreensão das transformações no processo de leitura decorrente da

migração do suporte tradicional impresso para o meio eletrônico. O método utilizado para avaliação dessas mudanças foi a mensuração dos indicadores do formato eletrônico (interatividade, hipertextualidade e hipermediação) do portal de periódico da CAPES, que revelaram o processo de Comunicação Extensiva (SIMEÃO, 2000, p. 19).

As variáveis incorporadas ao estudo de Simeão foram identificadas por meio das características, produtos e serviços disponibilizados pelos diferentes formatos, de modo a indicar cada um dos atributos (perfil e contexto, interatividade, hipertextualidade e hipermediação). Essas variáveis foram relacionadas com o formato impresso e as específicas do formato eletrônico (SIMEÃO, 2006, p. 169).

### 3.3.1.1 Interatividade

A potencialização da capacidade de interatividade nos meios de comunicação baseados em computadores vulgarizou o termo, que passou a ser utilizado para nomear as mais variadas situações, principalmente as relacionadas às questões mercadológicas. Para Lévy (2000), o termo é invocado tantas vezes como se todos soubessem do que se trata. “O termo interatividade em geral ressalta a participação ativa do beneficiário de uma transação de informação” (LÉVY, 2000, p.79).

Outra questão, observada por Guissani (1999), é que, com o incremento da interatividade, nota-se a diminuição da distância entre os pólos integrantes do processo de comunicação.

Um dos fenômenos observados na interatividade digital é o progressivo desaparecimento da divisão entre o produtor e o consumidor da informação. Em Internet, todos são potencialmente e simultaneamente escritores e jornalistas, editores e leitores. (GUISANI, 1999 apud FLIZIKOWSKI, 2003, p. 39)

No contexto da Comunicação Extensiva, Miranda e Simeão (2005) entendem a interatividade como a possibilidade de diálogo entre o usuário (interpretante) e o sistema, entre o sistema e o usuário e dos usuários entre si através do sistema. A principal característica desse indicador, segundo os autores, é a interação do sistema com os usuários, sejam eles emissores ou receptores.

A interatividade nas mídias digitais implica, ainda, no aumento de participação do receptor, que adquire o “poder” de escolher as mensagens e tornar-se, também, produtor de novos conteúdos. Em outras palavras, o usuário passa a acessar à rede nos seus próprios termos, atribuindo um caráter pessoal à sua forma de navegar.

### 3.3.1.2 Hipertextualidade

O segundo indicador da Comunicação Extensiva trata da possibilidade da interconexão de conteúdos múltiplos, tendo com principal fator o direcionamento intertextual através de *links* conceituais e de deslocamento (MIRANDA; SIMEÃO, 2005, p. 188-189).

Resgatando a definição de Theodor Nelson, autor do termo “hipertexto”, verificamos que este constitui-se em:

uma escrita não seqüencial, num texto que se bifurca, que permite que o leitor escolha e que se leia melhor numa tela interativa. De acordo com a noção popular, trata-se de uma série de blocos de texto conectados entre si por nexos, que formam diferentes itinerários para o usuário (NELSON apud LANDOW, 1995, p. 15).

A partir desta definição verificamos três elementos – blocos de textos; ligações entre eles; e meio digital/tela do computador – que vão dar sustentação a uma dinâmica particular de funcionamento do hipertexto no que diz respeito à organização das informações (escrita) e ao acesso das mesmas (leitura).

Na prática, a utilização do hipertexto nas redes de computador permite em um mesmo ambiente a coexistência de textos, sons e imagens, tendo como elemento inovador a possibilidade de interconexão quase instantânea através de *hiperlinks*, não só entre partes de um mesmo texto, mas entre textos fisicamente dispersos, localizados em diferentes suportes e arquivos integrantes da teia de informação constituída pela Web.

Acredita-se que o *hiperlink* é o elemento diferencial apresentado pelo hipertexto em suporte digital que potencializa os aspectos de não-linearidade e de intertextualidade, associada às características da multimedialidade. A forma de

organização (léxias) e formatação (interconexão) da informação ganha novo sentido com a hipertextualidade, proporcionando a flexibilidade necessária para a navegação de um ponto a outro em uma estrutura de rede.

### 3.3.1.3 Hipermediação

O uso de *hiperlinks* em conteúdo multimídia (áudio, vídeo, fotos, animações) é chamado de hiperídia. Tecnicamente, não há diferenças em fazer *links* em texto ou em imagens. O processo de hipermediação, portanto, caracteriza-se essencialmente pelo emprego de recursos de áudio e imagens em movimento, podendo fazer uso de *hiperkinks* ou não. Para Negroponte (1995, p. 66), “A hiperídia é um desenvolvimento do hipertexto, designando a narrativa com alto grau de interconexão, a informação vinculada”.

Existe uma discussão corrente acerca da inclusão de textos sonoros e visuais no conceito de hipertexto. Para alguns autores não há distinção entre hipertexto e hiperídia, podendo o primeiro abarcar, além de textos escritos, também os conteúdos multimídia.

Segundo a visão Extensiva, a hipermediação é a

combinação da informação em suas múltiplas dimensões. Texto, imagem e áudio são utilizados na construção do conteúdo numa lógica discursiva não-linear que obedece aos comandos do usuário (MIRANDA; SIMEÃO, 2005, p. 189).

De acordo com os autores, esse conceito se distingue dos anteriores por concentrar-se na capacidade de promoção da construção em bases metatextuais. Além disso, a hiperídia é uma das particularidades mais sofisticadas no processo de adaptação das estruturas de informação no contexto da comunicação extensiva devido à exigência de um conhecimento mais complexo sobre o uso de cada mídia (SIMEÃO, 2005).

### 3.3.2 Comunicação Extensiva e RSS

Observa-se que a tecnologia RSS incorporada aos documentos eletrônicos contempla o modelo da Comunicação Extensiva, por representar uma expansão na maneira com que os conteúdos digitais são disseminados. Como veremos mais detalhadamente no Capítulo 4, os formatos de sindicância de conteúdos Web estabelecem um padrão de comunicação diferenciado, invertendo o fluxo da disseminação como normalmente estamos acostumados. Em vez das informações serem “empurradas” como acontece com os meios tradicionais (intensivos), no caso do RSS, cabe ao usuário “puxar” o conteúdo a partir das fontes de informação sobre os quais deseja se manter informado. Além disso, este recurso tecnológico oferece os meios para que esta informação recebida seja também reutilizada, favorecendo o seu compartilhamento.

Essa característica, por si, indica um alto grau de participação do usuário que deixa de ser apenas um consumidor de informação para se tornar um fornecedor, caso deseje. Assim, percebemos o RSS contemplando o primeiro – e talvez mais importante – indicador da Comunicação Extensiva, a interatividade, que, segundo as palavras de Simeão e Miranda (2004),

é conquistada através de linguagens mais abertas e flexíveis, com a disponibilidade de um conjunto de ferramentas, produtos e serviços que significarão um maior espaço de armazenagem em servidores e bases e uma maior habilidade de editores e autores”.

Compreendendo o funcionamento da tecnologia RSS, percebe-se que estamos lidando com um poderoso instrumento capaz de modificar não somente a perspectiva do usuário em termos do acesso, mas, principalmente o seu hábito de “navegação” das páginas Web. O conceito de hipertextualidade apresenta-se de forma ainda mais ampla, pois a leitura via RSS acontece não somente com uma, mas com várias fontes de informação simultaneamente. Assim, o processo da Comunicação Extensiva novamente se faz notar devido ao leque de possibilidades (rede) para se proceder à absorção da informação que, apesar de recuperada de modo seletivo, não oferece limites e coerções para o usuário, possibilitando uma ação comunicativa mais aberta e espontânea (MIRANDA; SIMEÃO, 2006).

Desde a sua origem, os formatos RSS foram pensados para servirem como “resumos” dos sítios da Web, facilitando a localização de informação atualizada por parte do usuário. Por isso, a leitura de páginas “sindicadas”, em geral, acontece de forma dinâmica e superficial. A maioria dos canais RSS disponibiliza somente os resumos dos textos para que o usuário tenha uma noção do assunto tratado. O acesso ao documento completo acontece em um segundo momento, por meio de um clique sobre o título.

Outra característica do formato eletrônico é a hipermediação, também presente nos formatos RSS por meio de elementos próprios que representam e identificam conteúdos hipermédia. Qualquer programa ou sistema “agregador de conteúdo” possui a capacidade de interpretar e exibir diversas mídias (áudio, vídeo, animação), independente do suporte. Tecnicamente, chama-se *podcasting* a distribuição de documentos multimídia por meio dos formatos RSS, os quais podem ser executados em páginas Web ou mesmo por *Ipods* e *handlets* (dispositivos móveis pessoais) em geral.

### 3.4 Acesso Aberto: Um Novo Modelo para a Comunicação Científica

De acordo com Oliveira (2006, p. 37), a abertura comercial da Internet juntamente com o surgimento da Web, nos anos 90, impulsionou as editoras científicas comerciais a disponibilizarem seus títulos em formato eletrônico. Como exemplo, tem-se *Elsevier ScienceDirect* (cerca de 1.500 títulos de periódicos), *SpringerLink* (2.380 títulos), *Blackwell Synergy* (700 títulos) e *Wiley Interscience* (cerca de 1 mil títulos).

Apesar do aumento considerável no número de periódicos eletrônicos, entretanto, o preço das assinaturas das publicações impressas manteve-se também em alta. O acesso à versão eletrônica das revistas, em alguns casos, ocorre de forma gratuita apenas como complemento da assinatura do impresso. Esse e outros problemas motivaram um grupo de pesquisadores a pensar em um novo formato de

comunicação científica, utilizando novos ambientes digitais de acesso livre e irrestrito, com tecnologia de baixo custo.

A primeira iniciativa neste sentido teve início em 1991, com a criação de um repositório de *pre-prints* (artigos não revisados pelos pares), cobrindo a área de física teórica do Laboratório Nacional de Los Alamos. O projeto de Paul Ginsparg, chamado ArXiv.org, foi seguido por outras ações desenvolvidas isoladamente. Destaca-se o Cogprints (*Cognitive Sciences Eprint Archive*), da Universidade de Southampton, o qual abrange as áreas de psicologia, lingüística e neurociências, e o NCSTRL (*Networked Computer Science Technical Reference Library*), voltado para a área de Ciências da Computação.

Um segundo momento importante foi a chamada proposta “subversiva”, do pesquisador britânico Stevan Harnad, ao sugerir a criação de espaços alternativos para divulgação de pesquisas científicas, livre das cobranças pelo acesso (WEITZEL, 2005, p. 175).

Em 1999, uma nova ordem para a comunicação científica começou a se estabelecer, com o suporte da tecnologia, após a realização de uma convenção em Santa Fé, Novo México, organizada por Paul Ginsparg, Rick Luce e Herbert Van de Sompel. Essa reunião foi o embrião do movimento conhecido como *Open Archives Initiative* (OAI). Além do auto-arquivamento realizado pelos autores, o modelo OAI é baseado em ferramentas e procedimentos, tais como: metadados padronizados para descrição, acesso livre à produção científica e um protocolo de interoperabilidade para diferentes modelos de arquivos públicos, o OAI-PMH (*Open Archives Initiative – Protocol of Metadata Harvesting*). A integração promovida por esse instrumental permite novas formas de disponibilização e compartilhamento de informações na Internet e favorece diferentes instâncias para publicação, variando da edição de revistas científicas e repositórios de *electronic-prints* (*e-prints*) a repositórios institucionais e sistemas de gerenciamento de eventos (WEITZEL, 2005, p. 162-163).

Como resultado da Convenção de Santa Fé foi publicado um documento que estabelece as características técnicas e arquitetônicas das soluções de

arquivamento, além de recomendar os mecanismos que deverão ser utilizados para promover o intercâmbio e compartilhamento de informações, que também será responsável pela agregação de valor à informação que será disponibilizada.

Uma das principais críticas à essa iniciativa, porém é o mito da extinção da revisão por pares que, de acordo com os editores comerciais, comprometeria o controle de qualidade e a legitimidade da informação científica. Sobre isso, fica para consideração que:

As publicações eletrônicas, em geral, parecem romper com os modelos funcionalistas do fluxo de informação científica, reduzindo o intervalo de tempo destinado à geração, ao uso e à divulgação do conhecimento registrado. Porém, ao contrário do mito, a OAI não transgride os requisitos de legitimação que orientam as publicações científicas impressas, inerentes à C&T (WEITZEL, 2005, p. 180).

Harnad (2000), por sua vez, reforça que a avaliação dos pares é o trunfo da liberdade para a produção científica, representando uma “mão invisível” que sustenta a qualidade da ciência. Porém deve-se distinguir o momento em que ocorrerá tal processo. Para os defensores da OAI, a revisão *a priori* provoca atraso na publicação dos artigos e rejeições sob critérios restritos à comissão editorial, enquanto a avaliação *a posteriori* permite novas oportunidades de certificação da qualidade dos textos, oferecendo transparência no processo.

Nos dias atuais, além dos arquivos abertos, tem crescido também o número de títulos de periódicos eletrônicos científicos disponibilizados livremente na Web, seguindo a filosofia do movimento de acesso aberto (*open access*). Márdero Arellano, Ferreira e Caregnato (2005, p. 205) definem a expressão “acesso aberto” como “a disponibilização na Internet de literatura de caráter acadêmico ou científico, permitindo a qualquer pessoa ler, descarregar (*download*), copiar, distribuir, imprimir, pesquisar ou referenciar (*links*) o texto integral dos documentos”.

Dados do *Directory of Open Access Journals* (DOAJ), da *Lund University Libraries*, indicam a existência de mais de 3 mil periódicos<sup>10</sup> de acesso aberto. No entanto, é importante observar que a disponibilização livre da publicação científica não significa que a mesma seja gratuita. Isso vai depender do modelo de negócio

---

<sup>10</sup> Dado obtido em 02/01/2008 por meio do sítio <<http://www.doaj.org>>

implementado nas publicações, que pode ser gratuita para os leitores e autores ou gratuita para os leitores, mas paga pelos autores ou suas instituições.

Em todo caso, o modelo de acesso aberto oferece vantagens em relação ao sistema tradicional de comunicação científica, por garantir o padrão de qualidade por meio da revisão de pares, porém sem transferência de *copyright* (direitos autorais) para o editor. Outro aspecto favorável, segundo Oliveira (2006, p. 42), é a diminuição das diferenças geográficas e financeiras dos pesquisadores localizados em regiões distantes dos grandes centros de pesquisa mais avançados, promovendo a “equidade de acesso à informação entre países desenvolvidos e em desenvolvimento”.

Merece destaque as iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), que investiu na tradução para o idioma Português dos *softwares E-prints* e *Open Journal System* (OJS), que viabilizam a implementação de repositórios de arquivos abertos e a edição/publicação de revistas eletrônicas. No Brasil, esses projetos são conhecidos, respectivamente, pelo nome de Diálogo Científico e Sistema Eletrônico de Editoração de Revistas (SEER).

### 3.5 Blogs Científicos

Uma revisão sobre comunicação científica, com ênfase no papel das tecnologias de informação na disseminação do conhecimento científico, ficaria incompleta se deixássemos de abordar o papel dos *blogs* científicos nos dias atuais. Verifica-se este meio como importante canal da comunicação informal entre cientistas e pesquisadores e mesmo como canal de comunicação formal com o público em geral.

Surgido há pouco mais de uma década, o termo *weblog* (rapidamente abreviado para *blog*) era inicialmente empregado para se referir aos arquivos de atividades (*logs*) gerados pelos servidores Web, mas seu conceito foi ampliado para abranger o conjunto de aplicações capaz de gerenciar conteúdos de páginas

atualizadas com determinada frequência e cujas entradas (no jargão técnico, *posts*) são publicadas segundo uma ordenação cronológica (SAUERS, 2006, p. 1).

Em sua essência, os *blogs* são publicações pessoais que se distinguem por uma característica marcante: a interatividade proporcionada entre comentários publicados (e registrados) pelos autores e seus leitores. Como idéia, sua origem remonta ao nascimento da própria Web, fazendo parte da rotina de muitos indivíduos como uma espécie de diário eletrônico. Porém, como fenômeno específico, apenas recentemente vem sendo reconhecido como um típico veículo de comunicação sob o aspecto de sua capacidade de disseminação de novas idéias, servindo como importante instrumento para comunidades com interesses em comum e até mesmo por organizações, seja para fins pessoais ou informacionais.

A popularização dos *blogs* vem ocorrendo de maneira acelerada, e boa parte desse sucesso se deve às facilidades oferecidas pelas ferramentas de publicação, tais como *Blogger*<sup>11</sup> e *Wordpress*<sup>12</sup>, que despontam entre as mais utilizadas. Vale mencionar que todos os serviços de criação de *blogs* permitem a geração automática de canais RSS à medida que um novo *post* (ou comentário) é publicado. Esse detalhe é especialmente importante já que as informações disponibilizadas nos *blogs* geralmente são atualizadas com frequência. Na opinião de Davison-Turley (2005, p. 28), “os *blogs* são recursos de informação extraordinários, mas é o RSS que realmente os tornam poderosos”.

De acordo com o sítio *Technorati* – um motor de busca especializado em *blogs* – atualmente existem cerca de 112,8 milhões<sup>13</sup> de *blogs* ativos no mundo, um dado que pode ser interpretado como um indicativo de impacto cultural para a nossa sociedade da informação e do conhecimento. Não tão raros, os *blogs* mantidos por cientistas começam a reforçar as estatísticas, demonstrando que estes profissionais começam a querer mudar certas tradições, com a divulgação irrestrita dos resultados de suas pesquisas, que antes circulavam apenas nos ambientes acadêmicos ou nos

---

<sup>11</sup> <<http://www.blogger.com>>

<sup>12</sup> <<http://wordpress.com>>

<sup>13</sup> Dado obtido em 10/01/2008 por meio da página <<http://www.technorati.com/about>>

laboratórios. Em uma busca realizada neste mesmo sítio, encontramos 24,2 mil *blogs* classificados pelo termo *science*.<sup>14</sup>

O assunto continua polêmico no meio científico. Ainda assim, em julho de 2006, a revista *Nature* solicitou ao *Technorati* que vasculhasse os sítios do gênero para descobrir quais deles possuíam o maior tráfego entre aqueles cujo assunto principal é ciência. O resultado da pesquisa apontou cinco *blogs* científicos entre os 3,5 mil mais visitados do *ranking* geral, os quais são identificados no Quadro 3.5.

Nome do <i>blog</i>	Posição geral
Pharyngula <a href="http://www.scienceblogs.com/pharyngula">http://www.scienceblogs.com/pharyngula</a>	179 <sup>o</sup>
Panda's Thumb <a href="http://www.pandasthumb.org">http://www.pandasthumb.org</a>	1.647 <sup>o</sup>
Realclimate <a href="http://www.realclimate.org">http://www.realclimate.org</a>	1.884 <sup>o</sup>
Cosmic Variance <a href="http://www.cosmicvariance.com">http://www.cosmicvariance.com</a>	2.174 <sup>o</sup>
Scientific Activist <a href="http://www.scienceblogs.com/scientificactivist">http://www.scienceblogs.com/scientificactivist</a>	3.429 <sup>o</sup>

Fonte: Nature, 442

**Quadro 3.5** – *Ranking* de *blogs* científicos no mundo

De posse desta lista, a revista entrevistou os autores dos *blogs* para saber o que eles achavam desse sucesso. “Na maioria das vezes, faço apenas um texto sumarizado dos conceitos apresentados em sala de aula. Claro que se fosse para um periódico científico não teria o mesmo sucesso”, comentou Phil Myers, criador do *blog Pharyngula* (sobre biologia evolutiva) à revista *Nature*. Professor da Universidade de Minnesota, Myers acredita que os *blogs* são uma espécie de conversa informal, como as que ocorrem no intervalo dos eventos científicos (BLOGOSFERA CIENTÍFICA, 2006).

Um amostra bastante representativa do universo dos *blogs* criados por cientistas e/ou suas instituições, além dos mantidos por jornalistas especializados

<sup>14</sup> <<http://www.technorati.com/blogs/tag/science>>, consultado em 10/01/2008

em ciência, pode ser verificada no portal *ScienceBlogs*<sup>15</sup>, que reúne mais de 70 *blogs* especializados e categorizados por temas. O crescimento desse segmento tem sido tão intenso que, em janeiro de 2007, foi organizada a primeira conferência sobre *blogs* científicos, sediada na Carolina do Norte, Estados Unidos<sup>16</sup>. Dentre os resultados desse evento, destaca-se a publicação de um livro (em versões impressa e eletrônica) intitulado *The Open Laboratory: The Best Writing on Science Blogs 2006*, uma coletânea com 50 *posts* selecionados que mostram a qualidade e a diversidade dos escritos em *blogs* científicos até 2006.

No Brasil, os *blogs* científicos ou de divulgação científica também começam a marcar presença no ciberespaço e, o mais importante, todos eles disponibilizam seus conteúdos por meio dos formatos RSS. Entre os escritos por cientistas, destacam-se o *Por dentro da Ciência*<sup>17</sup>, do físico Adilson Oliveira, e o *Via Gene*<sup>18</sup>, liderado pela doutora em genética molecular, Ana Cláudia Lessinger, da Universidade de Campinas (Unicamp). Entre os *blogs* de não-cientistas, são bons exemplos os *blogs Ciência em Dia*<sup>19</sup>, do jornalista Marcelo Leite, e o *Inova Brasil*<sup>20</sup>, especializado em divulgação de informação sobre inovação científica e tecnológica.

Em artigo publicado no sítio de notícias G1, o presidente da Sociedade Brasileira de Neurociências e Comportamento (SBNeC), Stevens Rehen, sugere que o movimento dos *blogs* científicos pode ser uma alternativa ao processo tradicional de avaliação e publicação de artigos científicos relacionados. Como exemplo, o pesquisador comenta os resultados do Fórum Mundial de Ciências da Saúde da Próxima Geração, que reuniu centenas de editores de revistas científicas e jovens cientistas. Segundo Rehen (2007), entre as recomendações propostas pelos participantes estão a disponibilização imediata de todos os artigos científicos submetidos em *blogs* especializados (que funcionariam como as atuais revistas científicas): nos quais ocorreria a avaliação pública e certificação de seu conteúdo por revisores especializados, além da inclusão de comentários e sugestões não

---

<sup>15</sup> <<http://www.scienceblogs.com>>

<sup>16</sup> <<http://wiki.scienceblogging.com>>

<sup>17</sup> <<http://pordentrodaciencia.blogspot.com>>

<sup>18</sup> <<http://viagene.blogspot.com>>

<sup>19</sup> <<http://cienciaemdia.zip.net>>

<sup>20</sup> <<http://inovabrasil.blogspot.com>>

anônimas por membros da comunidade científica. “Com isso, 100% da produção científica mundial passaria a estar disponível *on-line*. Ao mesmo tempo, injustiças, duplicações de dados e fraudes seriam mais facilmente e rapidamente identificadas” (REHEN, 2007).

A proposta dos *blogs* científicos, portanto, se identifica com o princípio da publicidade do conhecimento defendido por Ziman (1979, p. 155) quando afirma que a ciência é uma forma de conhecimento público. Assim sendo, a comunicação científica torna-se condição necessária para a validação e socialização desse conhecimento, o qual é construído a partir de um ciclo constante e auto-regenerativo da comunicação: conhecimento → publicidade → opinião pública → novo conhecimento (BARRETO, 1998, p. 123).

Ferramentas tecnológicas, como *blogs* e RSS, portanto, podem contribuir significativamente com a fase mais importante do ciclo da comunicação científica: a disseminação das informações geradas a partir do resultado de pesquisas, as quais formam a base para a construção do conhecimento científico que, por sua vez, deve (ou deveria ser) de domínio público.

## CAPÍTULO 4

### Procedimientos Metodológicos

---

Neste capítulo serão detalhados os procedimentos utilizados durante a coleta dos dados necessários para a verificação dos objetivos, conforme declarados no capítulo introdutório desta dissertação.

Identificam-se como atividades preliminares para o desenvolvimento da pesquisa a realização de um levantamento bibliográfico da literatura técnica e especializada acerca dos tópicos centrais a serem estudados, mais especificamente o papel da comunicação científica, os sistemas de recuperação e filtragem de informação, o conceito de Disseminação Seletiva de Informação e a tecnologia de sindicância de conteúdos na Web/RSS. Esses temas correspondem aos pressupostos teóricos que nortearam o estudo, destacando-se as interferências verificadas pelas tecnologias de informação e comunicação (conceito de comunicação extensiva) e as aplicações dos formatos de sindicância de conteúdos Web, tendo o RSS como principal recurso tecnológico.

#### **4.1 Tipo de Pesquisa**

O presente trabalho pode ser classificado como uma pesquisa descritiva por relatar as características de um determinado fenômeno sem a interferência ou manipulação do pesquisador (KÖCHE, 1982, p. 79). Neste sentido, considera-se como objeto de investigação os variados aspectos da tecnologia RSS e suas possíveis aplicações e utilização por uma determinada comunidade científica.

Com relação aos seus objetivos, trata-se de um estudo exploratório, por delinear uma aproximação inicial para a investigação de um tema ainda pouco conhecido, no caso o potencial dos formatos de sindicância RSS em sistemas de notificação e agregação de conteúdos para aplicações no âmbito da comunicação científica, especialmente nos periódicos científicos eletrônicos.

Um estudo exploratório visa proporcionar ao pesquisador maior familiaridade com o tema tratado. Caracterizamos essa pesquisa, portanto, como descritiva-

exploratória, aspecto justificado por Babbie (1996), que resume as finalidades desse tipo de pesquisa:

os estudos exploratórios são feitos, mais tipicamente, com três finalidades: (1) simplesmente para satisfazer a curiosidade do investigador, bem como seu desejo por uma melhor compreensão, (2) para testar a viabilidade de empreender um estudo mais cuidadoso, e (3) desenvolver os métodos a serem empregados em um estudo mais cuidadoso (BABBIE, 1986).

De acordo com a literatura, estudos descritivos-exploratórios, em geral, buscam co-relação entre variáveis por meio de inferência de dados sobre fenômenos já ocorridos, procurando o refinamento de conceitos e o desenvolvimento de hipóteses. Segundo Gil (1999, p. 45), esse tipo de pesquisa tem com objetivo fundamental “proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito”.

## 4.2 Etapas da Pesquisa

Após o levantamento bibliográfico, seguiu-se a etapa da pesquisa documental, realizada por meio de consultas a diversos sítios Web e bases de dados *on-line* disponibilizadas pelo portal de periódicos da CAPES. A literatura não especializada também foi relacionada, ainda que esta não tenha sido determinante para o aspecto teórico do trabalho, mas pelo fato de que parte da informação sobre o tema “RSS” é encontrada muitas vezes nos próprios sítios ou *blogs* que utilizam a tecnologia.

Nota-se que a temática sobre sindicância de conteúdo e RSS, apesar de relativamente nova, vem despertando o interesse de profissionais da informação (geralmente bibliotecários estrangeiros), provocando um aumento da produção bibliográfica dos títulos relacionados à área da Ciência da Informação. Essa constatação pode ser verificada a partir de uma consulta feita na base de dados LISA (*Library and Information Science Abstracts*), que indexa a literatura corrente da área em 68 países, reunindo textos de 20 idiomas diferentes. Foi realizada uma busca pelo termo “RSS” nesta base considerando o período de 1997 (surgimento da tecnologia) até 2007. Como resultado, foram encontradas 224 referências de

documentos publicadas nos últimos 10 anos (Tabela 4.1), sendo um único artigo de revisão disponibilizado por uma revista no idioma Português (*Informação e Sociedade: estudos*), em 2007. Porém, trata-se da tradução de um texto originalmente publicado em inglês intitulado *Library 2.0 theory: Web 2.0 and its implications for libraries*<sup>21</sup>.

**Tabela 4.1** – Evolução da literatura sobre RSS segundo a base de dados LISA

<b>Período</b>	<b>Total de referências</b>	<b>Variação percentual</b>
1997-1999	3	-
2000-2002	12	+300%
2003-2005	73	+508%
2006-2007	140	+91,7%
<b>Total</b>	<b>228</b>	-

Fonte: LISA, 13/12/07

Os dados acima indicam um incremento acentuado na publicação de documentos relacionados ao tema estudado durante os anos de 2003 a 2005, em comparação com o período observado anteriormente. Curiosamente, este período engloba o ano de 2004, considerado por muitos autores como o ponto de partida para a popularização definitiva dos *blogs* em todo o mundo.

Após a consolidação da base teórica pertinente, outra etapa importante foi a realização de um experimento prático visando o aprofundamento das características da tecnologia RSS. Desse modo, identificou-se na própria Web um conjunto de serviços e ferramentas específicas para o melhor aproveitamento de suas funcionalidades. O resultado foi a construção de um serviço voltado para a agregação de conteúdos da área de Ciência da Informação (ver item 4.6). Essa etapa foi desenvolvida paralelamente à fase de elaboração dos instrumentos de coleta e ajudou a testar algumas das impressões iniciais a respeito das possibilidades da tecnologia.

<sup>21</sup> Maness, J. Teoria da biblioteca 2.0: Web 2.0 e suas implicações para as bibliotecas. **Informação & Sociedade: Estudos**, João Pessoa, v. 17, n. 1, p. 43-51, jan./abr., 2007. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/831/1464>>. Acesso em: 13 jan. 2008.

O Quadro 4.1 apresenta uma relação dos objetivos específicos, com a indicação dos respectivos capítulos nos quais são tratados e os métodos utilizados.

Objetivos Específicos	Capítulos	Método	Variáveis
1. descrever as características e funcionamento da tecnologia de sindicção de conteúdos Web	5	pesquisa bibliográfica	-
2. identificar como os formatos RSS vêm sendo empregados no contexto da comunicação científica	6	pesquisa documental	-
3. Identificar o comportamento informacional de leitores de periódicos científicos eletrônicos em CI	7	levantamento ( <i>survey</i> )	frequência de leitura uso de alerta tipo de alerta utilizado satisfação com o serviço de alerta uso de ferramenta de busca nos periódicos profundidade de leitura
4. Levantar as possibilidades de aplicação da tecnologia RSS como serviço de disseminação seletiva de informação científica	4-7	experimento	-

**Quadro 4.1** – Relacionamento entre objetivos específicos, capítulos e variáveis

No caso do objetivo 3, identifica-se, também, as variáveis utilizadas para mapear o comportamento informacional de um grupo de usuários de periódicos eletrônicos – particularmente da área da Ciência da Informação – a fim de verificar a viabilidade (ou não) na adoção da tecnologia proposta. As variáveis analisadas foram:

**frequência de leitura** – visa medir a quantidade média de acessos pelo usuário a um periódico em um determinado período de tempo (semana, mês, quinzena, semestre).

**uso de alerta** – procura descobrir se o usuário faz uso de algum tipo de serviço de notificação sobre a publicação de uma nova edição corrente.

**tipo de alerta utilizado** – diz respeito ao modo pelo qual a notificação é realizada, seja pelo envio de uma mensagem de correio eletrônico contendo o sumário do periódico, *link* para a publicação, ou outro tipo de aviso.

**satisfação com o serviço de alerta** – procura avaliar se o serviço de alerta utilizado atende às expectativas do usuário quanto a estar atualizado sobre a publicação de artigos recentes.

**uso de ferramenta de busca nos periódicos** – refere-se à utilização de mecanismos automáticos para recuperação de informação dentro do próprio periódico.

**profundidade de leitura** – refere-se à profundidade de leitura normalmente dedicada aos artigos em um primeiro momento. Assim, o usuário pode se manter inicialmente informado apenas a partir da leitura dos sumários dos periódicos ou ter a necessidade de ir além, com a leitura dos resumos dos artigos. Também é possível que esse usuário tenha o hábito de fazer uma leitura mais profunda, necessitando do texto integral.

Utilizou-se a técnica de levantamento (*survey*) como estratégia de coleta de dados. A escolha deste método se deu, principalmente, pelo fato de poder ser aplicado a uma amostra de indivíduos de uma população conhecida com o uso de instrumentos padronizados.

O recolhimento dos dados foi feito mediante o preenchimento de um questionário eletrônico (Apêndice A) disponibilizado na Internet por meio do serviço gratuito Survey Gizmo<sup>22</sup>. A opção pelo questionário *on-line* em vez do envio de arquivos anexados via correio eletrônico pareceu mais indicada devido ao número elevado da população a ser estudada que, apesar de seus aspectos homogêneos, encontra-se dispersa geograficamente. Além disso, a utilização da mídia Internet se mostra particularmente propícia para avaliações desta natureza.

Finalmente, identificou-se um conjunto de características comuns aos artigos científicos a partir de um estudo comparativo realizado junto a alguns periódicos com

---

<sup>22</sup> <<http://www.surveygizmo.com/>>

o propósito de verificar o tipo de metainformação necessária para representar a estrutura funcional de um artigo científico, tais como título, nome do autor, data da publicação, tipo do documento (PDF/HTML), etc. O instrumento de coleta utilizado durante esta etapa foi um segundo questionário aplicado a um grupo reduzido de pesquisadores. O principal objetivo deste instrumento foi identificar o grau de importância conferido por esses usuários aos elementos funcionais dos artigos científicos.

### 4.3 População Estudada

Como sujeitos desta investigação foram selecionados os pesquisadores da área da Ciência da Informação por percebermos que esses profissionais constituem um público receptivo ao uso de novas tecnologias, principalmente quando estas favorecem aspectos relacionados à divulgação e uso da informação. Além disso, acredita-se que os profissionais da informação podem ser usuários potenciais dos formatos de sindicância de conteúdos Web, tanto como consumidores quanto produtores e disseminadores dessa tecnologia.

A amostra de estudo correspondeu aos pesquisadores de alguma maneira vinculados à Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB), que reúne os cursos e Programas de Pós-Graduação brasileiros da área, além de sócios individuais (profissionais, pesquisadores, estudantes de pós-graduação e profissionais egressos dos programas). A escolha desta amostra se justifica pelo fato desta comunidade estar diretamente envolvida com a comunicação científica, tanto como usuários (leitores dos periódicos científicos) como produtores de informação científica (autores, avaliadores e, em alguns casos, editores científicos). Por adotarmos um critério de intencionalidade, conveniência e representatividade, a amostra pode ser caracterizada como não-probabilística intencional.

Na época em que foi realizado o estudo o total de sócios-pesquisadores da ANCIB girava em torno de 300, de acordo com dados obtidos pela secretaria da

associação<sup>23</sup>. No entanto, a estratégia de divulgação do questionário eletrônico previu o envio de mensagens via correio eletrônico para uma base de 460 endereços cadastrados<sup>24</sup> no grupo de discussão da ANCIB na Internet (ancib@yahoogroups.com.br).

O segundo questionário foi encaminhado apenas para os 7 (sete) coordenadores dos Grupos de Trabalho da ANCIB, que foram contatados pessoalmente por ocasião do 8º Encontro Nacional de Pesquisa em Ciência da Informação (Enancib), realizado em outubro de 2007, em Salvador. Além da apresentação sobre os propósitos da pesquisa, foi dada uma breve explanação sobre a tecnologia RSS.

Essa amostra não-aleatória em particular se justifica por considerarmos a representatividade desse grupo em relação ao universo dos pesquisadores da ANCIB e pelo fato de constituírem um quadro de pesquisadores comprovadamente experientes e, portanto, aptos para fornecerem respostas para as questões específicas investigadas (Apêndice B). Todos os coordenadores dos GTs colaboraram com a pesquisa, sendo que a maioria preferiu encaminhar o questionário posteriormente, por meio de correio eletrônico. Os entrevistados desse grupo foram:

- Coordenadora do GT 1 (Estudos Históricos e Epistemológicos da Informação) – Dra. Lena Vania Ribeiro Pinheiro (IBICT)
- Coordenadora do GT 2 (Organização e Representação do Conhecimento) – Dra. Rosali Fernandez de Sousa (IBICT)
- Coordenadora do GT 3 (Mediação, Circulação e Uso da Informação) – Dra. Regina Maria Marteleto (UniRio)
- Coordenador do GT 4 (Gestão da Informação e do Conhecimento nas Organizações) – Dr. Ricardo Rodrigues Barbosa (UFMG)

---

<sup>23</sup> Dado obtido em 23/11/2007. Nota: É difícil precisar um número exato, uma vez que os associados filiam-se a todo o momento.

<sup>24</sup> Dado obtido em 23/11/2007

- Coordenadora do GT 5 (Política e Economia da Informação) – Dra. Sarita Albagli (IBICT)
- Coordenador do GT 6 (Informação, Educação e Trabalho) – Dr. Francisco das Chagas de Souza (UFSC)
- Coordenadora do GT 7 (Produção e Comunicação da Informação em CT&I) – Dra. Ida Regina Chittó Stumpf (UFRGS)

#### 4.4 Instrumentos de Coleta

Como já informado, foram elaborados dois instrumentos de coleta distintos, a saber: um questionário destinado aos pesquisadores em geral (Apêndice A) e um segundo questionário dirigido apenas aos 7 (sete) coordenadores dos Grupos de Trabalho da ANCIB (Apêndice B). O primeiro instrumento, composto de 14 questões, visou monitorar as principais características de uso dos periódicos eletrônicos da área da Ciência da Informação por meio de questões fechadas, em sua maioria. Ao mesmo tempo, procurou-se investigar o conhecimento dos pesquisadores a respeito da tecnologia RSS.

O questionário respondido pelos coordenadores dos GTs foi constituído por 11 itens, com um bloco de questões comuns sobre acesso e uso de periódicos. Este instrumento se distingue fundamentalmente do primeiro pela adição de duas Escalas Lickert, respectivamente de 4 e de 5 pontos. Esse recurso permite que os respondentes indiquem mais facilmente o grau de preferência em relação aos elementos investigados. A primeira escala tem o propósito de monitorar a importância dos atributos (características) dos artigos de periódicos eletrônicos. A segunda escala, por sua vez, visa identificar os tipos de conteúdos que, segundo esses pesquisadores, poderiam ser disseminados como canais RSS em periódicos científicos eletrônicos.

#### 4.4.1 Características do questionário eletrônico

**Parte I – Dados de identificação:** são seis itens que reúnem os dados demográficos e acadêmicos da população estudada, tais como: cidade/estado, faixa etária, titulação mais elevada e perfil de atuação em relação às revistas científicas.

**Parte II – Comportamento Informacional:** formado pelo grupo de questões de 2 a 8, busca verificar a frequência e uso dos periódicos científicos eletrônicos, uso de serviços de alerta, uso de ferramenta de busca disponibilizada pelos periódicos eletrônicos e a forma de leitura dos artigos.

**Parte III – Tecnologia RSS:** formado pelo grupo de questões de 9 a 12, elaborado para levantar o grau de conhecimento da população estudada em relação à tecnologia RSS. Em caso positivo, identifica os tipos de *feeds* RSS preferidos pelos cientistas da informação e questiona sobre eventuais barreiras em relação a adoção de RSS em periódicos científicos. Por fim, a questão 13 tem o objetivo de verificar se os cientistas da informação gostariam de realizar buscas simultâneas em mais de uma fonte de informação (característica dos serviços de agregação) e a questão 14 (aberta) visa colher elementos sobre as principais mudanças percebidas pelos respondentes nos procedimentos de acesso e uso da informação a partir da utilização de periódicos científicos eletrônicos.

#### 4.4.2 Características do questionário aplicado aos coordenadores de GTs:

**Parte I – Atributos funcionais dos artigos científicos:** Escala de Lickert elaborada para analisar, em grau de importância, os principais elementos que caracterizam o artigo científico eletrônico.

**Parte II – Acesso e uso:** são três questões criadas para verificar a utilização ou não de serviços de alerta, a forma de acesso aos periódicos eletrônicos, e a

freqüência de uso de ferramentas de busca disponibilizada pelos periódicos eletrônicos.

**Parte III – Tecnologia RSS:** são seis questões que buscam avaliar aspectos relacionados à utilização do padrão RSS nos artigos científicos de periódicos eletrônicos em Ciência da Informação.

#### 4.5 Pré-Teste

Antes da elaboração definitiva do questionário destinado aos pesquisadores da ANCIB realizamos uma fase de pré-teste a fim de identificar as melhores questões e também corrigir eventuais inconsistências neste instrumento. Esta etapa foi importante uma vez que pudemos perceber a necessidade de efetivar alguns ajustes antes da publicação do questionário *on-line*, como o agrupamento de questões e mesmo a mudança na redação de enunciados.

Os 11 questionários de pré-teste foram aplicados durante os meses de setembro e outubro de 2007, encaminhados via correio eletrônico e, em alguns casos, em formato impresso para alunos e professores de programas de pós-graduação em Ciência da Informação de diferentes universidades, entre elas a Universidade Federal da Paraíba, Universidade Federal de Minas Gerais, Universidade Federal Fluminense (RJ), além da própria Universidade de Brasília.

A partir da análise parcial das respostas recebidas, percebemos que um número elevado de pesquisadores (64%) considerou extremamente importante o recebimento de notificação automática para manterem-se atualizados sobre o lançamento de uma nova edição de um periódico da área. No entanto, 45% dos respondentes não costumam assinar tais serviços de alerta. Outros dados que nos chamaram a atenção foi o fato de que 64% dos respondentes entrevistados na fase de pré-teste conhecem ou já ouviram falar na tecnologia RSS. Todos eles afirmaram que assinariam *feeds* RSS, caso fossem disponibilizados pelos periódicos científicos eletrônicos.

## 4.6 Protótipo de Agregador de Conteúdos em CI

Trata dos procedimentos gerais adotados para a construção de um serviço de “agregação de conteúdo” temático desenvolvido especialmente para fins de demonstração da tecnologia RSS aplicada às fontes de informação de interesse da área da Ciência da Informação. A idéia do experimento surgiu a partir das pesquisas iniciais realizadas sobre as especificações RSS com o intuito de exercitar o processo de criação e disponibilização de *feeds*.

Durante os meses de fevereiro e março de 2007 foi realizado um levantamento das fontes de informação relacionadas à área da Ciência da Informação que, eventualmente, possuísem *feeds* RSS. Inicialmente, foram reunidas informações de alguns periódicos eletrônicos estrangeiros, *blogs* e repositórios digitais. Os conteúdos das revistas nacionais foram adicionados depois já que, na época, não existiam *feeds* disponibilizados pelas publicações científicas brasileiras da área.

Como exercício, optou-se pela criação manual desses arquivos. Para a construção do protótipo foi adotada a especificação RSS 2.0, já que não havia a preocupação com a representação dos metadados específicos para os artigos científicos. Se fosse este o caso, seria necessário o uso da flexibilidade do formato RDF, dando preferência à especificação RSS 1.0 (ver Capítulo 6).

### 4.6.1 Amostra

Os *feeds* dos periódicos nacionais foram desenvolvidos para um grupo de 10 periódicos eletrônicos da área da Ciência da Informação que utilizam o Sistema Eletrônico de Editoração de Revista (SEER). Trata-se de uma ferramenta voltada para administração do processo editorial e publicação de periódicos eletrônicos adaptada pelo IBICT do software desenvolvido pelo *Open Journal System* (OJS), desenvolvido pelo *Public Knowledge Project* (PKP) da Universidade de British Columbia, Canadá.

A preferência pelas revistas baseadas no OJS se justifica pelo fato deste ser um sistema já consolidado para publicação e gerenciamento de periódicos revisados por pares. Em março de 2007, os periódicos usando o OJS reuniam mais de 900 títulos<sup>25</sup>, incluindo as principais revistas nacionais, graças à iniciativa do IBICT em oferecer treinamento especializado aos editores científicos brasileiros. Outro ponto forte que contribuiu para a escolha dos periódicos baseados no SEER/OJS é que este sistema já dispõe de um *plug-in* para RSS desde sua versão 1.x, permitindo a geração automática de *feeds* na medida em que um novo artigo é publicado. A partir das versões 2.x, o suporte a RSS foi definitivamente incorporado pelo *software* como um módulo de sindicância. Uma vez ativado, o conteúdo dos artigos podem ser acessados a partir de três formatos de sindicância: RSS 1.0, RSS 2.0 e Atom.

Os periódicos nacionais selecionados nesta amostra foram: *Arquivística.net*, *Ciência da Informação*, *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia* (Instituto Brasileiro de Informação em Ciência e Tecnologia), *Em Questão* (Universidade Federal do Rio Grande do Sul), *Informação e Sociedade: Estudos* (Universidade Federal da Paraíba), *Informação e Informação* (Universidade Estadual de Londrina), *Perspectivas em Ciência da Informação* (Universidade Federal de Minas Gerais), *Revista ACB* (Associação Catarinense de Bibliotecários), *Revista Digital de Biblioteconomia e Ciência da Informação* (Universidade Estadual de Campinas) e *Transinformação* (Pontifícia Universidade Católica de Campinas).

É importante ressaltar que, à época em que este estudo foi realizado, ainda não havia sido publicadas no SEER outras revistas da área, as quais poderiam ter sido adicionadas à amostra, tais como a tradicional *Revista de Biblioteconomia de Brasília* (que teve o conteúdo das edições anteriores digitalizadas pelo Departamento de Ciência da Informação e Documentação da Universidade de Brasília) e a estreante *Ponto de Acesso*, lançada em junho de 2007 pelo departamento de Ciência da Informação da Universidade Federal da Bahia. Esta última, inclusive, possui o módulo de sindicância ativo.

---

<sup>25</sup> Uma lista dos periódicos suportados pelo sistema OJS encontra-se disponível na página: <http://pkp.sfu.ca/ojs-journals>

O Quadro 4.2 traz um levantamento das revistas eletrônicas da área construídas sob a plataforma SEER/OJS<sup>26</sup>, mesmo as que não tiveram seus *feeds* agregados no protótipo ou que foram publicadas posteriormente à realização desta etapa.

Como é possível verificar, metade das revistas em Ciência da Informação listadas utiliza versões superiores a 2.x da plataforma SEER/OJS, que simplifica o processo de sindicância. No entanto, apenas duas delas (Informação e Sociedade: estudos e Ponto de Acesso) habilitaram o módulo específico para distribuição de conteúdos por meio de RSS/Atom. Já a revista Ciência da Informação, ainda na versão 1.1.10, conseguiu disponibilizar seu *feed* RSS por meio da instalação de um *plug-in*.

Os resultados obtidos com o experimento e as análises preliminares verificadas encontram-se detalhadas nos itens 7.6 (p. 159) e 7.7 (p. 162) do Capítulo 7 – Análise dos Resultados.

---

<sup>26</sup> Levantamento realizado em 05/01/2008

Título	Editora	Versão SEER	URL da publicação	Possui feed?	Feed RSS
1. Arquivística.net		1.1.10	<a href="http://www.arquivistica.net/ojs/index.php">http://www.arquivistica.net/ojs/index.php</a>	Não	
2. Biblionline	UFPB	2.1.0.1	<a href="http://periodicos.ufpb.br/ojs2/index.php/biblio">http://periodicos.ufpb.br/ojs2/index.php/biblio</a>	Não	
3. Brazilian Journal of Information Science	UNESP (Marília)	1.1.10	<a href="http://www.portalppgci.marilia.unesp.br/bjis/index.php">http://www.portalppgci.marilia.unesp.br/bjis/index.php</a>	Não	
4. Ciência da Informação	IBICT	1.1.10	<a href="http://www.ibict.br/cienciadainformacao">http://www.ibict.br/cienciadainformacao</a>	RSS 2.0	<a href="http://www.ibict.br/cienciadainformacao/rss.php">http://www.ibict.br/cienciadainformacao/rss.php</a>
5. Em Questão	UFRGS	2.1.0	<a href="http://www.seer.ufrgs.br/index.php/EmQuestao">http://www.seer.ufrgs.br/index.php/EmQuestao</a>	Não	
6. Encontros Bibli	UFSC	2.1.1.0	<a href="http://periodicos.ufsc.br/index.php/eb/index">http://periodicos.ufsc.br/index.php/eb/index</a>	Não	
7. Informação e Cognição	UNESP (Marília)	1.1.7	<a href="http://www.portalppgci.marilia.unesp.br/reic">http://www.portalppgci.marilia.unesp.br/reic</a>	Não	
8. Informação e Informação	UEL	1.1.10	<a href="http://www.uel.br/revistas/informacao/">http://www.uel.br/revistas/informacao/</a>	Não	
9. Informação e Sociedade: Estudos	UFPB	2.1.0.1	<a href="http://www.ies.ufpb.br/ojs2/index.php/ies">http://www.ies.ufpb.br/ojs2/index.php/ies</a>	RSS 1.0 RSS 2.0 Atom	<a href="http://www.ies.ufpb.br/ojs2/index.php/ies/feed/rss">http://www.ies.ufpb.br/ojs2/index.php/ies/feed/rss</a>
10. LIINC	UFRJ/IBICT	1.1.10	<a href="http://www.ibict.br/liinc/">http://www.ibict.br/liinc/</a>	Não	
11. Perspectivas em Ciência da Informação	UFMG	2.1.1.0	<a href="http://www.eci.ufmg.br/pcion-line/">http://www.eci.ufmg.br/pcion-line/</a>	Não	
12. Pesquisa Bras. em CI e Biblioteconomia	IBICT	1.1	<a href="http://www.ibict.br/pbcib/index.php">http://www.ibict.br/pbcib/index.php</a>	Não	
13. Ponto de Acesso	UFBA	2.1.1.0	<a href="http://www.portalseer.ufba.br/index.php/revistaici/index">http://www.portalseer.ufba.br/index.php/revistaici/index</a>	RSS 1.0 RSS 2.0 Atom	<a href="http://www.portalseer.ufba.br/index.php/revistaici/feed/rss">http://www.portalseer.ufba.br/index.php/revistaici/feed/rss</a>
14. Revist@CRB-7	CRB-7	2.1.1.0	<a href="http://www.revistacr7.org.br/index.php/net">http://www.revistacr7.org.br/index.php/net</a>	Não	
15. Revista ACB	ACB	1.1.9	<a href="http://www.acbsc.org.br/revista/ojs/">http://www.acbsc.org.br/revista/ojs/</a>	Não	
16. Revista Brasileira de Biblioteconomia e Documentação	FEBAB	2.1.1.0	<a href="http://www.febab.org.br/rbbd/ojs-2.1.1/index.php/rbbd/index">http://www.febab.org.br/rbbd/ojs-2.1.1/index.php/rbbd/index</a>	Não	
17. Revista Digital de Biblioteconomia e Documentação	Unicamp	1.1	<a href="http://server01.bc.unicamp.br/seer/ojs/">http://server01.bc.unicamp.br/seer/ojs/</a>	Não	
18. Revista de Biblioteconomia de Brasília	CID/UnB	2.1.1.0	<a href="http://164.41.105.3/portalseer/ojs-2.1.1/index.php/RBB/index">http://164.41.105.3/portalseer/ojs-2.1.1/index.php/RBB/index</a>	Não	
19. Separatas para Cibertecários	Unesp (Rio Claro)	2.1.0.1	<a href="http://cecemca.rc.unesp.br/ojs/index.php/separatas">http://cecemca.rc.unesp.br/ojs/index.php/separatas</a>	Não	
20. Transinformação	PUC-Campinas	1.1.8	<a href="http://revistas.puc-campinas.edu.br/transinfo/index.php">http://revistas.puc-campinas.edu.br/transinfo/index.php</a>	Não	

**Quadro 4.2** – Periódicos científicos eletrônicos em Ciência da Informação disponíveis na plataforma SEER (12/2007)

ALMEIDA, R.L. (2008) *Disseminação de conteúdos na Web...* Dissertação (Mestrado) – CID/UnB.

# **CAPÍTULO 5**

## **Tecnología RSS**

---

## 5.1 Apresentação

O aumento vertiginoso de conteúdos na Web tem levado não somente o usuário comum, mas até mesmo pesquisadores e cientistas a utilizarem esta fonte de informação como recurso cada vez mais presente em seu dia-a-dia. No entanto, além dos problemas relacionados à confiabilidade da informação publicada na Web, verifica-se um aumento descontrolado do número de páginas. Assim, faz-se necessária a utilização de mecanismos de busca, recuperação e filtragem cada vez mais sofisticados.

O volume de informação disponível nos sítios da Internet também se configura como um problema recorrente para aqueles que fazem uso constante dessa tecnologia. Mesmo o usuário especializado encontra dificuldade em manter-se bem informado ou identificar eventuais atualizações feitas em determinado conteúdo, como o acompanhamento das edições eletrônicas de um periódico eletrônico e a leitura de novos comentários sobre os artigos publicados, por exemplo. O problema tende a aumentar ainda mais na medida em que se diversificam as fontes de informação utilizadas.

Diante desse cenário, verifica-se a emergência de serviços capazes de notificar automaticamente os usuários sobre novos conteúdos ou atualização recentes que surgem nas páginas Web previamente selecionadas, oferecendo facilidades à leitura de conteúdos disponibilizados digitalmente e promovendo uma gestão mais eficiente do tempo do usuário (BERNARDINO, 2006, p. 49).

É assim que a tecnologia RSS<sup>27</sup> – acrônimo de *RDF Site Summary*, *Rich Site Summary* ou, ainda, *Really Simple Syndication* – surge como uma solução para que os usuários tomem conhecimento de novos conteúdos provenientes das fontes de informação na Web sem precisarem acessar diretamente suas respectivas páginas. Inicialmente, a tecnologia foi projetada para permitir a distribuição e divulgação de notícias agrupadas de um conjunto diversificado de fontes de informação, mas

---

<sup>27</sup> A aplicação correta do significado da sigla RSS depende da especificação a ser utilizada, conforme explicado no item 4.2, sobre a história da evolução dos formatos de sindicância de conteúdos.

posteriormente a aplicação do RSS ampliou-se para qualquer tipo de informação digital, passível de atualizações.

Os tópicos seguintes tratarão das particularidades dessa tecnologia, tal como o desenvolvimento dos formatos de sindicância de conteúdos Web, a arquitetura do padrão RSS, os meios para sua utilização e eventuais barreiras na adoção deste recurso.

## 5.2 O Que é RSS?

O RSS é um conjunto de especificações voltadas para agregação e distribuição de conteúdos da Web, que facilita o processo de consulta e partilha de informação proveniente de diversas fontes de informação, periodicamente sujeitas a alterações ou atualizações (PILGRIM, 2002).

Tecnicamente, um dos principais trunfos dessa tecnologia reside em sua simplicidade, já que RSS nada mais é do que um arquivo-texto codificado dentro de um padrão compatível com o formato XML (*eXtensible Markup Language*). Este arquivo também é conhecido pelo nome de *feed* já que é “alimentado” constantemente, na medida em que ocorre alguma atualização no conteúdo. Assim, um arquivo RSS (*feed*) é, na realidade, uma lista constituída pelos elementos essenciais que descrevem uma determinada informação da Web: o título do documento, sua URL (*Uniform Resource Locator*, o endereço que localiza os sítios na Web) e uma breve descrição de seu conteúdo. Outros dados poderão ser acrescentados, melhorando a qualidade da informação referenciada.

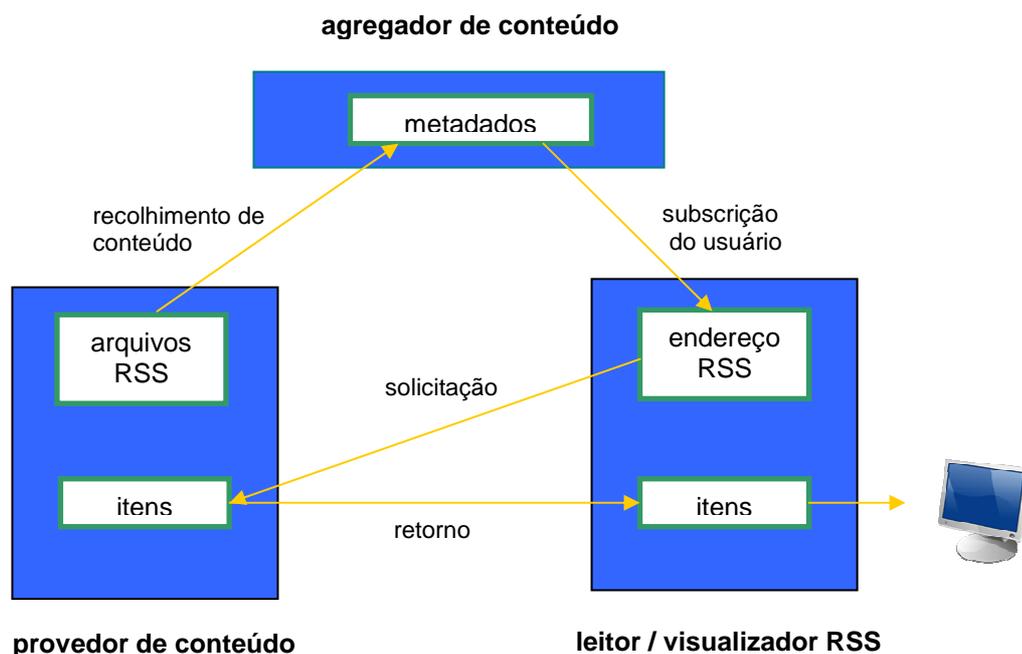
Uma vez formatado segundo o vocabulário RSS específico, o conteúdo Web está pronto para ser compartilhado e lido, juntamente com outras fontes de informação, por meio de uma categoria de *software* ou aplicações capazes de interpretar os documentos RSS. Esse processo de disponibilização de dados para a sua recuperação posterior recebe o nome de sindicância (*syndication*) de conteúdos, que é a chave para entender a tecnologia. O termo “sindicância” é comumente identificado pelas indústrias de entretenimento e de mídia como o método utilizado

para compartilhar um determinado conteúdo através de vários meios de comunicação de forma simultânea. Uma coluna de jornal publicada em diversos veículos espalhados em diferentes regiões do país é um exemplo de syndicação. No contexto da Web, trata-se do conjunto de atividades de coleta e replicação de conteúdos dinâmicos em ambientes digitais, promovendo a troca regular de informação atualizada entre diferentes páginas (ALMEIDA, 2007).

A economia de tempo com a leitura de informações praticamente personalizadas confere uma grande vantagem à utilização dos programas leitores de RSS, também conhecidos como “agregadores de conteúdo”, devido à sua característica de reunir, em um único ambiente, conteúdos variados produzidos por diversas fontes sem a necessidade de o usuário ter que acessar cada sítio individualmente. Por meio desses programas é possível fazer uma espécie de “assinatura” dos conteúdos desejados e examinar rapidamente as notícias de maneira condensada. Quando encontrar um item que desperte interesse, basta clicar no título para abrir a página com o texto integral.

Verificamos o uso intensivo dessa tecnologia principalmente em páginas cuja atualização ocorre com frequência, como acontece com os *blogs*, serviços de notícias e as tradicionais formas de publicação, como jornais e revistas. Porém, qualquer outra atividade que envolva atualização dinâmica de conteúdos pode se beneficiar dos recursos do RSS, notadamente serviços de notificação automática em geral, como os alertas.

A Figura 5.1 apresenta um esquema representativo da dinâmica de funcionamento de um sistema baseado em RSS.



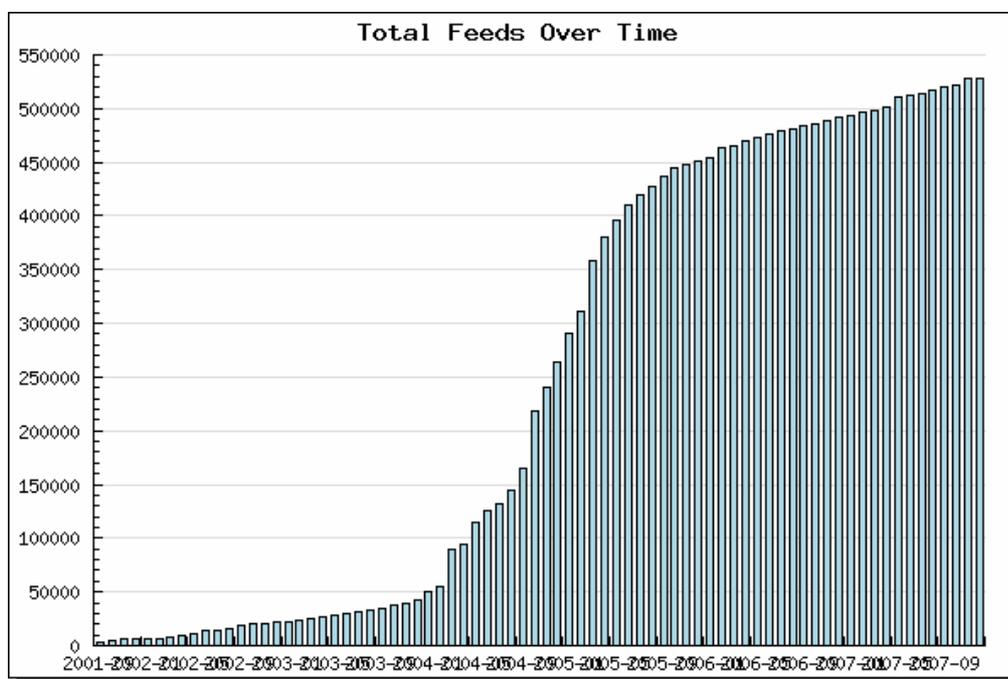
Fonte: adaptado de GUTIEREZ (2004)

**Figura 5.1** – Esquema de funcionamento de um sistema baseado em RSS

No Brasil, por enquanto o potencial do RSS ainda não é totalmente explorado, mas alguns jornais de circulação nacional, como a *Folha de S. Paulo*, o *Estado de S. Paulo* e *O Globo* e os grandes portais (*Globo.com*, *Terra* e *UOL*) começaram a disponibilizar *feeds* RSS para a maior parte de suas editorias. Em outros países, ao contrário, centenas de publicações eletrônicas já oferecem este recurso, apostando na agilidade deste novo tipo de leitura que inverte o fluxo normal da disseminação, já que, em vez do leitor procurar o que quer, são os assuntos dos quais ele gosta que vão até ele. Além de escolher os temas, o leitor também poderá definir a periodicidade com que desejará receber as atualizações e classificar as notícias de acordo com suas preferências, por meio de filtros. Uma outra característica interessante desse sistema é que os conteúdos lidos são apenas referenciados pelo documento RSS, ou seja, não existe cópia de informações nem apropriação indevida. O conteúdo continuará publicado em seu sítio original.

A Figura 5.2 mostra a proporção do crescimento total de *feeds* RSS desde setembro de 2001 até dezembro de 2007, de acordo com o Syndic8, um dos

diretórios de *feeds* mais importantes na Internet. Em dezembro, somente esse diretório contabilizava cerca de 528,5 mil *feeds* registrados.



Fonte: Syndic8 (<http://www.syndic8.com/stats.php#Actions>, acesso em: 30/12/2007)

**Figura 5.2** – Crescimento do número de *feeds* RSS entre 09/2001 a 12/2007

Além de facilitar a leitura das páginas Web, dispensando o acesso constante aos sítios em busca de atualizações, outra razão para a crescente popularidade da tecnologia RSS está relacionada à facilidade na implantação e utilização deste recurso sem nenhum custo, tanto para leitura quanto para a replicação de conteúdos. Incorporar um *feed* RSS em uma página para que esta também forneça informação agregada de outras fontes é algo relativamente simples de se fazer. Basta inserir um *script* (geralmente disponíveis nas linguagens de programação *JavaScript* ou PHP) que execute o arquivo XML do sítio de origem. Por se tratar de um documento que contém apenas texto ASCII (ao contrário de formatos binários), um arquivo RSS é facilmente interpretado pela maioria das aplicações de computador. Esta é uma prática comum das agências de notícias, como a CNN,

BBC, *Reuters*, etc, que permitem outras páginas Web incorporarem suas notícias e resumos por meio de acordos visando ampliar o número de leitores.

É importante frisar que a tecnologia RSS não é restrita apenas à distribuição de informação textual. Na verdade, as especificações RSS podem ser utilizadas para levar qualquer tipo de informação de um lugar a outro, não importando o tipo dos dados. Isso permite aplicações interessantes como, por exemplo, associar um *feed* a uma coleção de fotos do Flickr<sup>28</sup>. Nesse caso, toda vez que um conjunto de imagens for atualizado, o mesmo acontecerá com o *feed* correspondente. Da mesma forma, também podemos utilizar RSS nas aplicações de *podcasting* (transmissão de arquivos de áudio e vídeo) por meio da *tag* <enclosure>. O funcionamento é semelhante ao envio de uma mensagem de correio com anexo, porém, enquanto o anexo contém um arquivo, um *enclosure* irá oferecer uma ligação ao conteúdo, permitindo que o mesmo seja descarregado no próprio programa agregador ou em dispositivos portáteis (HAMMOND et al., 2004).

Tamanha flexibilidade e simplicidade encontrada no padrão RSS motivou a aplicação da tecnologia de sindicância de conteúdo em um número cada vez maior de portais e demais sítios da Web. Esse assunto, aliás, foi objeto de análise do pesquisador Paul Miller em artigo publicado na revista *Ariadne* intitulado *Syndicated content: it's more than just some file formats* (Conteúdos sindicados: muito mais do que simples formatos de arquivos). Miller (2003) aborda questões relacionadas à adoção da tecnologia de sindicância como meio de reutilização de recursos digitais e traz uma lista de boas práticas para criação e manutenção de *feeds* RSS por parte dos produtores de conteúdo.

### 5.3 Breve História da Sindicância de Conteúdos na Web

Na segunda metade dos anos 90, quando a Internet no Brasil ainda era novidade, algumas experiências com o envio de informações personalizadas passaram a ser realizadas, inicialmente com a utilização do correio eletrônico.

---

<sup>28</sup> <<http://www.flickr.com>>

Posteriormente surgiram os serviços de personalização e recorte de notícias na Web, a exemplo do *Crayon*<sup>29</sup> e o recurso “Active Channel”, desenvolvido pela Microsoft para o seu navegador Internet Explorer 4.0 a partir da especificação *Channel Definition Format* (CDF) para a descrição de páginas Web.

Esses dois projetos tinham em comum a missão de reunir em um só local (agregar) conteúdos variados e dispersos na Web por meio de uma tecnologia batizada como *push*. A idéia era “empurrar” informações selecionadas pelos usuários dos serviços em vez de esperar que eles visitassem os sítios para “puxar” (*pull*) os conteúdos que necessitavam. Embora tenha sido uma grande promessa na época, a tecnologia *push* não emplacou. Empresas especializadas, como a norte-americana PointCast, bem que tentaram popularizar esta tecnologia, mas o problema é que se tratava de serviços distintos e os programas utilizados eram complicados demais para os usuários em fase de adaptação ao ambiente Web, recém criado.

Em dezembro de 1997, porém, uma idéia aparentemente simples do programador Dave Winer, da UserLand, iniciou uma revolução no processo de personalização e distribuição de conteúdos Web, a partir do desenvolvimento do formato *ScriptingNews*. Este recurso permitiu reunir informações vindas de diferentes fontes para publicação em um único lugar: o pioneiro *blog* do próprio Winer<sup>30</sup>, aliás, em atividade até os dias de hoje. O *ScriptingNews* é considerado o formato de sindicância mais antigo da Web, tendo sido adotado por diversos sítios de notícias.

Nesta mesma época, Ramanathan Guha e outros pesquisadores da Apple Computer desenvolveram o *Meta Content Framework* (MCF), que consiste essencialmente em uma especificação de um formato para estruturar informação de metadados sobre sítios Web. Em março de 1999, quando o projeto foi descontinuado, Guha trocou a Apple pela Netscape, onde fez uma adaptação do

---

<sup>29</sup> Crayon (<http://crayon.net>) é abreviatura de *Create Your Own Newspaper*, um serviço de informação personalizada que oferece ao leitor a possibilidade de criar o seu próprio jornal com *links* para mais de uma centena de fontes de notícias disponíveis na Web.

<sup>30</sup> O *ScriptingNews* (<http://www.scripting.com>) teve seu primeiro post publicado em 1º de abril de 1997. É considerado um dos primeiros *blogs* de que sem tem notícia.

MCF para trabalhar com XML. E assim foi criada a primeira versão do formato *Resource Description Framework* (RDF), que viria a se tornar a base para a criação do primeiro protocolo de sindicância na Web, o RSS, um acrônimo para *RDF Site Summary*, em sua versão 0.90.

Em julho do mesmo ano, Dan Libby, também da Netscape, simplificou e aperfeiçoou o formato com o objetivo de utilizá-lo no serviço *My Netscape* como padrão para construção de sistemas de publicação de *headlines* (manchetes) em páginas pessoais dos usuários da Internet. A idéia básica era construir um resumo a partir do conteúdo de uma determinada página, com os respectivos *links* para as fontes de informação originais. Por isso, a nova versão RSS 0.91, baseada em XML, foi chamada de *Rich Site Summary*, numa referência a um de seus objetivos, que era sumarizar os sítios da Web.

No ano seguinte, a equipe de desenvolvedores da Netscape abandonou o projeto por achá-lo complexo e inviável para seus propósitos. Mas uma empresa de menor porte, a UserLand, de Dave Winer, decidiu dar continuidade ao desenvolvimento do RSS para aplicá-lo às suas ferramentas de diários eletrônicos, precursores dos *weblogs* ou *blogs* (PILGRIN, 2002).

Em agosto de 2000, paralelamente à iniciativa da UserLand, um outro grupo de programadores independentes (RSS-DEV *Working Group*), liderados por Rael Dornfest, da editora O'Reilly, projetou uma nova especificação, tirando partido da extensibilidade da sintaxe RDF e da utilização dos *namespaces*<sup>31</sup> que asseguravam a não coalizão entre elementos. Surgia, então, a versão RSS 1.0, que trazia de volta a maioria dos elementos das versões anteriores e baseava-se nos mesmos princípios que deram origem ao RSS versão 0.90 (antes da simplificação para a versão 0.91).

O grupo da UserLand, liderado por Winer, continuou seu trabalho com o desenvolvimento de outras versões para RSS como 0.92, 0.93, 0.94 até chegar à versão 2.0, em setembro de 2002 (a especificação 1.0 pertence a outro formato, não

---

<sup>31</sup> *namespaces* é um método usado pela linguagem XML que provê a qualificação de elementos e atributos usados na linguagem. Por meio deste recurso é possível incorporar a utilização de outros elementos na estrutura do RSS.

compatível com a série e, portanto, não se trata de uma atualização técnica). Com a especificação 2.0, a sigla RSS passou, então, a ter outro significado, o de *Really Simple Syndication* (sindicação realmente simples), já que seu foco é a simplicidade na distribuição de conteúdos.

Com a saída de Winer da UserLand, a responsabilidade pelo desenvolvimento do RSS 2.0 ficou à cargo do *Berkman Center* para Internet & Sociedade da Universidade de Harvard, que disponibilizou a tecnologia para domínio público, sob uma licença *Creative Commons*, em julho de 2003.

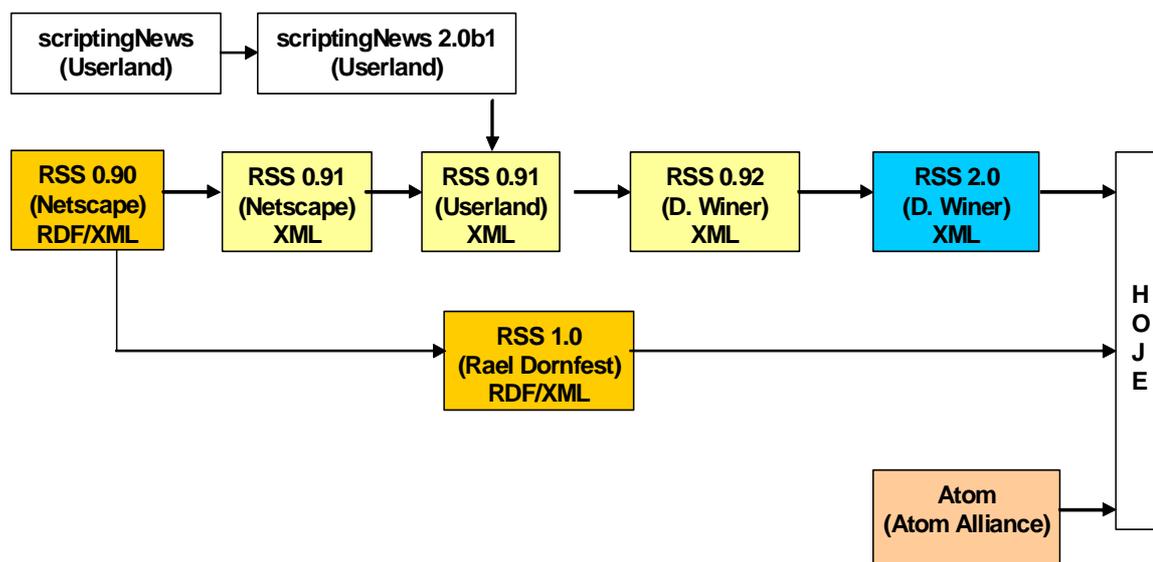
Em 2005, um grupo de desenvolvedores e provedores de serviços associados ao IETF – *Internet Engineering Task Force*, preocupado com as idas e vindas do desenvolvimento das especificações RSS, decidiu criar um novo formato para syndicação de conteúdo, denominado Atom (originalmente chamado *Echo* e *Pie*). Tal como o RSS, o Atom também procurou ser 100% neutro, aberto e extensível, facilitando sua implementação por qualquer desenvolvedor (WITTENBRINK, 2005, p. 34).

O Atom também é baseado na linguagem XML, mas seu desenvolvimento é tido como mais sofisticado. De acordo com especialistas, trata-se de uma proposta de unificação do padrão RSS 1.0 com o RSS 2.0 e poderá até substituí-lo, já que conta com o apoio de grandes corporações, como o Google, que o acolheu como padrão para o seu serviço de *blogs*. Porém, o fato é que o Atom desempenha as mesmas funções básicas do RSS e muitas pessoas usam genericamente o termo “RSS” como sinônimo de syndicação de conteúdos na Web, independente do padrão utilizado ser RSS ou Atom.

Conhecendo o histórico do desenvolvimento dos formatos de syndicação na Web compreendemos melhor a razão para os diferentes significados do acrônimo RSS. Assim, *RDF Site Summary* está associado às versões RSS 0.90 e RSS 1.0, enquanto *Really Simple Syndication* refere-se às especificações RSS 0.91, RSS 0.92, RSS 0.93, RSS 0.94. Já a especificação RSS 2.0 é conhecida pelo nome *Rich Site Summary*. A diferença básica entre essas versões está no fato de que tanto RSS 0.90 quanto a RSS 1.0 são aplicações desenvolvidas em conformidade com o

padrão RDF, recomendado desde 1999 pelo *W3 Consortium*, entidade internacional responsável pela gestão de protocolos da Web. Já as outras especificações de RSS baseiam-se exclusivamente na linguagem XML.

A escolha da versão a ser implementada depende do tipo de aplicação desejada. As versões baseadas em RDF são mais úteis para integração com outros sistemas, inclusive em aplicações de Web Semântica, devido à sua capacidade de modularização. No entanto, as versões baseadas em XML são mais simples e fáceis de serem aplicadas para diversas finalidades. A Figura 4.3 mostra, por meio de um esquema de blocos, a evolução do desenvolvimento das especificações RSS e Atom.



Fonte: SAUERS (2006, p. 126)

**Figura 5.3** – História do desenvolvimento dos *feeds* RSS e Atom

O Quadro 5.1, por sua vez, traz informações sobre os formatos de sindicância de conteúdos Web mais importantes, com os respectivos endereços de suas especificações.

<b>Formato</b>	<b>Data</b>	<b>Autor</b>	<b>URL da especificação</b>
MCF ( <i>Metadata Content Format</i> )	1995	Ramanathan Guha / Apple Computer	<a href="http://mcf.sourceforge.net/Docs/MCF/">http://mcf.sourceforge.net/Docs/MCF/</a>
CDF ( <i>Channel Definition Format</i> )	março 1997	Castedo Ellerman / Microsoft	<a href="http://msdn.microsoft.com/workshop/delivery/cdf/reference/CDF.asp">http://msdn.microsoft.com/workshop/delivery/cdf/reference/CDF.asp</a>
<i>Meta Content Format</i> (usando XML)	junho 1997	R. V. Guha/ Netscape  Tim Bray/ Textuality	<a href="http://www.w3.org/TR/NOTE-MCF-XML">http://www.w3.org/TR/NOTE-MCF-XML</a>
<i>Scripting News</i>	dezembro de 1997	Dave Winer/  UserLand	<a href="http://davenet.scripting.com/1997/12/15/scriptingNewsInXML">http://davenet.scripting.com/1997/12/15/scriptingNewsInXML</a>
RSS 0.90	março 1999	Netscape	<a href="http://www.purplepages.ie/RSS/netscape/rss0.90.html">http://www.purplepages.ie/RSS/netscape/rss0.90.html</a>
RSS 0.91	julho 1999	Dan Libby / Netscape	<a href="http://www.purplepages.ie/RSS/netscape/rss0.91.html">http://www.purplepages.ie/RSS/netscape/rss0.91.html</a>
RSS 0.91 (UserLand)	abril 2000	Dan Libby/ Netscape  Dave Winer/ UserLand	<a href="http://backend.userland.com/rss091">http://backend.userland.com/rss091</a>
RSS 1.0	agosto 2000	Rael Dornfest/ O'Reilly et al.	<a href="http://Web.resource.org/rss/1.0/">http://Web.resource.org/rss/1.0/</a>
OPML ( <i>Outline Processor Markup Language</i> )	setembro 2000	Dave Winer/ UserLand	<a href="http://www.opml.org/spec">http://www.opml.org/spec</a>
RSS 0.92	dezembro 2000	Dave Winer/ UserLand	<a href="http://backend.userland.com/rss092">http://backend.userland.com/rss092</a>
RSS 0.93 ( <i>draft</i> )	abril 2001	Dave Winer/ UserLand	<a href="http://backend.userland.com/rss093">http://backend.userland.com/rss093</a>
RSS 2.0	setembro 2002	Dave Winer	<a href="http://blogs.law.harvard.edu/tech/rss">http://blogs.law.harvard.edu/tech/rss</a>
Atom 1.1	agosto 2005	Mark Nottingham, Richard Sayre et al.	<a href="http://www.ietf.org/rfc/rfc4287.txt">http://www.ietf.org/rfc/rfc4287.txt</a>

**Quadro 5.1** – Principais formatos de sindicaco de contedo (WITTENBRINK, 2005)

## 5.4 Arquitetura de um *feed* RSS

Independente de como são chamados ou do número de suas versões, *feeds* RSS são todos baseados nas regras da linguagem XML. Para aquelas pessoas já familiarizadas com o código HTML (*HyperText Markup Language*), utilizada na construção de páginas Web, a estrutura de um arquivo RSS lhe parecerá familiar (SAUERS, 2006). Porém, diferentemente da linguagem HTML – que se limita a fornecer um formato universal para representar a informação sem acrescentar referência quanto à estrutura e significado (semântica) dos dados contidos – o padrão RSS é capaz de representar informação sobre recursos Web.

No entendimento de Nottingham (2005), as especificações RSS são linguagens de representação de metadados essenciais que permitem a distribuição de resumos de sítios e listas de *links* em tempo real com o propósito de fornecer aos usuários informações rápidas para ajudar a decidirem se querem ou não ter acesso ao conteúdo completo. Em sua formatação mais simples, o conjunto de metadados utilizado por qualquer versão de RSS consiste no título, seguido de um *link* (localização da página) e uma breve descrição (resumo) do conteúdo a ser disponibilizado.

O uso de *tags* (rótulos) específicos, como `<title>`, `<link>` e `<description>` permite tratar cada unidade de informação como um objeto distinto, possibilitando estruturar um determinado conteúdo para que esse seja interpretado e tratado por *scripts* ou programas especiais de modo que esses dados se transformem em objetos qualificados como atributos. Tem-se, então, a possibilidade de reutilização automatizada da informação, podendo compartilhá-la mais facilmente com os usuários e/ou com outros sistemas de informação (interoperabilidade), organizá-las em bancos de dados e realizar pesquisas automáticas. Como uma aplicação XML, a idéia geral é que tudo aquilo que seja passível de ser descrito por meio de `<tags>`, é candidata a ser integrada via *feed* RSS.

Tais características se identificam com os princípios da Web Semântica, que, segundo Bernes-Lee et al. (2001) reúne um conjunto de modelos conceituais que visam propor soluções para a melhor compreensão e gerenciamento dos conteúdos armazenados na Web atual, independente da forma em que estes se apresentem, seja texto, som, imagem e gráficos. Para o autor, considerado “pai” da Web e um dos precursores da Web Semântica, o princípio deste novo paradigma está localizado na valoração dos conteúdos e na capacidade de novas funções, tais como a interpretação e processamento de informações, além do compartilhamento de resultados com outros programas. O fato é que a concepção original de Berners-Lee sobre a Web estabelecia a partilha estruturada de informação, em vez de “um caleidoscópio de leitura”, como acabou acontecendo (HAMMOND et al., 2004).

Enquanto as pesquisas sobre o tema seguem em plena ebulição, há quem afirme que a tecnologia RSS já represente a prática da Web Semântica mais difundida fora do círculo acadêmico (DOWNES, 2005). De acordo com o pesquisador, RSS é o que surgiu, de fato, para descrição de conteúdos em tempo real, sendo utilizado por milhões de sítios na Web para descrever não somente recursos, mas até mesmo pessoas, lugares e objetos. Por ser baseado no padrão XML, o RSS também permite a troca de informações entre diversas plataformas e possibilita a descrição de dados em arquivos-texto.

A arquitetura de um *feed* RSS segue a especificação da linguagem XML 1.0 (WINER, 2003), e seu vocabulário é composto de elementos, sub-elementos, atributos, etc. Sua estrutura tem como elemento principal `<rss>`, juntamente com o atributo *version*, que indica a versão da especificação RSS.

Os conteúdos, por sua vez, são organizados por meio de uma hierarquia de dois níveis de informação, designadamente as coleções e itens individuais. As coleções correspondem ao sub-elemento denominado `<channel>` (canal), enquanto os itens individuais de informação, inseridos dentro da coleção, correspondem ao elemento `<item>`, que representa um objeto individual de informação que compõe o recurso a ser descrito. Ou seja, um canal consiste na descrição de um recurso que, por sua vez, pode ser constituído por um ou vários itens.

Um *feed* RSS poderá reunir vários sub-elementos *<item>*, representando informações relacionadas a seu conteúdo, tais como *<title>* (título), *<description>* (descrição), *<language>* (idioma), entre outros, podendo opcionalmente ser adicionada metainformação no processo de descrição do recurso.

O modelo básico comum a qualquer especificação RSS segue as seguintes linhas de desenvolvimento, segundo Kyrnyn (2004):

1. Trata-se de um documento baseado nas regras de XML, logo o mesmo deve ser bem formado (*well-formed*). Todas as *tags* devem ser fechadas;
2. O primeiro elemento de um documento RSS é o *<rss>*. Isso inclui um atributo que especifica a versão utilizada, que é obrigatório;
3. O próximo elemento é o *<channel>*, que contém os metadados que descrevem o próprio canal, em geral um título, uma breve descrição e a URL do recurso descrito. Normalmente a URL definida pode ser a mesma da página Web, ou endereço do *feed* RSS;
4. O elemento *<title>* contém a informação sobre o título do recurso descrito ou mesmo de um *Item* corrente;
5. O elemento *<link>* indica a URL da página Web que corresponde ao *feed* RSS. Se estiver sendo utilizado dentro do elemento *<item>*, então *<link>* refere-se a URL de um conteúdo específico.
6. O elemento *<description>* serve para descrever todo o *feed* RSS ou um determinado *Item*.
7. O elemento *<item>* especifica cada artigo ou conteúdo dentro do documento RSS. Os sub-elementos necessários para descrição deste elemento são: *<title>*, *<description>* e *<link>*, podendo ser adicionadas

opcionalmente mais metainformação. As especificações das duas variantes do RSS exigem que exista pelo menos um *Item*.

Independentemente da versão RSS adotada, as considerações acima devem ser observadas na criação de documentos RSS. Usando esses elementos básicos, temos a estrutura de um *feed* RSS em sua forma mais simples, como no código exibido na Figura 5.4.

```
<?xml version= "1.0"?>
<rss version = "0.92">
  <channel>
    <title> Agregação de conteúdo na Web</title>
    <link> http://www.cid.unb.br/exemplo_feed.htm </link>
    <description>
      Este é um exemplo de um feed RSS básico
    </description>

    <item>
      <title>Esse é o primeiro item</title>
      <link>http://www.cid.unb.br/exemplo_feed.htm </link>
      <description>
        Esse é o primeiro item
      </description>
    </item>

    <item>
      <title>Esse é o segundo item</title>
      <link>http://www.cid.unb.br/exemplo_feed2.htm </link>
      <description>
        Esse é o segundo item
      </description>
    </item>

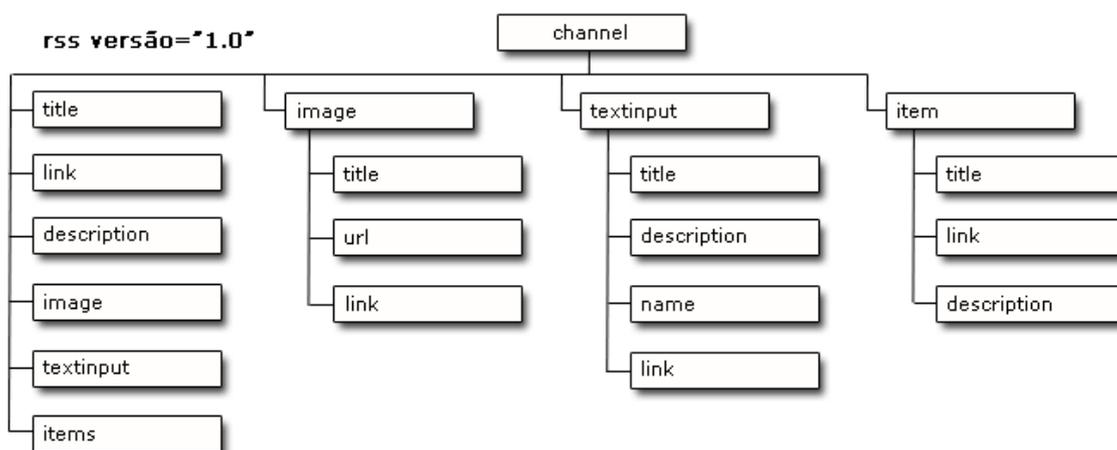
  </channel>
</rss>
```

**Figura 5.4** – Estrutura básica de um *feed* RSS (versão 0.92)

Apesar dos pontos em comum, existem diferenças estruturais entre as especificações RSS 1.0 e RSS 2.0, assim como nas suas versões descendentes. A criação manual de conteúdos na versão 2.0 é bem mais fácil, porém é considerada semanticamente pobre. Ao contrário, o que torna o RSS 1.0 mais interessante, é o fato desta especificação permitir interoperabilidade com outras linguagens

compatíveis com RDF/XML, facilitando o seu processamento por meio de outras máquinas, além de oferecer a possibilidade de extensão com outros vocabulários de metadados, proporcionando descrição semântica mais rica para recursos Web.

Para marcarmos bem a diferença em termos de arquitetura das duas principais correntes de RSS, detalhamos, a seguir, os modelos representativos dos elementos e a descrição dos sub-elementos requeridos e opcionais dessas versões. Os elementos básicos do RSS 1.0 encontram-se representados por meio de um modelo visualizado na Figura 5.5. Seus elementos requeridos e opcionais são detalhados nos Quadros 5.2 e 5.3.



**Figura 5.5** – Modelo representativo da especificação RSS 1.0 (HAMMERSLEY, 2003)

Sub-Elementos Requeridos do RSS 1.0	
<title>	Descrição do título do <i>feed</i> . Deve possuir, no máximo, 40 caracteres.
<link>	A URL relacionada ao site. Deve possuir, no máximo, 500 caracteres.
<description>	Palavras que descrevem o <i>channel</i> . A descrição tem que ser clara e objetiva.
<items>	Um índice de RDF contendo “sub-elementos” de um feed. Esses sub-elementos são geralmente organizados através do elemento <i>seq</i> do RDF com uma lista de recursos.

Sub-elementos Item	
<item>	Este elemento representa o conteúdo disponibilizado pelo <i>feed</i> . É permitida somente a utilização de 15 itens.
<title>	Título do conteúdo. Deve possuir, no máximo, 100 caracteres.
<link>	URL para o conteúdo. Deve possuir, no máximo, 500 caracteres.
<description>	Descrição ou resumo sobre o conteúdo. Deve possuir, no máximo, 500 caracteres.

**Quadro 5.2** – Descrição dos sub-elementos requeridos do RSS 1.0

**Obs.** No elemento *Item* foram mantidos somente os sub-elementos que descrevem informações sobre o título, *link* e descrição. É possível, no entanto, agregar informações adicionais por meio de módulos, como os metadados *Dublin Core*.

Sub-elementos opcionais do RSS 1.0	
<image>	Imagem associada ao <i>feed</i> . Ex: <image rdf:resource="http://xml.com/universal/images/xml_tiny.gif" />
<textinput>	Associa o elemento <i>textInput</i> a um <i>feed</i> particular.
<image>	Este elemento possui sub-elementos que representam informações sobre a imagem do <i>feed</i> .
<url>	A URL de arquivos de imagens GIF, JPG ou PNG correspondentes ao <i>feed</i> . Deve possuir, no máximo, 500 caracteres.
<title>	Descrição da imagem. É utilizado para o atributo ALT da tag HTML <img>. Deve possuir, no máximo, 40 caracteres.
<link>	URL para qual a imagem deve ser redirecionada. Deve possuir, no máximo, 500 caracteres.
<textInput>	Informações utilizadas para formulário. É utilizado, por exemplo, para um formulário de busca.
<title>	Descrição do botão de submissão. Deve possuir, no máximo, 100 caracteres.
<description>	Texto que descreve a utilização do <i>textInput</i> . Deve possuir, no máximo, 500 caracteres.
<name>	Nome do campo que possui os dados a serem submetidos para um script (PERL, PHP, ASP...). Deve possuir, no máximo, 20 caracteres.
<link>	URL do script para onde será submetido. Deve possuir, no máximo, 500 caracteres.

**Quadro 5.3** – Descrição dos sub-elementos opcionais no RSS 1.0

Os elementos básicos da versão RSS 2.0 (uma atualização da série de especificações RSS 0.9x), estão representados no modelo visualizado na Figura 5.6. Os elementos que diferem do RSS 0.9x aparecem em destaque. Os Quadros 5.4 e 5.5 apresentam os elementos requeridos e opcionais do RSS 2.0.

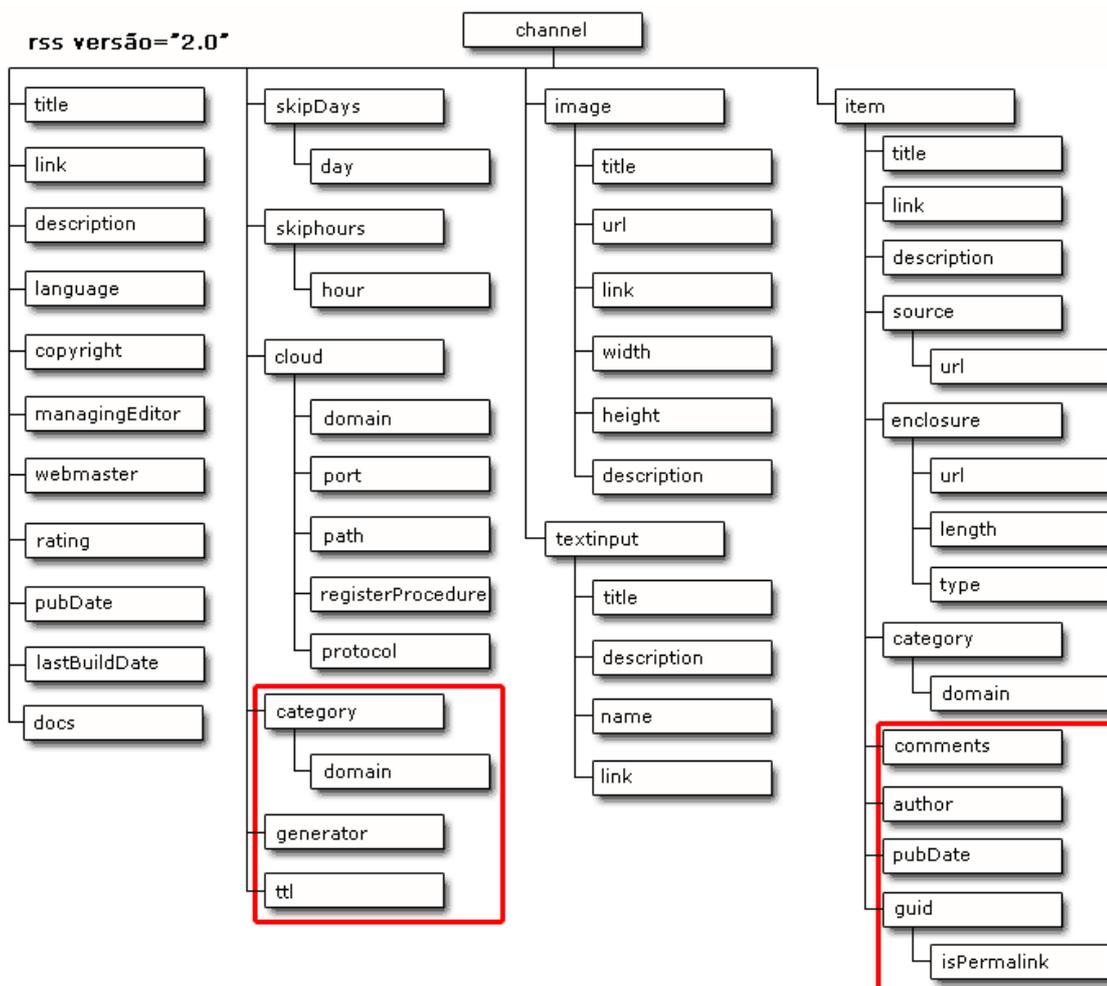


Figura 5.6 – Modelo representativo da especificação RSS 2.0 (HAMMERSLEY, 2003)

Sub-elementos requeridos do RSS 2.0	
<title>	Descrição do título do <i>feed</i> . Deve possuir no máximo 100 caracteres.
<link>	A URL relacionada ao site. Deve possuir no máximo 500 caracteres.
<description>	Palavras que descrevem o <i>channel</i> . A descrição tem que ser clara e objetiva.
<language>	Linguagem na qual o <i>feed</i> foi escrito.

<image>	Este elemento possui sub-elementos que representam informações sobre a imagem do <i>feed</i> .	
	<url>	A URL de arquivos de imagens GIF, JPG ou PNG correspondentes ao <i>feed</i> . Deve possuir no máximo 500 caracteres.
	<title>	Descrição da imagem. É utilizado para o atributo ALT da tag HTML <img>. Deve possuir no máximo 100 caracteres.
	<link>	URL para qual a imagem deve ser redirecionada. Deve possuir no máximo 500 caracteres.
	<width> e <heigth>	Informa a altura ( <i>heigth</i> ) e largura ( <i>width</i> ) da imagem.
	<description>	Descrição adicional sobre a imagem.
<b>Sub-elementos Item</b>		
<item>	Este elemento representa o conteúdo disponibilizado pelo <i>feed</i> . É permitido somente a utilização de 15 itens.	
	<title>	Título do conteúdo. Deve possuir no máximo 100 caracteres.
	<link>	URL para o conteúdo. Deve possuir no máximo 500 caracteres.
	<description>	Descrição ou resumo sobre o conteúdo. Deve possuir, no máximo, 500 caracteres.
<source>	Nome do <i>site</i> onde o conteúdo ( <i>item</i> ) foi retirado. Ex: <source url="http://scriptingnews.userland.com/xml/scriptingNews2.xml">Scripting News</source>	
	<url>	URL do <i>feed</i> .
<enclosure>	Elemento utilizado para disponibilizar arquivos adicionais ao <i>item</i> . Este elemento possui somente atributos. Ex: <enclosure url="http://www.scripting.com/mp3s/weatherReportSuite.mp3" length="12216320" type="audio/mpeg" />	
	<url>	URL do arquivo.
	<length>	Tamanho em <i>bytes</i> do arquivo.
	<type>	Formato ( <i>MIME type</i> ) do arquivo.
<category>	Nomes das categorias em que o <i>item</i> faz parte. Ex: <category domain="http://www.cid.unb.br/">CID</category>	
	<domain>	URL que identifique as categorias

Quadro 5.4 – Descrição dos sub-elementos requeridos do RSS 2.0

<b>Sub-elementos opcionais do RSS 2.0</b>		
<copyright>	Nome do responsável que detém os direitos autorais sobre as informações do <i>feed</i> . Deve possuir, no máximo, 10 caracteres	
<managingEditor>	O endereço de e-mail do editor. Deve possuir, no máximo, 100 caracteres.	
<webMaster>	O endereço de e-mail do <i>webmaster</i> . Deve possuir, no máximo, 100 caracteres.	

<rating>	A taxa de PICS ( <i>Platform for Internet Content Selection</i> ) para o <i>feed</i> . Deve ter, no máximo, 500 caracteres. Mais informações sobre PICS podem ser encontradas em <a href="http://www.w3c.org/PICS">http://www.w3c.org/PICS</a> .	
<pubDate>	Data de publicação do <i>feed</i> , no padrão RFC 822.	
<lastBuildDate>	Data e hora, no estilo RFC 822, da última alteração do <i>feed</i> .	
<docs>	URL que disponibiliza informações sobre o padrão utilizado para uma referência futura (Ex: <a href="http://backend.userland.com/rss091">http://backend.userland.com/rss091</a> ).	
<skipDays> e <skipHours>	O <i>skipDays</i> representa o dia da semana e o <i>skipHours</i> representa a hora (1-24 GMT ( <i>Greenwich Mean Time</i> )) em que o <i>feed</i> foi lido.	
<textInput>	Informações utilizadas para formulário. É utilizado, por exemplo, para um formulário de busca.	
	<title>	Descrição do botão de submissão. Deve possuir, no máximo, 100 caracteres.
	<description>	Texto que descreve a utilização do <i>textInput</i> . Deve possuir, no máximo, 500 caracteres.
	<name>	Nome do campo que possui os dados a serem submetidos para um <i>script</i> (PERL, PHP, ASP...). Deve possuir, no máximo, 20 caracteres.
	<link>	URL do <i>script</i> . Deve possuir, no máximo, 500 caracteres.
<category>	Nomes das categorias em que o <i>feed</i> faz parte.	
	<domain>	URL que identifique as categorias.
<generator>	Nome do programa que gerou o <i>feed</i> .	
<ttl>	Valor em minutos que representa o tempo em que um <i>feed</i> permanecerá sem ser alterado.	
<b>Sub-elementos Item</b>		
<comments>	URL de uma página para comentários relativos ao <i>item</i> .	
<author>	Endereço de E-mail do autor do <i>item</i> .	
<pubDate>	Data de publicação do <i>item</i> , no padrão RFC 882.	
<guid>	URL que referencie o <i>item</i> , ou seja, um identificador único que pode ser usado para verificar se é um novo <i>item</i> . Esta URL pode ou não ser o <i>link</i> do <i>item</i> .	
	<isPermaLink>	É uma condição ( <i>true</i> ou <i>false</i> ) que indicará se a URL utilizada é o <i>link</i> do <i>item</i> ou não.

Quadro 5.5 – Descrição dos sub-elementos opcionais do RSS 2.0

## 5.5 Funcionamento Básico da Tecnologia RSS

A maneira mais simples para que qualquer usuário se beneficie da tecnologia RSS é dispor de um *software* “agregador de conteúdo”, um tipo de aplicativo capaz de ler e interpretar documentos RSS, mostrando os itens novos ou atualizados. Tais programas geralmente são gratuitos, de fácil instalação e, na maior parte, se

assemelham a um leitor de correio eletrônico. Também é possível ler e gerenciar os *feeds* RSS a partir de serviços baseados na Web. Neste caso, basta cadastrar-se, criar uma conta (*login* e senha) e acessar o serviço toda vez que desejar receber as atualizações sobre os assuntos escolhidos. A vantagem dessa modalidade em relação aos programas para *desktop* é que o usuário poderá acessar toda a sua coleção de *feeds* a partir de qualquer computador, independente do local onde esteja.

Nos dois casos, porém, é possível fazer uma espécie de “assinatura” dos conteúdos de diversas fontes de informação e examinar rapidamente as novidades ou atualizações realizadas em determinada página que disponha de um *feed* RSS, de acordo com a periodicidade desejada. As páginas Web que já fornecem esse serviço, em geral, utilizam um ícone padronizado de cor laranja  para mostrar que ali existe conteúdo passível de ser assinado. Contudo, é possível encontrar outros ícones estilizados com as letras RSS, XML acompanhados (ou não) do número de sua versão como, por exemplo: .

Alguns navegadores, como o Mozilla Firefox, Internet Explorer (a partir da versão 7.0), Opera 7.5 e Safari 2.0 são capazes de reconhecer automaticamente quando uma página Web possui um ou mais canais RSS. A Figura 5.7 mostra a página da CAPES visualizada pelo Mozilla Firefox, que indica a existência de conteúdo RSS.

A subscrição em um canal RSS normalmente é feita utilizando um método quase manual: copiando-se a URL do *link* completo do *feed* (geralmente com extensão .xml ou .rss) e colando-o de volta no aplicativo leitor de RSS. No entanto, navegadores que suportam RSS, como o Mozilla Firefox, oferecem um jeito mais amigável. Ao acessar um *feed* RSS, o navegador formata uma página especial e exibe um botão que dá ao usuário a opção de inscrever-se naquele *feed* ou não (Figura 5.8).

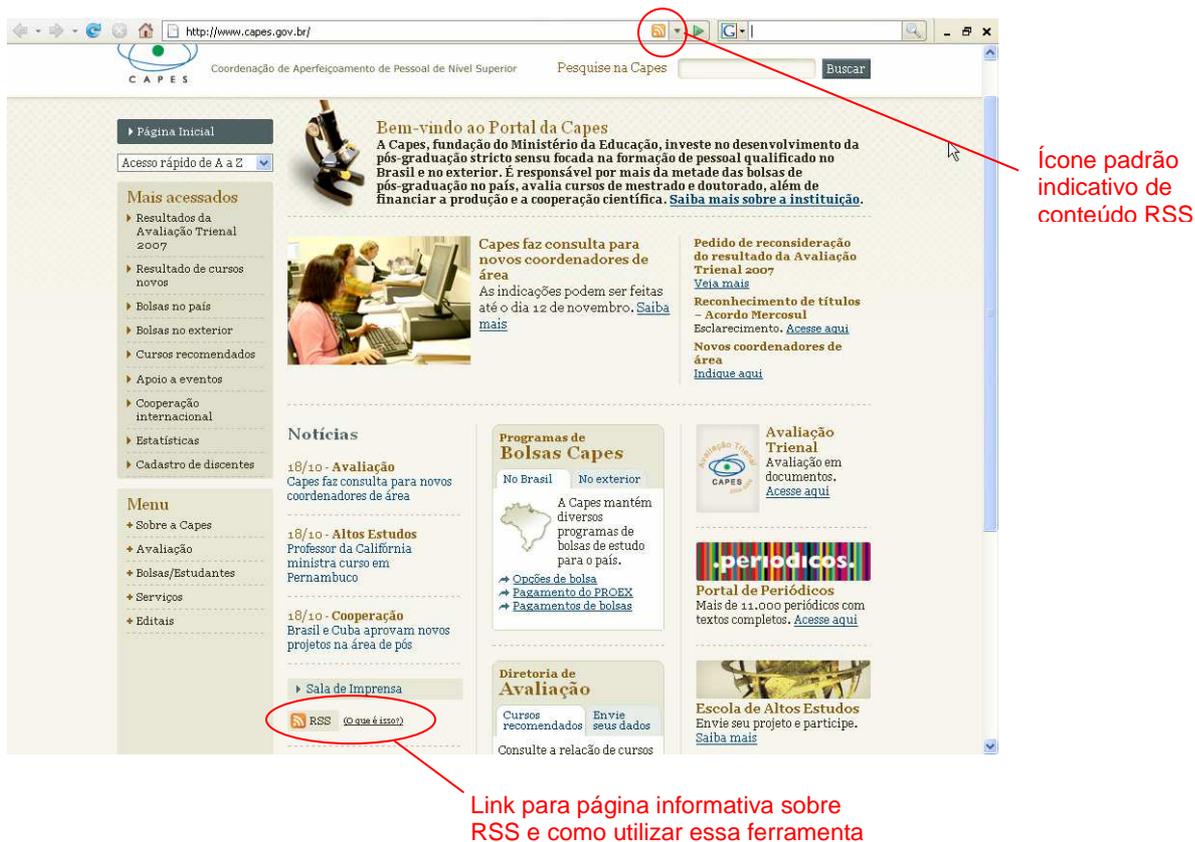


Figura 5.7 – página inicial da CAPES visualizada pelo navegador Mozilla Firefox

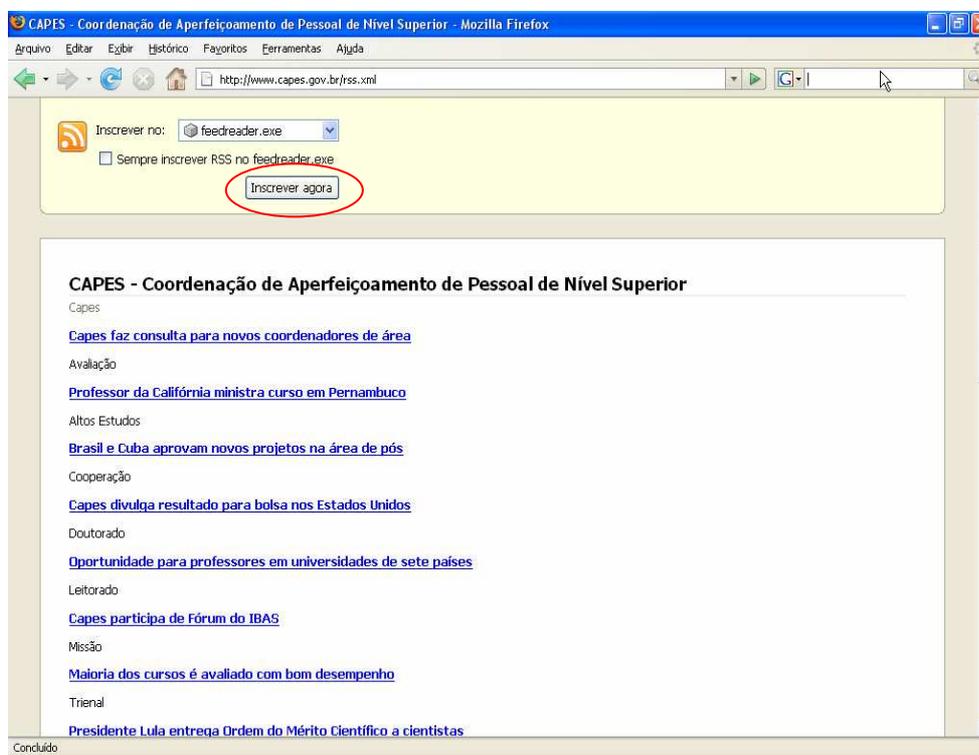


Figura 5.8 – Exibição do feed RSS da CAPES pelo navegador Mozilla Firefox

A Figura 5.9 mostra uma tela típica de um programa leitor de RSS para *desktop* (computador pessoal), em que é possível ver, na coluna da esquerda, todos os *feeds* adicionados, estruturados em forma de árvore (a organização depende exclusivamente do usuário). A leitura dos *feeds* é feita do lado direito da tela, que é dividida em duas partes: acima se localizam os tópicos, ou as manchetes das notícias (*headlines*) enquanto na parte inferior lê-se um resumo ou, em alguns casos, o texto integral do item selecionado. O ato de “clique” duas vezes no título fará com que, na janela inferior, apareça a informação original, do modo com que foi publicada na Web.

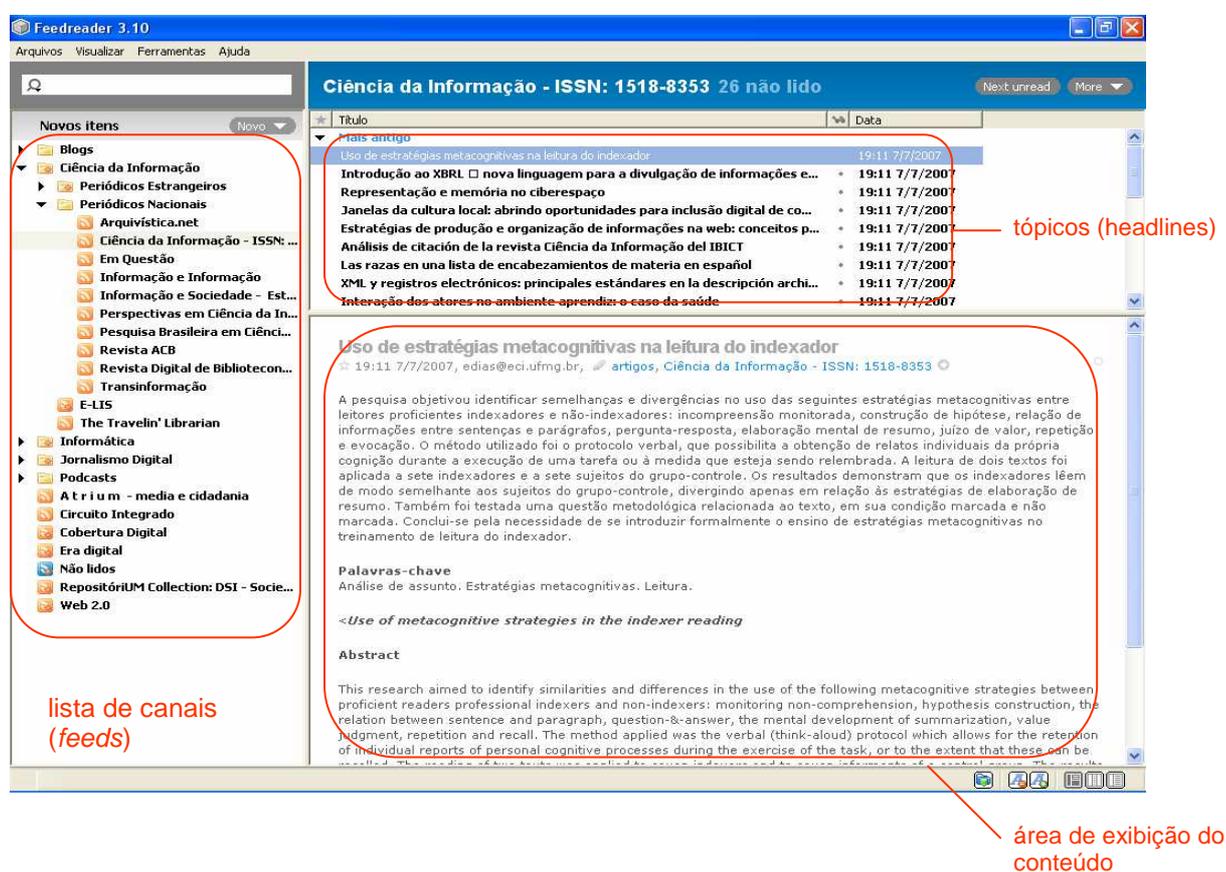


Figura 5.9 – tela do FeedReader, exemplo de um aplicativo agregador de conteúdo

Atualmente, existe uma variedade de aplicações voltadas para a leitura e gerenciamento de *feeds* RSS em *desktops*. Em seguida são apresentados alguns exemplos:

### Para Windows

- RSSReader  
<http://www.rssreader.com>
- FeedReader  
<http://www.feedreader.com>
- FeedDemon  
<http://www.newsgator.com/Individuals/FeedDemon>

### Para Macintosh

- NewsMac  
<http://www.thinkmac.co.uk/newslife>
- NetNewsWire  
<http://www.newsgator.com/Individuals/NetNewsWire>

### Para Linux

- Straw  
<http://www.gnome.org/projects/straw>
- Syndicator  
<http://syndicator.sourceforge.net>

Conforme já comentado, outra maneira de fazer a leitura de *feeds* RSS se dá por meio de aplicações baseadas na Web, tais como os serviços *Google Reader*<sup>32</sup>, cuja tela pode ser vista na Figura 5.10, *Bloglines*<sup>33</sup>, *NewsisFree*<sup>34</sup>, entre outros. Tais serviços, inclusive, vêm ganhando a preferência dos usuários como método de recebimento e leitura de *feeds*.

---

<sup>32</sup> <<http://www.google.com.br/reader>>

<sup>33</sup> <<http://www.bloglines.com>>

<sup>34</sup> <<http://www.newsisfree.com>>



Figura 5.10 – Tela do Google Reader, leitor de RSS baseado na Web

Para Çelikbas (2005), a assinatura de *feeds* RSS oferece uma série de vantagens para o usuário em relação ao recebimento de outros tipos de alertas originados a partir de correio eletrônico, por exemplo:

1. privacidade: não é necessário o envio de mensagens para inscrever-se em um *feed* RSS;
2. proteção contra spam: é impossível o envio de spam (informações não solicitadas pelo usuário) com o uso de RSS. Se as informações recebidas pelo *feed* não forem úteis ou começarem a incomodar, o usuário pode facilmente cancelar a inscrição do canal;
3. fácil cancelamento: quando desejar cancelar o recebimento de informações a partir de um *feed*, o usuário não precisará enviar uma

mensagem por correio eletrônico como acontece com as listas de discussão. Para cancelar um *feed*, basta apagar o mesmo de sua lista;

4. gerenciamento de conteúdo atualizado: *feeds* RSS proporcionam um tipo de serviço de alerta eficiente sobre novos conteúdos recém-publicados. O usuário pode configurar seu leitor de RSS para notificá-lo sempre que desejar ser informado sobre a publicação de uma nova atualização.

Uma questão que deve interessar ao usuário iniciante da tecnologia é descobrir como e onde encontrar *feeds* RSS. É sabido que muitos sítios na Internet com *feeds* associados oferecem *links* com informações sobre os canais disponíveis. Geralmente estes estão localizados nos menus ou no rodapé da página inicial e, na maior parte das vezes, acompanhados por um ícone indicativo de RSS. Mas, no meio de milhares de sítios da Web, como saber quais oferecem *feeds*? A resposta é simples: da mesma maneira que existem os mecanismos de busca de conteúdos na Web, podemos contar com diversos diretórios de indexação (Quadro 5.6) que auxiliam o usuário a pesquisar os *feeds* RSS disponíveis, bem como os assuntos que cobrem.

Diretório	URL	Comentários
Syndic8	<a href="http://www.syndic8.com">http://www.syndic8.com</a>	É o maior e mais completo diretório de <i>feeds</i> RSS. Disponibiliza mecanismos de pesquisa e navegação de canais organizados por áreas específicas.
Find RSS	<a href="http://www.findrss.net">http://www.findrss.net</a>	Bem organizado. Possui uma seção denominada <i>Latest Additions</i> , que disponibiliza o acesso aos <i>feeds</i> mais recentes adicionados à base de dados.
RSS Feeds.com	<a href="http://www.rssfeeds.com">http://www.rssfeeds.com</a>	Publica uma lista de <i>feeds</i> , organizada por categorias temáticas.
RSS Network	<a href="http://www.rss-network.com">http://www.rss-network.com</a>	Catálogo organizado por temas e assuntos específicos. Oferece ferramenta de busca rápida.
Feeds4All	<a href="http://www.feeds4all.com">http://www.feeds4all.com</a>	Catálogo simples. Publica uma breve descrição do <i>feed</i> .
Feed See	<a href="http://www.feedsee.com">http://www.feedsee.com</a>	Permite realizar buscas em <i>feeds</i> RSS por termos livres ou temas pré-definidos pelo serviço.

**Quadro 5.6** – Diretórios de indexação e divulgação de *feeds* RSS

Especificamente em relação ao conteúdo científico, destacamos o diretório SciencePORT<sup>35</sup>, por meio do qual é possível localizar cerca de 2 mil *feeds* utilizando termos livres ou navegando pelas categorias e subcategorias do sítio, tais como Tecnologia, Educação, Saúde, Humanidades, etc. Outros diretórios voltados para localização de *feeds* científicos são o eFeeds<sup>36</sup>, que reúne informações relativas aos serviços RSS oferecidos pelas revistas científicas, e os serviços Medworm<sup>37</sup>, especializado em *feeds* próprios da área de Saúde, e LibWorm<sup>38</sup>, dirigido a bibliotecários e profissionais da informação em geral.

Caso um sítio não disponibilize conteúdos RSS, é possível criar um ou mais *feeds* artificialmente por meio de serviços gratuitos, tais como o Dapper<sup>39</sup>, Feed43<sup>40</sup>, Page2RSS<sup>41</sup>, entre outros. Contudo, vale o alerta de que os *feeds* “sintéticos” podem não funcionar adequadamente no caso de mudanças no *layout* do sítio original.

## 5.6 Barreiras para Utilização de RSS

Mesmo com tantas vantagens e facilidades proporcionadas pela tecnologia, verificamos alguns entraves que podem dificultar a ampla difusão deste tipo de recurso. O principal empecilho, em nossa opinião, é o próprio desconhecimento dos formatos RSS, tanto da parte dos responsáveis pelo conteúdo de sítios quanto por seus usuários. No Brasil, de uma maneira geral, ainda são poucos os sítios que oferecem canais RSS. Mesmo no ambiente acadêmico, existe um desconhecimento generalizado a respeito de termos como *feeds*, *syndication* (sindicação) ou *aggregator* (agregador).

Em outros países, como os Estados Unidos e Canadá, a adoção de RSS em serviços de informação é bem mais comum. Mesmo assim, nota-se resistências por

---

<sup>35</sup> <<http://scienceport.org>>

<sup>36</sup> <<http://www.public.iastate.edu/~CYBERSTACKS/eFeeds.htm>>

<sup>37</sup> <<http://www.medworm.com>>

<sup>38</sup> <<http://www.libworm.com>>

<sup>39</sup> <<http://www.dapper.net>>

<sup>40</sup> <<http://www.feed43.com>>

<sup>41</sup> <<http://page2rss.com>>

parte da comunidade de editores científicos. A própria equipe de desenvolvimento do OJS admite que o módulo para geração de *feeds* RSS não é muito conhecido por seus usuários e são poucos os periódicos eletrônicos que disponibilizam *feeds* em suas edições.

Diferentemente da facilidade da leitura de um correio eletrônico, o método atual de assinar um *feed* – copiando a URL de um *link* (normalmente com extensão .xml ou .rss.) e colando-o de volta no aplicativo leitor de RSS – não é algo tão óbvio para o usuário não iniciado em informática. Além disso, caso este usuário esteja utilizando um navegador Web sem suporte à RSS (como é o caso do Internet Explorer, em versões anteriores a 7.0), ao entrar com o endereço de um *feed* no campo *location* (endereço) do navegador, o aspecto desse documento será o de um código XML, incompreensível para o usuário leigo. E o que é pior: provavelmente esta pessoa achará muito complicada a navegação por meio de *feeds* e talvez desista de vez de experimentá-la.

Vozes contrárias à tecnologia apontam outro problema para consideração: será que o tráfego da Web vai agüentar quando todo mundo passar a usar os agregadores? Explica-se: com os navegadores convencionais, os usuários visitam os sítios da Web somente quando desejam acessar determinado conteúdo e enquanto estiverem diante das telas de seus computadores. O funcionamento de um agregador é diferente já que, na maioria das vezes, ele permanece aberto, varrendo constantemente os sítios que contenham canais RSS em busca de atualização. A periodicidade do processo de coleta depende da configuração do programa. Mas é certo que, em demasia, os acessos quase simultâneos vindos de milhares de clientes ao mesmo tempo podem comprometer a largura de banda no caso de sítios muito visitados, provocando um efeito que Hedlund (2004) denominou RSS *overload* (sobrecarga de RSS).

Para minimizar o problema, algumas idéias vêm sendo empregadas visando a redução do tráfego de *feeds*. O serviço *Bloglines*, por exemplo, disponibilizou um servidor Web para armazenar as atualizações dos canais mais visitados,

---

funcionando como uma espécie de *cache*<sup>42</sup>. Contudo, a melhor solução parece ser mesmo a adoção de boas práticas por parte do usuário final, como o ajuste dos agregadores para buscarem por novidades nos *feeds* em longos intervalos de tempo. Essa frequência deveria ser baseada na regularidade de atualização dos conteúdos, ou seja, se um determinado sítio possui poucas alterações em um longo período de tempo – como acontece com os periódicos científicos – não há a menor necessidade de que o seu *feed* seja monitorado a todo instante.

---

<sup>42</sup> Na área de computação, *cache* é um método de armazenamento temporário para dados acessados com frequência com a finalidade de agilizar seu processamento.

## **CAPÍTULO 6**

### **A Tecnologia RSS Aplicada à Comunicação Científica**

---

Como já comentado no Capítulo 3, a comunicação científica tem se beneficiado dos avanços oferecidos pelas tecnologias de informação, permitindo mais agilidade no desenvolvimento das pesquisas e colocando em contato pesquisadores localizados em diferentes partes do mundo (OLIVEIRA e NORONHA, 2005, p.1). Além disso, as ferramentas utilizadas para o cumprimento de sua missão – troca de informações entre membros de uma mesma comunidade científica – acompanharam essa evolução e rapidamente adaptaram-se aos meios eletrônicos, produzindo novos formatos e funcionalidades, do suporte ao processo de publicação e comunicação científica.

Dito isso, parece natural que a tecnologia RSS fosse também incorporada aos periódicos eletrônicos científicos, promovendo agilidade principalmente no que diz respeito à fase de disseminação da informação científica. No Brasil estamos dando os primeiros passos com a disponibilização de *feeds* RSS para o sumário das revistas hospedadas pelo portal SciELO (*Scientific Electronic Library On-line*)<sup>43</sup>, desde abril de 2007. Em outros países, no entanto, o número de editoras científicas que já aderiram à tecnologia cresce a cada dia, com aplicações que vão muito além da simples notificação para os novos artigos disponibilizados a partir de sumários de periódicos.

Um dos primeiros trabalhos voltados para o estudo da tecnologia RSS sob a perspectiva acadêmica foi desenvolvido por Bernardino (2006), que aponta vantagens significativas na aplicação desse importante recurso para disseminação de informação científica e especializada. Afinal, o fato dos editores científicos disponibilizarem *feeds* RSS em suas revistas permite, por exemplo, que o resultado de uma pesquisa chegue ao conhecimento de um pesquisador interessado no tema, sem que o mesmo precise se preocupar em saber se os periódicos que costuma ler tiveram uma nova edição publicada, pois tão logo saia uma nova edição, o pesquisador será notificado.

Desse modo, o editor estará facilitando a vida do seu leitor na medida em que esse poderá localizar exatamente aquilo que procura ou que demonstre interesse,

---

<sup>43</sup> <<http://www.scielo.org>>

de modo seletivo, já que o uso do RSS pressupõe a intencionalidade do usuário de “puxar” a informação a partir de suas fontes de interesse. De acordo com Sá (2005, p.8), é um elemento importante, pois permite ao usuário “vigiar a sua dieta informativa, sem intermediários e sem vínculos”.

E o que parece ser uma desvantagem, em um primeiro momento, na realidade é um modo eficiente de aumentar a visibilidade dos periódicos eletrônicos, tornando mais fácil a sua disseminação independentemente da ação do usuário. Como existem muitas fontes de interesse do usuário, a maioria não retorna ao mesmo sítio todos os dias para verificar possíveis atualizações. Ao oferecer canais RSS para as suas publicações, os editores estarão permanentemente em contato com os usuários que escolheram navegar nas páginas Web por meio de *feeds*, sendo bastante provável que eles visitem a página da revista quando forem informados da publicação de algum artigo de seu interesse.

Hammond (2003) ressalta que os *feeds* RSS utilizados pelas revistas científicas eletrônicas se distinguem dos encontrados em sítios de notícias e *blogs* em geral devido a uma característica particular da comunicação científica: a capacidade de citar e produzir citações para um determinado documento. Assim, surge a necessidade de representar a informação para além dos elementos básicos utilizados na descrição de recursos, tais como *<title>*, *<link>* e *<description>*. O ideal, portanto, seria adicionar um conjunto de metadados capaz de melhorar a descrição semântica dos artigos científicos.

## 6.1 Conjuntos de Metadados

O termo “metadado” (*metadata*) foi cunhado por Jack Myres, em 1969, para denominar os dados cuja função era a descrição de arquivos convencionais. Começou a aparecer mais freqüentemente na literatura sobre Sistemas Gerenciadores de Banco de Dados (SGBD) nos anos 80, para descrever as características das informações armazenadas nos bancos de dados (VELLUCCI, 1998 apud THOMAZ; SANTOS, 2003).

No domínio da Ciência da Computação, “metadado” é entendido como sendo “dados sobre dados” (WEILBEL, 1995), sendo basicamente utilizados para prover informações referenciais, contextos e características sobre determinado recurso. MILSTEAD & FIELDMAN (1999, p. 34) defendem a idéia de que, na verdade, os metadados antecedem a existência do próprio termo e citam como um exemplo típico o conhecido formato MARC – *Machine Readable Cataloging* para descrever documentos bibliográficos. Em termos de serviços e aplicações Web, um dos principais objetivos do uso de metadados, segundo Marcondes (2006, p. 96), é:

não somente descrever documentos eletrônicos e informações em geral, possibilitando sua avaliação de relevância por usuários humanos, mas também permitir agenciar computadores e programas especiais, robôs e agentes de software, para que eles compreendam os metadados associados a documentos e possam então recuperá-los, avaliar sua relevância e manipulá-los com mais eficiência.

Deste modo, é desejável que possamos incluir metadados adicionais nos documentos RSS/XML com o propósito de enriquecer as aplicações e a descrição semântica dos recursos informacionais, como já mencionado no Capítulo 4. No caso das aplicações científicas, especialistas sugerem a adoção da especificação RSS 1.0 por estar em conformidade com a sintaxe RDF, considerada ideal para a inclusão suplementar de metainformação por meio de outros vocabulários de metadados, sendo bastante útil para a integração em sistemas de Web Semântica. (NOTTINGAM, 2005).

### 6.1.1 Vocabulário Dublin Core

Um dos fortes candidatos para a adição de elementos suplementares de informação dentro de um *feed* RSS é o vocabulário-padrão proposto pela *Dublin Core Metadata Initiative* (DCMI)<sup>44</sup>, formada pela OCLC – *On-line Computer Library Center* e NCSA – *National Center for Supercomputing Applications*. O objetivo da iniciativa *Dublin Core* (DC) é descrever informações sobre documentação, em geral, para resolver o problema de localização de informação na Internet.

---

<sup>44</sup> <<http://dublincore.org>>

O padrão é composto por dois níveis de especificações: simples e qualificado. Em seu modo simples, o *Dublin Core* (DC) é formado por 15 elementos descritivos, enquanto o modo qualificado inclui três elementos adicionais (audiência, proveniência e detentor de direitos). Também é possível refinar cada elemento por meio de qualificadores que garantem mais precisão ao significado desses elementos. Por exemplo, o elemento *Date* tem como qualificadores os termos *Issued*, *Available*, *Modified* e *Valid*.

Em seguida, listamos os 15 elementos do conjunto de metadados *Dublin Core* Simples (*Core Metadata Element Set*, DCMES):

1. *Title* (título)
2. *Creator* (autor, responsável pelo documento)
3. *Subject* (assunto)
4. *Description* (descrição, resumo, sumário)
5. *Publisher* (publicador, responsável pela disponibilização do documento)
6. *Contributor* (outros colaboradores)
7. *Date* (data de publicação)
8. *Type* (tipo de recurso: página da internet, *software*, áudio, vídeo)
9. *Format* (formato do arquivo que contém o documento eletrônico: texto, PDF, HTML, DOC, etc.)
10. *Identifier* (identificador, geralmente a URL de um documento eletrônico)
11. *Source* (fonte de origem)
12. *Language* (idioma)
13. *Relation* (relacionamento com outros documentos, como versões)
14. *Coverage* (abrangência espacial ou temporal, sobre lugar ou época)
15. *Rights* (especificação sobre direitos autorais)

Vale dizer que cada elemento DC é opcional e pode ser repetido. Também não há ordem para apresentação ou uso dos elementos. Outro aspecto interessante é que o conjunto DC provê uma base comum que permite o intercâmbio semântico com outros formatos de metadados mais complexos e consolidados na forma de padrões internacionais de descrição de informação eletrônica, como o *Machine Readable Cataloging* (MARC 21).

### 6.1.2 Vocabulário PRISM

O conjunto de metadados PRISM<sup>45</sup> – *Publishing Requirements for Industry Standards Metadata* – é uma especificação baseada no padrão *Dublin Core*, criada e mantida pela IDEAlliance – *International Digital Enterprise Alliance*, um grupo composto por editoras científicas e organizações envolvidas com a produção de conteúdos editoriais para Web. Sua finalidade é garantir um vocabulário XML voltado para as necessidades editoriais, ou seja, trata-se de um formato específico para o gerenciamento, agregação e processamento de notícias, catálogos, livros e periódicos em geral.

A intenção da especificação PRISM (atualmente na versão 1.0) é prover a infra-estrutura necessária para a) o intercâmbio e preservação de conteúdos; b) uma coleção de elementos para descrição desse conteúdo e, ainda, c) um conjunto de vocabulários controlados com os respectivos valores para esses elementos.

Assim, além do padrão genérico *Dublin Core*, o conjunto PRISM permite a extensão em *feeds* RSS no que se refere à informação bibliográfica dos artigos, tais como: *issn* (número serial da publicação), *publicationName* (nome da publicação), *volume* (volume da edição), *number* (número da edição), *startingPage* (página de início do artigo), etc. (HAMMOND et al., 2004).

O vocabulário básico oferecido pelo PRISM<sup>46</sup> reúne 48 elementos, os quais podem ser incorporados aos níveis *<channel>* e *<item>* da especificação RSS 1.0 por meio de um módulo próprio. Em alguns casos, é preciso consultar as regras da especificação para verificar o grau de compatibilidade entre os elementos PRISM e DC. No caso do elemento *dc:relation*, por exemplo, não é recomendado que o mesmo seja usado juntamente com o vocabulário PRISM.

De acordo com Hammond (2003), a utilização conjunta de elementos dos vocabulários DC e PRISM é suficiente para cobrir a descrição dos termos requeridos

---

<sup>45</sup> <<http://www.prismstandard.org>>

<sup>46</sup> <[http://www.prismstandard.org/resources/mod\\_prism.html](http://www.prismstandard.org/resources/mod_prism.html)>

pelos sumários de publicações seriadas graças ao mecanismo de extensão modular da especificação RSS 1.0. A Figura 6.1 mostra um exemplo de utilização de metadados Dublin Core (identificados na cor azul) e PRISM (vermelho), incorporados a um *feed* RSS 1.0.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns="http://purl.org/rss/1.0/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:prism="http://prismstandard.org/namespaces/1.2/basic/"
>
  <channel rdf:about="http://www.sciencedirect.com/science/issue/5290-2003-999559997-369185">
    <title>Tetrahedron Letters 44(2)</title>
    <link>http://www.sciencedirect.com/science/issue/5290-2003-999559997-369185</link>
    <description>Tetrahedron Letters Volume 44, Issue 2</description>
    <dc:publisher>Elsevier</dc:publisher>
    <prism:coverDisplayDate>6 January 2003</prism:coverDisplayDate>
    <prism:distributor>ScienceDirect</prism:distributor>
    <prism:copyright>Copyright (c) 2003, ScienceDirect. All rights reserved.</prism:copyright>
    <items>
      <rdf:Seq>
        <rdf:li rdf:resource="http://www.sciencedirect.com/science/article/7/1/a0b575" />
      </rdf:Seq>
    </items>
  </channel>
  <image rdf:about="http://www.sciencedirect.com/science/page/static/science/logo.gif">
    <title>ScienceDirect</title>
    <url>http://www.sciencedirect.com/science/page/static/science/logo.gif</url>
    <link>http://www.sciencedirect.com/science/journal/00404039</link>
  </image>

  <item rdf:about="http://www.sciencedirect.com/science/article/7/1/42081f47191a0b575">
    <title>Unexpected reactivity of oxygenated 2-acyl-1,3-dithianes with electrophiles</title>
    <link>http://www.sciencedirect.com/science/article/7/1/81f47191a0b575</link>
    <description>Unexpected reactivity of oxygenated with electrophiles, Pages 213-216</description>
    <dc:title>Unexpected reactivity of oxygenated 2-acyl-1,3-dithianes with electrophiles</dc:title>
    <dc:creator>Michael Smietana</dc:creator>
    <prism:publicationName>Tetrahedron Letters</prism:publicationName>
    <prism:volume>44</prism:volume>
    <prism:number>2</prism:number>
    <prism:startingPage>213</prism:startingPage>
  </item>
</rdf:RDF>
```

**Figura 6.1** – Exemplo de *feed* RSS utilizando vocabulários de metadados DC e PRISM

Alguns editores científicos, nomeadamente a *BioMed Central*, *Ingenta Connect* e *Nature Publishing Group* já fazem uso de metadados *Dublin Core* e *PRISM* juntamente com os *feeds* RSS de suas revistas (Quadro 6.1).

Editora	Versão RSS	Tipos de conteúdos
1. BioMed Central <a href="http://www.biomedcentral.com">http://www.biomedcentral.com</a>	RSS 1.0	Sumários, artigos recentes
2. British Medical Journal <a href="http://www.bmj.com">http://www.bmj.com</a>	RSS 1.0	Sumários, pré-prints
3. Ingenta Connect <a href="http://www.ingentaconnect.com">http://www.ingentaconnect.com</a>	RSS 1.0	Sumários, documentos recentes
4. Int. Union of Crystallography <a href="http://journals.iucr.org">http://journals.iucr.org</a>	RSS 1.0	Sumários e artigos de acesso livre
5. Nature Publishing Group <a href="http://www.nature.com/webfeeds">http://www.nature.com/webfeeds</a>	RSS 1.0	Sumários, artigos, notícias, lista de empregos
6. Oxford University Press <a href="http://www.oxfordjournals.org">http://www.oxfordjournals.org</a>	RSS 1.0	Sumários, artigos recentes
7. Wiley InterScience <a href="http://www.interscience.wiley.com">http://www.interscience.wiley.com</a>	RSS 1.0	Sumários

**Quadro 6.1** – Editoras científicas que utilizam elementos de metadados dos vocabulários DC e PRISM na criação de *feeds* RSS

Em levantamento realizado a partir dos *feeds* dos periódicos dessas 7 (sete) editoras, identificamos os metadados dos vocabulários *Dublin Core* e *PRISM* que complementam a informação no nível dos elementos *<channel>* (dados sobre a publicação) e *<item>* (dados sobre os artigos) dos formatos RSS 1.0. Os metadados encontrados são apresentados, respectivamente, nos Quadros 6.2 e 6.3.

Metadados	Editoras						
	<i>British Medical Journal</i>	<i>Nature</i>	<i>Int. Union of Crystallography</i>	<i>Biomed</i>	<i>Ingenta Connect</i>	<i>Oxford University Press</i>	<i>Wiley InterScience</i>
	<b>&lt;CHANNEL&gt;</b>						
D C	<dc:creator>			√			
	<dc:date>			√		√	√
	<dc:description>			√		√	
	<dc:format>			√			
	<dc:identifier>			√		√	
	<dc:language>		√	√			√
	<dc:publisher>		√	√		√	√
	<dc:rights>		√	√		√	√
	<dc:source>			√			
	<dc:title>			√		√	
	<dc:type>			√			
			√			√	
P R I S M	<prism:copyright>			√			
	<prism:coverDate>			√			
	<prism:coverDisplayDate>			√		√	
	<prism:distributor>				√		
	<prism:elssn>	√				√	
	<prism:startingPage>			√			
	<prism:issn>	√	√	√		√	√
	<prism:publicationName>	√	√			√	
		√	√				
			√				

**Quadro 6.2** – Metadados DC e PRISM utilizados juntamente com o elemento *<channel>* de feeds RSS de editoras científicas

Metadados/Elementos	Editoras						
	<i>British Medical Journal</i>	<i>Nature</i>	<i>Int. Union of Crystallography</i>	<i>Biomed</i>	<i>Ingenta Connect</i>	<i>Oxford University Press</i>	<i>Wiley InterScience</i>
	<i>&lt;ITEM&gt;</i>						
D C	<dc:creator>	√	√	√	√	√	√
	<dc:date>	√	√	√	√	√	√
	<dc:description>			√			
	<dc:format>			√			
	<dc:identifier>	√	√	√	√	√	√
	<dc:language>			√			
	<dc:publisher>	√		√		√	√
	<dc:rights>			√			√
	<dc:source>		√	√			
	<dc:subject>			√			
	<dc:teaser>			√			
	<dc:title>	√	√	√			√
	<dc:type>			√			
P R I S M	<prism:copyright>			√		√	
	<prism:coverDisplayDate>					√	
	<prism:endingPage>			√		√	
	<prism:isPartOf>					√	
	<prism:issn>				√	√	
	<prism:number>		√	√	√	√	√
	<prism:publicationDate>	√	√	√	√		√
	<prism:publicationName>		√	√	√	√	
	<prism:section>	√		√			√
<prism:startingPage>			√	√		√	
<prism:volume>		√	√		√	√	

**Quadro 6.3** – Metadados DC e PRISM utilizados juntamente com o elemento *<item>* de feeds RSS de editoras científicas

Observamos que 11 dos 15 elementos do conjunto *Dublin Core* são utilizados pelas editoras científicas junto ao elemento *<channel>* do RSS, que define os metadados da revista como um todo, enquanto o conjunto PRISM colabora com 10 elementos de um total de 48. Com relação ao complemento de informação junto ao elemento *<item>*, notamos o uso de 13 dos 15 elementos DC, sendo todos eles aproveitados nos *feeds* dos periódicos da *International Union of Crystallography*. 11 metadados PRISM foram utilizados pelas revistas das editoras analisadas.

A partir dos dados indicados nos Quadros 6.2 e 6.3, notamos que os artigos dos periódicos da *International Union of Crystallography* são os que dispõem de mais recursos oferecidos pela adição de metainformação. Curiosamente, três das sete editoras analisadas (*British Medical Journal*, *Oxford University Press* e *Biomed*) não fazem uso de metadados DC. A editora *Biomed*, aliás, também não usa os elementos PRISM, pelo menos no que se refere ao elemento *<channel>*.

Em relação à utilização de informação complementar ao elemento *<item>*, ou seja, metadados relacionados aos artigos propriamente ditos, percebemos um uso razoável do conjunto de metadados PRISM, específico para a representação de informação de artigos científicos. A exceção ficou por conta dos periódicos da *Wiley InterScience*. Dentre as editoras analisadas, aliás, esta é a que menos se apóia no uso de metadados de uma maneira geral.

Outras editoras científicas, como a *American Chemical Society*, *Blackwell Publishing* e *National Geographic*, por exemplo, optaram por fornecer *feeds* sem a inclusão de outros módulos de metadados (Quadro 6.4). Neste caso, como é de se esperar, a maioria das editoras preferiu utilizar a versão RSS 2.0, que torna mais simples a tarefa de sindicância de seus conteúdos.

Editora	Versão RSS	Tipos de conteúdos
1. American Chemical Society <a href="http://pubs.acs.org">http://pubs.acs.org</a>	RSS 2.0 / Atom	sumários
2. American Institute of Physics <a href="http://journals.aip.org">http://journals.aip.org</a>	RSS 1.0	sumários, notícias, listas de empregos, eventos, divulgação de novos produtos
3. Blackwell Publishing <a href="http://www.blackwellpublishing.com/press/rss.asp">http://www.blackwellpublishing.com/press/rss.asp</a>	RSS 2.0	notícias de imprensa
4. National Geographic <a href="http://news.nationalgeographic.com">http://news.nationalgeographic.com</a>	RSS 2.0	notícias
5. Institute of Physics Publishing <a href="http://journals.iop.org">http://journals.iop.org</a>	RSS 2.0	sumários, notícias, listas de empregos, eventos, divulgação de novos produtos
6. Springer <a href="http://www.springer.com">http://www.springer.com</a>	RSS 2.0	sumários, novos produtos

**Quadro 6.4** – Editoras científicas que não utilizam elementos de metadados na descrição de suas publicações

## 6.2 Aplicações Mais Comuns

No âmbito da comunicação científica, a aplicação do padrão RSS verificada com mais frequência é mesmo a sindicância de sumários de periódicos. As editoras que aderiram à tecnologia costumam publicar um *feed* para cada uma de suas revistas, de modo que seus leitores possam ser avisados sobre a disponibilidade de uma nova edição tão logo ela seja publicada. Como estratégia de divulgação, a maioria das editoras atualiza seus *feeds* com o título e uma descrição dos artigos (em geral, os resumos), além do respectivo *link* para a página onde cada artigo encontra-se hospedado no sítio da revista. Algumas dessas editoras também vêm utilizando os canais RSS para oferecerem um conjunto diversificado de serviços de alerta para seus usuários, informando sobre o lançamento de novos produtos, eventos e até mesmo oportunidades de empregos em áreas de pesquisa.

### 6.2.1 Bibliotecas e Unidades de Informação

Em artigo publicado no *blog HigherEd BlogCon*<sup>47</sup>, o professor Jay Bhatt, da Universidade Drexel (Pensilvânia, Estados Unidos), traz algumas sugestões de uso da tecnologia RSS para incrementar a utilização de recursos eletrônicos por usuários de bibliotecas acadêmicas. Bhatt (2006) comenta que as bibliotecas deveriam ampliar o escopo de seus serviços com a criação de *feeds* voltados para o atendimento de demandas específicas, como as mostradas no Quadro 6.5.

A adoção do RSS por unidades de informação, aliás, também foi tema de um artigo publicado pela *Library Hi Tech News*, no qual Holvoet (2006, p.33) conclama as bibliotecas e os profissionais da área a adotarem uma postura pró-ativa em relação ao uso dessa tecnologia emergente. Para a autora, o RSS simplifica consideravelmente as tarefas de busca e fornecimento de informação, constituindo-se como uma importante ferramenta de gerenciamento para os bibliotecários.

Sugestão de uso	Exemplos
Anunciar as novas aquisições de livros e outros materiais, segundo áreas de interesse	Biblioteca da Universidade de Oklahoma <a href="http://libraries.ou.edu/rss">http://libraries.ou.edu/rss</a>
Anunciar a disponibilidade de novas bases de dados	Biblioteca da Universidade do Tennessee <a href="http://www.lib.utc.edu/databases/results.php?LIST=new">http://www.lib.utc.edu/databases/results.php?LIST=new</a>
Promover eventos organizados pela biblioteca	Biblioteca da Universidade de Binghamton <a href="http://library.lib.binghamton.edu/mt/librarynews">http://library.lib.binghamton.edu/mt/librarynews</a>
Divulgar oportunidades de trabalho oferecidas pelos diversos departamentos da instituição	Departamento de Engenharia Civil, Arquitetura e Meio Ambiente da Universidade Drexel <a href="http://www.drexel-civil.org">http://www.drexel-civil.org</a>

Fonte: adaptado de Bhatt (2006)

**Quadro 6.5** – Utilização de *feeds* RSS por bibliotecas universitárias

<sup>47</sup> <<http://www.higheredblogcon.com>>. Trata-se de um *blog* voltado à promoção de ferramentas sociais no contexto das comunidades acadêmicas.

Desse modo, além das sugestões mencionadas anteriormente, as bibliotecas universitárias e/ou especializadas poderiam criar serviços igualmente úteis, colecionando *feeds* e tornando-os disponíveis aos seus usuários. Wusterman (2004, p. 405-406) recomenda dois caminhos possíveis:

1. Prover diretórios de *feeds*

Neste caso, temos uma espécie de catálogo criado para facilitar os usuários quanto à localização de páginas com conteúdos descritos em RSS/Atom. A biblioteca da Universidade de Nevada, Reno (Estados Unidos), por exemplo, criou um índice<sup>48</sup> reunindo mais de 7 mil periódicos que disponibilizam canais RSS para seus artigos. Outra boa referência é a biblioteca de *feeds*<sup>49</sup> publicada na página oficial do governo americano, organizada por categorias distribuídas em eixos temáticos.

2. Incorporar novas informações ao sítio da biblioteca a partir de *feeds* de outras fontes

Esta é uma das aplicações mais comuns relacionada ao uso de RSS: agregar informações provenientes de *feeds* previamente selecionados, exibindo-os como conteúdo complementar em um sítio Web via protocolo HTML. Existem várias ferramentas gratuitas que facilitam esse trabalho, mas destacamos o serviço RSSxpress-Lite<sup>50</sup>, desenvolvido pela Universidade de Bath (Reino Unido). Seu funcionamento é baseado na geração de um *JavaScript* que pode ser facilmente copiado para dentro do código HTML de qualquer sítio. O programa ainda traz dezenas de *feeds* de interesse para as bibliotecas.

Tendo reunido uma quantidade significativa de *feeds* úteis, percebemos uma valiosa coleção de fontes de informação, que poderá ser facilmente reutilizada e compartilhada com outros usuários por meio do formato OPML (*Outline Processor*

---

<sup>48</sup> <<http://www.library.unr.edu/ejournals/alphaRSS.aspx>>

<sup>49</sup> <[http://www.usa.gov/Topics/Reference\\_Shelf/Libraries/RSS\\_Library.shtml](http://www.usa.gov/Topics/Reference_Shelf/Libraries/RSS_Library.shtml)>

<sup>50</sup> <<http://rssxpress.ukoln.ac.uk>>

*Markup Language*)<sup>51</sup>. Trata-se de uma aplicação XML voltada para intercâmbio de listas de canais RSS inscritas a partir da geração (exportação) ou leitura (importação) de um arquivo-texto com extensão .opml, suportado por *softwares* agregadores ou qualquer sistema compatível com o formato. O objetivo de um arquivo OPML é, portanto, garantir acesso simultâneo, seletivo e organizado a uma coleção de *feeds*, oferecendo os meios para criação de várias fontes de informação sobre determinado assunto. Em suma, o que o RSS faz para os sítios Web, o OPML faz para os *feeds* RSS.

Unidades especializadas de informação, como as bibliotecas temáticas e universitárias poderão fornecer aos seus usuários a possibilidade de assinar diversas fontes de informação personalizadas de uma só vez (LEAL, 2007). Seguindo essa filosofia, temos o *BlogBridge Library*<sup>52</sup>, uma poderosa e intuitiva maneira de criar grandes e bem organizadas bibliotecas de *feeds* para várias aplicações a partir de listas OPML. Esse mesmo recurso pode ser utilizado para agregar *feeds* de uma coleção inteira de periódicos, como ocorre com o BaRf<sup>53</sup> (*Bioinformatics aggregated RSS feeds*), que disponibiliza a lista OPML dos *feeds* das principais revistas da área de Bioinformática.

### 6.2.2 Bases de Dados

No que diz respeito aos benefícios do RSS para a comunidade acadêmica em geral, outra aplicação que parece ter bastante futuro é a geração automática de *feeds* relacionados aos resultados de buscas efetivadas em bases de dados *on-line* (WUSTERMAN, 2004, p. 406). A idéia é permitir que o usuário assine um *feed* personalizado, criado sob demanda para atender às suas necessidades de informação, já que este *feed* estará relacionado à expressão utilizada na pesquisa. A partir daí, este mesmo usuário passará a ser notificado toda vez que um novo registro for incluído na base de dados, desde que atenda às condições estabelecidas por ele durante a elaboração da estratégia de busca.

---

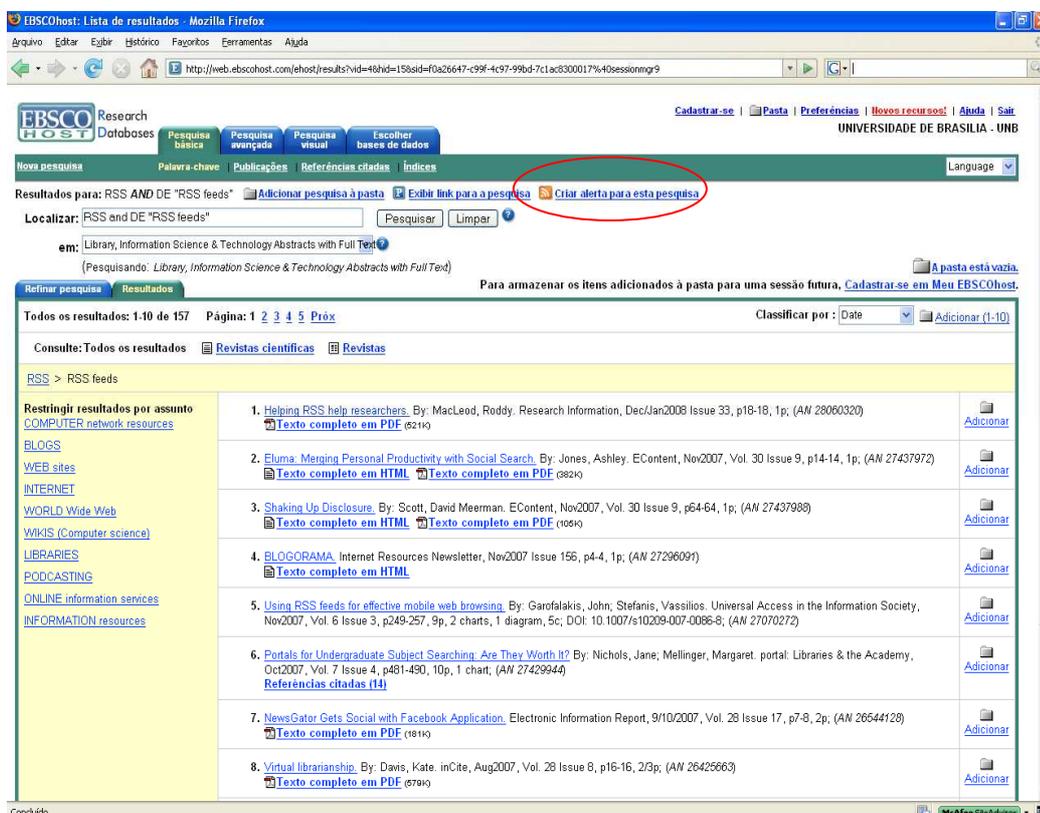
<sup>51</sup> <<http://www.opml.org>>

<sup>52</sup> <<http://library.blogbridge.com>>

<sup>53</sup> <<http://barf.jcowboy.org>>

Para ilustrar melhor este processo, tomemos como exemplo a base de dados LISTA (*Library/Information Science & Technology Abstracts*) com texto completo, suportada pelo provedor EBSCO e disponibilizada no portal da CAPES. Realizamos uma busca em toda a base pelo termo “RSS”, restringindo os documentos classificados segundo o assunto “RSS feeds”.

Como resultado, a base retornou 157 registros relacionados, com a opção de “criar um alerta para esta pesquisa” via *feed* RSS (indicado pelo ícone padrão). Ao clicarmos sobre este *link* (Figura 6.1), recebemos uma informação sobre como fazer a inscrição no *feed* correspondente (Figura 6.2). Assim, se desejarmos continuar recebendo resultados desta pesquisa, independentemente do acesso à base de dados, tudo o que precisamos fazer é confirmar a assinatura no canal personalizado para recebermos uma notificação assim que um novo documento relevante for adicionado à base.



The screenshot shows the EBSCOhost search results page for the query "RSS AND DE 'RSS feeds'". The page is displayed in a Mozilla Firefox browser window. The search results are filtered to show 1-10 of 157 results. The results list includes several articles related to RSS feeds, such as "Helping RSS help researchers", "Efluma: Merging Personal Productivity with Social Search", and "Shaking Up Disclosure". Each result includes a title, author, publication information, and options to view the full text in HTML or PDF. A red circle highlights the "Criar alerta para esta pesquisa" button in the search interface.

Restringir resultados por assunto	Resultados
COMPUTER network resources	1. <a href="#">Helping RSS help researchers</a> . By: MacLeod, Roddy. Research Information, Dec/Jan2008 Issue 33, p18-18, 1p; (AN 28060320) <a href="#">Texto completo em PDF</a> (6214)
BLOGS	2. <a href="#">Efluma: Merging Personal Productivity with Social Search</a> . By: Jones, Ashley. EContent, Nov2007, Vol. 30 Issue 9, p14-14, 1p; (AN 27437972) <a href="#">Texto completo em HTML</a> <a href="#">Texto completo em PDF</a> (6214)
WEB sites	3. <a href="#">Shaking Up Disclosure</a> . By: Scott, David Meerman. EContent, Nov2007, Vol. 30 Issue 9, p64-64, 1p; (AN 27437968) <a href="#">Texto completo em HTML</a> <a href="#">Texto completo em PDF</a> (6064)
INTERNET	4. <a href="#">BLOGORAMA</a> . Internet Resources Newsletter, Nov2007 Issue 156, p4-4, 1p; (AN 27296091) <a href="#">Texto completo em HTML</a>
WORLD Wide Web	5. <a href="#">Using RSS feeds for effective mobile web browsing</a> . By: Carofalakis, John; Stefanis, Vassilios. Universal Access in the Information Society, Nov2007, Vol. 6 Issue 3, p249-257, 9p, 2 charts, 1 diagram, 5c; DOI: 10.1007/s10209-007-0086-8; (AN 27070272) <a href="#">Texto completo em HTML</a> <a href="#">Texto completo em PDF</a> (6214)
WIKIS (Computer science)	6. <a href="#">Portals for Undergraduate Subject Searching: Are They Worth It?</a> By: Nichols, Jane, Mellinger, Margaret. portal: Libraries & the Academy, Oct2007, Vol. 7 Issue 4, p481-490, 10p, 1 chart; (AN 27429344) <a href="#">Referências citadas (14)</a>
LIBRARIES	7. <a href="#">NewsCater Gets Social with Facebook Application</a> . Electronic Information Report, 9/10/2007, Vol. 28 Issue 17, p7-8, 2p; (AN 26544128) <a href="#">Texto completo em PDF</a> (6214)
PODCASTING	8. <a href="#">Virtual librarianship</a> . By: Davis, Kate. inCite, Aug2007, Vol. 28 Issue 6, p16-16, 2/3p; (AN 26426663) <a href="#">Texto completo em PDF</a> (6794)

Figura 6.2 – Documentos que tratam de “RSS”, segundo a base LISTA/EBSCO

The screenshot shows the EBSCOhost interface in a Mozilla Firefox browser. The page title is "EBSCOhost: Lista de resultados". The browser address bar shows the URL: <http://web.ebscohost.com/ehost/results?vid=8&hid=4&sid=f0a26647-c9f-4c97-99bd-7c1ac8300017%40sessionmgr9>. The EBSCOhost logo and navigation menu are visible at the top. A modal window titled "Seu aviso foi criado" (Your alert has been created) is open, displaying the following information:

- Nome do aviso:** RSS and DE "RSS feeds" on 2008-01-22 09:35 PM
- Base de dados:** Library, Information Science & Technology Abstracts with Full Text
- Estratégia de pesquisa:** RSS and DE "RSS feeds"
- O alerta expira depois de:** Dois meses de inatividade [Novos Feeds serão removidos se não forem acessados dentro de (1) semana da criação.]
- Frequência:** Daily
- Syndication Feed:** <http://ss.ebscohost.com/AlertSyndicationService/Syndication.aspx?GetFeed?guid=1244892>

Below the modal window, there is a list of search results with "Adicionar" buttons next to each item:

4. [BLOGORAMA](#). Internet Resources Newsletter, Nov2007 Issue 156, p4-4, 1p; (AN 27296091) [Texto completo em HTML](#)
5. [Using RSS feeds for effective mobile web browsing](#). By: Garofalakis, John; Stefanis, Vassilios. Universal Access in the Information Society, Nov2007, Vol. 6 Issue 3, p249-257, 9p, 2 charts, 1 diagram, 5c; DOI: 10.1007/s10209-007-0086-8; (AN 27070272)
6. [Portals for Undergraduate Subject Searching. Are They Worth It?](#) By: Nichols, Jane; Mellinger, Margaret. portal: Libraries & the Academy, Oct2007, Vol. 7 Issue 4, p481-490, 10p, 1 chart; (AN 27429944) [Referências citadas \(1\)](#)
7. [NewsGator Gets Social with Facebook Application](#). Electronic Information Report, 9/10/2007, Vol. 28 Issue 17, p7-8, 2p; (AN 26644128) [Texto completo em PDF](#) (181k)
8. [Virtual librarianship](#). By: Davis, Kate. inCite, Aug2007, Vol. 28 Issue 8, p16-16, 2/3p; (AN 26426663)

Figura 6.3 – Criação dinâmica de *feed* RSS correspondente a um resultado de busca

Além do EBSCO, provedores comerciais de informação consagrados, como Scopus (Elsevier) e Proquest, também incorporaram o RSS como um serviço de valor adicionado para disseminação de seus conteúdos, permitindo que os seus usuários recebam notificações sobre os assuntos de seu interesse de modo seletivo. Algumas bases de dados *on-line* comerciais que já dispõem deste recurso estão relacionadas no Quadro 6.6.

Base de dados	Descrição
AccessScience <a href="http://www.accessscience.com">http://www.accessscience.com</a>	Enciclopédia de Ciência e Tecnologia da McGraw-Hill. Os <i>feeds</i> RSS fornecem informações sobre os verbetes, porém é preciso ser assinante do serviço para acessar o texto completo
Highwire Press Journals <a href="http://highwire.stanford.edu">http://highwire.stanford.edu</a>	Base de dados dos periódicos disponibilizados pelo projeto Highwire Press, da Universidade de Stanford. Requer cadastro gratuito. Os <i>feeds</i> RSS encontram-se disponíveis na seção “Alerta”
HubMed <a href="http://www.hubmed.org">http://www.hubmed.org</a>	Interface alternativa da base de dados PubMed, de literatura médica. Disponibiliza <i>feeds</i> RSS para os resultados das buscas
Knovel <a href="http://www.knovel.com/">http://www.knovel.com/</a>	Base de dados de referências para assuntos relacionados às engenharias. Os <i>feeds</i> RSS fornecem os conteúdos atualizados
Lexis/Nexis <a href="http://www.lexisnexis.com/presscenter/rss/">http://www.lexisnexis.com/presscenter/rss/</a>	Fornecer informações relacionadas às áreas jurídica, corporativa e governamental. Os <i>feeds</i> RSS são usados como alertas para informar sobre atualizações na base de dados
Proquest <a href="http://www.proquest.com/syndication/rss/">http://www.proquest.com/syndication/rss/</a>	Reúne vários tipos de coleções destinadas a diversas áreas de conhecimento. Disponibiliza <i>feeds</i> para teses e dissertações e também para especialidades da área de Saúde

Fonte: adaptado de Bhatt (2006)

**Quadro 6.6** – Exemplos de bases de dados *on-line* que disponibilizam *feeds* RSS

### 6.2.3 Outras Aplicações

Mais do que divulgar informações recentes, o RSS também pode ser empregado com outras finalidades, tais como a integração de repositórios de dados estruturados por meio de *feeds* via protocolo OAI-PMH. Graças à compatibilidade entre os metadados, bibliotecas digitais e demais serviços especializados de informação podem recolher os dados provenientes de vários *feeds*, cobrindo informações sempre atualizadas e procedendo a sindicância de seus conteúdos.

Uma iniciativa bem sucedida nesse sentido é o *IMesh Toolkit Project*<sup>54</sup>, apoiado pela *International Digital Libraries Programme* (JISC/NSF). Basicamente, trata-se de um módulo criado para acessar registros de repositórios de dados por

<sup>54</sup> <<http://www.imesh.org/toolkit>>

meio do protocolo OAI-PMH, permitindo a geração de uma lista de conteúdos no formato RSS 2.0. Caso seja necessário, esses arquivos (*feeds*) poderão ser modificados posteriormente por meio de programas editores, como o RSSexpress<sup>55</sup>. Adicionalmente é possível incluir comentários ou *links* para recursos relacionados, bem como converter o arquivo XML, por meio de um *script*, para que o *feed* seja exibido em qualquer página Web (DUKE, 2003).

Outro bom exemplo é o *Bibliorandum*<sup>56</sup>, um sistema de informação corrente em Biblioteconomia e Ciência da Informação que reúne um motor de busca customizado do *Google* juntamente com um agregador de conteúdos temático de fontes relevantes da área. Por meio da integração entre diferentes protocolos, permite ao usuário realizar a inscrição em canais RSS gerados dinamicamente a partir das diversas categorias de informação reunidas pelo serviço, tais como sítios de notícias, periódicos, *blogs* e repositórios digitais.

Funcionalmente, o *Bibliorandum* atua tanto como consumidor quanto produtor de RSS. Na opção *blogosfera*, recebe, agrega e redistribui conteúdos em RSS, enquanto que no recurso *Akademya* (repositórios digitais) recolhe informações por meio do protocolo OAI-PMH, agrega e redistribui como formato RSS (ANJOS, 2007). A Figura 6.3 mostra a tela da opção *Akademia*, por meio da qual é possível recuperar os 25 documentos mais recentes publicados por 10 repositórios digitais de acesso aberto, entre eles o E-LIS, *Library Student Journal*, *Open Research On-line*, etc. A apresentação dos textos depositados é feita segundo ordenamento decrescente da data de publicação. O serviço também oferece ao usuário a opção de pesquisar os documentos pelo título, autor e resumo, além de filtrar os textos segundo o seu idioma de origem.

O projeto foi desenvolvido como trabalho final de uma disciplina do Instituto Politécnico do Porto, Portugal, e tem-se mantido ativo e em constante desenvolvimento com o apoio da Associação Portuguesa para a Gestão da Informação (INCITE)<sup>57</sup>.

---

<sup>55</sup> <<http://rssxpress.ukoln.ac.uk>>

<sup>56</sup> <<http://www.bibliorandum.net>>

<sup>57</sup> <<http://www.incite.pt>>

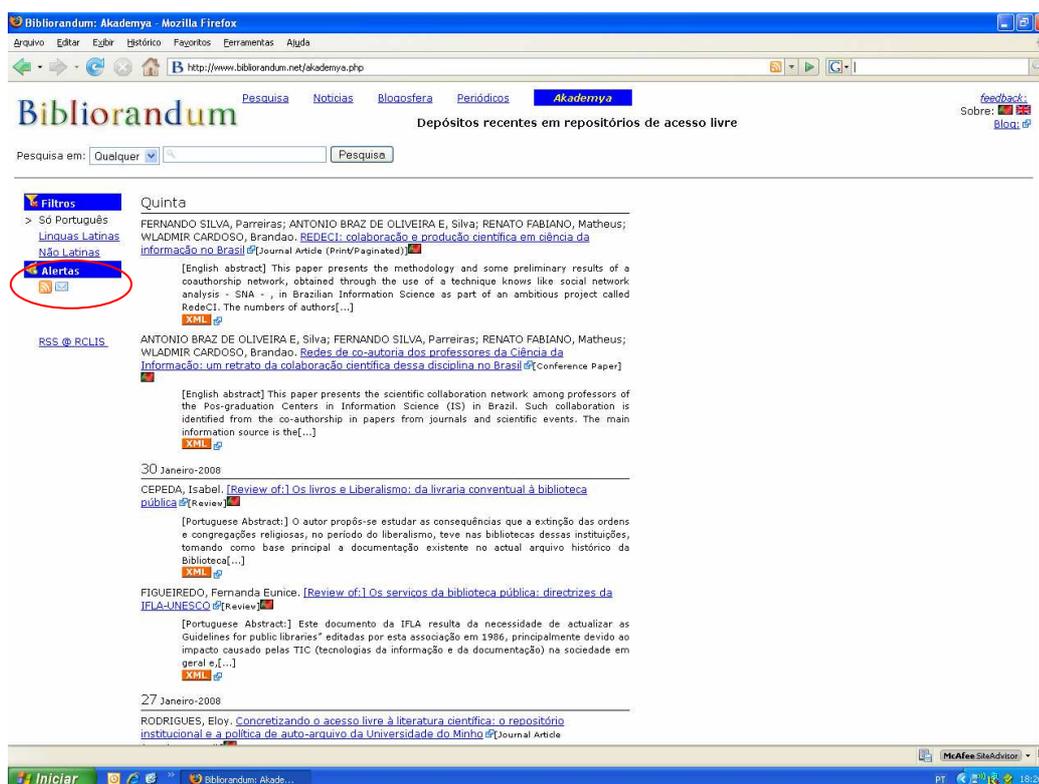


Figura 6.4 – Tela do *Bibliorandum*: feed RSS da produção científica recente

Cabe ressaltar que, apesar de compartilharem a mesma estrutura tecnológica na base de sua implementação (tanto o RSS quanto o OAI-PMH utilizam documentos XML que são transportados através do protocolo HTTP), os dois possuem objetivos distintos. Enquanto o RSS é predominantemente utilizado para disseminação de informação, o OAI-PMH está focado no processo de coleta (*harvesting*) de metadados (HAMMOND et al., 2004; BERNARDINO, 2006, p. 71).

A versatilidade do RSS também pode ser verificada em aplicações originais, como a transmissão e interoperabilidade de vocabulários de metadados científicos baseados em RDF/XML, tais como CML<sup>58</sup> – *Chemical Markup Language*, usado para representação de dados químicos, e MathM<sup>59</sup> – *Mathematical Markup Language*, especializado na exibição de caracteres e expressões matemáticas, entre outros.

<sup>58</sup> <http://www.xml-cml.org>

<sup>59</sup> <http://www.w3.org/TR/MathML2>

## **CAPÍTULO 7**

### **Análise dos Resultados**

---

O presente capítulo visa agrupar e sistematizar os dados obtidos junto à amostra de usuários de periódicos científicos eletrônicos da área da Ciência da Informação após a aplicação dos instrumentos de coleta. Do mesmo modo, procura-se analisar os resultados observados com o experimento de criação do serviço de “agregação de conteúdo” temático.

Conforme explicado no Capítulo 4 – Procedimentos Metodológicos, foram utilizados dois instrumentos de coleta distintos. Assim, quando for possível, tentaremos agrupar as respostas enviadas pelos pesquisadores por meio do questionário *on-line* juntamente com as obtidas pelos coordenadores dos GTs da ANCIB, permitindo a complementação de informações.

### **7.1 Características da População Estudada**

O questionário *on-line* foi disponibilizado por meio do serviço SurveyGizmo entre os dias 26 de dezembro de 2007 e 14 de janeiro de 2008. Nesse período, foram enviadas duas mensagens para o grupo de discussão da ANCIB ([ancib@yahoogroups.com.br](mailto:ancib@yahoogroups.com.br)) com o objetivo de convidar os pesquisadores a responderem o questionário. A primeira mensagem foi enviada no dia 26/12/07 pela própria presidente da Associação, Profa. Dra. Marisa Brascher de Medeiros. A segunda mensagem foi encaminhada no dia 08/01/08 pelo pesquisador, reforçando o convite aos membros da lista.

Ao final dos 20 dias em que o questionário esteve no ar, contabilizamos 149 respostas (32,4% da amostra estimada em 460 endereços de e-mail cadastrados pelo grupo de discussão), sendo 104 válidos e 45 com respostas inválidas ou incompletas. Acreditamos que o retorno obtido foi significativo, se levarmos em conta a época da coleta de dados, coincidindo com o período de férias de docentes e discentes. Além disso, temos que considerar o fato de que muitos indivíduos podem possuir mais de um endereço eletrônico cadastrado no grupo e/ou alguns desses endereços podem não ser mais válidos.

Já o retorno dos formulários preenchidos pelos coordenadores dos grupos de trabalho da ANCIB correspondeu à totalidade do grupo pesquisado. As informações sobre os totais de questionários recebidos e analisados podem ser verificadas na Tabela 7.1.

**Tabela 7.1 – Total de respostas**

Usuários	Válidos		Inválidos		Total		Total estimado	
	Nº	%	Nº	%	Nº	%	Nº	%
Pesquisadores em geral	104	22,6	45	9,8	149	32,4	460	100
Coordenadores dos GTs	7	100	0	0	7	100	7	100

De acordo com a Tabela 7.2, a maior parte dos respondentes do questionário *on-line* (17,3%) é constituída por jovens pesquisadores de 26 a 30 anos, seguido pelos pesquisadores da faixa etária entre 41 a 45 anos (15,4%). Entre os usuários que nos responderam destacamos, também, um número significativo de pesquisadores experientes, com idade superior a 60 anos (13,5%), o que indica uma polaridade interessante na amostra estudada.

**Tabela 7.2 – Faixa etária**

Faixa etária	Nº	%
20 a 25	10	9,6
26 a 30	18	17,3
31 a 35	15	14,4
36 a 40	7	6,7
41 a 45	16	15,4
46 a 50	8	7,8
51 a 55	7	6,7
56 a 60	9	8,6
mais de 60	14	13,5
<b>Total</b>	<b>104</b>	<b>100</b>

Com relação à formação acadêmica (Tabela 7.3), verificamos um equilíbrio entre os pesquisadores com grau de doutor (38,5%) e de mestre (37,5%). Apenas uma pequena parte dos respondentes declarou ter apenas a graduação como titulação mais elevada. Estamos lidando com uma população qualificada por sua

natureza, portanto, já esperávamos esse tipo de participação. Em relação às áreas de formação, em sua maior parte o público investigado tem origem na própria Ciência da Informação (35,6%), seguido pela Biblioteconomia (32,7%) e Ciência da Comunicação (12,5%).

**Tabela 7.3** – Titulação acadêmica

<b>Titulação</b>	<b>Nº</b>	<b>%</b>
Graduação	25	24
Mestrado	39	37,5
Doutorado	40	38,5
<b>Total</b>	<b>104</b>	<b>100</b>

O questionário *on-line* foi respondido, em sua maioria, por pesquisadores residentes no estado de Minas Gerais (26,9%), seguido pelo Distrito Federal (16,3%). Rio de Janeiro e São Paulo, empatadas com 13,5%, aparecem em terceiro lugar entre as unidades da federação que apresentaram o maior número de colaboradores. O estudo, no entanto, contou com a participação de usuários de 15 estados brasileiros diferentes, conforme discriminado na Tabela 7.4.

**Tabela 7.4** – Estado de origem

<b>Estado</b>	<b>Nº</b>	<b>%</b>
MG	28	26,9
DF	17	16,3
RJ	14	13,4
SP	14	13,4
BA	6	5,77
PR	5	4,81
RS	5	4,81
SC	4	3,85
ES	3	2,88
PB	2	1,92
PE	2	1,92
GO	1	0,96
MA	1	0,96
MT	1	0,96
PA	1	0,96
<b>Total</b>	<b>104</b>	<b>100</b>

Quanto ao perfil de atuação dos pesquisadores em relação às publicações científicas, verificamos pela Tabela 7.5 que a totalidade dos respondentes é formada por leitores das revistas, sendo que 64,4% declararam ser também autores de artigos científicos e uma boa parte da amostra (42,3%) aparece envolvida no processo de avaliação por pares desses artigos.

**Tabela 7.5** – Perfil de atuação\*

<b>Atuação</b>	<b>Nº</b>	<b>%</b>
Leitor	104	100
Autor	67	64,4
Avaliador	44	42,3
Editor	7	6,7

\*Respostas múltiplas

Finalmente, a Tabela 7.6 indica que a maioria da nossa amostra (25,6%) é formada por pesquisadores cujos trabalhos possuem afinidades com o grupo sobre Organização e Representação do Conhecimento (GT 2) da ANCIB. Os pesquisadores vinculados ao GT 7 (Produção e Comunicação da Informação em CT&I), aparecem em seguida com 24,3%.

Essa constatação nos despertou para o fato de que a presente pesquisa pode estar inserida no contexto de interesse desses dois Grupos de Trabalho, indicando uma afinidade, mesmo que não expressa, desta parcela de pesquisadores com o tema aqui estudado.

**Tabela 7.6** – Vínculo aos GTs da ANCIB\*

<b>GT</b>	<b>Nº</b>	<b>%</b>
GT 2	20	25,6
GT 7	19	24,3
GT 4	18	23,0
GT 3	14	17,9
GT 1	10	12,8
GT 5	4	5,1
GT 6	4	5,1

\*Respostas múltiplas

## 7.2 Comportamento Informacional

Uma vez caracterizada a amostra, procuramos avaliar o comportamento junto aos periódicos eletrônicos da área de modo a satisfazer às necessidades de informação do grupo. Assim, analisamos algumas características desse comportamento por meio das variáveis identificadas no item 4.2 (Etapas da Pesquisa) do Capítulo 4 – Metodologia.

### 7.2.1 Freqüência de Uso de Periódicos em Ciência da Informação

A Tabela 7.7 apresenta um panorama sobre a freqüência de utilização de 10 periódicos da área da Ciência da Informação selecionados como grupo de amostra para o estudo. A partir dos dados tabulados, observamos que a revista mais utilizada é a tradicional *Ciência da Informação*, editada pelo IBICT, que foi apontada na pesquisa por 102 usuários, dos quais 20 (19,6%) afirmaram acessar a revista com periodicidade semanal.

Outras revistas bastante utilizadas, segundo o levantamento, são a *Perspectivas em Ciência da Informação* (UFMG) e *Transinformação* (PUC-Campinas). Ambas foram indicadas por 97 usuários, porém a *Perspectivas* leva vantagem em termos da freqüência de acesso: 15 pessoas disseram fazer uso da revista a cada sete dias, em média, contra 8 visitas semanais verificadas à revista *Transinformação*.

Em outro extremo, a revista eletrônica *Arquivística.net* foi lembrada por 76 pessoas, das quais 56 (73,7%) revelaram que, apesar de conhecer a publicação, não costumam acessar a revista. Outros baixos índices de utilização foram verificados nas revistas *Em Questão* (Universidade Federal do Rio Grande do Sul) e na Revista ACB, editada pela Associação Catarinense de Biblioteconomia.

Tabela 7.7 – Frequência de utilização das revistas em Ciência da Informação\*

Revistas	Frequência					Total
	Conhece, mas não usa	Semanalmente	Quinzenalmente	Mensalmente	Semestralmente	
Arquivística.net	56 73,7 %	3 3,9%	4 5,3%	5 6,6%	8 10,5%	<b>76</b>
Ciência da Informação	2 2,0%	20 19,6%	28 27,5%	36 35,3%	16 15,7%	<b>102</b>
Em Questão	41 51,9%	3 3,8%	4 5,1%	12 15,2%	19 24,1%	<b>79</b>
Informação e Sociedade: Estudos	13 14,3%	10 11,0%	14 15,4%	30 33,0%	24 26,4%	<b>91</b>
Informação e Informação	25 28,7%	4 4,6%	10 11,5%	22 25,3%	26 29,9%	<b>87</b>
Perspectivas em Ciência da Informação	4 4,1%	15 15,5%	22 22,7%	39 40,2%	17 17,5%	<b>97</b>
Revista ACB	59 72,8%	-	2 2,5%	6 7,4%	14 17,3%	<b>81</b>
Rev. Bras. de Bibliot. e Documentação	36 39,6%	4 4,4%	9 9,9%	18 19,8%	24 26,4%	<b>91</b>
Rev. Digital de Biblioteconomia e Ci	33 37,1%	8 9,0%	7 7,9%	15 16,9%	26 29,2%	<b>89</b>
Transinformação	10 10,3%	8 8,2%	21 21,6%	37 38,1%	21 21,6%	<b>97</b>
<b>Média (%)</b>	<b>31,3</b>	<b>8,40</b>	<b>13,6</b>	<b>24,7</b>	<b>21,9</b>	

\*Respostas múltiplas

Em uma observação mais atenta, verificamos que a maior parte dos usuários costuma acessar esse tipo de publicação em períodos mais longos de tempo, o que pode ser justificado pela periodicidade demorada na atualização das revistas científicas. Assim, mesmo as revistas mais utilizadas são mais frequentemente consultadas em períodos superiores a um mês ou, ainda, a cada seis meses, conforme verificado nas médias percentuais.

### 7.2.2 Utilização de Serviços de Alerta

De acordo com a Tabela 7.8, a maioria dos respondentes (72,1%) considera “extremamente importante” o recebimento de notificações sobre a atualização dos conteúdos em um periódico científico, porém um percentual menor (64,4%) declarou fazer uso de serviços de alerta para manterem-se informados sobre o lançamento de uma nova edição, por exemplo. (Tabela 7.9).

**Tabela 7.8** – Importância da notificação de conteúdo

<b>Grau de importância</b>	<b>Nº</b>	<b>%</b>
Extremamente importante	75	72,1
Importante	25	24
Irrelevante	4	3,9
<b>Total</b>	<b>104</b>	<b>100</b>

**Tabela 7.9** – Uso de serviços de alerta

<b>Uso de Alerta</b>	<b>Nº</b>	<b>%</b>
Sim	68	65,4
Não	32	30,8
Não responderam	4	3,8
<b>Total</b>	<b>104</b>	<b>100</b>

Constatamos pela Tabela 7.10 que o formato de alerta mais freqüente, segundo 58,8% dos respondentes que declararam fazer uso desse tipo de serviço, é o recebimento de mensagem de correio eletrônico contendo o sumário do periódico. O encaminhamento de um *link* para a publicação aparece em segundo lugar (25%), seguido pelo recebimento de mensagem avisando sobre o lançamento da publicação (10,3%).

**Tabela 7.10** – Tipo de serviço de alerta\*

<b>Tipo</b>	<b>Nº</b>	<b>%</b>
e-mail com sumário da revista	40	58,8
e-mail com <i>link</i>	17	25
e-mail com aviso simples	7	10,3
mais de um tipo	2	2,9
serviços de alerta de bibliotecas	1	1,5
outros	1	1,5
<b>Total</b>	<b>68</b>	<b>100</b>

\*Respostas múltiplas

Já a Tabela 7.11 mostra que, do total de 68 usuários de alerta, 35 (51,5%) classificaram o serviço utilizado como “satisfatório”, atendendo parcialmente suas expectativas. Outros 42,7%, por outro lado, garantem que se sentem totalmente atendidos, considerando o serviço “totalmente satisfatório” quanto ao envio de notificações tão logo as edições dos periódicos sejam publicadas.

**Tabela 7.11** – Classificação do serviço de alerta

<b>Classificação</b>	<b>Nº</b>	<b>%</b>
Satisfatório	35	51,5
Totalmente satisfatório	29	42,7
Insatisfatório	2	2,9
Não responderam	2	2,9
<b>Total</b>	<b>68</b>	<b>100</b>

Os índices verificados até aqui demonstram que, de uma maneira geral, os pesquisadores desejam receber notificações sobre as edições correntes e que boa

parte deles já fazem uso de serviços de alerta, principalmente contendo o sumário das revistas. Também notamos que esses pesquisadores consideram o serviço utilizado satisfatório ou totalmente satisfatório, ou seja, atendendo às expectativas de uso. Os resultados apresentados até aqui indicam a preocupação dos pesquisadores em manterem-se sempre atualizados em relação às suas fontes. Estar informado o quanto antes sobre a publicação de um periódico de seu interesse pode ser um diferencial importante na atividade de pesquisa.

### 7.2.3 Profundidade de Leitura e Recuperação de Informação

Em nossa investigação procuramos verificar a profundidade com que os usuários realizam a leitura de informações em periódicos eletrônicos (Tabela 7.12). Para a maioria dos respondentes (70,3%), apenas a leitura do resumo dos artigos é capaz de mantê-los atualizados. Mas para 23,7%, além dos resumos, faz-se importante também a leitura dos sumários do periódico.

Tabela 7.12 – Hábito de leitura\*

Hábito de leitura em periódicos eletrônicos	Nº	%
Na maioria das vezes, apenas o resumo me mantém atualizado. Mas em alguns casos, procuro ler os textos integrais dos artigos	71	70,3
Além dos sumários, tenho hábito de ler os resumos dos artigos	24	23,7
Costumo ler apenas os sumários das revistas	11	10,8
Sempre leio os textos integrais de todos os artigos dos periódicos	2	1,98

\*Respostas múltiplas

O fato dos pesquisadores realizarem a leitura prioritariamente dos resumos dos artigos para se atualizarem indica que, mais do que desejar estar bem informado sobre o maior número de assuntos possíveis, de uma maneira geral, o usuário de um serviço eletrônico de informação não costuma perder tempo com a leitura de um texto longo na tela do computador. Caso um determinado artigo lhe

desperte interesse, o usuário tem a possibilidade de imprimir o documento para ser lido posteriormente.

Sob esse tema, consideramos bastante oportuno o comentário deixado pela pesquisadora Lena Vania Ribeiro Pinheiro, coordenadora do GT 1 (Estudos Epistemológicos de Informação). Sua opinião representa a situação enfrentada e o comportamento da maioria dos pesquisadores, independente de sua área de atuação: “o resumo é muito importante por que representa o conteúdo do artigo e é por este meio que sei se é pertinente e devo ler o artigo na íntegra. O pesquisador não tem tempo para ler tudo que é publicado de sua área, então o resumo é um excelente indicador. Dispensio o resumo somente quando tenho interesse direto no autor ou no artigo. Resumos são indispensáveis em sistemas de alerta ou serviço de DSI”.

A leitura praticamente “dinâmica” a partir dos resumos de um determinado conteúdo caracteriza um comportamento típico de uma navegação por meio de *feeds* RSS, indicando que os nossos respondentes possuem características de usuários potenciais desta tecnologia.

Recordamos que os canais RSS, de uma maneira geral, disponibilizam apenas uma breve descrição dos documentos (o resumo, no caso dos artigos científicos). Assim, no caso do usuário se interessar pela leitura do texto completo, tudo o que precisará fazer é clicar sobre o título deste documento para ser direcionado ao sítio no qual se localiza a informação da forma original.

Outro hábito importante diz respeito à utilização das ferramentas de busca disponibilizadas pelos periódicos. Nesse sentido, os resultados agrupados na Tabela 7.13 mostram que 59,6% dos respondentes do questionário fazem uso eventual dessas ferramentas, enquanto a maioria do grupo de coordenadores dos GTs da ANCIB (71,4%) apontou o uso freqüente destes mecanismos.

Essa aparente contradição pode ser interpretada como um indicativo de que o uso mais intensivo dos sistemas de recuperação por parte de pesquisadores notadamente com mais experiência está relacionado com um maior grau de

necessidade de informação que, em última instância, pode até implicar no conseqüente aumento de sua produção científica.

**Tabela 7.13** – Utilização de ferramentas de busca

Usa ferramenta de busca	Pesquisadores em geral		Coordenadores dos GTs	
	Nº	%	Nº	%
Eventualmente	62	59,6	2	28,6
Sim, frequentemente	38	36,5	5	71,4
Desnecessária	1	1	-	-
Não responderam	3	2,9	-	-
<b>Total</b>	<b>104</b>	<b>100</b>	<b>7</b>	<b>100</b>

Ainda em relação ao aspecto da recuperação de informação, questionamos a respeito do desejo dos pesquisadores na realização de buscas simultâneas por termos livres em mais de uma fonte de informação (Tabela 7.14). A maioria absoluta dos respondentes do questionário (93,2%) demonstrou ser favorável a esse tipo de recurso. Essa quase unanimidade não pode ser observada entre os coordenadores dos GTs, sendo que um deles chegou a afirmar que prefere acessar os periódicos de maneira individualizada.

**Tabela 7.14** – Necessidade de buscas simultâneas

Gostaria de fazer buscas simultâneas em várias fontes?	Pesquisadores em geral		Coordenadores dos GTs	
	Nº	%	Nº	%
Sim	97	93,6	3	42,8
Não	7	6,4	4	57,2
<b>Total</b>	<b>104</b>	<b>100</b>	<b>7</b>	<b>100</b>

Também merecem destaque os argumentos dos respondentes do questionário que manifestaram opinião contrária. Para 3 (três) pesquisadores, as buscas por termos livres, em geral, não apresentam bons resultados. O ideal seria a utilização de um vocabulário controlado. Já outro usuário indicou um serviço de meta-busca, o Holmes (<http://www.holmes.feudo.org>), por meio do qual é possível

realizar consultas por termos livres em mais de uma fonte na área da Ciência da Informação utilizando a ferramenta PKP *Harvester*.

Sobre esse ponto, merece consideração o fato de que o funcionamento dos mecanismos de busca dos periódicos científicos pode variar dependendo da tecnologia e do *software* utilizado pelo sistema de publicação. A maioria dos periódicos eletrônicos em Ciência da Informação utiliza a plataforma SEER, mas os respondentes poderiam estar se referindo a outras publicações e, conseqüentemente, às suas *interfaces* de busca. A multiplicidade destes sistemas representa um problema para o usuário, obrigando-o a ter que aprender diferentes técnicas para realizar pesquisas em revistas distintas.

Consideramos pertinente a observação de Oliveira (1996) ao identificar uma maneira de minimizar este problema com o desenvolvimento de uma *interface* única capaz de agregar os diferentes títulos, permitindo a realização de buscas simultâneas e o uso de apenas uma única ferramenta.

Para que isso seja possível, os títulos devem possuir protocolos de intercâmbio de dados, o que nem sempre ocorre, dificultando iniciativas de consórcios que disponibilizam títulos de diversas editoras e também alguns de livre acesso (OLIVEIRA, 1996, p. 96)

Esta constatação nos faz invocar, mais uma vez, as funcionalidades da tecnologia RSS no que diz respeito à possibilidade de integração de dados entre os diferentes sistemas de recuperação das revistas científicas eletrônicas e, mais ainda, uma vez que esses títulos disponham de *feeds* RSS, a simples utilização de um *software* ou sistema agregador de conteúdo facilitará a recuperação de conteúdos disponibilizados por diferentes fontes, por meio de uma *interface* única.

### 7.3 Conhecimento e Uso da Tecnologia RSS

Por estarmos tratando de um assunto ainda pouco difundido no meio acadêmico, um de nossos objetivos específicos foi justamente descobrir o grau de conhecimento dos usuários de periódicos científicos eletrônicos a respeito da

tecnologia de sindicaco de contedo, representada pelos formatos RSS. Para isso utilizamos os seguintes parâmetros:

**conhecimento de RSS** – procura averiguar se o usurio possui conhecimento (ou pelo menos tem noo) da tecnologia RSS.

**uso de RSS** – identifica se o usurio faz uso de *feeds* RSS em seu dia a dia. Caso afirmativo, procura levantar qual(is) tipo(s) de fontes de informao normalmente utiliza.

**assinatura de *feeds* em peridico cientfico** – refere-se à inteno do usurio em utilizar a tecnologia por meio de subscrio de *feeds* RSS de peridicos cientficos eletrnicos

**identificao de barreiras para uso de *feeds*** – refere-se aos meios e processos que, segundo os usurios, atrapalhariam a utilizao de *feeds* RSS no âmbito dos peridicos cientficos eletrnicos.

Como pode ser observado na Tabela 7.15, a maioria dos pesquisadores da rea da Cincia da Informao (60,6%) afirmou ter conhecimento (ou pelo menos j ter ouvido falar) a respeito da tecnologia RSS. Entre os coordenadores de GTs da ANCIB, esse percentual ficou em 57,2%.

**Tabela 7.15** – Conhecimento da tecnologia RSS

Resposta	Pesquisadores em geral		Coordenadores dos GTs	
	Nº	%	Nº	%
Sim	63	60,6	4	57,2
No	41	39,4	3	42,8
<b>Total</b>	<b>104</b>	<b>100</b>	<b>7</b>	<b>100</b>

Apesar do conhecimento declarado sobre a tecnologia, curiosamente a maior parte dos respondentes (46%) afirmou no cultivar o hbito de usar RSS. Entre os coordenadores de GTs esse nmero  elevado para 71,4%.

Procuramos saber se a faixa etária do usuário poderia exercer alguma influência quanto a uma possível resistência à tecnologia, supondo que pesquisadores de mais idade seriam menos favoráveis às novas tecnologias do que os mais novos (MEADOWS, 1999). Porém, este fato não pôde ser comprovado. A adoção da tecnologia é observada em todas as idades, porém existe uma concentração maior de uso tanto na faixa compreendida entre os 31 quanto na dos 56 anos. Por outro lado encontramos um alto índice de pessoas nas faixas de 20 a 30 anos que, apesar de conhecerem, não utilizam RSS. O mesmo aconteceu com os pesquisadores na faixa de 51 a 55 anos (Tabela 7.16).

**Tabela 7.16** – Relação entre faixa etária e uso de RSS

Faixa etária	Uso de RSS			
	Sim		Não	
	Nº	%*	Nº	%*
20 a 25	4	50	4	50
26 a 30	6	50	6	50
31 a 35	7	58,4	5	41,6
36 a 40	4	80	1	20
41 a 45	5	50	5	50
46 a 50	2	40	3	60
51 a 55	1	15	3	75
56 a 60	2	66,6	1	33,3
mais de 60	3	75	1	15

\*em relação ao total de respondentes da mesma faixa etária

Já entre os usuários que afirmam assinar *feeds* RSS, a preferência com relação ao tipo de conteúdo caiu sobre os jornais e revistas (42,8%). De fato, esses veículos são os que disponibilizam a grande maioria dos *feeds* existentes. Os *blogs*, com 28,6% aparecem em segundo lugar. Os (poucos) *feeds* oferecidos pelos serviços de informação aparecem em último lugar dentre as opções apresentadas, com 19%, conforme indicado na Tabela 7.17.

Tabela 7.17 – Hábito de assinatura de *feeds* RSS\*

Assina <i>feeds</i> RSS?	Pesquisadores em geral		Coordenadores dos GTs	
	Nº	%	Nº	%
Não	29	46	5	71,4
Sim, jornais/revistas	27	42,8	1	14,3
Sim, <i>blogs</i>	18	28,6	0	-
Sim, serviços / unidades de informação	12	19	1	14,3

\*Respostas múltiplas

Quando questionados se estes usuários estariam dispostos a assinar *feeds* RSS de periódicos científicos eletrônicos (caso fossem disponibilizados), mais uma vez a maioria dos usuários que afirmou conhecer a tecnologia (n=63) respondeu afirmativamente. De acordo com a Tabela 7.18, o índice verificado foi de 86,1% entre os pesquisadores em geral e de 57,1% entre os coordenadores dos GTs.

Tabela 7.18 – Assinatura de *feeds* RSS em periódicos científicos eletrônicos

Assinaria <i>feeds</i> em periódicos científicos?	Pesquisadores em geral		Coordenadores dos GTs	
	Nº	%	Nº	%
Sim	56	88,9	4	57,1
Ainda não é uma necessidade	2	3,3	1	14,2
Depende do preço	1	1,6	1	14,2
Falta paciência	1	1,6	0	-
O período entre as edições é grande, acredito que por e-mail é suficiente	1	1,6	0	-
Os alertas por e-mail são satisfatórios	1	1,6	0	-
Quando quero uma informação eu vou atrás, não gosto de ser importunado	1	1,6	1	14,2
<b>Total</b>	<b>63</b>	<b>100</b>	<b>7</b>	<b>100</b>

Os motivos apresentados pelos usuários que se manifestaram contrariamente à inscrição de *feeds* em periódicos científicos estiveram basicamente focados na simples falta de necessidade no momento, segundo dois pesquisadores (n=63) e um coordenador de GT (n=7). Outras questões abordadas foram a autonomia do usuário em relação à busca de informação. “Quando eu quero uma informação eu vou atrás,

não gosto de ser importunado”, afirmaram dois usuários (um pesquisador e um coordenador de GT). Também verificamos a preocupação de outros dois usuários quanto à questão dos custos envolvidos, o que demonstra a falta de familiaridade com a tecnologia RSS, uma vez que o uso dos *feeds* em si é totalmente gratuito.

De todas as opiniões, no entanto, a que nos chamou mais a atenção foi uma observação a respeito da demora na atualização da informação de uma revista científica, algo que, segundo este respondente, não justificaria o uso de *feeds* em periódicos desta natureza. “O período entre as edições é grande, acredito que por e-mail é suficiente”. Outro pesquisador, aliás, também se mostrou conformado com o recebimento de notificações de novos conteúdos de periódicos via correio eletrônico. “Os alertas por e-mail são satisfatórios”, afirmou.

Sobre isso, temos a comentar que a adoção da tecnologia RSS em periódicos científicos eletrônicos não se limita apenas à sua aplicação em serviços de notificação automática. Para tanto, os serviços tradicionais de alerta via correio eletrônico, principalmente, vêm cumprindo bem o seu papel e atendem à necessidade do usuário de uma maneira geral, como esta mesma pesquisa pôde constatar. Acreditamos, porém, que os formatos RSS aplicados à comunicação científica podem vir a representar uma função bem mais nobre, seja por meio de sua utilização como disseminação seletiva, mas, principalmente, como recurso tecnológico que promove o intercâmbio de informações e reutilização por outras fontes primárias (agregação e distribuição de conteúdo) ou mesmo via integração com as fontes secundárias, como as bases de dados, bibliotecas e repositórios digitais.

A Tabela 7.19 relaciona as opiniões dos pesquisadores quanto às eventuais barreiras identificadas para a implantação desta tecnologia. Para cerca de 87% dos respondentes que afirmaram conhecer a tecnologia (n=63) não existe qualquer impedimento em relação à utilização de RSS em periódicos científicos eletrônicos. Entre os coordenadores dos GTs, esse índice foi de 75% (n=4).

Tabela 7.19 – Barreiras identificadas na utilização de feeds RSS

Barreiras	Pesquisadores em geral		Coordenadores dos GTs	
	Nº	%	Nº	%
Nenhuma	55	87,2	3	75
“A depender do usuário de RSS, pode haver uma grande quantidade de itens a serem lidos e as revistas podem se misturar a estes”	1	1,6	0	-
A frequência com que novas informações são disponibilizadas nos periódicos	1	1,6	0	-
“A informação deve ser relevante ao usuário no momento que ele a recebe. Isto é bem complexo já que ele não explicita sua necessidade”	1	1,6	0	-
“Caso seja apenas um serviço de alerta, não vejo utilidade pois o e-mail supre esta necessidade de forma adequada”	1	1,6	0	-
“Excesso de informação na caixa postal eletrônica do assinante”	1	1,6	0	-
“Não existe um API que possibilite o <i>download</i> direto de um artigo através do <i>feed</i> ”.	1	1,6	0	-
“Tempo para aprender e para preparar”	1	1,6	0	-
“Usuário resistente ao uso dessa tecnologia por falta de treinamento nesse tipo de <i>software</i> ”	1	1,6	1	25
<b>Total</b>	<b>63</b>	<b>100</b>	<b>4</b>	<b>100</b>

Apenas 8 (oito) pesquisadores (12,7% do total dos que conhecem o RSS) apontaram eventuais barreiras em relação à aplicação da tecnologia. O interessante é que apenas as três últimas observações constantes na Tabela 8.19 dizem respeito ao aspecto tecnológico, o que, para nós, é considerado o principal obstáculo, quer seja pela necessidade de capacitação do usuário no uso da tecnologia ou mesmo pelo simples fato do RSS ainda não ser um recurso transparente para o usuário comum, ou seja, reconhecido e suportado pelas ferramentas com as quais este usuário está acostumado a lidar (programas de correio eletrônico, navegadores, pacotes de escritório, etc.).

Os outros cinco comentários, por outro lado, identificam aspectos relacionados à maneira pela qual o recurso será disponibilizado, ou seja, a solução depende do uso racional e bom senso por parte do editor ou produtor de conteúdo. Concordamos que o uso de RSS apenas como serviço de alerta é praticamente inútil, como apontou um de nossos respondentes. E que a distribuição de vários

itens sem a devida marcação (por meio do uso de *tags*) colaborará por aumentar a confusão e o excesso de informação em vez de minimizar o efeito de sua sobrecarga. Por isso, o estudo de usuário e da gestão da informação, a adoção de boas práticas para criação e manutenção de *feeds* RSS, além da escolha da especificação adequada ao tipo de aplicação desejada devem fazer parte do planejamento antes da implementação deste recurso em qualquer serviço de informação.

#### 7.4 Atributos dos Artigos de Periódicos

Após um levantamento prévio das características mais comuns dos artigos de periódicos científicos em Ciência da Informação, solicitamos aos 7 (sete) coordenadores dos GTs da ANCIB que conferissem um grau de importância aos atributos considerados essenciais à representação da informação científica. Os resultados apresentados no gráfico da Figura 7.1 indicam que os elementos mais relevantes são o título do artigo, nome do(s) autor(es) e resumo, indicados por toda a amostra como “muito importante”. Em seguida, aparecem os atributos palavras-chave e data de publicação, também classificados como “muito importante” por 5 (cinco) pesquisadores.

Outros atributos relacionados à referência da publicação, como volume, número, e página inicial e final dos artigos também foram classificados pela maioria dos respondentes como “muito importante”. Já as informações sobre o texto (formato, idioma e dimensão) foram votadas com um grau menor de importância.

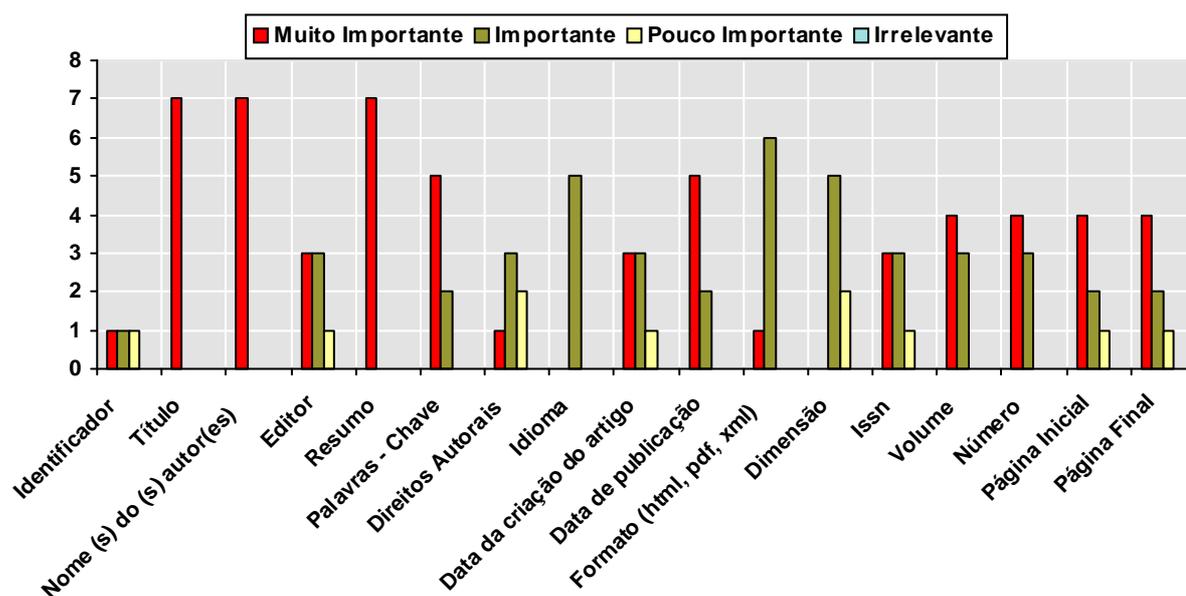


Figura 7.1 – Atributos funcionais dos artigos científicos

Além dos 17 atributos identificados – relacionados no questionário encaminhado aos pesquisadores – recebemos sugestões de outros dois elementos, considerados igualmente relevantes para representação do artigo científico. São esses: local de publicação e filiação dos autores. Já a coordenadora do GT 1, Lena Vania Pinheiro, lembrou das informações que compõem na denominada “legenda bibliográfica” que devem constar em todas as páginas do periódico impresso: título abreviado, local de publicação, volume fascículo, páginas e ano.

A partir deste panorama, podemos perceber que todos os elementos ligados à representação da informação científica estão contemplados, principalmente, pelo conjunto de metadados dos módulos Dublin Core e PRISM (ver Capítulo 5), credenciando a especificação RSS 1.0 como a ideal para ser utilizada em aplicações de sindicância de artigos científicos. Porém, acreditamos que mesmo a versão RSS 2.0 pode ser adotada no caso de aplicações mais simples de disseminação, haja vista que dois elementos considerados essenciais (título e resumo) são contemplados por esta especificação.

Ressaltamos que os principais sistemas utilizados pelas revistas científicas eletrônicas nacionais – SEER e SciELO – oferecem a possibilidade de sindicância dos artigos científicos. No caso do SEER, os conteúdos são disponibilizados nas versões RSS 1.0, RSS 2.0 e Atom. Porém, o serviço precisa ser habilitado pelo editor da revista. Já o SciELO disponibiliza todos os artigos das revistas hospedadas, mas apenas no formato RSS 2.0.

### **7.5 Conteúdos RSS em Periódicos Científicos**

Como já demonstrado, a tecnologia de sindicância de conteúdos Web pode ser utilizada sob diversas perspectivas, dependendo do tipo de aplicação desejada pelo produtor de informação (editor científico, por exemplo), além do tipo de conteúdo a ser disseminado.

Baseados nas experiências das editoras científicas analisadas no Capítulo 6 – A Tecnologia RSS Aplicada à Comunicação Científica, decidimos verificar, junto à nossa amostra selecionada de coordenadores de GTs, quais os conteúdos potencialmente úteis para serem fornecidos no contexto da disseminação de informação científica.

Dentre as opções oferecidas (Figura 7.2), a disponibilização dos “sumários” das revistas e a divulgação de “publicações recentes” foram consideradas as mais importantes na opinião de 5 (cinco) pesquisadores. Em seguida aparecem a divulgação de “eventos” e de “pré-prints”, empatados com 3 (três) votos.

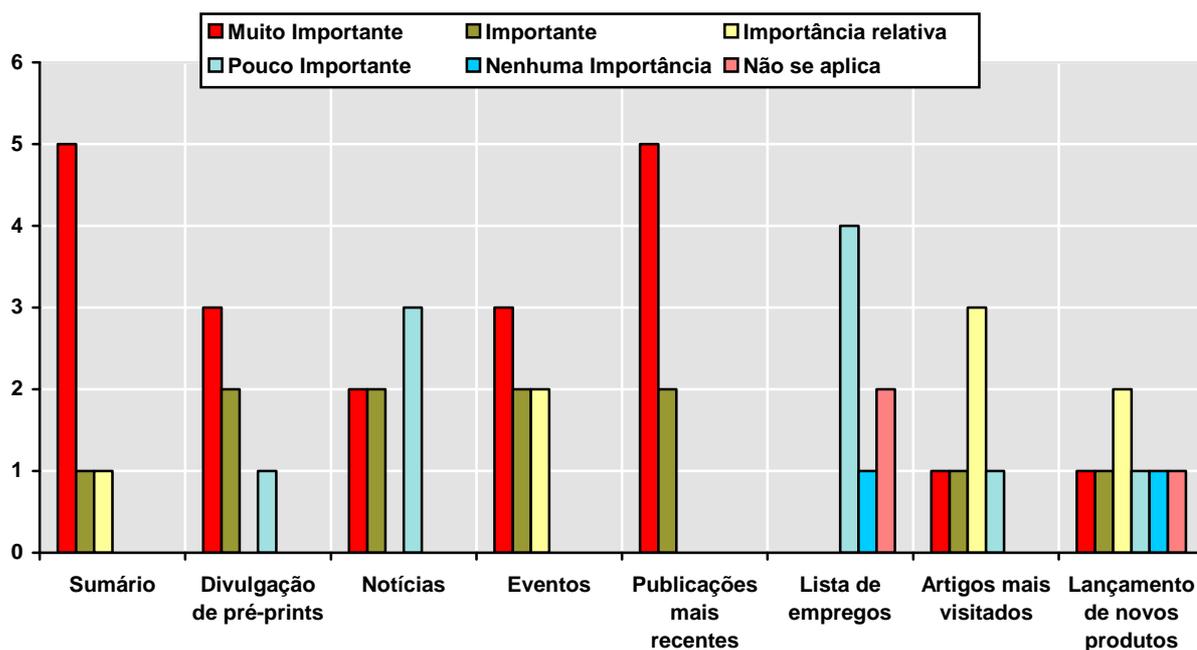


Figura 7.2 – Possíveis aplicações de RSS em periódicos científicos eletrônicos

É interessante constatar que os conteúdos mais cotados coincidem com as aplicações mais freqüentemente encontradas quanto ao uso de *feeds* RSS em periódicos científicos eletrônicos, o que demonstra a vocação deste recurso como importante tecnologia para disseminação, principalmente no que diz respeito à divulgação dos sumários das revistas.

## 7.6 Criação dos *feeds* de Periódicos Nacionais

Quanto ao experimento realizado, inicialmente procurou-se criar os *feeds* RSS dos periódicos nacionais da área de Ciência da Informação. Para isso, foi utilizada uma ferramenta de autoria para facilitar o processo de construção dos documentos (arquivos .xml). A escolha recaiu sobre uma ferramenta proprietária chamada FeedForAll<sup>60</sup> pelo desconhecimento, naquele momento, de outra

<sup>60</sup> <<http://www.feedforall.com>>

alternativa de *software* livre ou de código aberto. Posteriormente, porém, identificou-se uma outra ferramenta gratuita, o RSS Builder<sup>61</sup>.

Independentemente do *software* adotado, as informações básicas do canal (títulos) precisam ser preenchidas por meio de uma espécie de formulário. As informações requeridas são: `<title>` (título do canal), `<link>` (URL da página correspondente ao canal) e `<description>` (uma breve descrição sobre o conteúdo do *feed*). Depois que o canal é criado, faz-se necessário alimentar os conteúdos dos *itens* propriamente ditos, que correspondem às informações básicas dos artigos: `<title>` (do artigo), `<link>` (localização da página do artigo na revista) e `<description>` (resumo do artigo).

O método utilizado para a inserção dos conteúdos foi o acesso individual das páginas correspondentes na revista, copiando e colando as informações pertinentes, tanto para a criação do canal quanto para os artigos. No caso do campo `<description>` do *item* de cada artigo, aproveitamos o resumo fornecido pelo próprio autor. O elemento `<pubDate>` (data de publicação) foi preenchido automaticamente pela ferramenta.

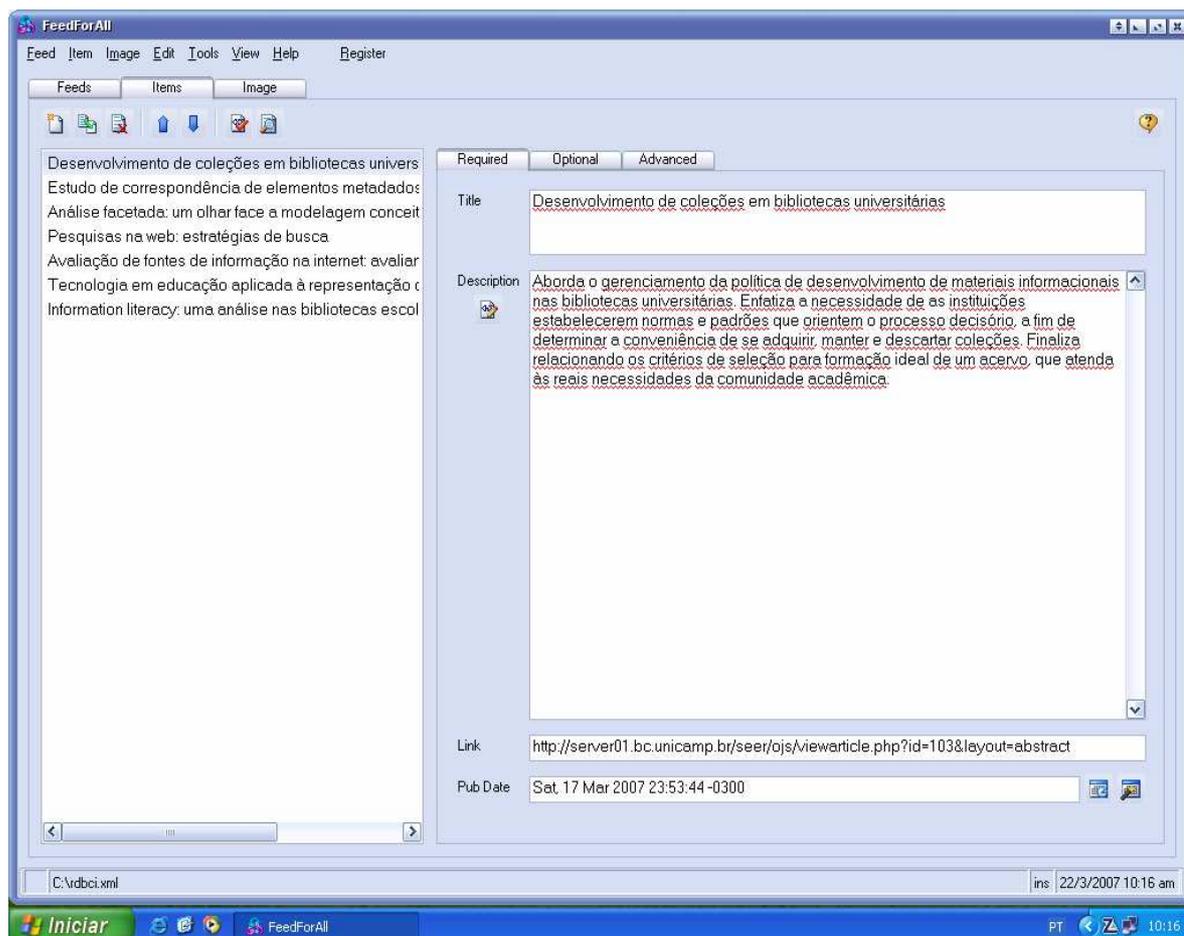
Outras informações adicionais e relevantes, como autoria dos artigos, por exemplo, não foram representadas, devido à limitação da especificação RSS 2.0. Como já explicado, o objetivo naquele momento foi experimentar a tecnologia e, por isso, não houve preocupação com a inserção de metadados adicionais, mas reconhecemos a sua importância. A Figura 7.3 mostra a tela do aplicativo FeedForAll, utilizado em nosso experimento, ilustrando o preenchimento dos campos requeridos dos elementos dos artigos.

Os *feeds* tiveram como base a última edição de cada publicação, atualizados até o mês de maio de 2007. Como esses arquivos foram criados manualmente e não por meio do sistema da revista, na medida em que novas edições foram atualizadas, o mesmo não se procedeu com os *feeds*, não havendo atualizações desde então. Posteriormente, os arquivos foram transferidos para um serviço de hospedagem com o propósito de formar uma base para que estes pudessem ser facilmente

---

<sup>61</sup> `<http://home.hetnet.nl/mr_2/43/bsoft/rssbuilder>`

distribuídos. Mais recentemente, tivemos conhecimento de outras tecnologias<sup>62</sup> para a geração de *feeds* automáticos a partir da varredura e extração de informações de conteúdos Web. Se tivéssemos usado esses serviços, teríamos conseguido produzir um *feed* ativo, que se manteria sempre atualizado.



**Figura 7.3** – Tela do FeedForAll, ferramenta de autoria para criação de *feeds* RSS

<sup>62</sup> Exemplos de ferramentas para geração automática de *feeds* são o Openkapow (<http://openkapow.com>) e Dapper (<http://www.dapper.net>).

## 7.7 Montagem do Protótipo do Agregador Temático

Depois de um breve período de testes, todos os *feeds* foram reunidos (ver Quadro 7.2) em uma aplicação especialmente criada por meio do serviço Netvibes<sup>63</sup>. A ferramenta é capaz de gerenciar módulos criados a partir de *feeds* RSS. Trata-se de uma aplicação gratuita tecnicamente definida como *mashup*<sup>64</sup>.

Fontes	Feed RSS
<b>periódicos nacionais</b>	
1. Arquivística.net	<a href="http://www.rlalmeida.correiovip.com.br/arqnet/arqnet.xml">http://www.rlalmeida.correiovip.com.br/arqnet/arqnet.xml</a>
2. Ciência da Informação	<a href="http://www.ibict.br/cienciainformacao/rss.php">http://www.ibict.br/cienciainformacao/rss.php</a>
3. Em Questão	<a href="http://www.rlalmeida.correiovip.com.br/emquestao/emquestao.xml">http://www.rlalmeida.correiovip.com.br/emquestao/emquestao.xml</a>
4. Informação e Informação	<a href="http://www.rlalmeida.correiovip.com.br/iei/iei.xml">http://www.rlalmeida.correiovip.com.br/iei/iei.xml</a>
5. Informação e Sociedade	<a href="http://www.ies.ufpb.br/ojs2/index.php/ies/feed/rss">http://www.ies.ufpb.br/ojs2/index.php/ies/feed/rss</a>
6. Perspectivas em Ciência da Informação	<a href="http://www.rlalmeida.correiovip.com.br/pci/pci.xml">http://www.rlalmeida.correiovip.com.br/pci/pci.xml</a>
7. Revista ACB	<a href="http://www.rlalmeida.correiovip.com.br/acb/acb.xml">http://www.rlalmeida.correiovip.com.br/acb/acb.xml</a>
8. Revista Brasileira de Bibliotec. e Documentação	<a href="http://www.rlalmeida.correiovip.com.br/rbbd/rbbd.xml">http://www.rlalmeida.correiovip.com.br/rbbd/rbbd.xml</a>
9. Revista Digital de Bibliotec. e Documentação	<a href="http://www.rlalmeida.correiovip.com.br/rdbci/rdbci.xml">http://www.rlalmeida.correiovip.com.br/rdbci/rdbci.xml</a>
10. Transinformação	<a href="http://www.rlalmeida.correiovip.com.br/transinfo/transinfo.xml">http://www.rlalmeida.correiovip.com.br/transinfo/transinfo.xml</a>
<b>periódicos estrangeiros</b>	
1. Ariadne	<a href="http://www.ariadne.ac.uk/rss.xml">http://www.ariadne.ac.uk/rss.xml</a>
2. D-Lib Magazine	<a href="http://www.dlib.org/rss/dlib.rss">http://www.dlib.org/rss/dlib.rss</a>
3. Information Development	<a href="http://idv.sagepub.com/rss/current.xml">http://idv.sagepub.com/rss/current.xml</a>
4. JASIST	<a href="http://www3.interscience.wiley.com/rss/journal/76501873">http://www3.interscience.wiley.com/rss/journal/76501873</a>
5. Journal of Information Science	<a href="http://jis.sagepub.com/rss/current.xml">http://jis.sagepub.com/rss/current.xml</a>
6. Library Journal	<a href="http://feeds.feedburner.com/LibraryJournalNews">http://feeds.feedburner.com/LibraryJournalNews</a>
<b>blogs</b>	
1. A Informação	<a href="http://a-informacao.blogspot.com/atom.xml">http://a-informacao.blogspot.com/atom.xml</a>
2. Bibliotecários Sem Fronteiras	<a href="http://bsf.org.br/feed">http://bsf.org.br/feed</a>
3. Biblioteconomia de Babel	<a href="http://bibbabel.wordpress.com/feed">http://bibbabel.wordpress.com/feed</a>
4. ExtraLibris	<a href="http://feeds.feedburner.com/revistaextralibris">http://feeds.feedburner.com/revistaextralibris</a>
5. Librarians' Internet	<a href="http://lii.org/ntw.rss">http://lii.org/ntw.rss</a>

<sup>63</sup> <<http://www.netvibes.com>>

<sup>64</sup> Segundo a Wikipédia (2007), *mashup* é definido como “um website ou uma aplicação web que usa conteúdo de mais de uma fonte para criar um novo serviço completo”.

6. O Ser Bibliotecário	<a href="http://feeds.feedburner.com/oserbibliotecario/">http://feeds.feedburner.com/oserbibliotecario/</a>
7. Novasinapse.com	<a href="http://feeds.feedburner.com/novasinapse/">http://feeds.feedburner.com/novasinapse/</a>
8. RABCI e BiblioBlogs 2.0	<a href="http://infocultura.info/rabci/rss.xml">http://infocultura.info/rabci/rss.xml</a>
9. Web 2.0 pt	<a href="http://web20pt.wordpress.com/feed">http://web20pt.wordpress.com/feed</a>
<b>repositórios</b>	
1. E-LIS	<a href="http://eprints.rclis.org/last.xml">http://eprints.rclis.org/last.xml</a>
2. EServer TC Library	<a href="http://tc.eserver.org/recent_rss.xml">http://tc.eserver.org/recent_rss.xml</a>

\* após a conclusão do protótipo, a própria revista passou a disponibilizar seus *feeds*

### Quadro 7.2 – Feeds RSS das fontes de informação reunidas pelo agregador temático em CI

Cada módulo criado no âmbito do serviço Netvibes representa uma fonte de informação distinta, cujo conteúdo é proveniente de um *feed* RSS/Atom. No caso do serviço proposto, adicionamos um módulo para cada *feed* criado (no caso dos periódicos nacionais) e *feeds* pesquisados (periódicos estrangeiros, *blogs* e repositórios digitais), como mostrado na Figura 7.4.

Figura 7.4 – Tela inicial do agregador temático em Ciência da Informação

A função geral do sistema é realizar uma varredura periódica em todas as fontes cadastradas, de acordo com a preferência do usuário. A aplicação se encarregará de identificar as atualizações nos conteúdos dos sumários dos periódicos, *posts* dos *blogs* e depósitos mais recentes dos repositórios. Se desejarmos acessar o texto completo de um determinado documento, basta clicar sob o seu título para que sejamos remetidos diretamente à página correspondente à informação nos termos como foi publicada originalmente em seu sítio.

O modelo apresentado demonstra uma das principais aplicações referentes à utilização de RSS. No caso, apresentamos a dinâmica de um mecanismo capaz de reunir, em única página Web, as referências dos artigos (com respectivo resumo) dos periódicos nacionais e estrangeiros, além de *blogs* e repositórios em Ciência da Informação. O modelo poderia ser estendido para agregar, ainda, base de dados ou outras fontes relacionadas, desde que disponham de *feeds* associados, configurando-se como uma espécie de “biblioteca digital pessoal”.

Uma vez tendo todo esse conteúdo agregado, é possível realizar buscas simultâneas em toda a coleção. Se desejarmos fazer uma pesquisa pelo termo “gestão” em todas as fontes, por exemplo, o resultado da busca trará todos os documentos que contenham a expressão “gestão” em seu título ou na sua descrição (Figura 7.5).

Esse recurso, proporcionado pela característica de agregação dos serviços baseados em RSS, configura-se como um diferencial diante de alguns sistemas de recuperação de informação tradicionais, permitindo uma “centralidade no resgate”, obtida pela possibilidade de recuperação simultânea em diversas fontes de informação.

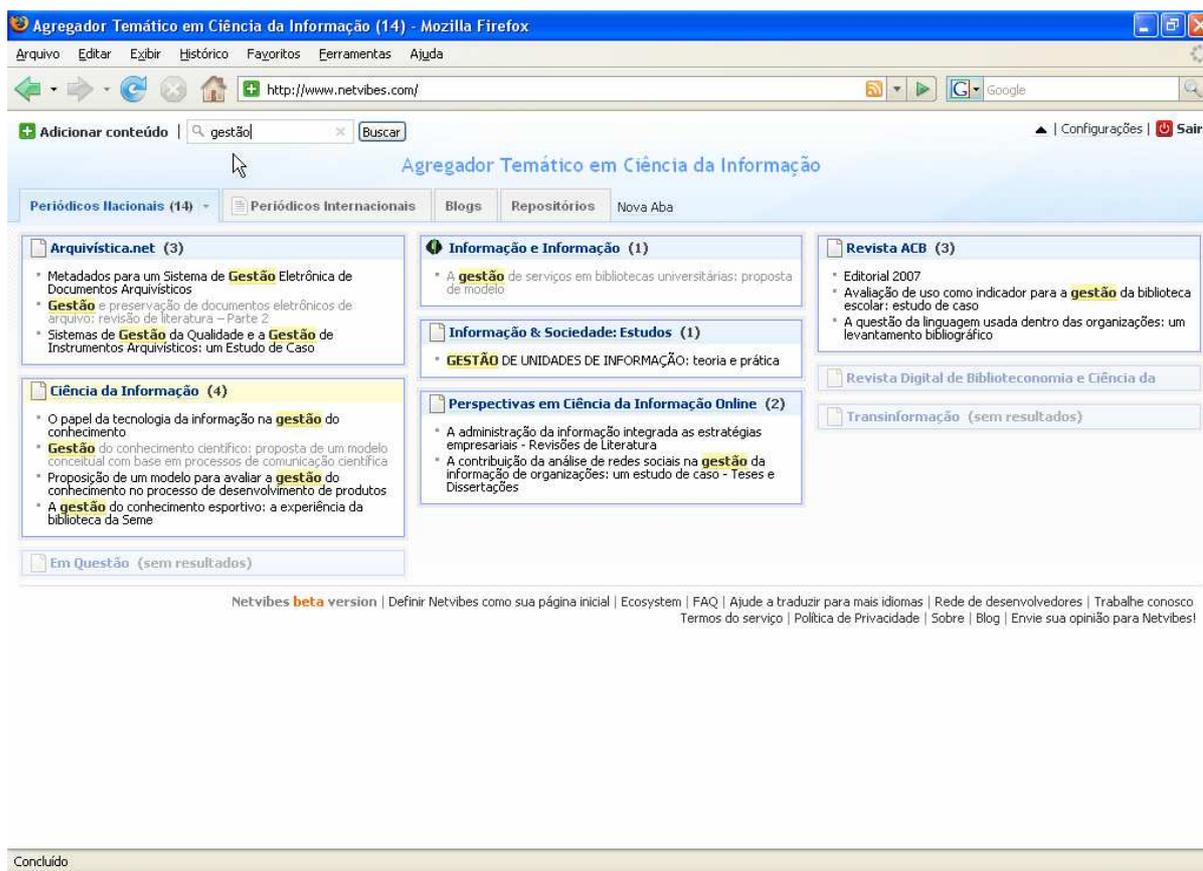


Figura 7.5 – Resultado da busca pelo termo “gestão” em fontes simultâneas

Em um segundo experimento, procurou-se aplicar a dinâmica de utilização de *feeds* como serviço de disseminação seletiva de informação. Para isso, adotamos uma ferramenta gratuita recém-lançada pela *Yahoo!* denominada *Pipes*<sup>65</sup>, um ambiente de criação de *mashups* programável visualmente e, até o momento, sem precedentes. A inspiração do nome vem do comando *pipe* (cano), muito usado pelo sistema operacional Unix para concatenar outros comandos. A analogia de um serviço de disseminação seletiva está no fato de que a aplicação permite a criação de uma “tubulação” por meio da qual é possível juntar várias fontes de informação e filtrar e personalizar o resultado que sai do outro lado.

O trabalho realizado foi o de agregar as 10 revistas nacionais selecionadas na amostra e filtrar os conteúdos dos artigos de modo que pudessem atender a um critério pré-estabelecido. Sem entrarmos no detalhamento da ferramenta,

<sup>65</sup> <<http://pipes.yahoo.com>>

demonstraremos apenas o resultado do serviço *pipe* criado para filtrar o conteúdo das fontes selecionadas que atendam ao critério “comunicação científica”, ao mesmo tempo em que exibe esses artigos segundo a ordem cronológica de publicação (Figura 7.6).

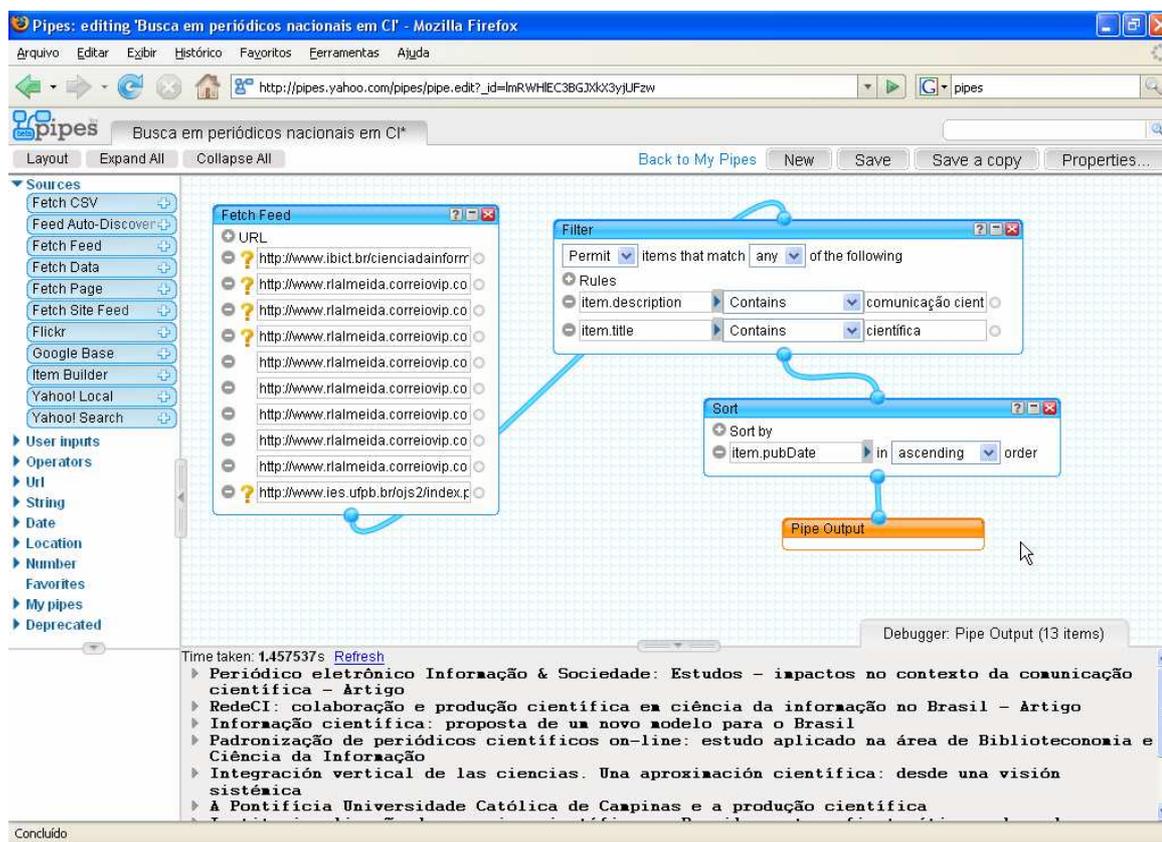


Figura 7.6 – Serviço *pipe* criado para filtragem dos *feeds* agregados

Em termos práticos, este serviço oferece a possibilidade de filtrar os conteúdos a partir de uma ou várias fontes de informação, atendendo aos critérios pré-determinados pelo usuário e encaminhar os resultados por meio de correio eletrônico ou mesmo gerando um *feed* personalizado. A Figura 7.7 mostra a lista completa dos artigos recuperados (com os respectivos *links* para os textos originais) a partir das fontes agregadas que atendem o critério determinado pelo serviço *pipe*, ou seja, segundo o termo “comunicação científica”, ordenando-os por data de publicação.

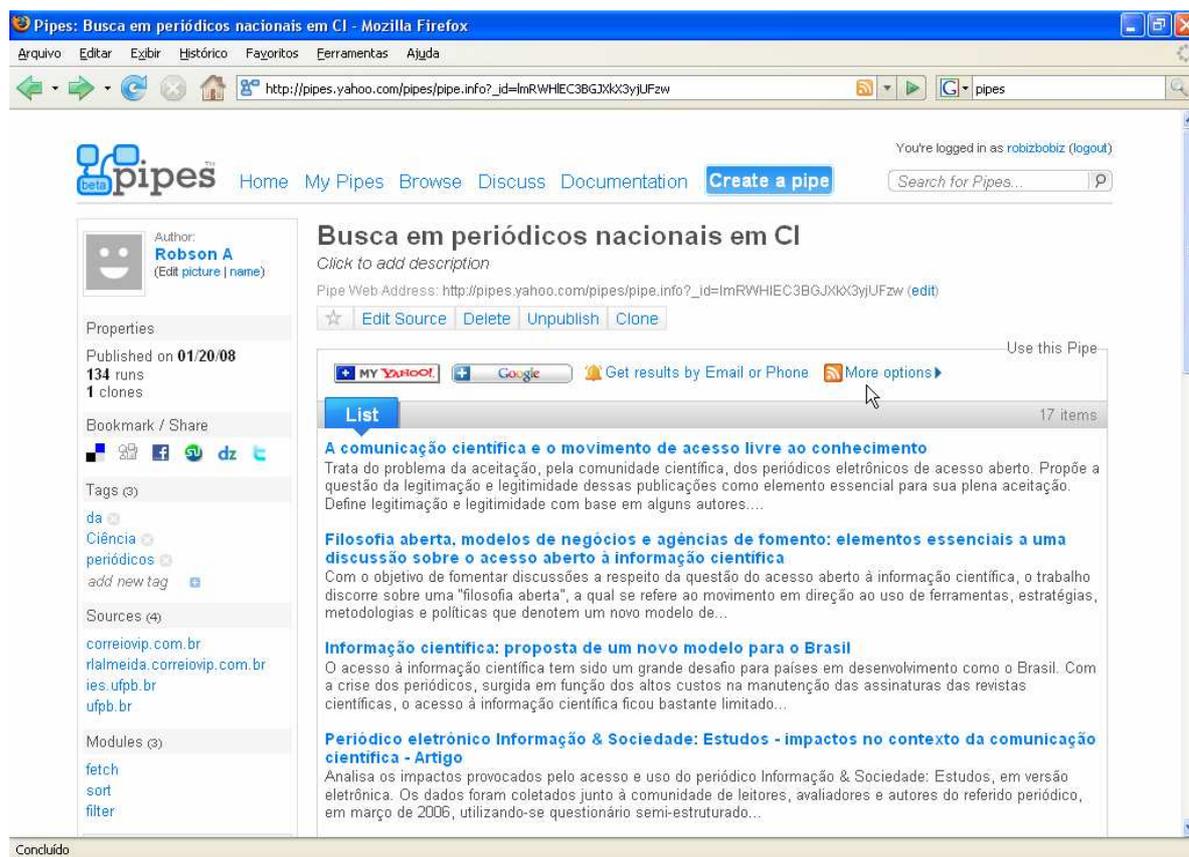


Figura 7.7 – Filtragem de conteúdos a partir de *feeds* agregados

Entendemos que o tipo de aplicação descrita anteriormente pode ser considerado um exemplo, ainda que simplificado, de um serviço de disseminação seletiva de informação, no qual o “perfil” é definido primeiramente, pela própria seleção das fontes consultadas, mas, principalmente, pelo mecanismo de filtragem definido no serviço *pipe*. Sob este aspecto, os resultados observados nos permitem verificar um de nossos objetivos específicos, o de levantar possibilidades da tecnologia RSS como serviço de disseminação seletiva de informação científica.

## **CAPÍTULO 8**

### **Conclusões e Recomendações**

---

No decorrer desta dissertação procurou-se apresentar e discutir as características da tecnologia de sindicância de conteúdos na Web, viabilizada principalmente pelos formatos RSS. Nossa primeira observação, antes de analisarmos algumas tendências que serão relatadas mais adiante, diz respeito ao aspecto inovador deste recurso tecnológico emergente, que o diferencia dentre tantos outros, devido ao vasto leque de possibilidades para sua aplicação. Percebemos que a tecnologia RSS se tornará cada vez mais popular graças à demanda crescente por serviços capazes de fornecer informações personalizadas para os seus usuários.

Considerando a nossa proposta inicial, de estudar as possibilidades de aplicação do RSS como agente facilitador da disseminação de informação científica, acreditamos que cumprimos nossos objetivos, tomando como base a pesquisa documental realizada, que balizou a construção dos capítulos de revisão, mas, principalmente, pelos resultados verificados junto à amostra de pesquisadores envolvidos com os periódicos eletrônicos científicos.

Deste modo, quanto aos objetivos específicos 1 e 2, apresentados no Capítulo 1, item 1.4 (p. 22), dir-se-ia que encontram-se contemplados, respectivamente pelo Capítulo 5, que descreve as características, histórico, arquitetura e modo de operação dos recursos RSS, e no Capítulo 6, que aborda a tecnologia aplicada no contexto acadêmico e científico.

A partir desta cobertura descritiva, podemos concluir que os agentes envolvidos com a comunicação científica (autores, editores, avaliadores, leitores, bibliotecários e profissionais da informação em geral) constituem um público potencialmente beneficiário das oportunidades oferecidas pelo RSS, tanto com relação à absorção de informação por meio da leitura de *feeds* (ponto de vista do usuário) quanto à possibilidade de disseminar informação sob demanda (ponto de vista do produtor de informação).

O levantamento de dados junto aos pesquisadores da ANCIB, inclusive na amostra selecionada dos coordenadores dos Grupos de Trabalho, serviu para atendermos aos objetivos 3 (identificar o comportamento informacional dos usuários

dos periódicos eletrônicos da área da Ciência da Informação). Nossa intenção foi verificar as características básicas do uso da informação por esses pesquisadores em relação aos periódicos científicos eletrônicos (frequência de leitura, profundidade de leitura, uso de mecanismo de busca, etc.) para buscar possíveis correspondências ao estilo da navegação em conteúdos digitais por meio de *feeds* RSS.

Os resultados analisados confirmam as nossas suspeitas de que, pelo menos entre os profissionais da informação, existem características de uso da informação compatíveis com os requisitos estabelecidos pelo RSS, especialmente quanto ao hábito de leitura em baixa profundidade, no nível dos resumos de artigos. A possibilidade de realizar buscas simultâneas em fontes de informação agregadas é outra característica da tecnologia que atende às necessidades explicitadas pela maioria dos pesquisadores, segundo o nosso estudo. Também nos chamou atenção o fato de que um percentual elevado (60,6%) afirmou possuir conhecimento (ou pelo menos ter ouvido falar) a respeito da tecnologia, mostrando pré-disposição para a assinatura de canais RSS em aplicações acadêmicas/científicas, desde que disponibilizadas.

Vale lembrar que o fato de estarmos lidando com uma amostra intencional não permite que sejam feitas generalizações. Talvez se o estudo fosse realizado com pesquisadores de outras áreas teríamos obtido dados diferentes, com uma população mais ou menos resistente à adoção da tecnologia proposta, por exemplo. Porém, suspeitamos que, de uma maneira geral, a utilização do recurso pode ser bem-vinda, principalmente com relação a sua característica de permitir a simplificação da disseminação da informação no contexto da Web.

Falando em disseminação, lembramos que o nosso objetivo 4 foi justamente levantar as possibilidades de utilização da tecnologia RSS como serviço de disseminação seletiva de informação científica. Tratamos deste tópico no Capítulo 6, item 6.2.2 (p. 132-134), quando falamos da criação dinâmica de *feeds* personalizados a partir dos resultados de busca em bases de dados *on-line*, e também no Capítulo 7, item 7.7 (p. 165-167), o qual detalha as etapas da montagem do protótipo do agregador temático, particularmente quanto ao experimento com a

aplicação *Pipe*, utilizada para filtrar conteúdos a partir de fontes de informação distintas, segundo um critério de seleção previamente definido.

Apesar das (poucas) barreiras identificadas neste trabalho, inclusive algumas apontadas pelos pesquisadores da área de Ciência da Informação, o estudo detalhado da tecnologia RSS aponta que esta é plenamente viável de ser incorporada em qualquer sistema de informação baseado na Web, desde as bases de dados *on-line*, bibliotecas e repositórios digitais e, ainda, nos sistemas de publicação de periódicos eletrônicos, como o SEER/OJS, também contemplado em nosso trabalho.

Os periódicos eletrônicos que utilizam a plataforma SEER/OJS já dispõem de um serviço de notificação que oferece aos seus usuários a possibilidade de se cadastrarem para receber, via correio eletrônico, um aviso com o sumário das novas edições à medida que forem publicadas. Se este leitor desejar acompanhar os trabalhos publicados sobre uma determinada área, terá que repetir o procedimento de cadastro em cada um dos periódicos, o que significa dizer que o mesmo receberá uma notificação diferente para cada atualização. Usando RSS, não há necessidade de preenchimento de cadastros e o usuário ainda tem a vantagem de manter a sua caixa postal menos cheia.

Assim, no que se refere à aplicação de *feeds* em periódicos científicos eletrônicos, parece adequada a criação de canais RSS como serviço de alerta, atualizando o leitor quando da publicação de um artigo de seu interesse. Porém, defendemos a utilização do RSS, principalmente, pelo fato de poder representar a informação científica, permitindo a integração com outros recursos tecnológicos de informação (bases de dados, bibliotecas digitais, *blogs*, páginas Web em geral, etc.) por meio de protocolos compatíveis com a linguagem XML, como é o caso do padrão de acesso aberto OAI-PMH, utilizado pela maioria dos repositórios digitais. Neste caso, sugerimos a adoção da especificação RSS 1.0, por permitir a adição de metainformação aos elementos básicos para sindicância de conteúdo.

A fim de aumentar a base de utilização do RSS por parte dos usuários de periódicos científicos, recomenda-se que os editores científicos habilitem, sempre que possível, o módulo de sindicância de seus sistemas de publicação, como no

caso da plataforma SEER/OJS. Sugere-se que também seja criada uma página explicativa do tipo “sobre RSS”, explicando pelo menos o básico a respeito da tecnologia, funcionamento e vantagens para o usuário, além de disponibilizar *links* para programas ou sistemas leitores de RSS.

No que diz respeito à contribuição da literatura estudada, constatamos que a tecnologia RSS, se usada adequadamente, pode constituir-se num eficiente sistema de Disseminação Seletiva de Informação. Também verificamos a aderência do RSS com os ideais da função básica da comunicação científica: o paradigma do compartilhamento da informação, incentivada também pelo movimento de acesso aberto. Além disso, a utilização de RSS configura-se como exemplo prático do conceito de Comunicação Extensiva.

A partir dos resultados obtidos nessa pesquisa, podemos resumir algumas tendências a respeito do potencial da tecnologia de sindicância de conteúdo Web como recurso capaz de incrementar as funcionalidades da comunicação científica com base nos seguintes tópicos:

- o conjunto dos recursos propostos pelos formatos RSS representa uma inovação tecnológica no campo dos serviços digitais de referência, bem como no desenvolvimento de novas funcionalidades para periódicos eletrônicos, bases de dados *on-line*, bibliotecas digitais, etc.;
- verifica-se o RSS como uma ferramenta que agrega valor aos serviços automatizados de notificação (alerta) de conteúdos em ambientes Web, mas principalmente, como recurso para representação da informação, permitindo a reutilização e interoperabilidade com outras fontes de informação, tais como os repositórios digitais;
- o comprometimento dos produtores de conteúdo e desenvolvedores é de fundamental importância para o desenvolvimento de produtos e serviços que façam uso e sejam suportados pela tecnologia RSS;
- vocabulários de metadados, como *Dublin Core* e PRISM são adequados para a descrição dos atributos funcionais dos artigos científicos e podem ser facilmente incorporados à tecnologia RSS, por meio da especificação 1.0;

- mais estudos são necessários para ampliar a discussão sobre o tema, buscando-se novos enfoques e aplicações.

As vantagens da adoção da tecnologia RSS em aplicações acadêmicas parecem ser imensas, a começar pelos benefícios para os usuários, que passarão a dispor de uma poderosa ferramenta por meio da qual poderão realizar buscas simultâneas em cima de várias fontes de informação, aumentando a relevância dos termos recuperados. Além disso, esses usuários terão em mão um “poder” até hoje pouco imaginado: o de personalizar a informação que deseja consumir e de produzir novos conteúdos utilizando-se dos recursos permitidos pelos formatos de sindicância, de maneira rápida e sem a necessidade de intermediários.

Já o produtor de informação, ao disponibilizar seus conteúdos a partir desses formatos, ampliará o processo de divulgação da informação científica, independente da ação do usuário. Ao prover um *feed* para a sua publicação, por exemplo, os editores de revistas científicas estarão permanentemente em contato com seus usuários, aumentando as chances de que um determinado artigo desperte a sua atenção.

A idéia de sindicância de conteúdos na Web existe há mais de 10 anos, mas apenas recentemente a tecnologia RSS vem se difundindo graças ao fenômeno da explosão dos *blogs*. Assim, acreditamos que em um futuro próximo será bem possível que este recurso seja mais amplamente utilizado tanto da parte de quem publica quanto de quem consome informação na Web.

A disponibilização de *feeds* nas mais variadas aplicações, independente de sua natureza, será algo tão natural que o usuário não estranhará mais os ícones indicativos de que um determinado sítio possui um canal RSS. Algumas tarefas, como a localização da agulha de um *feed* dentro do palheiro da Web e a sua respectiva inscrição em um aplicativo leitor ainda precisam ser simplificadas, tornando-se algo transparente para o usuário.

## 8.1 Sugestões para Estudos Futuros

A partir das considerações apresentadas, como resultado de uma investigação exploratória, propomos a seguinte hipótese, passível de ser testada em um estudo futuro: “produtos e serviços de informação que disponibilizam *feeds* RSS tenderão a ampliar a visibilidade e acesso dos conteúdos fornecidos, sendo mais divulgados e disponibilizados por meio de sistemas que realizam a recolha de metadados (*harvesting*), por exemplo”.

Uma vez que a presente pesquisa foi desenvolvida com tempo e escopo limitados, sugerimos que o levantamento realizado seja posteriormente aprofundado, podendo incluir outras áreas do conhecimento, de modo que a nossa contribuição seja apenas um ponto de partida para um estudo ainda mais fundamentado, em nível de doutoramento, com base em outros parâmetros e cruzamento de mais variáveis e indicadores a serem estabelecidos.

Percebe-se, também, um leque de oportunidades a serem aproveitadas em estudos futuros que podem ter como base as muitas possibilidades levantadas e discutidas neste trabalho, tais como:

- a) estudos mais aprofundados sobre a integração dos metadados com os serviços de disseminação de informação;
- b) desenvolvimento de modelos de agregadores de conteúdos temáticos, customizados segundo a área de conhecimento, a exemplo do *Bibliorandum*, voltado para a área de Ciência da Informação;
- c) a realização de um estudo de usuário ou grupo focal com a finalidade de verificar eventuais mudanças no comportamento informacional de utilizadores de *feeds* RSS em relação aos que não fazem uso da tecnologia;
- d) estudo de medição de impacto da efetividade na absorção de informação recebida/utilizada pelos usuários de recursos RSS;

- e) pesquisa sobre o desenvolvimento de novas aplicações e tipos de conteúdos passíveis de serem disseminados por meio da tecnologia de sindicância de conteúdo Web.

Acreditamos que os estudos decorrentes podem ser ainda muito mais amplos do que as sugestões ora apresentadas, pois, tal como a revolução silenciosa iniciada pelo correio eletrônico, estamos presenciando apenas o início de uma Era de serviços e aplicações da chamada Web 2.0 (há quem fale até em Web 3.0), com características bem marcantes, como usuários produzindo conteúdos, compartilhamento e interoperabilidade de dados e informações e serviços que combinam mais de uma fonte de informação como uma experiência integrada. Recursos tecnológicos como RSS fazem parte dessa nova abordagem, cada vez mais colaborativa e centrada no usuário.

## REFERÊNCIAS

---

- ALMEIDA, R. L. de. Da disseminação seletiva à web syndication: uma proposta para a comunicação científica. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB, 8. , 2007, Salvador. **Anais eletrônicos...** Salvador: ANCIB, 2007. Disponível em: <<http://www.enancib.ppgci.ufba.br/artigos/GT7--157.pdf>>. Acesso em: 03 nov. 2007.
- ANDERSON, B.. Keeping up: SDI to RSS. **Behavioral & Social Sciences Librarian**, v. 24, n. 2, p.113-117, 2006.
- ANJOS, J dos. **Bibliorandum** (Relatório final de Projeto em Implementação de Sistemas de Comunicações e Informação), 2007. Disponível em: <<http://www.bibliorandum.net/relatorio/PISCI-TPFinal-9040193.pdf>> Acesso em: 05 jan. 2008.
- BABBIE, E. **The Practice of Social Research**. 4th ed. Belmont: Wadsworth Publ., 1986.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. New York: ACM Press, 1999. 511p.
- BARROS, M. A. de. Blogs e bibliotecários. In: ENCONTRO NACIONAL DOS ESTUDANTES DE BIBLIOTECONOMIA, DOCUMENTAÇÃO, CIÊNCIA E GESTÃO DA INFORMAÇÃO, 27., 2004, Recife. **Anais...** Recife, 2006. Disponível em: <[http://eprints.rclis.org/archive/00009331/01/blogs\\_e\\_bibliotec%C3%A1rios.pdf](http://eprints.rclis.org/archive/00009331/01/blogs_e_bibliotec%C3%A1rios.pdf)>. Acesso em: 18. nov. 2007.
- BARRETO, A. de A. A condição da informação. **São Paulo em Perspectiva**, São Paulo, v.16, n.3, p. 67-74, jul. 2002.
- BARRETO, A. de A. Mudança estrutural no fluxo do conhecimento: a comunicação eletrônica. **Ciência da Informação**, Brasília, v.27, n.2, 1998, p. 122-127.
- BARBALHO, C. R. S. Periódico científico: parâmetros para avaliação de qualidade. In: FERREIRA, S. M. S. P; TARGINO, M. G. (Org.). **Preparação de Revista científica: Teoria e Prática**. São Paulo: Reichmann & Autores, 2005, v. 1, p. 123-160.
- BAX, M. P.; ALVARENGA, L.; PARREIRAS, F. S.; BRANDÃO, W. C. Sistema automático de Disseminação Seletiva. In: IFLA M&M, 2004, São Paulo, **Anais...**, São Paulo: USP, 2004. Disponível em: <[http://www.fernando.parreiras.nom.br/publicacoes/dsi\\_ifla.pdf](http://www.fernando.parreiras.nom.br/publicacoes/dsi_ifla.pdf)> Acesso em: 28 out. 2006.
- BELKIN, N. J.; CROFT, W. B. Information filtering and information retrieval: two sides of the same coin?. **Communications of the ACM**, v.35, n.12, p. 29-38, dez. 1992.
- BERNARDINO, T. S. M. P. **Perspectiva sobre a Utilização da Tecnologia RSS no Contexto da Comunicação Científica**. 2006. 180 f. Dissertação. (Mestrado). Departamento de Sistemas de Informação, Universidade do Minho, Guimarães Portugal. 2006.
- BETTIOL, E. M. Necessidades de informação: uma revisão. **Revista de Biblioteconomia de Brasília**, Brasília, v. 18, n. 1, p. 59-69, 1990.

- BHATT, J. Using RSS to increase user awareness of e-resources in academic libraries. **Higheredblogcon2006**. Disponível em: <<http://www.higheredblogcon.com/index.php/using-rss-to-increase-user-awareness-of-e-resources-in-academic-libraries>>. Acesso em: 02. jan. 2008.
- BLOGOSFERA CIENTÍFICA. **Agência Fapesp**, São Paulo, 06 ago. 2006. Disponível em: <[http://www.fapesp.br/agencia/boletim\\_dentro.php?data%5Bid\\_materia\\_boletim%5D=5735](http://www.fapesp.br/agencia/boletim_dentro.php?data%5Bid_materia_boletim%5D=5735)>. Acesso em: 09 dez. 2007.
- BRIQUET DE LEMOS, A. **Periódicos eletrônicos: problema ou solução?** Disponível em: <[http://www.briquetdelemos.com.br/briquet/briquet\\_lemos7.htm](http://www.briquetdelemos.com.br/briquet/briquet_lemos7.htm)>. Acesso em: 16 mar. 2007.
- CARDOSO, J. C.; OLIVEIRA, J. B. **Uma proposta de Interface para bibliotecas digitais configuráveis ao perfil do usuário**. Porto Alegre, 2000. [Plano de Estudo e Pesquisa].
- CARLSON, C. N. (2003) **Information overload, retrieval strategies and Internet user empowerment**. In: HADDON, Leslie, Eds. *Proceedings The Good, the Bad and the Irrelevant (COST 269) 1(1)*, pp. 169-173. Helsinki, Finland, 2003. Disponível em: <<http://eprints.rclis.org/archive/00002248/>>. Acesso em: 13 mai. 2007.
- CHAN, L. Electronic journals and academic libraries. **Library Hi Tech**, v. 17, n.1, p. 10-16, 1999.
- COSTA, S. M. S. **The impact of computer usage on scholarly communication amongst academic social scientists**. 1999. 291p. Tese (Doctoral Thesis). Department of Information Science, Loughborough University, Loughborough, 1999.
- COSTA, S. M. S. Mudanças no processo de comunicação científica: o impacto do uso de novas tecnologias. In: MUELLER, S. P. M.; PASSOS, E. J. L. (Org.). **Comunicação Científica**. v. 1. Brasília: Departamento de Ciência da Informação e Documentação, Universidade de Brasília, 2000. p. 85-106.
- COSTA, S. M. S.; SILVA, W. A. A.; COSTA, M. B. Publicações científicas eletrônicas no Brasil: mudanças na comunicação formal, também?. **Revista de Biblioteconomia de Brasília**, Brasília, v.25, n.1, p. 57-76, jan./jun. 2001.
- CRUZ et al. O impacto dos periódicos eletrônicos em bibliotecas universitárias. **Ciência da Informação**, Brasília, v. 32, n.2, p.47-53, maio/ago., 2003.
- CUNHA, L. Publicações científicas por meio eletrônico: critérios, cuidado, vantagens e desvantagens. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 2, n. 1, p. 77-92, jan./jun. 1997.
- ÇELIKBAS, Z. **What is RSS and how can it serve libraries?**. Istanbul Technical University. Faculty of Aeronautics and Astronautics, Istanbul, Turkey. Disponível em: <[http://eprints.rclis.org/archive/00002531/01/RSS\\_and\\_libraries\\_EN3.pdf](http://eprints.rclis.org/archive/00002531/01/RSS_and_libraries_EN3.pdf)>. Acesso em: 07 nov. 2006.
- DAVISON-TURLEY, W. Blogs and RSS: Powerful Information Management Tools. **Library Hi Tech News**, v. 22, n. 10, p.28-29, 2005.

DIAS, C. **Comunicação Científica**. Disponível em:

<<http://www.geocities.com/clauidiad/comunica.pdf>>. Acesso em: 03. mar. 2007.

DIAS, G. et al. Periódico eletrônico Informação & Sociedade: Estudos – impactos no contexto da comunicação científica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 3. 2006.

DIAS, G. **Periódicos Científicos eletrônicos brasileiros na área da Ciência da Informação: análise das dinâmicas de acesso e uso**. 2003. 190f. Tese (Doutorado). Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2003.

DUKE, M. Delivering OAI records as RSS: an IMesh Toolkit module for facilitating resource sharing. **Ariadne**, Issue 37, abr. 2003. Disponível em: <<http://www.ariadne.ac.uk/issue37/duke>>. Acesso em: 13 out.2007.

FACHIN, G. R. B. **Modelo de avaliação para periódicos científicos on-line: proposta de indicadores bibliográficos e telemáticos**. 2002. 210 f. Dissertação (Mestrado em Engenharia de Produção)- Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002.

FLIZIKOWSKI, M.R. Distribuição personalizada de informações pela Internet. In: do 26. CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 26. , 2003, Belo Horizonte. **Anais...** São Paulo: Intercom, 2003 (CD-ROM).

GARVEY, W. D. **Communication, the essence of science**. Oxford: Pergamon International Library, 1979.

GIL, A. C.; **Métodos e técnicas de pesquisa social**. 5. ed. São Paulo: Atlas, 1999.

GOMES, S. H. A. **Inovação tecnológica no sistema formal de comunicação científica: os periódicos eletrônicos nas atividades de pesquisa dos acadêmicos de cursos de pós-graduação brasileiros**. 1999. 465f. Tese (Doutorado). Faculdade de Estudos Sociais Aplicados, Universidade de Brasília, Brasília, 1999.

GÜNTER, H. Pesquisa qualitativa versus pesquisa quantitativa: esta é a questão? **Psicologia: Teoria e Pesquisa**, Brasília, v. 22 n. 2, p. 201-210, mai./ago. 2006. Disponível em: <<http://www.scielo.br/pdf/ptp/v22n2/a10v22n2.pdf>>. Acesso em: 20 nov. 2007.

GUTIERREZ, S. Distribuição de Conteúdos e Aprendizagem On-line, **RENOTE – Revista Novas Tecnologias na Educação**. Porto Alegre, CINTED, v.2, n.2, nov. 2004. Disponível em: <[http://www.cinted.ufrgs.br/renote/nov2004/artigos/a6\\_distribuicao\\_conteudos.pdf](http://www.cinted.ufrgs.br/renote/nov2004/artigos/a6_distribuicao_conteudos.pdf)>. Acesso em: 16 jul. 2007.

HAMMOND, T. **Why choose RSS 1.0?**. XML.com. Disponível em:

<<http://www.xml.com/pub/a/2003/07/23/rssone.html>>. Acesso em: 18. fev. 2007.

HAMMOND, T. et al. The Role of RSS in Science Publishing Syndication and Annotation on the Web. **D-Lib Magazine**, 10 (12). 2004. Disponível em:

<<http://www.dlib.org/dlib/december04/hammond/12hammond.html>>. Acesso em: 13. fev. 2007.

HARNAD, S. The invisible hand of peer review. **Exploit Interactive**, v.5, apr. 2000.

Disponível em: <<http://www.exploit-lib.org/issue5/peer-review>>. Acesso em: 12 ago. 2007.

HARRISON, T. M.; STEPHEN, T. D. The electronic journal as the heart of an on-line scholarly community. **Library Trends**, v.48, n.4, p.592-608, 1995.

HEDLUND, M. The New Bloglines Web Service. **O'Reilly Network**, 2004, Disponível em: <<http://www.oreillynet.com/pub/a/network/2004/09/28/bloglines.html>>. Acessado em: 12. nov. 2007.

HENRY, J. Annual report of the board of regents of the Smithsonian Institution...during the year 1851. Washington: 1852, apud KRZYZANOWSKI, Rosaly Fávero; TARUHN, Rosane. Biblioteca eletrônica de revistas científicas internacionais: projeto de consórcio. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 193-197, maio/ago. 1998.

HOLVOET, K. What Is RSS and How Can Libraries Use It to Improve Patron Service?. **Library Hi Tech News**, n.8, 2006, pag.32-33.

HOUSMAN, E. M. Selective dissemination of information. In: **Annual Review of Information Science and Technology**. Washington, EUA: American Society of Information Science, 1973. p. 221-241.

HURD, J. M. Models of scientific communication systems. In: CRAWFORD, Suzany; HURD, J. M; WILIER, A. C. **From print to electronic: the transformation of scientific communication**. Medford, NJ: Information Today, 1996. p. 9-33.

HURD, J. M. The transformation of scientific communication: a model for 2020. **Journal of the American Society for Information Science**, v.51, n. 14, p. 1279-1283, 2000.

KÖCKE, J. C. Fundamentos de Metodologia Científica. 7. ed. Caxias do Sul: Vozes, 1982.

KYRNIN, J. **What is RSS and how do you use it**. Disponível em: <<http://webdesign.about.com/cs/rss/a/aa052603a.htm>>. Acesso em: 14 mar. 2007

KLING, R.; McKIM, G. Scholarly communication and the continuum of electronic publishing. **Journal of the American Society of Information Science**, v. 50, n.10, p. 890-896, 1999.

KRZYZANOWSKI, R. F.; TARUHN, R. Biblioteca eletrônica de revistas científicas internacionais: projeto de consórcio. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 193-197, maio/ago. 1998.

LANCASTER, F. W. The evolution of electronic publishing. **Library Trends**. Urbana, v. 43, n. 4, p. 518-527, 1995.

LANCASTER, F. W; FAYEN, E.G. **Information Retrieval On-Line**. Los Angeles: Melville Publishing Company, 1973.

LANDOW, G. **Hipertexto: la convergência de la teoría crítica contemporánea y la tecnología**. Buenos Aires: Paidós, 1995.

LANGSCHIELD, L. Electronic Journal Forum: VPIEJ-L: An On-line Discussion Group for Electronic Journal Publishing Concerns. **Serials Review**, v. 20, n. 1, p. 89-94, 1994.

LEAL, D. **Notícias empresariais e listas de leitura OPML: BlogBridge RSS Feed Library**. Disponível em:

<[http://www.masternewmedia.org/pt/entrega\\_e\\_distribuicao\\_de\\_conteudos/OPML-rss-feeds/noticias-empresariais-e-listas-de-leitura-OPML-BlogBridge-RSS-Feed-Library-Analise-20071003.htm](http://www.masternewmedia.org/pt/entrega_e_distribuicao_de_conteudos/OPML-rss-feeds/noticias-empresariais-e-listas-de-leitura-OPML-BlogBridge-RSS-Feed-Library-Analise-20071003.htm)>. Acesso em: 09.dez. 2007

LE COADIC, Y. **A Ciência da Informação**. Brasília: Brique de Lemos, 2ª ed., 2004.

LÉVY, P. **Cibercultura**. São Paulo: Editora 34, 2000.

LEWIN, J. **An introduction to RSS news feeds. Using open formats for content syndication**, nov., 2000. Disponível em: <<http://www-28.ibm.com/developerworks/library/w-rss.html>>. Acesso em: 03 nov. 2006.

LONGO, R. M. J. Disseminação seletiva da informação (SDI): “estado da arte” e tendências futuras. **Revista de Biblioteconomia de Brasília**, 6 (2), p.101-120, jul./dez. 1978.

LONGO, R. M. J. **Sistema de Recuperação da Informação: Disseminação Seletiva da Informação e Bases de Dados**. Brasília: Thesaurus, 1979, 276p.

LUHN, H. P. (1961) Selective dissemination of new scientific information with the aid of electronic processing equipment. **American Documentation**, v. 12, p.131-138.

LYMAN, P.; VARIAN, H. R. **How Much Information**, 2003. Disponível em: <<http://www.sims.berkeley.edu/how-much-info-2003>>. Acesso em: 08 fev. 2007.

MARCONDES, C. H. Metadados: descrição e recuperação de informações na Web. In: MARCONDES, C. H. et al. (Orgs). **Bibliotecas Digitais: saberes e práticas**. Salvador: EDUFBA; Brasília: IBICT, 2006, p. 95-111.

MARCONDES, C. H.; GOMES, S. L. R. O impacto da internet nas bibliotecas brasileiras. **Transinformação**, Campinas, v. 9, n. 2, p. 57-68, 1997.

MÁRDERO ARELLANO, M. A. ; FERREIRA, S. M. S. P.; CAREGNATO, S. E. Editoração eletrônica de revistas científicas com suporte ao protocolo OAI. In: FERREIRA, S. M. S. P; TARGINO, M. G. (Org.). **Preparação de Revista Científica: Teoria e Prática**. São Paulo: Reichmann & Autores, 2005, v.1, p. 123-160.

McKNIGHT, C. Electronic journals - past, present... and future?. **Aslib Proceedings**, v. 45, n.1, p. 7-10, 1993.

McMURDO, G. Electric writing: changing contexts of communication. *Journal of Information Science*, v.21, n. 2, p. 140-146, 1995 apud OLIVEIRA, E. B. P. M. ; NORONHA, D. P. A Comunicação científica e o meio digital. **Informação & sociedade: estudos**, João Pessoa, v. 15, n. 1, p. 1-12, 2005.

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

MILLER, Paul. Syndicated Content: it's more than just some file formats. **Ariadne**, Issue 35, abr.2003. Disponível em: <<http://www.ariadne.ac.uk/issue35/miller>>. Acesso em: 08 nov.2007.

MILSTEAD, J.; FELDMAN, S.. Metadata: project and standards. **On-line**, v. 23, p. 32-40, jan./feb. 1999.

MIRANDA, A. L. C.; SIMEÃO, E. L. M. S. Estrutura da informação e modelo extensivo: uma abordagem para a Ciência da Informação. In: \_\_\_\_\_ . **Informação e tecnologia: conceitos e recortes**. 1. ed. Brasília: Departamento de Ciência da Informação e Documentação, 2005. p. 177-199.

MIRANDA, A. L. C.; SIMEÃO, E. L. M. S. Arquitetura e implementação de conteúdos através da tecnologia EVM.net no portal do Departamento de Ciência da Informação e Documentação da Universidade de Brasília (CID/UnB). In: Colóquio Internacional de Bibliotecários, 13, Guadalajara. **Anais...** Guadalajara, 2006. Disponível em: <[http://www.antoniomiranda.com.br/ciencia\\_informacao/Arquitetura\\_implementacao.pdf](http://www.antoniomiranda.com.br/ciencia_informacao/Arquitetura_implementacao.pdf)>. Acesso em: 13. jan. 2008.

MIRANDA, D. B.; PEREIRA, M. N. F. O periódico científico como veículo de comunicação: uma revisão de literatura. **Ciência da Informação**, Brasília, v.25, n.3, 1996.

MOOERS, C. N. Zatoncoding applied to mechanical organization of knowledge. **American Documentation**, Washington, v.2, p.20-32, 1951.

MOURA SPERONI, R. et al. Uma biblioteca digital aberta com serviços personalizados. In: WORKSHOP DE BIBLIOTECAS DIGITAIS, 2. , 2006. **Anais...** Florianópolis: Renata de Matos Galante, Ronaldo dos Santos Mello, 2006. p. 11-20.

MUELLER, S. P. M.; PASSOS, E.J.L. (org.). **Comunicação Científica**. Brasília: Departamento de Ciências da Informação e Documentação da UnB, 2000, 144p., v. 1.

MUELLER, S. P. M. O periódico científico e as bibliotecas universitárias: velhos problemas, novas soluções. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 8. ,Campinas. **Anais...** Campinas: Universidade de Campinas, 1994. p. 80-102.

MUELLER, S. P. M. O impacto das tecnologias de informação na geração do artigo científico: tópicos de estudo. **Ciência da Informação**, Brasília, v.23, n.3, p.309-317, set./dez.1994.

NEHMY, R. M; PAIM, I. A desconstrução do conceito de "qualidade da informação". **Ciência da Informação**, Brasília, v.27, n.1, 1998. Disponível em: <<http://www.ibict.br/cienciadainformacao/viewarticle.php?id=385>>. Acesso em: 11 mar. 2007.

NOCETTI, M. **A disseminação seletiva da informação: teoria e prática**. Brasília: ABDF, 1980.

NOTTINGAM, M. **RSS tutorial for content publishers and webmasters**. Disponível em: <<http://www.mnot.net/rss/tutorial>>. Acesso em: 03 nov. 2006.

OLIVEIRA, E.B.P.M. **Uso de periódicos científicos eletrônicos por docentes e pós-graduandos do Instituto de Geociências da USP**. 2006. 139 f. Dissertação (Mestrado) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.

OLIVEIRA, E. B. P. M. ; NORONHA, D. P. A comunicação científica e o meio digital. **Informação & sociedade: estudos**, João Pessoa, v. 15, n. 1, p. 1-12, 2005.

ORNELAS, M L.; ARROYO, G. C. Las revistas académicas electrónicas en Internet. **Revista Mexicana de Comunicación**. Disponível em:  
<<http://www.mexicanadecomunicacion.com.mx/Tables/RMC/rmc83/revistas.html>>. Acesso em 18 mar.2007.

PACKER et al. SciELO: uma metodologia para publicação eletrônica. **Ciência da Informação**, Brasília, v.27, n.2, p.109-121, mai./ago. 1998.

PILGRIM, M. **What is RSS?**. XML.com, 2002. Disponível em:  
<<http://xml.com/pub/a/2002/12/18/dive-into-xml.html>>. Acesso em: 02 set. 2006.  
REHEN, Stevens. **Blog pode ser futuro da publicação científica**. G1, Disponível em:  
<<http://g1.globo.com/Noticias/Ciencia/0,,MUL23599-5603,00.html>>. Acesso em: 02 dez. 2007.

PINHEIRO, L. V. R. Evolução da comunicação científica até as redes eletrônicas e o periódico como instrumento central deste processo In:  
Conferência Iberoamericana de Publicações Eletrônicas no Contexto da Comunicação Científica, 1, 2006, Brasília. **Anais...** Brasília, 2006. Disponível em:  
<<http://portal.cid.unb.br/CIPECCbr/viewpaper.php?id=48>> . Acesso em: 13 set. 2007.

PINHEIRO, L. V. R.; GOMES, S. L. R. Redes eletrônicas e seus impactos na comunicação científica de pesquisadores brasileiros. In: Encontro Nacional de Pesquisas em Ciência da Informação, 5, 2003, Belo Horizonte. **Anais...** Belo Horizonte, 2003.

**RSS Specifications**. Disponível em: <<http://www.rss-specifications.com/history-rss.htm>>. Acesso em: 05 jan. 2007.

RUSSO, M. ; SANTOS, E. T. G. ; SANTOS, M. J. V. C. Produção Científica Brasileira: da comunicação à recuperação via Web. In: Fórum Nacional de Padronização e Divulgação da Produção Científica, 1., 2001, Fortaleza. **Anais...** Fortaleza, 2001.

SÁ, A. O código oculto das notícias electrónicas. In: Congresso da Associação Portuguesa de Ciências da Comunicação, 4, 2005, Aveiro: Universidade do Aveiro. **Anais...** Aveiro, 2005. Disponível em:  
<[https://repositorium.sdum.uminho.pt/bitstream/1822/45111/1/albertosa2005\\_codigoculnotic.pdf](https://repositorium.sdum.uminho.pt/bitstream/1822/45111/1/albertosa2005_codigoculnotic.pdf)>. Acesso em: 28. set. 2007.

SABBATINI, M. **Publicações eletrônicas na Internet**: Yendis, São Caetano do Sul, 2005, 319p.

SARACEVIC, T. The concept of "relevance" in Information Science: a historical review. In: SARACEVIC, Tefko, Ed. **Introduction to Information Science**. New York: R. R. Bowker Co., 1970. p. 111-154.

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

SAUERS, M. P. **Blogging and RSS: a librarian's guide**: Information Today Inc., New Jersey, 2006, 272 p.

- SHETH, B. D. **A learning approach to personalized information filtering**. 1994. 76f. Dissertação (Mestrado). Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Boston, 1994. Disponível em: <<http://agents.media.mit.edu/publications/sheth-thesis.pdf>>. Acesso em: 01 jun. 2007.
- SILVA, E.L.; MENEZES, E.M.; BISSANI, M. A Internet como canal de comunicação científica. **Informação & Sociedade: Estudos**, João Pessoa, v. 10, n.2, 2000. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/163>> Acesso em: 04 jul.2007
- SILVA FILHO, A. M.; DELGADO, M. V. M. A sobrecarga da informação na era da Internet, **Revista Espaço Acadêmico**, n. 23, abr. 2003. Disponível em: <<http://www.espacoacademico.com.br/023/23amsf.htm>>. Acesso em: 13 jun. 2007.
- SILVA, G. B. **Uma proposta de criação perfis de usuários da Internet para filtragem de informação personalizada**. 2002. 75f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação). Universidade Federal de Lavras, Lavras. Disponível em: <[http://www.comp.ufla.br/monografias/ano2002/Uma\\_proposta\\_de\\_criacao\\_de\\_perfis\\_de\\_usuarios\\_da\\_internet\\_para\\_filtragem\\_de\\_informacao\\_personalizada.pdf](http://www.comp.ufla.br/monografias/ano2002/Uma_proposta_de_criacao_de_perfis_de_usuarios_da_internet_para_filtragem_de_informacao_personalizada.pdf)>. Acesso em: 27 mai. 2006.
- SIMEÃO, E. L. M. S. **Comunicação Extensiva e Informação em Rede**. 1 ed. Brasília: CID/UnB, 2006. v. 1. 277p.
- SIMEÃO, E. L. M. S.; MIRANDA, A. L. C. Comunicação Extensiva e o formato do periódico científico eletrônico. In: International Conference on Electronic Publishing, 8, Brasília. **Anais...** Brasília, 2004. Disponível em: <<http://eprints.rclis.org/archive/00002445/01/elpubport.pdf> >. Acesso em: 13. jan. 2008.
- SOUZA, T. E. R; ALBUQUERQUE, M. E. B. C. Periódicos científicos em biblioteconomia e Ciência da Informação: consulta por alunos concluintes do curso de Biblioteconomia da UFPB. **Bibliion-line**, v.1, n.2, 2005. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/biblio/article/viewFile/587/425>> Acesso em: 03 ago. 2007
- STUMPF, I. R. C. Passado e futuro das revistas científicas. **Ciência da Informação**, Brasília, v. 25, n. 3, 1996. Disponível em: <<http://www.ibict.br/cion-line/viewarticle.php?id=504&layout=abstract>>. Acesso em: 03 mai. 2007.
- TARGINO, M. G. Comunicação científica: uma revisão de seus elementos básicos. **Informação & Sociedade: Estudos**, João Pessoa, v. 10, n. 2, 2000. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/326>>. Acesso em: 04 jul. 2007.
- THOMAZ, K. P.; SANTOS, V. M. Metadados para o gerenciamento eletrônico de documentos de caráter arquivístico - GED/A: estudo comparativo de modelos e formulação de uma proposta preliminar. **DataGramZero - Revista de Ciência da Informação**, v. 4, n. 4, 2003. Disponível em: <[http://www.dgz.org.br/ago03/Art\\_04.htm](http://www.dgz.org.br/ago03/Art_04.htm)>. Acesso em: 13 out. 2007.
- WEILBEL, S. Metadata: the foundations of resource description. **D-Lib Magazine**, July, 1995. Disponível em: < <http://www.dlib.org/dlib/July95/07weibel.html>>. Acesso em: 15 out. 2007.

WEITZEL, S. R. E-prints: modelo da comunicação científica em transição. In: FERREIRA, S. M. S. P; TARGINO, M. G. (Org.). **Preparação de Revista Científica: Teoria e Prática**. São Paulo: Reichmann & Autores, 2005, v.1, p. 161-193.

WIKIPÉDIA. Mashup. Disponível em: <<http://pt.wikipedia.org/wiki/Mashup>>. Acesso em: 13. dez 2007.

WILSON, T. D. Information Needs and Uses: fifty years of progress? In: VICKERY, B. C. (Ed.) **Fifty Years of Information Progress: a Journal of Documentation review**. London: Aslib, 1994. p. 15-51.

WILSON, T. D. Models in Information Behavior Research. **Journal of Documentation**, London, v. 55, n. 3, p. 249-271, jun. 1999.

WILSON, T.D. **Human information behavior**. Informing Science, v.3, n.2, p. 49-55, 2000. Disponível em: <<http://inform.nu/Articles/Vol3/v3n2p49-56.pdf>>. Acesso em: 14 jan. 2007.

WINER, D. **RSS 2.0 Specification**, 2003, Disponível em: <<http://blogs.law.harvard.edu/tech/rss>>. Acesso em: 03 out. 2007.

WITTENBRINK, H. **RSS and ATOM Understanding and Implementing Content Feeds and Syndication**. PACKT Publishing, 2005. Disponível em: <[http://www.packtpub.com/files/RSS\\_and\\_Atom\\_Book\\_Chapter1\\_what\\_are\\_newsfeeds.pdf](http://www.packtpub.com/files/RSS_and_Atom_Book_Chapter1_what_are_newsfeeds.pdf)>. Acesso em: 06 mai. 2007.

WUSTEMAN, J. RSS: the latest feed. **Library Hi Tech News**, v. 22. n. 4, p. 404-413, 2004.

ZIMAN, J. **Conhecimento Público**. São Paulo: Editora Itatiaia, 1979.

---

**APÊNDICE A – Modelo do questionário on-line disponibilizado aos pesquisadores da ANCIB**

**UNIVERSIDADE DE BRASÍLIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

**Comportamento Informacional dos usuários de periódicos eletrônicos em CI:  
Análise das funcionalidades e ferramentas**

Prezado (a) colega,

O presente questionário tem como finalidade realizar um levantamento sobre o uso dos periódicos eletrônicos, especialmente os relacionados à área da Ciência da Informação editados sob a plataforma SEER/OJS.

Lembramos que essa pesquisa está sendo conduzida com finalidade exclusivamente acadêmica no âmbito do mestrado em Ciência da Informação da Universidade de Brasília (UnB). Em caso de dúvidas ou esclarecimentos adicionais, por favor entre em contato com o pesquisador.

Sua participação é fundamental para que esta pesquisa seja bem-sucedida em seus objetivos. Assim sendo, solicitamos a sua colaboração no sentido de responder esse questionário eletrônico com a maior brevidade possível.

Sinceros agradecimentos,

Robson Lopes de Almeida  
Mestrando em Ciência da Informação  
Departamento de Ciência da Informação e Documentação

O questionário abaixo foi disponibilizado eletronicamente a partir do endereço <<http://rss.periodicos.sgizmo.com>> por meio da aplicação Survey Gizmo.

## QUESTIONÁRIO

### 1) Dados de identificação

a) Titulação mais elevada:

Graduação

Mestrado

Doutorado

Área de formação:

b) Sua faixa etária:

20-25

26-30

31-35

36-40

41-45

46-50

51-55

56-60

> 60

c) Cidade onde mora:

Estado:

d) Áreas de atuação em Ciência da Informação:

e) Perfil de atuação em relação às revistas científicas [pode responder mais de uma opção]:

Autor

Leitor

Avaliador

Editor

f) GT vinculado à ANCIB

GT 1 – Estudos Históricos e Epistemológicos da Informação

GT 2 – Organização e Representação do Conhecimento

GT 3 – Mediação, Circulação e Uso da Informação

GT 4 – Gestão da Informação e do Conhecimento nas Organizações

GT 5 – Política e Economia da Informação

GT 6 – Informação, Educação e Trabalho

GT 7 – Produção e Comunicação da Informação em CT&I

**2) Com que frequência utiliza os periódicos abaixo em busca de atualização?**

	Conhece		Usa			
	e Não usa	toda	quinzenalmente	mensalmente	semestralmente	
		0	1 semana	2	3	4
a) Arquivística.net	<input type="checkbox"/>					
b) Ciência da Informação	<input type="checkbox"/>					
c) Em Questão	<input type="checkbox"/>					
d) Informação e Sociedade	<input type="checkbox"/>					
e) Informação e Informação	<input type="checkbox"/>					
f) Perspectivas em Ciência da Informação	<input type="checkbox"/>					
g) Revista ACB	<input type="checkbox"/>					
h) Revista Brasileira de Biblioteconomia e Documentação	<input type="checkbox"/>					
i) Revista Digital de Biblioteconomia e Ciência da Informação	<input type="checkbox"/>					
j) Transinformação	<input type="checkbox"/>					

**3) Na sua opinião, o recebimento de uma notificação automática sobre a atualização de conteúdo em um periódico científico é:**

- Extremamente importante, pois procuro me manter atualizado (a) o quanto antes
- Importante, porém não tenho pressa em consultar os novos artigos
- Irrelevante, pois já tenho o hábito de acessar os periódicos regularmente
- Desagradável, prefiro pesquisar por conta própria

**4) Faz uso de algum tipo de serviço de alerta para se manter informado sobre o lançamento de uma nova edição?**

- Sim
- Não (pular para a questão 7)

**5) Como costuma receber a notificação do serviço de alerta?**

- e-mail contendo o sumário da revista
- e-mail com o link para a publicação
- e-mail avisando sobre a atualização de uma nova edição da revista
- Outros. Qual?

**6) Como classifica o serviço de alerta utilizado:**

- Totalmente satisfatório (atende plenamente minhas expectativas)
- Satisfatório (atende parcialmente minhas expectativas)
- Insatisfatório (não atende minhas expectativas)

**7) Você utiliza as ferramentas de busca disponibilizadas nos periódicos para recuperar determinada informação?**

- Sim, todas as vezes em que acesso uma revista
- Sim, eventualmente
- Não. Por quê?

**8) Em relação ao seu hábito de leitura de artigos de periódicos eletrônicos, escolha a alternativa que melhor se enquadre ao seu comportamento:**

- Costumo ler apenas os sumários das revistas
- Além dos sumários, tenho hábito de ler os resumos dos artigos
- Na maioria das vezes, apenas o resumo me mantém atualizado. Mas em alguns casos, procuro ler os textos integrais dos artigos
- Sempre leio os textos integrais de todos os artigos dos periódicos

**9) Já ouviu falar na tecnologia RSS para distribuição automática de conteúdo?**

- Sim
- Não (pular para a questão 13)

**10) Costuma fazer assinatura de canais RSS (*feeds*)? [pode responder mais de uma opção]:**

**Sim**

- jornais/revistas
- blogs
- serviços e/ou unidades de informação

**Não**

**11) Assinaria *feeds* RSS de periódicos científicos se fossem disponibilizados?**

- Sim
- Não. Por quê?

**12) Você identifica alguma barreira com relação à utilização de *feeds* RSS em periódicos científicos?**

- Sim. Qual (is)?
- Não

**13) Gostaria de realizar buscas por termos livres em mais de uma revista simultaneamente?**

- Sim
- Não. Por quê?

**14) Por favor, descreva as mudanças ocorridas em seus procedimentos de acesso e uso da informação, a partir da utilização de periódicos científicos eletrônicos.**

## APÊNDICE B – Formulário aplicado aos coordenadores de GTs da ANCIB

### UNIVERSIDADE DE BRASÍLIA PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

**Análise de funcionalidades e ferramentas de disseminação de periódicos científicos eletrônicos em Ciência da Informação**

**Nome:**

**GT:**

**PARTE I – Levantamento de atributos funcionais de artigos de periódicos eletrônicos para representação da informação**

**1. Como o Sr.(a) classificaria, em grau de importância, os atributos dos artigos de periódicos científicos eletrônicos citados abaixo:**

Atributos	Muito Importante	Importante	Pouco Importante	Irrelevante
	4	3	2	1
identificador				
título				
nome do(s) autor(es)				
editor				
resumo				
palavras-chave				
direitos autorais				
idioma				
data de criação do artigo				
data de publicação				
formato (html, pdf, xml)				
dimensão (tamanho do artigo)				
issn				
volume				
número				
página inicial				
página final				

**2. O Sr.(a) destacaria outro elemento relevante para representação em um artigo científico na área de CI? Qual?**

## **PARTE II – Acesso e uso**

**3. Como o Sr. (a) se mantém atualizado em relação à publicação de uma nova edição de um periódico da área de CI? (faz uso de serviço de alerta?)**

**4. Quais as formas de acesso o Sr. (a) costuma utilizar?**

- diretamente pelo sítio da revista
- diretório das revistas do SEER/OJS
- SciELO
- Outros? Qual(is)?

**5. Faz uso de ferramentas de busca disponibilizadas pelos periódicos/portais para recuperar informação de seu interesse?**

- Não
- Sim,  com frequência
- eventualmente

## **PARTE III – Tecnologia RSS**

**6. Qual(is) característica(s) a seguir o Sr.(a) gostaria de ver implementada nos periódicos eletrônicos que costuma acessar? (pode responder a mais de uma)**

- acesso direto a um documento (artigo científico) sem ter que, necessariamente, passar pelo resumo das revistas.
- Possibilidade de realizar buscas simultâneas em mais de uma fonte de informação.
- Possibilidade de distinguir os documentos já lidos e/ou marcação de documentos considerados importantes.
- Todas as anteriores
- Nenhuma das anteriores

**7. Conhece ou já ouviu falar na tecnologia RSS (*Really Simple Syndication*) para agregação e distribuição de conteúdo na Web?**

- Não (fazer uma breve explanação e saltar para a pergunta 9)  
 Sim

**8. Tem o hábito de assinar canais RSS?**

- Não  
 Sim,  jornais/revistas  
 blogs  
 serviços e/ou unidades de informação

**9. Assinaria canais RSS de periódicos científicos se esses fossem disponibilizados?**

- Não Por quê?  
 Sim

**10. O Sr.(a) identifica alguma barreira em relação à utilização de canais RSS em periódicos científicos?**

- Não  
 Sim Qual?

**11. Quais as aplicações o Sr.(a) indicaria para utilização de canais RSS em periódicos científicos? Atribua um valor de acordo com o nível de importância, segundo a escala.**

Escala de pontuação por grau de importância	5 – muito importante	2 – pouco importante
	4 – importante	1 – nenhuma importância
	3 – importância relativa	0 – Não se aplica

Aplicações RSS	5	4	3	2	1	0
----------------	---	---	---	---	---	---

sumário (Tabela de conteúdo)						
divulgação de pré-prints						
notícias						
eventos						
publicações mais recentes						
lista de empregos						
artigos mais visitados						
lançamento de novos produtos						