

RAFAELA BOEIRA GIORDANO

Da necessidade ao conhecimento:
recuperação da informação na web em Ciência da Informação

Dissertação de mestrado
Março de 2011





UNIVERSIDADE FEDERAL DO RIO DE JANEIRO – UFRJ



INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E
TECNOLOGIA–IBICT



PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO –
PPGI

RAFAELA BOEIRA GIORDANO

**DA NECESSIDADE AO CONHECIMENTO:
recuperação da informação na web em Ciência da Informação**

RIO DE JANEIRO

2011

Rafaela Boeira Giordano

**DA NECESSIDADE AO CONHECIMENTO:
recuperação da informação na web em Ciência da Informação**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação, Convênio Instituto Brasileiro de Informação em Ciência e Tecnologia e Universidade Federal do Rio de Janeiro/Faculdade de Administração e Ciências Contábeis, como requisito parcial para a obtenção do título de Mestre em Ciência da Informação.

Orientador: Prof. Dr. Jorge Calmon de Almeida Biolchini

Rio de Janeiro

2011

G494 Giordano, Rafaela Boeira.

Da necessidade ao conhecimento: recuperação da informação na web em Ciência da Informação / Rafaela Boeira Giordano. – Rio de Janeiro: UFRJ / IBICT, 2011.

145 f. : il., 30cm.

Dissertação (Mestrado em Ciência da Informação) – Universidade Federal do Rio de Janeiro; Instituto Brasileiro de Informação em Ciência e Tecnologia, 2011. Orientador: Jorge Calmon de Almeida Biolchini.

1. Recuperação da informação. 2. Comportamento informacional. 3. Informação científica na web. 4. Letramento Informacional. I. Biolchini, Jorge Calmon de Almeida. II. Título

CDD

CDU

Rafaela Boeira Giordano

**DA NECESSIDADE AO CONHECIMENTO:
recuperação da informação na web em Ciência da Informação**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação, Convênio Instituto Brasileiro de Informação em Ciência e Tecnologia e Universidade Federal do Rio de Janeiro/Faculdade de Administração e Ciências Contábeis, como requisito parcial para a obtenção do título de Mestre em Ciência da Informação.

Aprovada por:

Prof. JORGE CALMON DE ALMEIDA BIOLCHINI
(Doutor em Ciência da Informação IBICT/UFRJ – IBICT)

Prof^a. LENA VANIA RIBEIRO PINHEIRO
(Doutora em Comunicação e Cultura ECO/UFRJ – IBICT)

Prof^a. MARIA LUIZA DE ALMEIDA CAMPOS
(Doutora em Ciência da Informação IBICT/UFRJ – UFF)

Suplente externo:

Prof. MARCOS LUIZ CAVALCANTI DE MIRANDA
(Doutor em Ciência da Informação IBICT/UFRJ – UNIRIO)

Suplente interno:

Prof^a. ROSALI FERNANDEZ DE SOUZA
(Ph. D. in Information Science Polytechnic of North London, England – IBICT)

À acolhedora cidade do Rio de Janeiro

Agradecimentos

Ao professor Jorge Biolchini, pela orientação paciente e delicada.

Aos professores do IBICT e da UFRJ, responsáveis pelo despertar do interesse pela Ciência da Informação.

Aos funcionários do IBICT e da UFRJ, pelo atendimento sempre atencioso.

Aos 52 colegas do PPGCI/IBICT-UFRJ e aos sete cientistas da informação que gentilmente contribuíram para a realização deste estudo.

Em especial, aos amigos Ana Rosa Pais Ribeiro, Angela Monteiro Bettencourt, Antonio Victor Botão, Neusa Cardim e Tatiana Almeida, companheiros de turma, de divagações e de encontros animados.

Aos colegas do DAAD, *vielen Dank für die Unterstützung.*

Ao professor Eduardo Ariel, da ESPM, pela sugestão do mestrado.

Aos queridos Ana Paula Lückman e Maurício Oliveira, pelos modelos de projetos.

À Luci Giordano, pelo que veio antes desta etapa.

Aos amigos Márcia Feijó e Helton Bertol, pelo incentivo inicial.

A Claudio Werneck, pelo incentivo constante.

*“As coisas mais importantes
nunca devem ficar à mercê
das menos importantes.”*

Johann Wolfgang von Goethe (1749-1832)

GIORDANO, R. B. Da necessidade ao conhecimento: recuperação da informação na web em Ciência da Informação. Rio de Janeiro. Universidade Federal do Rio de Janeiro, Faculdade de Administração e Ciências Contábeis – UFRJ-FACC; Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT, 2011. Dissertação (Mestrado em Ciência da Informação). Orientador: Jorge Calmon de Almeida Biolchini.

Resumo

Na origem constitutiva da Ciência da Informação, o processo de busca e recuperação da informação tem representado novos desafios para seu desenvolvimento, face às mudanças determinadas pelo advento e crescimento da web. Diante do volume exponencialmente crescente de informações nesta depositadas, do número e diversidade de usuários e do aumento crescente da utilização de mecanismos de busca, esta pesquisa contribui para ampliar o entendimento do comportamento informacional de usuários especializados neste cenário. O trabalho visa à identificação e à análise do comportamento adotado por pós-graduandos em Ciência da Informação ao realizarem busca e recuperação de informação científica na web. Aprofundou-se especialmente o estudo sobre assuntos relacionados a recuperação da informação para produção científica, recuperação da informação na web, letramento informacional e digital, bem como comportamento informacional dos usuários. Na pesquisa de campo, foi realizada a aplicação de um conjunto integrado de instrumentos, composto por um questionário e pela realização de uma entrevista, desenvolvidos sob a perspectiva da técnica do incidente crítico, metodologia pertinente ao campo de investigação de Análise da Tarefa Cognitiva (CTA). A partir dos dados coletados, mapeou-se o percurso adotado pela população-alvo, assim como foram identificadas as tomadas de decisão, as estratégias adotadas, as dificuldades encontradas e o aprendizado adquirido no processo de busca e recuperação de informação científica na web. Os resultados apontaram diferentes aspectos sobre o planejamento e a realização do processo de busca e recuperação, relevantes para o aprimoramento profissional em relação à obtenção e utilização da informação.

Palavras-chave: Recuperação da informação. Recuperação da informação na web. Informação científica. Ciência da Informação. Letramento informacional. Comportamento informacional de usuários.

Abstract

In the Information Science's constitutive origin, the information search and retrieval process has represented new challenges for its development, given the changes determined by the advent and growth of the web. Given the exponentially increase of information's volume on the web, the number and diversity of users, and the increasing use of search engines, this research contributes to enhancing the understanding of users' information behavior. This work is focused on the identification and analysis of the behavior adopted by postgraduate students in Information Science at an experience of scientific information search and retrieval on the web. Subjects related to information retrieval for scientific production, information retrieval on web, information and digital literacy, as well as user's information behavior are specially studied. The field research encompassed the application of an integrated set of instruments, such as questionnaire and interview, developed from the critical incident technique, methodology related to the investigation's field Cognitive Task Analysis (CTA). From the data collected was possible to identify the paths chosen by the target population, as well as the decision making, strategies selected, difficulties experienced, and the improvements recognized along the process of scientific information search and retrieval on the web. The results show essentially deficiency of planning and reflection during the process. The results showed different aspects of planning and completion of the search and recovery, relevant to the professional development in relation to the collection and use of information.

Keywords: Information retrieval. Information retrieval on the web. Scientific information. Information Science. Information literacy. Users' information behavior.

Lista de ilustrações

Figura 1 – Uso da informação, externalidades e internalidades (MENOU, 1995, p. 480)	26
Figura 2 – Espectro da informação (BELKIN; ROBERTSON, 1976, p. 198)	28
Figura 3 – Atividades de busca (MARCHIONINI, 2006, p. 42)	42
Figura 4 – Assuntos tratados nas perguntas inseridas no questionário, inspirados no processo de busca	63
Figura 5 – Número de pós-graduandos do PPGCI/IBICT-UFRJ ingressados em 2009 e 2010	72
Figura 6 – Proporção de participantes da pesquisa de campo entre pós-graduandos do PPGCI/IBICT-UFRJ	73
Figura 7 – Proporção de participantes que responderam ao questionário de forma presencial x por e-mail	73
Figura 8 – Porcentagem de mestrandos e doutorandos estudados em relação ao ano de ingresso .	74
Figura 9 – Porcentagem de participação de doutorandos e mestrandos por ano de ingresso	74
Figura 10 – Cronograma de coleta de dados para pesquisa de campo	75
Figura 11 – Porcentagem de pós-graduandos entrevistados, divisão por mestrandos e doutorandos	75
Figura 12 – Campos de origem dos mestrandos participantes da pesquisa de campo	76
Figura 13 – Campos de origem dos doutorandos participantes da pesquisa de campo	76
Figura 14 – Campos de origem dos pós-graduandos participantes da pesquisa de campo	77
Figura 15 – IES, por estado, frequentadas na graduação; no detalhe, IES do estado do Rio de Janeiro	77
Figura 16 – Instituições nas quais doutorandos das turmas 2009/2010 concluíram mestrados	78
Figura 17 – Ano de graduação dos pós-graduandos participantes desta pesquisa de campo	78
Figura 18 – Trajetória dos doutorandos, de acordo com ano de graduação, de conclusão do mestrado e de ingresso no PPGCI-IBICT/UFRJ	79
Figura 19 – Comparativo da divisão, por linha de pesquisa, de mestrandos e doutorandos	80
Figura 20 – Divisão dos pós-graduandos por linha de pesquisa	81
Figura 21 – Nuvem de etiquetas com campos e áreas do conhecimento inseridas nos projetos de mestrado das turmas de 2009 e 2010 do PPGCI/IBICT-UFRJ	82
Figura 22 – Nuvem de etiquetas com campos e áreas do conhecimento inseridas nos projetos de doutorado das turmas de 2009 e 2010 do PPGCI/IBICT-UFRJ	82
Figura 23 – Nuvem de <i>tags</i> com campos e áreas do conhecimento inseridas nos projetos de dissertação e tese dos pós-graduandos das turmas de 2009 e 2010 do PPGCI IBICT-UFRJ	83
Figura 24 – Comparação entre as áreas estudadas pelos atuais doutorandos para elaboração da dissertação e áreas inseridas na elaboração da tese	84
Figura 25 – Primeira e segunda maneiras pelas quais mestrandos adquiriram conhecimento sobre como realizar buscas na web	85
Figura 26 – Primeira e segunda maneiras pelas quais doutorandos adquiriram conhecimento sobre busca na web	86

Figura 27 – Primeira e segunda opções por meio das quais pós-graduandos adquiriram conhecimento sobre como realizar busca e recuperação na web	86
Figura 28 – Frequência com que as opções foram combinadas	87
Figura 29 – Autoavaliação da capacidade de buscar e recuperar informações na web - primeira etapa	89
Figura 30 – Autoavaliação da capacidade de buscar e recuperar informações na web - segunda etapa	90
Figura 31 – Quantidade de tempo transcorrida desde a realização da experiência relatada	92
Figura 32 – Mecanismos de busca mais utilizados pelos mestrandos	95
Figura 33 – Mecanismos de busca mais utilizados pelos doutorandos	95
Figura 34 – Mecanismos de busca mais utilizados pelos pós-graduandos	95
Figura 35 – Bases especializadas usadas pelos pós-graduandos	97
Figura 36 – Representação da escolha dos mestrandos por recurso padrão ou avançado	98
Figura 37 – Representação da escolha dos doutorandos por recurso padrão ou avançado	98
Figura 38 – Representação do resultado de expressão booleana conjuntiva (AND)	99
Figura 39 – Representação do resultado de uma busca booleana disjuntiva (OR)	99
Figura 40 – Resultado de busca negativa (NOT)	100
Figura 41 – Resultado de busca booleana com o operador NOT	100
Figura 42 – Resultado de busca booleana utilizando parênteses	100
Figura 43 – Utilização de operadores para refinamento da busca	101
Figura 44 – Número de termos utilizados nas consultas relatadas pelos pós-graduandos	104
Figura 45 – Avaliação dos mestrandos sobre o impacto causado pela informação recuperada ..	107
Figura 46 – Avaliação dos doutorandos sobre o impacto causado pela informação recuperada	107
Figura 47 – Avaliação dos pós-graduandos sobre o impacto causado pela informação recuperada	108
Figura 48 – Estratégias mais usadas para a filtragem do material recuperado	110
Figura 49 – Grau de satisfação dos mestrandos em relação à experiência de busca relatada	111
Figura 50 – Grau de satisfação dos doutorandos em relação à experiência de busca relatada .	112
Figura 51 – Grau de satisfação dos pós-graduandos em relação à experiência de busca relatada	112
Figura 52 – Cruzamento de resultados do impacto causado pela informação obtida com o grau de satisfação dos pós-graduandos com a busca empreendida	116
Figura 53 – Principais dificuldades identificadas pelos pós-graduandos	117

Lista de abreviaturas e siglas

AASL – American Association of School Librarians
ACRL – Association of College Research Libraries
ADI – American Documentation Institute
ALA – American Library Association
ARPA – Advanced Research Projects Agency
ASIS – American Society of Information Science
ASK – *Anomalous State of Knowledge*
CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CDU – Classificação Decimal Universal
CERN – Organização Europeia para a Investigação Nuclear
CI – Ciência da Informação
CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico
CTA – Cognitive Task Analysis
Ence – Escola Nacional de Ciências Estatísticas
FID – Federação Internacional de Documentação e Informação
FACC – Faculdade de Administração e Ciências Contábeis
FID – Federação Internacional de Documentação e Informação
IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia
IES – Instituições de Ensino Superior
IIB – Instituto Internacional de Bibliografia
HTML – Hypertext Markup Language
http – Hypertext Transfer Protocol
SIDC – Scientific Information Dissemination Center
OCLC – On-line Computer Library Center
PDA – Personal Digital Assistant
PPGCI – Programa de Pós-Graduação em Ciência da Informação
UCLA – Universidade da Califórnia Los Angeles
UERJ – Universidade Estadual do Rio de Janeiro
UFF – Universidade Federal Fluminense
UFRJ – Universidade Federal do Rio de Janeiro
UFRRJ – Universidade Federal Rural do Rio de Janeiro
UFSC – Universidade Federal de Santa Catarina
UNIRIO – Universidade Federal do Estado do Rio de Janeiro
URL – Universal Resource Locator

Sumário

1.	Introdução	16
2.	Objetivos	20
3.	Busca e recuperação da informação no contexto contemporâneo da web	22
3.1.	Trajetória inspirada da recuperação da informação	29
3.1.1.	Recuperação da informação	32
3.1.2.	Recuperação da informação na web	36
4.	Comportamento e letramento informacional	45
4.1.	Letramento informacional	46
4.1.1.	Surgimento da expressão information literacy	47
4.1.2.	Definições de letramento informacional	48
4.2.	Letramento digital	51
4.3.	Comportamento informacional	52
4.4.	Interseção entre Ciência da Informação e Ciências Cognitivas	55
5.	Investigação do comportamento e do letramento informacional	58
5.1.	População-alvo	59
5.2.	Elaboração do questionário	61
5.3.	A técnica do incidente crítico	66
5.4.	Elaboração da entrevista	68
5.5.	Teste do questionário	69
5.6.	Aplicação da pesquisa	70
5.7.	Análise dos dados	70
6.	Resultados da pesquisa de campo	72
6.1.	Formação dos pós-graduandos participantes	76
6.2.	Campos do conhecimento dos pós-graduandos	80
6.3.	Habilidades na busca e recuperação da informação	84
6.4.	Capacidade de buscar e recuperar informação na web	88

6.5.	A experiência de busca na web	91
6.6.	Identificação do problema	93
6.7.	<i>Google</i> , mecanismo de busca preferido	94
6.8.	Recurso padrão x recurso avançado	97
6.9.	Da informação específica à elaboração dos termos da consulta	101
6.9.1.	Delimitação da consulta	102
6.9.2.	Comportamento dos usuários na formulação da consulta	103
6.9.3.	Escolha da base de dados ou do mecanismo de busca	105
6.10.	Impacto provocado pela busca	107
6.11.	Influência nos passos seguintes	108
6.12.	Cenário alternativo	109
6.13.	Estratégias para filtrar o conteúdo	110
6.14.	Satisfação dos usuários	111
6.15.	Comparação entre autoavaliação e grau de satisfação	114
6.16.	Comparação entre impacto e grau de satisfação	115
6.17.	Dificuldades enfrentadas no processo de busca	116
6.18.	Aprendizado	118
7.	Conclusões e perspectivas	119
8.	Referências	129
9.	Apêndices	
	Apêndice A	133
	Apêndice B	134
	Apêndice C	137
	Apêndice D	140
	Apêndice E	143

1. Introdução

Sucessivas transformações, especialmente após a Segunda Guerra Mundial, alteraram sobremaneira o tratamento conferido à informação, sobretudo a informação enquanto objeto armazenável e recuperável. Na segunda metade do século XX, enquanto a Ciência da Informação desenvolvia-se como o campo do conhecimento que se ocupa da informação em seus diferentes estágios, a sociedade com bases industriais convertia-se em uma sociedade com características informacionais.

O avanço das tecnologias de informação e comunicação foi essencial na consolidação da atualmente chamada sociedade da informação. Também levou ao desenvolvimento de mecanismos, plataformas e ferramentas que alteraram a maneira como a informação vinha sendo utilizada – passando a ser manejada não apenas por especialistas, mas por diferentes tipos de usuários finais. São exemplos disso os sistemas informatizados de recuperação da informação, a Internet e a World Wide Web.

A web¹ – nome pelo qual a interface humano-máquina da rede mundial de computadores tornou-se conhecida e termo adotado neste trabalho – criou na sociedade conectada computacionalmente a possibilidade, bem como a expectativa, de acesso facilitado e rápido a diferentes tipos de informação. Para o atendimento dessa nova demanda surgiram mecanismos de busca, cada vez mais sofisticados, e tentativas de organizar as informações dispostas no ciberespaço em bancos de dados especializados, bibliotecas digitais, repositórios institucionais ou mesmo periódicos científicos.

De acordo com Lévy (2000, p. 17), ciberespaço trata-se do novo meio de comunicação que surge da interconexão mundial dos computadores. Também denominado rede ou web, “o termo especifica não apenas a infraestrutura material da comunicação digital, mas o universo oceânico de informações que ela abriga, assim como os seres humanos que navegam e alimentam esse universo.”

¹ Teia, no idioma inglês

Fato inegável, a web é hoje o maior acervo de informações do mundo – a ponto de ser impossível calcular seu tamanho real. Ao mesmo tempo em que oferece um vastíssimo volume de informações, pode-se afirmar que justamente “peca pelo excesso”. Consequência disso é, por exemplo, o aumento do tempo gasto para se localizar alguns conteúdos de informação desejada. Diferente das antigas buscas nos arquivos de fichas das bibliotecas, quando se podia estimar o tamanho do acervo, na web é impossível saber o quanto foi recuperado entre tudo que existe sobre o assunto buscado – ponto positivo é que buscas sobre determinado assunto podem ser repetidas quantas vezes forem necessárias, e sem que as “fichas catalográficas” se esgotem.

Utilizar o conteúdo existente no ciberespaço de maneira competente, a fim de atender a diferentes tipos de necessidade informacionais, apresenta-se como um dos grandes desafios contemporâneos – apesar de recuperação da informação estar entre as áreas-chave da Ciência da Informação desde os seus primórdios e ocupar profissionais da informação há décadas. Mesmo diferente no formato, a web requer, especialmente desses profissionais que trabalham com informação, conhecimentos associados a atividades clássicas, como indexação, classificação, organização e recuperação da informação, para que seu uso seja aperfeiçoado e mais completo.

Embora estejam indexação, classificação, organização importantes inseridos no escopo da Ciência da Informação e de áreas correlatas, no presente trabalho recebem destaque as questões relacionadas à recuperação da informação, em especial a recuperação da informação empreendida na web por usuários dispostos a buscar por informação científica para fins de produção acadêmica, e mais especificamente para suas respectivas teses ou dissertações. A partir desta premissa, a população-alvo da pesquisa de campo que integra esta dissertação foi composta por mestrandos e doutorandos que ingressaram nos anos de 2009 e 2010 no Programa de Pós-Graduação em Ciência da Informação (PPGCI) do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), em convênio com a Universidade Federal do Rio de Janeiro (UFRJ).

A recuperação da informação, como resume Saracevic (2010, p. 162), trata sobre o que pode ser feito para acessar, de maneira rápida e efetiva, a informação inserida em determinado repositório – neste caso, a web. Estudos na área evoluíram de acordo com as necessidades apresentadas ao longo das décadas. Em um primeiro momento, sistemas de recuperação da informação eram destinados a profissionais da informação habilitados, que atuavam como intermediários entre os usuários e o acervo no qual seria realizada a busca.

Sem o trabalho desses profissionais tornava-se difícil ao usuário – hoje chamado final – localizar o que estava procurando.

Alterações significativas foram percebidas nesse cenário com o surgimento da web, em princípios dos anos 1990, assim como em decorrência do crescimento desenfreado, tanto do uso quanto do tamanho da rede. De interação estática e baixo volume de informação na fase inicial, a web transformou-se progressivamente no veículo preferido de usuários – com diferentes perfis e níveis de conhecimentos – e local ideal para o armazenamento de todos os tipos de materiais, sejam vídeos, textos, músicas, filmes ou imagens. Acrescente-se a isso, o caráter interativo e dinâmico da rede, além do fato de ser hoje utilizada por aproximadamente dois bilhões de pessoas ao redor do mundo.

Em pouco mais de uma década, a web passou a fazer parte do cotidiano de indivíduos inseridos nos mais diferentes contextos, inclusive, desde a sua origem, o científico. Para esse público em especial, a rede abriga artigos acadêmicos, revistas especializadas, bases de dados de temas específicos, teses, dissertações e acervos de bibliotecas digitais, entre outros materiais – normalmente disponibilizados em formato digital e prontos para serem salvos no computador do usuário.

Sobre as necessidades individuais, Ranganathan (1967 *apud* CAMPOS, 1999, p. 4) afirma que nem tudo é interesse de todos, tendo cada indivíduo suas próprias precisões. Ranganathan escreve ser inequívoco respeitar os diferentes tipos de usuários, assim como oferecer bibliotecas distintas e preparar variadas formas de organização de acervos de acordo com cada perfil de cada usuário.

Considerando-se o cenário atual, no qual a web surge como a principal fonte de informação, e feitas as devidas atualizações, as cinco leis de Ranganathan mostram-se muito pertinentes mesmo 40 anos depois de terem sido estabelecidas. “As leis de Ranganathan pensadas para a Biblioteconomia servem de inspiração para todos os profissionais da informação, independente do campo a que estão ligados.” (CAMPOS, 1999, p. 5)

“Livros são para ser usados”, inicia Ranganathan. Da mesma forma, o infinito volume de informações disponível na web está ali para ser descoberto e utilizado. “A cada leitor o seu livro” remete ao dito anteriormente, cada indivíduo tem necessidades próprias. A terceira lei, “a cada livro o seu leitor”, sugere existir na web informações disponíveis para todo tipo de usuário. A quarta lei refere-se a poupar o tempo do leitor, ou seja, quanto mais rápido a informação for recuperada, melhor, como observa Campos. Por fim, na quinta lei,

Ranganathan afirma ser a biblioteca um organismo em crescimento, característica da web desde o surgimento, levando-se em consideração que “a produção do conhecimento é um ato contínuo e dinâmico do ser humano”. (CAMPOS, 1999, p. 5)

Dentro do contexto apresentado até aqui surgiram diversas questões que estimularam a realização deste trabalho, entre elas as seguintes. Para cada demanda existem respostas que podem ser consideradas mais adequadas? Está conseguindo-se chegar efetivamente até elas? Com a diversidade e o volume incalculável de informações disponíveis na web, é possível encontrar exatamente o que se necessita? É possível julgar como completo o que foi recuperado em uma busca? As técnicas empregadas surtem o efeito desejado? Até que ponto as buscas empreendidas por pós-graduandos em Ciência da Informação obtêm resultados mais plenos do que as realizadas por outros usuários finais já estudados por pesquisadores da área? A área de formação influencia no desempenho?

Com a pesquisa de campo realizada para este estudo foram obtidas respostas para parte significativa dessas indagações, assim como obtidos subsídios para análise do comportamento de pós-graduandos em Ciência da Informação ao buscar informações na web, observando-se dificuldades, padrões de uso, grau de satisfação, aprendizado, entre outros parâmetros.

Esta dissertação divide-se em duas partes principais. A primeira é formada por revisão bibliográfica de conteúdos relacionados à Ciência da Informação, à recuperação da informação, ao letramento digital, às Ciências Cognitivas e à competência informacional dos usuários. Na segunda está concentrada a pesquisa de campo. De forma detalhada, trata-se sobre a aplicação da mesma, as técnicas empregadas, como os resultados foram tabulados e, posteriormente e especialmente, a interpretação conferida ao material coletado.

A partir dos dados colhidos tanto na revisão bibliográfica quanto na pesquisa de campo pretendeu-se chegar a conclusões sobre o comportamento do grupo pesquisado ao realizar buscas na web, assim como identificar aspectos desencadeadores de tomadas de decisão ou que influenciam no tipo de estratégias adotadas. E, a partir disso, compreender melhor como os pós-graduandos pesquisados agiram durante a experiência de buscar e recuperar informações na web.

2. Objetivos

O presente trabalho compreende uma pesquisa exploratória, de características essencialmente qualitativas, associando elementos quantitativos que auxiliam nas análises qualitativas.

Esta dissertação apresenta-se dividida em duas partes: revisão bibliográfica e pesquisa de campo. Na primeira parte está incluída a produção do referencial teórico, para a construção do qual foram utilizados textos de autores de diferentes áreas, tratando sobre temas relacionados à informação, à Ciência da Informação, à recuperação da informação, à competência informacional, ao letramento digital, a questões cognitivas e ao comportamento informacional dos usuários.

A segunda parte compreende, inicialmente, duas ferramentas para levantamento de dados: questionário e entrevista semiestruturada. Com esses instrumentos é realizada a coleta de dados, os quais, após minuciosa análise, levam ao estabelecimento de considerações e conclusões, bem como de perspectivas para estudos futuros.

Todo o trabalho baseia-se em um objetivo geral e em três objetivos específicos, os quais são apresentados em seguida:

Objetivo geral:

Estudar o comportamento de pós-graduandos em Ciência da Informação ao buscar e recuperar informações científicas na web para o desenvolvimento de seus projetos de pesquisa, com o objetivo de analisar aspectos que levam a desempenhos bem-sucedidos ou a condutas críticas.

Objetivos específicos:

- Mapear o percurso adotado pelos pós-graduandos em Ciência da Informação na busca e recuperação de informações científicas na web, a fim de identificar as estratégias adotadas ao longo das etapas do processo.
- Identificar as decisões tomadas pelo pós-graduando em Ciência da Informação no decorrer do processo de busca e recuperação na web, a fim de analisar os pontos críticos do processo em relação aos resultados e dificuldades encontrados.
- Identificar o grau de conhecimento dos pós-graduandos da área de Ciência da Informação em relação a diferentes competências e habilidades necessárias na condução do processo visando a indicar estratégias que permitam maior efetividade na busca e recuperação de informações científicas na web.

3. Busca e recuperação da informação no contexto contemporâneo da web

Informação integra todas as atividades humanas, está implícita em tudo o que fazemos, independente da área de atuação. Justamente por estar implícita e presente em praticamente todos os campos do conhecimento, defini-la dentro de uma única disciplina nem sempre é tarefa simples. Na Ciência da Informação, porém, tanto por ser a informação o objeto central de estudo quanto por ter recebido atenção especial nas pelo menos últimas seis décadas, é possível encontrar significativo número de considerações respaldadas por pesquisadores renomados da área.

Capurro e Hjørland realizaram em 2003² ampla revisão do conceito de informação, destacando o papel central assumido na sociedade contemporânea. Não por acaso, no final do século XX, o sociólogo Manuel Castells afirmou vivermos em uma Sociedade da Informação. Evidências do papel diferenciado da informação, de acordo com Capurro e Hjørland, são o desenvolvimento e a disseminação do uso das redes de computadores desde a Segunda Guerra Mundial e a emergência da Ciência da Informação como disciplina nos anos 1950.

Atendo-se à revisão dos autores no âmbito da Ciência da Informação tem-se acesso a usos do conceito de informação ainda no início do século XX, quando a expressão *bureau de informação* era utilizada para indicar um escritório onde eram oferecidos serviços de referência. Em 1915, o termo foi citado pela bibliotecária especializada Ethel Johnson: “antes de qualquer coisa, a biblioteca especializada é um *bureau de informação*”. Essa afirmação levou Williams (1998, p. 174 *apud* CAPURRO; HJØRLAND, 2007, p. 177) a destacar que a função das bibliotecas é tornar a informação disponível.

² O artigo original foi publicado em 2003 no *Annual Review of Information Science and Technology*. Ed. Blaise Cronin. v. 37, cap. 8, p. 343-411. A referência utilizada nesta dissertação foi a tradução publicada na revista *Perspectivas em Ciência da Informação*, em jan/abr. 2007

Ainda de acordo com Williams, como destacam Capurro e Hjørland, bibliotecários especializados foram os primeiros documentalistas nos Estados Unidos e, pela definição de Rayward (1998), documentalistas são vistos como os primeiros cientistas da informação. A evolução do uso do conceito documentação para o conceito Ciência da Informação sedimentou-se em 1968, quando o *American Documentation Institute* (ADI), fundado em 1937, foi renomeado *American Society for Information Science* (ASIS). Com isso, dentro do contexto da Ciência da Informação – sem entrar nos méritos e deméritos enumerados pelos autores –, o termo documentação foi gradualmente substituído por informação.

No clássico *Information Science: What is it?*, Borko inicia o texto comentando exatamente sobre a mudança de nome do *American Documentation Institute* e a dificuldade de dar aos amigos e familiares explicações sobre do que se tratava afinal essa nova ciência, a Ciência da Informação. Logo na introdução, Borko divide com os leitores que decidiu escrever o artigo por prazer antes que fosse forçado a fazê-lo por obrigação ou solicitação de alguém.

É a partir do capítulo *Professional Aspects of Information Science and Technology*, escrito por Robert S. Taylor no *Annual Review* do *American Documentation Institute*, que Borko redige sua notória definição³:

Ciência da Informação é a disciplina que investiga as propriedades e o comportamento da informação, as forças que governam seu fluxo e os meios de processá-la para otimizar sua necessidade e uso. Relaciona-se com o corpo de conhecimento relativo a origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação. (BORKO, 1968, p. 3)⁴

Borko (1968, p. 4) destaca que a pesquisa em Ciência da Informação investiga as propriedades e o comportamento da informação, a utilização e a transmissão da informação, bem como o processamento da informação para ótima armazenagem e recuperação. O autor vê ainda a Ciência da Informação como uma disciplina que tem como objetivo prover um

³ Em uma pesquisa mais aprofundada sobre a origem da definição clássica de Borko, chega-se não somente ao texto de Robert S. Taylor, mas aos Anais da reunião do Georgia Institute of Technology, realizada em 1962 e durante a qual a Ciência da Informação foi formalmente criada (BRAGA, 1995, p. 1)

⁴ “*Information science is that discipline that investigates the properties and behavior of information, the forces governing the flow of information, and the means of processing information for optimum accessibility and usability. It is concerned with that body of knowledge relating to the origination, collection, organization, storage, retrieval, interpretation, transmission, transformation, and utilization of information.*” (BORKO, 1968, p. 3)

conjunto de informações que levará avanços a instituições por meio de procedimentos dedicados ao acúmulo e à transmissão de conhecimento.

Classificando a Ciência da Informação como um campo interdisciplinar, Borko ressalta que a Biblioteconomia e a Documentação são aplicações da área. Para o autor, técnicas e procedimentos usados por bibliotecários e documentalistas deveriam basear-se nos resultados teóricos da Ciência da Informação, assim como, inversamente, os teóricos deveriam estudar a técnicas já amplamente aplicadas pelos profissionais envolvidos com o tratamento da informação há longo tempo. (BORKO, 1968, p. 3-4).

A ideia de cruzar competências consolidadas com habilidades novas reaparece na definição de Foskett (1973, p. 56), para quem a Ciência da Informação é uma disciplina que surge da “fertilização cruzada” de ideias que incluem a velha arte da Biblioteconomia, a nova arte da Computação, as artes dos novos meios de comunicação e aquelas ciências como Psicologia e Linguística, que, em suas formas modernas, têm a ver diretamente com todos os problemas da comunicação – a transferência do conhecimento organizado.

Wersig e Neveling (1975), citados por Belkin e Robertson (1976, p. 1), destacam que historicamente a Ciência da Informação desenvolve-se devido a um fenômeno específico que sempre existiu e que naquele momento, na metade dos anos 1970, torna-se um novo objeto de estudo. O fenômeno sobre o qual tratam os pesquisadores é o problema da transmissão do conhecimento para aqueles que dele precisam. E, na consideração pela qual se tornariam sempre lembrados, Wersig e Nevelling afirmam que tratar sobre essa questão é uma responsabilidade social, a qual parece ser o real fundamento da Ciência da Informação.

A partir das contribuições de Borko (1968), Foskett (1973), Wersig e Neveling (1975) e Capurro e Hjørland (2003), percebe-se o papel de destaque conferido à Ciência da Informação na tarefa de contribuir com a comunicação, a transmissão ou a transferência de conhecimento. Relacionado a esse aspecto, Saracevic (1996, p. 47) afirma que a Ciência da Informação é um campo dedicado a questões científicas e à prática profissional voltadas para os problemas de efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das necessidades de informação.

Definição ratificada em 2010, quando Saracevic afirma que

definição e noção básicas da área ainda estão de pé: Ciência da Informação é o campo de prática profissional e de pesquisa científica que aborda a comunicação eficaz de informação e dos objetos de informação, principalmente os registros de conhecimento entre humanos em um contexto de necessidade social, organizacional e individual e de uso da informação. O território da Ciência da Informação era e ainda é a transmissão do universo de conhecimento registrado, centrado na manipulação desse conhecimento, ou seja, a representação, organização e recuperação da informação, ao invés de informações conhecidas. (SARACEVIC, 2010, p. 166)

Ingwersen e Järvelin (2005, p. 385) escrevem que a informação estudada pela Ciência da Informação é aquela que satisfaz a duas solicitações: por um lado, resulta na transformação da estrutura cognitiva do emissor; por outro, quando percebida, afeta e transforma o estado do conhecimento do receptor.

O conhecimento a ser emitido tem contorno de informação, a qual, ao ser adequadamente assimilada pelo receptor, leva à produção de novo conhecimento, isto é, “modifica o estoque mental de informações do indivíduo e traz benefícios ao seu desenvolvimento e ao desenvolvimento da sociedade em que vive”, como afirma Barreto (1994, p. 5) Mas, afinal, de que natureza é a informação de que trata a Ciência da Informação? Ao longo de décadas, este vem sendo tema bastante explorado nos estudos da área.

Em artigo de 2004, Pinheiro discute teórica e conceitualmente diferentes abordagens de informação, revisando resultados dos estudos de diversos cientistas da informação, como Mikhailov, Foskett, Belkin e Robertson. A autora observa que tradicionalmente a informação é relacionada a documentos impressos e a bibliotecas, e que todos os campos do conhecimento alimentam-se de informação, mas poucos são aqueles que a tomam por objeto de estudo. A informação de que trata a Ciência da Informação, explica Pinheiro, pode estar no diálogo entre cientistas, na comunicação informal, na inovação para a indústria, na fotografia, ou mesmo na Internet.

A pesquisadora afirma ainda que a informação “movimenta-se em um terreno multifacetado, podendo tanto ser informação em uma determinada área quanto sob determinada abordagem” (PINHEIRO, 2004, p. 2). A informação em um determinado campo depende do contexto em que está inserida – e por sua transversalidade, pode perpassar todos os campos. Quando se trata da abordagem, o enfoque pode ser, por exemplo, cognitivista, relacionando informação a conhecimento; gerencial, com a informação influenciando a tomada de decisões; ou econômico, quando a informação passa a ser vista como mercadoria e adquire valor agregado (PINHEIRO, 2004, p. 2). Nesta dissertação, receberá atenção maior a

informação relacionada ao conhecimento, a partir de abordagem cognitivista, mesmo que, em alguns momentos da pesquisa de campo, a perspectiva gerencial se faça presente quando há tomada de decisão por parte da população-alvo.

Menou (1980), como relata Pinheiro (2004, p. 8), faz uma tentativa de retratar o uso real da informação ao construir um quadro com as externalidades e as internalidades que interferem nesse uso. O autor procura mostrar em primeiro lugar que os passos no processo de solução de problema dependem da base interna de conhecimento – individual e coletivo – com a qual fontes internas têm que combinar. “O conhecimento interno opera sob influência de uma série de fatores, identificados, por enquanto, como personalidade, cultura, emoção, lógica (processo de raciocínio) e inteligência (qualidade de julgamento/decisão).” (MENOU, 1995, p. 480) Esses aspectos podem ser vistos no diagrama representado na Figura 1.

É baseado em Brookes (1980, p. 252 *apud* MENOU, 1995, p. 480) que o autor faz a distinção entre lógica e inteligência. De acordo com Brookes, problemas podem ser solucionados com um cálculo objetivo ou por meio de um julgamento/decisão. “Julgamentos têm grande importância social, me surpreende que tanto esforço seja devotado para o ensino de cálculos e tão pouco para o ensino de julgamentos.”

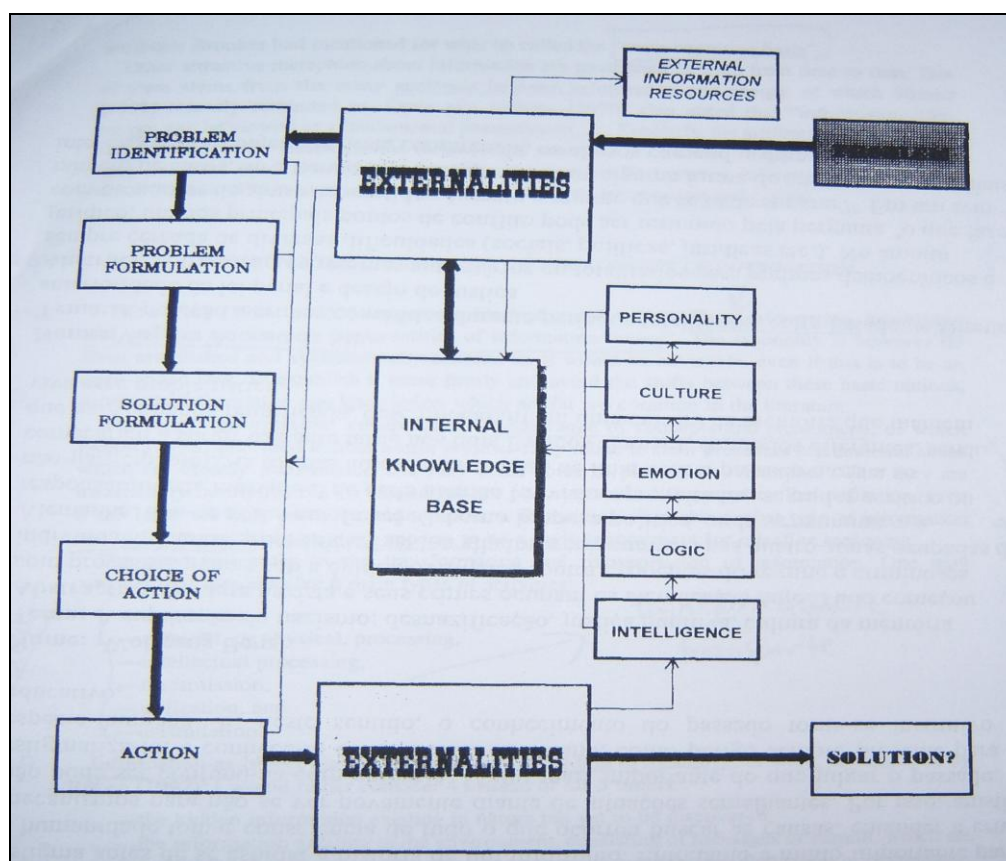


Figura 1 – Uso da informação, externalidades e internalidades (MENOU, 1995, p. 480)

Sobre a Figura 1, Menou descreve a alegria de descobrir que os fatores identificados em seu modelo – personalidade, cultura, emoção, lógica e inteligência – estão bem próximos dos atributos mencionados por Brookes no que o autor chama de “base metacognitiva”.

Menou (1995, p. 481) afirma que a palavra informação encapsula grande variedade de conceitos e fenômenos, os quais estão relacionados tanto a processos quanto a estados materiais, sendo que estes estão intimamente interligados. Essa conexão não impede a continuidade das diferenciações e definições – mesmo que no caso das definições, se resume a tentativas e tenha validade temporária. Nesse cenário, Menou destaca que permanecer alerta sobre com que tipo de informação está se lidando em determinado estágio parece ser o requerimento básico para um raciocínio efetivo.

Logo em seguida, o autor apresenta conceitos e fenômenos ligados a processos e a estados materiais. De acordo com Menou, processos referem-se à transformação e ao transporte da informação. Pode-se considerar que estejam envolvidos aí seis tipos de atividades: 1) aquisição; 2) processo material ou físico; 3) processo intelectual; 4) transmissão; 5) utilização; e 6) assimilação.

Em relação aos estados materiais, a “informação-como-objeto” é diferenciada de acordo com suas fontes ou armazenamento. A respeito disso, Menou considera quatro ordens:

1. A informação oculta existente na natureza, mas ainda não descoberta;
2. A informação existente no ambiente, que consiste em mensagens e estímulos, os quais são inteligíveis, mas ainda esperam para serem percebidos e interpretados, que também costuma ser chamada de informação latente;
3. A informação configurada por indivíduos; e
4. A informação configurada por grupos e sistemas sociais. (MENOU, 1995, p. 481)

No mesmo artigo, em 1995, Menou apontava a eminência de uma quinta ordem, referente à informação artificial contida nos sistemas de computador para um uso coletivo.

Menou declara que todas as dimensões – de processos a estados materiais – constantemente interagem e são interdependentes. Informação quando comunicada pode ser caracterizada concomitantemente por sua fonte, processo ou forma de cada canal, conteúdo expresso ou produto – podendo ser esta informação como coisa, objeto, fonte ou *commodity*. “A lógica primitiva, pela qual o pensamento ocidental tenta forçar a classificação da realidade de forma rígida, obriga a considerar se qualquer entidade está dentro ou fora de determinada categoria, enquanto estar dentro ou fora de uma categoria é claramente mais um impedimento do que qualquer outra coisa.” (MENOU, 1995, p. 481-2)

Alternativa para essa visão, de acordo com o autor, é proposta por Dervin e Nilan. Surge a ideia de que a informação atua quando faz sentido, quando ou encaixa-se em significados já conhecidos por quem a está recebendo ou integra-se a esses significados, transformando-os. “Quando se considera a natureza ou o papel da informação, sob qualquer perspectiva, torna-se central o significado ou sentido.” (MENOU, 1995, p. 482) O estudo de Dervin sobre como pessoas constroem sentido será retomado mais adiante, quando forem tratadas questões relacionadas às ciências cognitivas.

Essa discussão estimula um retorno ao texto clássico de Belkin e Robertson intitulado *Ciência da Informação e o fenômeno da informação*, de 1976. Para os pesquisadores, Ciência da Informação é uma disciplina orientada para o problema da transferência efetiva da informação desejada emitida por um gerador humano para um receptor humano. Também consideram que a noção única e comum a todos os conceitos de informação existentes é a da mudança de estrutura. “Em nosso ponto de vista, a *única* noção básica comum para a maioria dos usos de informação é a ideia da estrutura sendo alterada, o que definimos da seguinte maneira: *informação é aquilo que é capaz de alterar a estrutura.*” (BELKIN e ROBERTSON, 1976, p. 198, *grifo original*)

Ambos admitem ser a definição ampla demais e englobar noções para as quais a expressão nunca é usada. Sendo assim, decidem deixar o vocábulo informação indefinido, mas discutindo os vários usos do termo a partir da ideia de transformar a estrutura da mente. A maneira empregada foi escolher diversos contextos nos quais a informação é usada. Belkin e Robertson criaram o chamado espectro da informação (*Information Spectrum*), formatado como aparece na Figura 2.

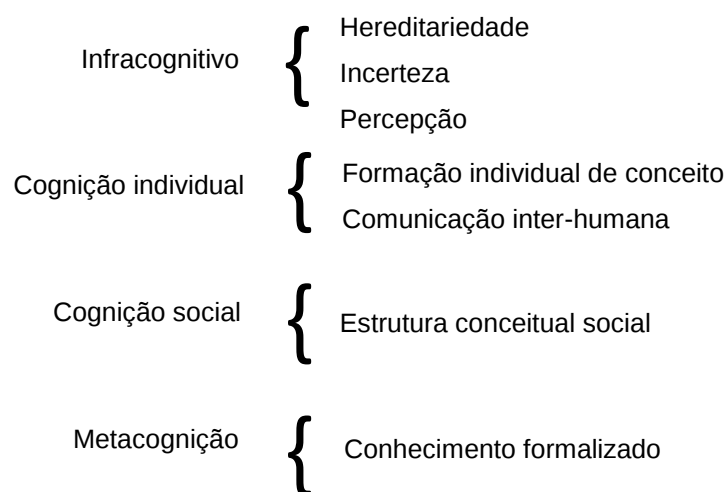


Figura 2 – Espectro da informação (BELKIN; ROBERTSON, 1976, p. 198)

Ao tomarem o espectro da informação como base, Belkin e Robertson afirmam que a informação para a Ciência da Informação começa com o advento das estruturas semióticas, na interface entre formação individual de conceito (*individual concept-forming*) e comunicação inter-humana (*inter-human communication*), e continua próxima à interface entre estrutura conceitual social (*social conceptual structures*) e conhecimento formalizado (*formalized knowledge*).

A parte do espectro da informação indicada como sendo de interesse da Ciência da Informação é caracterizada pelos autores como sendo a estruturação proposital da mensagem pelo remetente a fim de afetar a estrutura do conhecimento/imagem do receptor. “Isso implica que o remetente tenha conhecimento da estrutura do receptor.” (BELKIN e ROBERTSON, 1976, p. 200). Saracevic, em revisão de Pinheiro e Loureiro (1995, p. 47-48), corrobora com essa ideia ao afirmar que a efetividade da comunicação do conhecimento se dá quando de sua transmissão de um arquivo a outro, ocasionando mudanças.

Com o advento da web, porém, essa relação se complexifica à medida que o emissor, apesar de ter ideia do público que pode atingir, não sabe efetivamente em quem chegará a informação inserida na rede.

Em relação às discussões sobre o termo informação, há ainda a visão de Le Coadic (2004, p. 4), para quem informação é um conhecimento inscrito em um suporte em forma escrita, oral e audiovisual. Le Coadic afirma que a informação comporta um elemento de sentido, ou seja, é um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espacial-temporal. Exemplos são um impresso, um sinal elétrico, uma onda sonora ou, como visto mais recentemente, um texto na web.

Seja pelo simples prazer de conhecer, de estar informado sobre os acontecimentos políticos, os progressos da ciência e da tecnologia, ou pelo prazer menos simples de estar a par dos últimos temas e resultados das pesquisas, de acompanhar a vanguarda do conhecimento científico, o objetivo da informação permanece sendo a apresentação de sentidos ou seres em sua significação, ou seja, continua sendo o conhecimento; e o meio é a transmissão do suporte, da estrutura. (LE COADIC, 2004, p. 4-5)

3.1 Trajetória inspirada da recuperação da informação

Antes da web como a conhecemos hoje, existiram outras ideias de redes e equipamentos para armazenar as informações do mundo. Criativos, funcionais e completos, *Mundaneum* e *Memex* foram pensados por seus imaginativos criadores Paul Otlet e Vannevar

Bush, respectivamente, como soluções ideais não só para a organização, como para a recuperação da informação.

Esboçado em 1895 e desenvolvido parcialmente no início do século XX, o *Mundaneum* abrigaria o repertório mundial do conhecimento em fichas catalográficas, de acordo com a mente profícua de Otlet. Tudo seria desenvolvido no Instituto Internacional de Bibliografia (IIB), recém-inaugurado por Otlet junto com Henri La Fontaine após a Conferência Internacional de Bibliografia realizada em Bruxelas, na Bélgica, país de origem dos amigos advogados. No projeto de Otlet, os exemplares dos documentos e as microcópias dos livros arrolados no repertório seriam todos depositados no *Mundaneum*. (PEREIRA, 1995, p. 2)

Em 1914, o chamado *Répertoire Bibliographique Universel* de Otlet e La Fontaine consistia em mais de 11 milhões de fichas e o IIB tinha aproximadamente 700 membros, como relata Pereira (1995, p. 2) em artigo homenageando os 100 anos do IIB, que a partir de 1938 foi rebatizado como Federação Internacional de Documentação e Informação (FID). Para ordenar esse expressivo repertório, a dupla de pesquisadores desenvolveu a Classificação Decimal Universal (CDU), entre 1904 e 1907⁵, que vinha a ser uma adaptação da Classificação Decimal de Dewey, de 1876.

Atualmente, a CDU é utilizada em bibliotecas e centro de documentação de mais de 130 países em 40 idiomas. Com sua extravagante invenção, como descreve Pereira (1995, p. 3), Otlet pretendia tornar possível a classificação e recuperação automática de documentos, a manipulação mecânica de todos os dados registrados para obter novas combinações de fatos ou novas relações de ideias.

Um ano após a morte de Otlet e meio século depois de seu *Mundaneum*, outra invenção ganhava a atenção da comunidade científica: o *memex*. Idealizado pelo engenheiro Vannevar Bush, fora apresentado no artigo *As we may think*, publicado em julho de 1945 na revista científica *Atlantic Monthly*. No texto, Bush (1945, p. 104) define o *memex* como o “dispositivo no qual o indivíduo armazena todos os seus livros, registros e comunicações, mecanizado e que pode ser consultado com extrema velocidade e flexibilidade”.

Em seu artigo, Bush complementa dizendo que o *memex* é um suplemento ampliado e próximo da memória do indivíduo. Mais do que idealizar um equipamento que solucionaria

⁵ UDC Consortium. *What is the UDC?* Disponível em: <<http://www.udcc.org/about.htm>>. Acesso em: 6 dez. 2010

problemas relacionados à organização e, especialmente, à recuperação da informação, Bush ganhou notoriedade ao questionar os sistemas de indexação alfabéticos e numéricos adotados pelas bibliotecas até então. O pesquisador chamava a atenção para a discrepância que esse modelo tinha em relação à forma como a mente humana funcionava.

Como bem observou Bush, a mente opera por associação e seria esse também o modo como o *memex* trabalharia. Ao adotar esse princípio, afirma Pereira (1995, p. 4), o dispositivo permitiria ao usuário navegar por ele, como se fosse a extensão de sua própria mente. Ao longo dos anos seguintes, o texto visionário serviu de fonte de inspiração para muitos – e continua sendo, meio século depois, referência em diferentes áreas do conhecimento, especialmente nos estudos relacionados à Ciência da Informação.

Um dos pesquisadores inspirados por Bush foi Theodor Nelson, que em 1965 mencionou pela primeira vez o termo hipertexto no artigo *The Hypertext* publicado nos anais da Fundação Mundial de Documentação. Ao tratar sobre hipertexto, Nelson se referia à escrita não sequencial, um texto com uma bifurcação, permitindo ao leitor eleger qual o caminho a seguir, algo muito parecido com o já sugerido por Bush vinte anos atrás.

Para Morville e Rosenfeld, hipertexto é uma maneira não linear de estruturar informações. “O sistema do hipertexto envolve dois tipos primários de componentes: os itens ou grupos de informações interligados e as ligações (*links*) entre esses grupos de informações.” (MORVILLE; ROSENFELD, 2006, p. 77).

No período entre as ideias de Otlet e de Bush, na primeira metade do século XX, os sistemas de armazenamento e recuperação de informação tinham a forma de centenas, quando não milhares, de fichas catalográficas armazenadas em bibliotecas. No pós-guerra, porém, o cenário passou por mudanças significativas com o surgimento gradual de computadores – na época, máquinas volumosas que ocupavam salas inteiras, tinham alto custo, velocidade de processamento lenta e pequena memória interna para manipular dados (SANTIAGO, 2003, p. 27).

Em alta nos anos 1950, encontravam-se os cartões perfurados e as fitas magnéticas para armazenamento de dados. (PALMER, 1987 *apud* SANTIAGO, 2003, p. 28). O período também ficou marcado, como afirma Lesk (1995, p. 4), pelo lançamento do primeiro satélite artificial ao espaço. Com o Sputnik, a então União Soviética dava início à corrida ao espaço, ao mesmo tempo que despertava nos Estados Unidos uma preocupação: os norte-americanos pouco sabiam sobre os desenvolvimentos científicos na Rússia.

A euforia do lançamento do Sputnik e os boatos de que a União Soviética estaria desenvolvendo pesquisas importantes sobre transferência de informação levaram os Estados Unidos a investir em diferentes frentes, entre elas a recuperação da informação, tema que será abordado logo em seguida.

3.1.1 Recuperação da informação

A necessidade de informação de um indivíduo é o ponto de partida do processo de recuperação. Dos tradicionais sistemas usados em bibliotecas aos contemporâneos mecanismos de busca na web, todos são acionados e têm seus processos desencadeados tão logo o usuário formule e submeta sua demanda informacional. O atendimento a essa solicitação continua sendo atribuição de profissionais da informação, como, por exemplo, bibliotecários responsáveis pelos sistemas de referências das bibliotecas.

Porém, com os avanços tecnológicos, a disseminação do uso da web e as melhorias constantes nos mecanismos de busca, tarefas normalmente desempenhadas por profissionais habilitados passaram a ser executadas também pelos chamados usuários finais. Incluídos nesse grupo estão pesquisadores e cientistas – como os participantes da pesquisa de campo realizada para esta dissertação. “Recuperação de informação científica e literatura especializada, antes dominada por bibliotecários, está agora disponível diretamente para os cientistas.” (LAWRENCE e GILES, 1998, p. 98)

De acordo com Saracevic,

buscar informação tornou-se ubíquo – todo mundo é “buscador” hoje em dia. Entre outras razões, avanços em recuperação da informação fizeram isso possível. Nesse processo, recuperação da informação tornou-se também altamente comercial. Grandes mecanismos de busca contemporâneos são parte desta tendência. O campo da Ciência da Computação tem bastante envolvimento. A recuperação da informação contemporânea se propagou para várias áreas. Originalmente, recuperação da informação concentrava-se em textos. Isto se expandiu para vários outros tipos de mídias/suportes. Agora existem pesquisas e esforços pragmáticos devotados à recuperação da informação em música, vídeo, fotografias e imagens em movimento e multimídia. Enquanto originalmente a recuperação da informação era monolíngue, agora novos esforços são devotados para recuperação da informação interlíngue. No entanto, a grande questão da recuperação da informação continua sendo: “Como pode ser feito o acesso à informação gravada de maneira rápida e efetiva?” (SARACEVIC, 2010, p. 162)

O processo de recuperação da informação, tal qual foi cunhado por Calvin Mooers em 1951, engloba os aspectos intelectuais de descrição da informação e suas especificidades para a busca, além de quaisquer sistemas, técnicas ou máquinas que sejam empregadas para realizar esta operação (MOOERS, 1951 *apud* SARACEVIC, 1996, p. 44). Apesar de

formalizada nos anos 1950, a ideia da recuperação da informação já era sugerida por autores como Vannevar Bush na década anterior. Como já visto, ao apresentar o *memex* no artigo *As we may think*, Bush descreve como seria acessar automaticamente grandes quantidades de conhecimento armazenado no equipamento.

O final dos anos 1950 e o começo dos 1960, como descrito por Lesk (1995, p. 4), foram tempos de grande experimentação de sistemas de recuperação da informação, que funcionavam como interface entre o usuário e os documentos que compunham uma coleção. Saracevic (1996, p. 44) destaca que “nos anos 1950 uma massa crítica de cientistas, engenheiros e empreendedores começaram entusiasticamente a trabalhar o problema e a solução apontados por Bush. Nos anos 1960, esse trabalho tornou-se uma atividade relativamente ampla, bem financiada e organizada, que deu origem a debates estimulantes e à acalorada argumentação acerca das melhores e mais adequadas soluções (técnicas, conceitos, sistemas etc.)”.

As desejadas soluções surgiram tanto em nível teórico quanto em realizações práticas. Saracevic cita o que chama de exemplos históricos para mostrar a evolução de sistemas, técnicas e/ou máquinas utilizadas para a recuperação da informação. A variedade vai dos já citados cartões perfurados, nos anos 1950, aos CD-ROMS, surgidos nos anos 1980; dos sistemas não-interativos àqueles de múltiplas possibilidades de interação, com interfaces inteligentes, transformando a recuperação da informação em um processo altamente interativo; dos textos escritos às multimídias; e da recuperação de citações à recuperação de textos completos. “O trabalho determinado pela necessidade de recuperar informações suscitou questões e promoveu pesquisas exploratórias de fenômenos, processos e variáveis, bem como de causas, efeitos, comportamentos e manifestações relacionadas.” (SARACEVIC, 1996, p. 44)

As discussões geradas pelos testes na década de 1960, nas quais era questionada a indexação de todas as palavras em um documento e não somente das palavras-chave, levaram ao surgimento de medidas de avaliação, como revocação e precisão, desenvolvidas por Cyril Cleverdon, do então Cranfield College of Aeronautics, como aponta Lesk (1995, p. 5). Segundo Cleverdon, revocação (*recall*) refere-se à fração de documentos recuperados em relação ao total, enquanto que precisão diz respeito à fração de documentos recuperados que são relevantes.

Na área da Ciência da Informação, as definições de revocação e precisão foram disseminadas por Lancaster (1993, p. 306) em diversas obras. Para o autor, revocação é a extensão com que todos os itens numa base de dados considerados pertinentes são recuperados durante busca nessa base de dados. Uma busca com ‘revocação alta’ será aquela em que a maioria dos itens pertinentes, se não todos, forem recuperados. O coeficiente de revocação – medida da extensão com que se alcança a precisão desse processo – é o número de itens pertinentes recuperados dividido pelo número total de itens.

Precisão é a extensão com que todos os itens numa base de dados que são considerados relevantes são recuperados durante uma busca nessa base de dados. “O coeficiente de precisão – medida da extensão com que se alcança a precisão do processo de recuperação relevante – é o número de itens relevantes recuperados dividido pelo número total de itens recuperados”. (LANCASTER, 1993, p. 306)

Ao se tratar de sistemas fechados e com número conhecido dos itens que os compõem, as medidas de revocação e de precisão continuam válidas. No entanto, nos sistemas que compõem a web, as medidas convencionais mostram-se inadequadas, pois sem saber o tamanho do acervo, torna-se impossível precisar quanto dele foi recuperado ou mesmo quanto é relevante, como analisa Yang (2005, p. 33).

Na segunda metade da década de 1960, surgiram padrões permitindo o compartilhamento e a utilização de dados catalogados por meio de serviços bibliográficos. Exemplo é o desenvolvido pelo *On-line Computer Library Center* (OCLC), fundado em 1967 por iniciativa de instituições acadêmicas de Ohio, nos Estados Unidos, para desenvolver um sistema computadorizado por meio do qual as bibliotecas pudessem compartilhar recursos e reduzir custos. “O primeiro presidente do OCLC Frederick G. Kilgour vislumbrou a transformação do âmbito regional do centro para uma rede cooperativa internacional.” (SANTIAGO, 2004, p. 29) Os avanços computacionais colaboraram para o êxito da ideia. Atualmente, mais de 72 mil bibliotecas em 170 países e territórios em todo o mundo⁶ usam os serviços da OCLC para localizar, obter, catalogar, emprestar e preservar materiais de biblioteca.

A partir da década de 1970, computadores já conseguiam realizar cálculos mais complexos e gravar dados para serem usados posteriormente. Por conta disso, diferentes testes dos sistemas de recuperação *offline* puderam ser realizados. Se no sistema manual de busca de

⁶ Dados atualizados disponíveis em: <<http://www.oclc.org/americalatina/pt/default.htm>>. Acesso em: 21 jan. 2011

uma biblioteca, composto de fichas catalográficas, era praticamente impossível usar operações booleanas como “*and*”, “*or*” ou “*not*” – Lancaster (1993, p. 42) observa que as entradas lineares desses sistemas não permitiam as buscas de informações pela combinação de palavras –, na busca em um base de dados no computador já havia como recuperar materiais que mencionassem dois assuntos em particular, sem precisar ler todos os documentos que resultavam na busca.

Lancaster (1979, p. 67 *apud* SANTIAGO, 2004, p. 30) enumerou uma série de vantagens trazidas pelo sistema de recuperação da informação em computador:

1. Possibilidade de realizar diversas buscas ao mesmo tempo;
2. Habilidade em prover muitos pontos de acesso a um documento, de maneira extremamente econômica;
3. Habilidade em lidar com buscas complexas envolvendo um número grande de termos e suas complexas relações;
4. Habilidades em gerar uma saída (*output*) na forma de bibliografia impressa;
5. Habilidade em coletar, de forma sistemática, dados de gerenciamento sobre funcionamento do sistema;
6. Possibilidade de duplicar a base de dados de forma simples e barata, com o objetivo de ser utilizada na provisão de serviços de informação por um número diferente de centros. (LANCASTER, 1979, p. 67 *apud* SANTIAGO, 2004, p. 30)

A trajetória da interação entre novas tecnologias e recuperação da informação registrou um importante episódio em 29 de outubro de 1969. Em uma experiência pioneira, cientistas conseguiram criar uma rede de transmissão entre quatro computadores instalados em diferentes pontos dos Estados Unidos – na Universidade da Califórnia Los Angeles (UCLA), na Universidade da Califórnia em Santa Bárbara, na Universidade de Utah e no Instituto de Pesquisa de Stanford. O feito foi o primeiro passo para o desenvolvimento da Internet e, conseqüentemente, da web.

Em janeiro de 1975, surgia o primeiro computador pessoal (*personal computer*), o *Altair 8800*, mencionado em artigo do periódico *Popular Electronics*. Em meio a todos esses avanços, cresciam também as bases de dados e o desempenho dos sistemas de buscas. Santiago (2004, p. 31) aponta como fatores responsáveis por esse crescimento o papel preponderante dos produtores das bases de dados ao estabelecerem redes ou outras atividades de cooperação; o surgimento do *Scientific Information Dissemination Center* (SIDC) e do Centro de Serviço *On-line*, que funcionavam como serviços intermediários entre o produtor de bases de dados e o usuário final; e o reconhecimento gradual de que qualquer cientista ou outro profissional podia acessar qualquer base de dados que necessitasse, na hora que lhe

fosse mais conveniente. “No entanto o fator mais importante foi o aparecimento de habilidades de busca *on-line*, que tornaram as bases de dados amplamente acessíveis, representando verdadeira revolução na provisão de sistemas de informação.” (LANCASTER, 1979 *apud* SANTIAGO, 2004, p. 31)

Lancaster afirma que diferentes dos sistemas *off-line*, os sistemas de informação *on-line* permitiram a navegação (*browsing*), além de serem capazes de fornecer respostas rápidas. Nesses sistemas, o usuário passou a realizar a busca por conta própria, sem a intermediação de um especialista de informação. Santiago (2004, p. 31), baseando-se em Lancaster, aponta outros dois benefícios dos sistemas *on-line*: processamento multiusuário e operação em tempo real. “Essa indústria da informação tem suas raízes diretamente relacionadas aos trabalhos de recuperação da informação nos anos 1950 e 1960, que culminaram com a emergência de serviços *on-line* nos anos 1970 e com a viabilização internacional da indústria da informação nos anos 1980”, afirma Saracevic (1996, p. 45).

Sobre recuperação da informação e Ciência da Informação, Saracevic escreve também:

Resumindo, o trabalho com a recuperação da informação foi responsável pelo desenvolvimento de inúmeras aplicações bem-sucedidas (produtos, sistemas, redes, serviços). Mas, também, foi o responsável por duas outras coisas: primeiro, pelo desenvolvimento da Ciência da Informação como um campo onde se interpenetram os componentes científicos e profissionais. Certamente, a recuperação da informação não foi a única responsável pelo desenvolvimento da Ciência da Informação, mas pode ser considerada como principal; ao longo do tempo, a Ciência da Informação ultrapassou a recuperação da informação, mas os problemas principais tiveram sua origem aí e ainda constituem seu núcleo. Segundo, a recuperação da informação influenciou a emergência, a forma e a evolução da indústria informacional. Novamente, a recuperação da informação não foi o único fator, mas o principal. Como a Ciência da Informação, a indústria da informação atualmente não é apenas recuperação da informação, mas este é o seu componente mais importante. (SARACEVIC, 1996, p. 45).

A criação da web em 1990 e a disseminação do uso de mecanismos de busca por usuários finais incentivaram estudos na área de recuperação da informação no ciberespaço.

3.1.2 Recuperação da informação na web

Para tratar sobre recuperação da informação na web faz-se necessário retomar dois assuntos ligados à criação da web abordados anteriormente: o lançamento do satélite Sputnik ao espaço pela União Soviética em 4 de outubro de 1957 e a primeira transmissão em rede realizada em 29 de outubro de 1969. A primeira ação despertou a insegurança dos Estados Unidos frente à União Soviética, o que levou à criação do Departamento de Defesa (DoD) e

da *Advanced Research Projects Agency* (ARPA) – esta última, a responsável pelo desenvolvimento de novas tecnologias para uso militar.

O desafio dos cientistas da equipe da ARPA no final da década de 1960 era criar um sistema de comunicação que interligasse os computadores dos principais centros da agência. A partir das pesquisas dos cientistas da ARPA, nasceu a Arpanet, rede de computadores que continuaria funcionando mesmo se algum dos computadores em alguns dos centros da ARPA sofresse um ataque nuclear, maior preocupação em tempos de Guerra Fria. O primeiro teste da Arpanet, em outubro de 1969, envolveu quatro computadores instalados em lugares distintos nos Estados Unidos, como relatado anteriormente.

O êxito da operação inspirou novas experiências. Em 1973, ocorreu a primeira transmissão internacional, com a participação da University College de Londres e o Royal Radar Establishment da Noruega. Além dos departamentos de segurança, universidades passaram a fazer parte da nova rede. “A tecnologia desenvolvida para a Arpanet foi colocada à disposição das universidades e dos centros de pesquisa e formou o embrião da Internet.” (LÉVY, 2000, p. 251)

O crescimento levou à divisão da Arpanet em dois grupos, um formado pelas instituições militares, *Milnet*, e outro, pelas instituições não militares, espécie de nova Arpanet. O segundo grupo logo se transformou em um ambiente livre utilizado por professores e alunos em universidades, assim como por pessoas relacionadas a esse grupo.

O crescimento do uso da Internet nas duas décadas seguintes levou a pesquisas que culminaram com a criação da *World Wide Web* em 1990 pelo cientista Tim Berners-Lee do CERN (Organização Europeia para a Investigação Nuclear). A web⁷ é formada por um conjunto de unidades de informação denominadas “páginas”. Uma página é um arquivo de computador cuja dimensão (quantidade de caracteres) pode variar entre o tamanho de uma página de um livro e o tamanho de um livro inteiro. Ferneda (2003, p. 103) enumera as seguintes características contidas em páginas da web:

1. Esquema de endereçamento chamado *Universal Resource Locator* (URL);
2. Protocolo, o *Hypertext Transfer Protocol* (http), que permite que um programa no computador do usuário requisiute uma página (através da URL) ao computador onde a página está localizada (servidor ou host). O servidor responde à requisição enviando uma cópia da página ao computador do usuário;

⁷ No Brasil, a *World Wide Web* é mais comumente chamada de Internet, mesmo que a definição de Internet seja distinta de web. De acordo com Spink e Jansen (2004, p. 3), o termo web é usado para descrever todas as fontes de informação com base hipermídia disponíveis na Internet. Já o termo Internet engloba os mais diversos serviços: web, e-mail, FTP etc.

3. Padrão para a especificação da estrutura da página, *Hypertext Markup Language* (HTML), uma linguagem de marcação que permite definir diferentes componentes em uma página web. (FERNEDA, 2003, p. 103)

No campo da recuperação da informação, tão logo a web se consolidou, surgiram sistemas de busca de arquivos e sistemas de busca de informação utilizando menus e diretórios. Os sistemas *on-line* – assim como seus antecessores, os sistemas *off-line* – foram inspirados nos modelos de sistemas de recuperação da informação tradicionais.

Na primeira década da web sobressaíam-se os mecanismos de busca formados por diretórios, como o *Yahoo*. Vizine-Goetz (2002, p. 2) observa que esses mecanismos davam aos usuários a possibilidade de utilizar palavras e frases para recuperar itens de interesse. A autora aponta como bem-sucedidos os diretórios de pesquisa, ou seja, as listas de assuntos organizadas em categorias. O modelo para a criação das listas lembrava a clássica CDU. Nas palavras de Vizine-Goetz (2002), *Yahoo!* e *Infoseek*, dois dos mecanismos de busca mais utilizados na época, forneciam aos usuários a capacidade de navegar através de uma série de categorias de assunto e de descobrir documentos potencialmente relevantes.

No caso do *Yahoo*, a indexação era realizada manualmente, como nas bibliotecas. A eficiência do mecanismo, porém, como descreve Ferneda (2003, p. 107), dependia em grande parte de voluntários para obter URLs para sua base de dados. Diferente do *Yahoo*, o *Altavista* e, na virada do milênio, o *Google* indexavam o conteúdo automaticamente. A partir do conteúdo indexado, “como resultado da busca, o mecanismo apresentava uma lista ordenada de endereços de páginas (URLs) que atendiam à expressão de busca.” (FERNEDA, 2003, p. 112)

Baeza-Yates e Ribeiro-Neto (1999, p. 1) destacam que recuperação da informação na web lida com representação, armazenamento, organização e acesso à informação. A representação e a organização de itens de informação deveriam prover ao usuário o acesso facilitado à informação em que está interessado. Infelizmente, como salientam, caracterizar necessidade de informação é bastante complicado. Os autores citam um exemplo de necessidade de informação de um usuário qualquer:

Descobrir todas as páginas (documentos) contendo informação de times universitários de tênis, os quais devem 1) ser mantidos por universidades dos Estados Unidos e 2) participar do torneio de tênis NCAA. Para ser considerada relevante, a página recuperada deve incluir informação sobre o ranking nacional do time nos últimos três anos e conter o e-mail ou telefone do treinador. (BAEZA-YATES; RIBEIRO-NETO, 1999, p. 6)

A dificuldade começa logo no primeiro passo: Como transformar todos esses conceitos em linguagem formal adequada ao tipo de consulta melhor compreendido pelo

sistema de recuperação da informação em questão? A resposta padrão seria: tomando os conceitos que compõem a necessidade informacional para transformá-los no conjunto de palavras organizadas que a exprimem melhor. A esse conjunto de palavras dá-se o nome de consulta (no inglês, *query*), a qual pode ser inserida em um mecanismo de busca. O procedimento, porém, nem sempre se revela tão simples quando a orientação de como colocá-lo em prática.

Diante do tamanho da coleção de documentos reunida no ciberespaço – massiva em tamanho e variada em conteúdo, formato, propósito e qualidade –, começou-se a questionar, como aponta Yang (2005, p. 33), a validade das descobertas empregadas nos tradicionais sistemas de recuperação da informação e que foram transferidas para os sistemas na web. Falta salientar que ao serem criadas, essas maneiras de uso foram baseadas em coleções relativamente pequenas e homogêneas, como as de bibliotecas.

Outro aspecto observado por Yang – ressaltando que buscar informação na web é algo diverso em caráter e imprevisível por natureza – diz respeito especialmente à diversidade de usuários, com seus diferentes *backgrounds* e motivados pelos mais variados tipos de necessidades informacionais – semelhante ao que já ressaltava Ranganathan quase três décadas antes. Não era mais apenas profissionais habilitados que lidavam com os sistemas. “A variedade de experiências, conhecimento, motivação e propósitos significa que os usuários podem expressar diferentes tipos de necessidades de informação por meio de grande heterogeneidade de formas, com diferentes critérios, para satisfazer suas necessidades.” (YANG, 2005, p. 33)

Sobre os “novos” usuários de sistemas de recuperação, Manning (2009, p. 432) observou que na tradicional recuperação da informação usuários eram tipicamente profissionais com treinamento sobre como criar consultas para uma coleção cujo estilo e estrutura eles conheciam bem; já os usuários da web tendem a não saber – ou se importar – com a heterogeneidade do conteúdo e a sintaxe das linguagens de criação de consulta. “Ademais, nem é papel de uma plataforma como a web exigir esse tipo de conhecimento de bilhões de pessoas.” (MANNING, 2009, p. 432)

Com o aumento do uso da web cresceram o interesse e os estudos sobre seu uso. Yang (2005, p. 36) classifica os estudos sobre ambiente de busca na web em dois tipos: 1) aquele que tenta caracterizar a web a partir do conteúdo, analisando exemplos de conjuntos de documentos obtidos por um “rastejador da internet” (*web crawler*), como, por exemplo, os

realizados por Lawrence e Giles; e 2) aquele que tenta caracterizar o comportamento do usuário a partir de pesquisas realizadas diretamente com esses usuários ou analisando dados de mecanismos de buscas, como, por exemplo, os realizados por Jansen, Spink e Saracevic.

O grande problema de pesquisas relacionadas à web, inclusive no campo da recuperação da informação, como apontam Jansen, Spink e Saracevic (2000), é que tão logo são concluídas, já se tornam obsoletas, pois a web passa por um crescimento desenfreado desde que surgiu. “A web tem crescido e passado por mudanças desde a criação, sendo assim, descobertas apontadas nos primeiros estudos são ultrapassadas e pouco representativas para o atual ambiente de busca”, avalia também Yang (2005, p. 36) Apesar do ponto de vista um tanto negativo dos autores, curiosamente de pesquisadores que continuam a pesquisar sobre o ciberespaço, prefere-se aqui considerar que os estudos servem para mostrar a evolução e tornam-se um registro importante de um determinado momento da web.

Os primeiros estudos realizados ainda nos anos 1990 tomavam como base um pequeno número de exemplos e tinham escopo limitado, mas foram se expandindo ao longo da década seguinte. Em estudo realizado já em 2000, aplicado a número considerável de integrantes, Jansen, Spink e Saracevic analisaram 51.473 consultas formuladas por 18.113 usuários do *Excite*. A obsolescência de um estudo pode ocorrer pela desatualização dos dados, como também pela mudança de status de determinado objeto de estudo. Na época da pesquisa de Jansen, Spink e Saracevic, o *Excite* era um dos maiores mecanismos de busca dos Estados Unidos. Atualmente, o *Excite* ainda existe, mas perdeu consideravelmente o espaço. De qualquer forma, ao ser realizado, o estudo contribuiu como registro histórico para a área da recuperação da informação.

A pesquisa de Jansen, Spink e Saracevic resultou em três tipos diferentes de dados, que se relacionavam a: a) sessão de busca⁸ (mudanças nas consultas durante a sessão de busca, número visto de páginas e uso de retroalimentação por relevância – ‘*relevance feedback*’⁹); b) consultas¹⁰ (número de termos inseridos no campo de busca, uso de

⁸ Uma sessão é a série inteira de consultas utilizadas pelo usuário em determinado tempo. Uma sessão pode ser do tamanho de uma consulta, assim como conter várias consultas.

⁹ “*Relevance Feedback*”: “asking information seekers to make relevance judgments about returned objects and then executing a revised query based on those judgments”. (MARCHIONINI, 2006, p. 44) Também chamado de “*query modification*” (BELKIN, 1993, p. 2) / Tradução: Retroalimentação por relevância: pedir a buscadores de informação para fazerem julgamentos de relevância sobre objetos retornados e então executar uma revisão baseada nesses julgamentos.

¹⁰ Uma consulta consiste de um ou mais termos e possivelmente inclui operadores lógicos e modificadores.

operadores lógicos e modificadores, como aspas e asterisco); e c) termos¹¹ de distribuição (renque/frequência) e os termos mais usados. (JANSEN; SPINK; SARACEVIC, 2000, p. 207)

Entre os resultados apurados, Jansen, Spink e Saracevic constataram o desconforto dos usuários com os operadores booleanos, além de o costume de não passar da primeira página de resultados. A conclusão a que os pesquisadores chegaram nesta pesquisa de 2000 é que os conceitos de recuperação da informação na web precisavam passar por uma premente atualização.

Em 2001, estudo semelhante foi realizado. Curioso observar que já nessa época, Spink et al. formalizaram a constatação de que a web havia mudado rapidamente em seus primeiros seis anos de existência. Passada uma década dessa constatação, verifica-se que mudanças ainda mais significativas ocorreram, como a transformação da web estática para uma web mais dinâmica, com intensa participação dos usuários, disseminação de redes sociais e grandes avanços na abrangência dos sites de busca.

Assim como constataram as mudanças rápidas na web, Spink et al. observaram que o comportamento dos usuários não apresentou alterações significativas. As consultas continuaram curtas, pouco modificadas e com estrutura muito simples.

Poucas consultas incorporam recursos avançados de busca, e quando incluem, apresentam vários erros. Apesar de receberem longa lista de resultados, usuários olham apenas algumas poucas páginas, não vão além da primeira ou da segunda página de resultados. Também não estão interessados em retroalimentação por relevância. Em geral, um pequeno número de termos é usado com frequência muito alta, embora haja grande número de termos usados uma única vez. (SPINK et al., 2001, p. 234)

Apesar do crescimento e das transformações pelos quais passou a web, os estudos demonstram que o comportamento do usuário atual sofreu poucas alterações em relação ao comportamento observado uma década atrás. E mais, semelhante à época em que indivíduos buscavam as respostas para suas demandas informacionais nas bibliotecas mais próximas de suas casas, com a web nota-se que ainda perdura a lei do menor esforço.

Usuários utilizam o buscador mais popular ao invés de procurar bases mais apropriadas ao tipo de busca que será efetuada. As consultas demonstram falta de elaboração ou de organização lógica, são criadas como se a linguagem dos sistemas de recuperação da informação fosse pura linguagem natural, ou seja, sem o uso de recursos que ajudariam a

¹¹ Um termo é um conjunto de caracteres, sejam letras, números ou símbolos. Termos são palavras, abreviações, números, URLs e outras combinações.

refinar a busca: operadores booleanos ou modificadores, como aspas. Também parecem não se preocupar com a ordem em que as palavras serão inseridas no campo de busca.

Ao longo da última década, de forma natural, buscar na web passou a fazer parte do cotidiano dos indivíduos. Como observa Marchionini (2006, p. 41), “desde que computadores tornaram-se produtos acessíveis aos consumidores finais e a Internet tornou-se mídia de massa, realizar buscas na web transformou-se em atividade diária para todos, das crianças aos cientistas”. O autor destaca, no entanto, a necessidade de definir a hierarquia dos diferentes tipos de necessidades informacionais, desde aquelas relacionadas a fatos básicos que guiam ações de curto prazo – como previsões do tempo que determinarão se é necessário levar ou não uma sombrinha na bolsa – até aquelas sobre complexas redes de conhecimento tácito e implícito, que se transformam em expertise ao longo da vida.

Para estabelecer as diferenças entre os tipos de necessidades informacionais, Marchionini (2006, p. 41) desenvolveu um esquema contendo três atividades de busca nomeadas: “busca-breve” (*lookup*), aprender (*learn*) e investigar (*investigate*). Como se vê na Figura 3, as três categorias são representadas por nuvens que em alguns momentos se sobrepõem. A representação dá a entender que as pessoas podem engajar-se em mais de um tipo de busca paralelamente e um tipo pode estar incluído em outro em alguns casos.

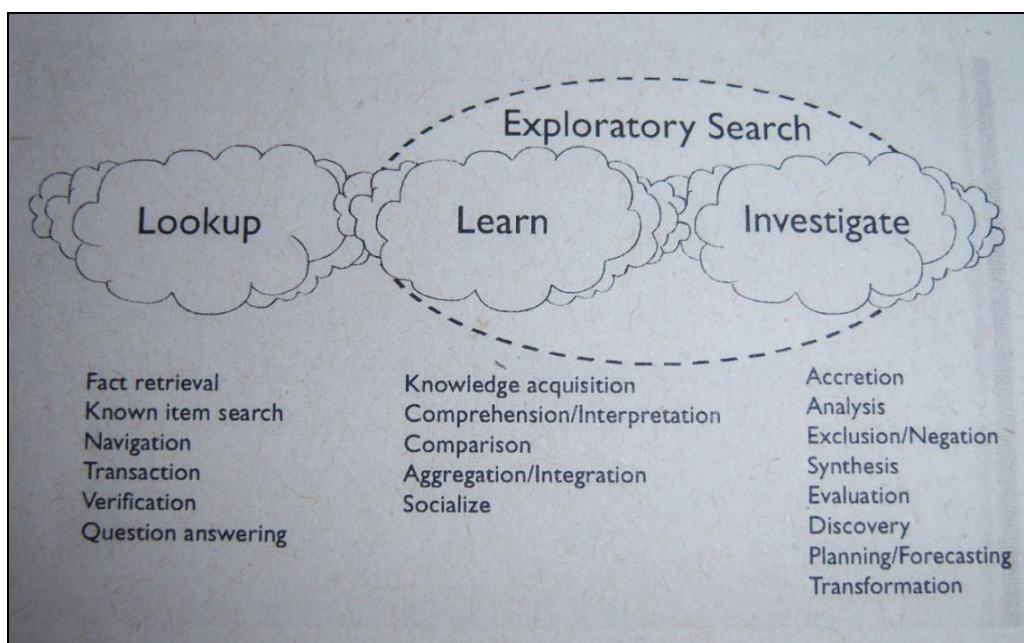


Figura 3 – Atividades de busca (MARCHIONINI, 2006, p. 42)

Observa-se também que aprender e investigar estão inseridos em uma categoria maior, a busca exploratória (*exploratory search*). Ao tratar sobre esse aspecto, Marchionini (2006, p. 42) relembra que buscar é atividade fundamental na vida de qualquer ser.

Todos os organismos procuram alimento e, como prevê a teoria de hierarquia de necessidades de Maslow, uma vez que o indivíduo tenha satisfeito suas necessidades fisiológicas básicas, ele começa a buscar a satisfação social e psicológica. Essas necessidades de nível elevado são normalmente informacionais e isso explica por que fontes de informação e facilidades de comunicação são tão sofisticadas em sociedades desenvolvidas. (MARCHIONINI, 2006, P. 42)

Ao abordar a busca-breve, Marchionini refere-se à busca do dia-a-dia. Como descreve, busca-breve é o tipo mais básico de busca, e serve de inspiração para o desenvolvimento de sistemas de gestão de bancos de dados e apoio para os mecanismos de busca da web. O que se recupera neste tipo de busca são números, nomes, frases curtas, arquivos específicos de texto ou outros suportes. Esse tipo de busca está entre as mais bem-sucedidas aplicações desenvolvidas para computadores e continua sendo tema de estudo e pesquisa. Porém, “assim que a web se tornou a primeira opção de fonte de informação para quem tem uma necessidade informacional, as pessoas, de modo geral, esperavam atender também a outros tipos de demandas, ir além da busca-breve”, destaca o autor.

Chega-se então ao segundo estágio, aprender. Com o crescente número de fontes primárias disponíveis *on-line*, é possível recuperar conjuntos de objetos que levam ao aprendizado, mas requerem processo cognitivo e interpretação. Os resultados apresentados neste tipo de busca determinam que o indivíduo gaste um tempo escaneando o material obtido, comparando e fazendo julgamentos qualitativos. Marchionini (2006, p. 42) chama atenção para o significado de aprender – neste caso, relacionado ao desenvolvimento de novo conhecimento. Faz parte do processo, reformular a busca o quanto for necessário até chegar à resposta mais adequada e que satisfaça a necessidade informacional.

Por fim, o autor trata sobre o terceiro estágio, investigar, no qual as buscas dão suporte a indagações envolvendo múltiplas interações durante longos períodos de tempo. Os resultados obtidos são avaliados criticamente antes de serem incorporados, tanto pessoal quanto profissionalmente. A modalidade envolve análise, síntese, avaliação e conhecimento pré-existente.

Marchionini (2006, p. 43) destaca a importância que pesquisadores da área de recuperação da informação vêm dando para a área de estudo de interação humano-computador. Quem sabe o que está buscando e que tipo de informação atende a sua demanda é o indivíduo. Por isso, tem-se estimulado a maior interação do usuário com o sistema. “Sob

esse ponto de vista, o desenho dos sistemas de busca precisa colocar as pessoas diretamente no processo de busca através de interfaces interativas que promovam continuamente o engajamento da pessoa com o processo de busca da informação.” Para esse indivíduo conseguir lidar com a situação, espera-se dela determinado grau de letramento informacional, como será visto agora.

4. Comportamento e competência informacional

Precisar o volume exato de informações disponíveis na web é tarefa praticamente impossível, mas em páginas da própria rede podem ser encontradas tentativas de medi-la. Porém, a disparidade entre os resultados e a forma como os métodos são utilizados para gerá-los os tornam difíceis de serem interpretados. A título de exemplo, foram tomadas duas pesquisas sobre o assunto. Em 2006, inspirado pela questão *Qual o tamanho da World Wide Web neste momento?*, o então mestrando em Business Communication e Mídia Digital pela universidade holandesa de Tilburg Maurice de Kunder desenvolveu um método¹² para medir a web indexada – em janeiro de 2011, segundo seus cálculos, o tamanho da web registrado era de pelo menos 13,58 bilhões de páginas¹³.

Em 2008, os engenheiros de software do *Google* Jesse Alpert e Nissan Hajaj divulgaram que os sistemas de busca da empresa haviam detectado um trilhão de URLs (localizador padrão de recursos) únicas. No mesmo texto, entretanto, eles mesmos admitem ser impossível definir o tamanho da web. A dupla de pesquisadores realizou a pesquisa acessando páginas iniciais com grande número de conexões e seguindo os links para novas páginas. Decidiram parar a pesquisa quando chegaram a um trilhão de páginas, pois ela nunca chegaria ao fim. Mesmo com a impossibilidade de precisão dos dados, Alpert e Hajaj acrescentaram ainda que, por meio de outras pesquisas, detectaram que em 1998 o *Google* tinha 26 milhões de páginas indexadas, e em 2000, um bilhão.

¹² O método de Maurice de Kunder baseou-se na frequência de palavras em diferentes coleções de documentos. Se uma palavra ocorre em 30 artigos em uma coleção de 30 mil artigos de jornais, então a frequência de documentos estimada daquela palavra é 1 em 1000. Quando o *Google* reporta que tem 9 milhões de webpages contendo aquela determinada palavra, então é possível estimar que o número de páginas indexadas no *Google* seja de 9 bilhões. O método foi usado em quatro mecanismos de busca: *Google*, *Yahoo Search*, *MSN* e *Ask*. Como esses mecanismos indexam muitas das mesmas webpages, isso significa que os valores encontrados não podem ser simplesmente somados. Para evitar esse erro, o pesquisador realizou um teste para estimar as sobreposições entre os quatro mecanismos.

¹³ Dado de 3 de janeiro de 2011, apresentado no site *WorldWideWebSize*. Disponível em: <<http://www.worldwidewebsize.com>>. Acesso: 3 jan. 2011

Mesmo diante de números tão distintos, da grande variação de uma pesquisa para outra, e, no caso do *Google*, de dados relativos apenas às páginas indexadas em um único mecanismo de busca, tratam-se de números grandes. Embora não existam dados oficiais, não é errada a afirmação de que na web está concentrado o maior volume de fontes de consulta disponíveis hoje no mundo.

Outro aspecto relevante relacionado às possibilidades de recuperação da informação propiciadas pela web diz respeito aos avanços dos mecanismos de busca. Se há pouco mais de uma década, o modelo existente compreendia diretórios de pesquisa, compostos por listas de assuntos organizadas manualmente em categorias (VIZINE-GOETZ, 2002, p. 3), na virada do milênio, passou a ser utilizada a busca por palavras-chave, como a oferecida pelo *Google*. Os resultados apresentados baseiam-se em rankings de relevância. Nos anos 2000, o mecanismo de busca popularizou-se a ponto de o vocábulo *Google*¹⁴ ganhar definição no dicionário *Merriam-Webster* e tornar-se inclusive um verbo, denotando assim o processo da própria busca, e não somente seu resultado, como na acepção original do termo.

Essa disponibilidade ilimitada de informações associada à capacidade cada vez mais alta de recuperação dos mecanismos de busca possibilita até mesmo imaginar que este seria o cenário perfeito para atender às mais variadas necessidades de informação dos indivíduos. Paradoxalmente, porém, como aponta Dudziak (2003, p. 23), justamente esses dois fatores contribuem para o aparecimento de barreiras ao acesso. Ao realizar uma busca sem conhecer ou sem utilizar de maneira apropriada os mecanismos de filtragem disponíveis, o indivíduo depara-se com um número tão volumoso de fontes, que se torna impossível analisar os resultados de forma rápida e fácil como se imaginava e se propunha poder.

4.1 Letramento informacional

É dentro do cenário relatado anteriormente que entra em evidência o letramento informacional (*information literacy*), ou competência em informação, denominação mais frequentemente adotada pelos autores de língua portuguesa, área de estudos que trata das habilidades fundamentais para que o indivíduo obtenha sucesso na sociedade da informação e das habilidades aplicáveis a todas as situações de resolução de um problema ligado à

¹⁴ *Google* – verb transitiv: to use the Google search engine to obtain information about (as a person) on the World Wide Web. Tradução da autora: *Google* – verbo transitivo: usar o motor de busca Google para obter uma informação na Web. Disponível em: <<http://www.merriam-Webster.com/dictionary/googling>>. Acesso em 19 mai. 2010

necessidade de informação, como define Hatschbach (2002, p. 14). No estudo sobre o tema, a autora também o descreve como a área de estudos e de programas educacionais que trata especificamente das questões relacionadas à capacitação do indivíduo para a resolução de problemas de informação, seja no âmbito escolar, profissional ou social.

Tentativas de tradução da expressão *information literacy* para o português já foram realizadas, mas autoras como Hatschbach e Dudziak seguem usando o termo em inglês. Nos trabalhos de Campello (2009), a autora utiliza a expressão *letramento informacional*, o mesmo adotado nesta dissertação. Outras variações encontradas são: *alfabetização informacional*, *literacia*, *fluência informacional*, *competência em informação* e *competência informacional*, de acordo com levantamento realizado por Dudziak (2002, p. 24).

4.1.1 Surgimento da expressão *information literacy*

Como mostram as revisões de literatura realizadas pelas pesquisadoras Hatschbach (2002), Dudziak (2003) e Campello (2003), a expressão *letramento informacional* foi registrada pela primeira vez em 1974 no relatório *The information service environment relationships and priorities* do bibliotecário Paul Zurkowski. O autor destaca que indivíduos treinados na aplicação de fontes de informação em seu trabalho podem ser chamados letrados informacionalmente (*information literates*), ou seja, são aqueles que aprenderam técnicas e habilidades para lidar com um grande número de ferramentas informacionais, bem como com fontes primárias, para encontrar informação que contribuísse na solução de seus problemas (ZURKOWSKI, 1974, p. 6). Mais adiante, entre os resultados da pesquisa de campo, será possível confirmar essa afirmação ao observar como bibliotecários aperfeiçoam suas habilidades de busca e recuperação na web.

Dudziak (2003, p. 24) cita outros dois momentos importantes na trajetória do *letramento informacional* ainda na década de 1970. O primeiro em 1976, quando Hamelink e Owens anteveem *letramento informacional* como instrumento de emancipação política, relacionado à cidadania, elevando o termo a um novo patamar, além da simples aquisição de habilidades e conhecimentos ligados à informação. Como destaca Campello (2003, p. 30), segundo Hamelink e Owens, cidadãos competentes no uso da informação teriam melhores condições de tomar decisões relativas a sua responsabilidade social. O segundo em 1979, quando é retomada a ênfase ao domínio de técnicas e habilidades de uso de ferramentas

informacionais na modelagem de soluções para problemas, requisito para a competência informacional. Ideias semelhantes seriam amplamente exploradas na década seguinte.

4.1.2 Definições de letramento informacional

Uma das definições amplamente usadas nos estudos sobre letramento informacional foi cunhada em 1989 pela *American Library Association* (ALA), como destaca Hatschbach (2002, p. 17). Duas décadas mais tarde, a definição continua sendo a definição oficial de letramento informacional da ALA, como se lê no site da associação¹⁵. “Letramento informacional é o conjunto de habilidades que o indivíduo deve ter para reconhecer quando uma informação é necessária e também que deve ter para localizar, avaliar e usar de forma efetiva a informação requerida” (ALA, 1989). Bem aceita, a definição desencadeou programas para impulsionar a disseminação do conceito, especialmente nos Estados Unidos e na Austrália.

Envolvidos neste processo estavam bibliotecários, os quais tiveram seu papel alterado por mudanças na sociedade e na educação – e recentemente com o surgimento de novas tecnologias –, como estuda Campello (2009, p. 68) em sua tese sob a perspectiva da biblioteca escolar. “Historicamente, os bibliotecários em escolas atuavam na promoção da leitura (...), com as mudanças na sociedade e na educação, nova função emerge, já que surge a demanda para a formação de pessoas competentes para lidar com a informação”. Campello ressalta que ações semelhantes têm sido sustentadas em pesquisas tanto nas áreas de Biblioteconomia e Ciência da Informação quanto da Educação. Com isso, fortaleceu-se o conceito de letramento informacional.

Sobre o comportamento do indivíduo considerado informacionalmente letrado, Doyle elaborou em 1994 a lista a seguir. O indivíduo letrado deve:

- Reconhecer que informação precisa e correta é a base para tomada de decisão inteligente;
- Reconhecer a necessidade de informação;
- Formular questões baseadas em necessidades de informação;
- Identificar fontes potenciais de informação;
- Desenvolver estratégias de pesquisa bem-sucedidas;

¹⁵ Definição de letramento informacional exposta no site da American Library Association. Disponível em: <<http://www.ala.org/ala/mgrps/divs/acrl/standards/informationliteracycompetency.cfm#ildef>> Acesso em: 3 jan. 2011

- Saber acessar diversas fontes de informação, incluindo o computador e outras tecnologias;
- Avaliar a informação;
- Organizar a informação para aplicação prática;
- Integrar informações novas a conhecimentos já adquiridos;
- Utilizar a informação de forma crítica e para a resolução de problemas. (DOYLE, 1994 *apud* Hatschbach, 2002, p. 20)

Bruce (1997 *apud* CAMPELLO, 2009, p. 76) focou-se no letramento informacional como fenômeno experimentado por pessoas que interagem com o universo informacional, ou seja, que usam informações de forma competente. A partir disso, o autor identifica sete experiências presentes nas diferentes relações entre usuário e informação: a experiência da tecnologia da informação; a experiência das fontes de informação; a experiência do processo de informação; a experiência do controle da informação; a experiência do processo de construção do conhecimento; a experiência da extensão do conhecimento; e a experiência da sabedoria”. (BRUCE, 1997, p. 117-151 *apud* CAMPELLO, 2009, p. 76)

Ainda sobre pessoas informacionalmente letradas, nos anos 2000, Kong et al. (2005) apresentaram uma versão considerando a globalização e a cultura digital.

A pessoa informacionalmente letrada na onipresente sociedade da informação deveria saber como determinar a natureza e a extensão da informação de que necessita, como encontrar informação efetiva e eficientemente e como usar informação para compreender questões econômicas, legais e sociais relacionadas com o uso da informação. Letramento informacional é a habilidade de dominar os processos de ser informado e constitui capacidade essencial, necessária aos cidadãos para se adaptarem à cultura digital, à globalização e à emergente sociedade baseada no conhecimento. (KONG et al., 2005 *apud* CAMPELLO, 2009, p. 69)

A partir das considerações referenciadas anteriormente, Dudziak (2003, p. 29), apresenta os objetivos do letramento informacional na formação de indivíduos que:

- Saibam determinar a natureza e a extensão de sua necessidade de informação como suporte a um processo inteligente de decisão;
- Conheçam o mundo da informação e sejam capazes de identificar e manusear fontes potenciais de informação de forma efetiva e eficaz;
- Avaliem criticamente a informação segundo critérios de relevância, objetividade, pertinência, lógica e ética, incorporando as informações selecionadas ao seu próprio sistema de valores e conhecimentos;
- Usem e comuniquem a informação, com um propósito específico, individualmente ou como membro de um grupo, gerando novas informações e criando novas necessidades informacionais;
- Considerem as implicações de suas ações e dos conhecimentos gerados, observando aspectos éticos, políticos, sociais e econômicos extrapolando para a formação da inteligência;
- Sejam aprendizes independentes; e
- Aprendam ao longo da vida. (DUDZIAK, 2003, p. 29)

A respeito desse último objetivo relacionado, “aprender ao longo da vida”, a *Association of College Research Libraries* (ACRL) ressalta que o letramento informacional é componente-chave para o aprendizado sem período determinado. No relatório *Information Power*, da *American Association of School Librarians* (AASL), reeditado em 1998 junto com a *Association for Educational Communications and Technology*, destaca-se que a habilidade de usar informação seja o ponto-chave para a aprendizagem permanente.

Como destaca Campello (2009, p. 72), desenvolver nas pessoas a capacidade de aprender ao longo da vida é central para a missão das instituições de ensino superior. Especialmente quando surgem novas tecnologias, como o vivenciado nas últimas décadas com o desenvolvimento da Internet e o surgimento da web. Campello (2009, p. 78) ressalta que o universo informacional na sociedade contemporânea, caracterizado por grande quantidade de informação em suportes variados, exige do estudante o domínio de habilidades informacionais que irão capacitá-lo para aprender.

O crescimento do interesse e do volume de conteúdo sobre letramento informacional pode ser observado em simples busca na web. Em 1998, como mostra Dudziak, ao inserir a expressão no *Altavista* (www.altavista.com), então o mais popular mecanismo de busca, eram gerados 9.150 resultados. Em janeiro de 2011, utilizando-se o mesmo *Altavista*, chega-se a 25.600.000 resultados. Ao realizar a mesma busca, com a expressão no *Google* (www.google.com), o resultado é 7.490.000 páginas¹⁶. Nas duas buscas, foi utilizada apenas a expressão, sem o uso de aspas.

Como apresentado nesta revisão, na sociedade da informação em que se vive atualmente torna-se essencial a disseminação de estudos e práticas que colaborem para a elevação do nível de letramento informacional dos indivíduos. Em tempos de avanços tecnológicos, além de letramento informacional, cresce a necessidade de letramento digital (*digital literacy*) por estar o indivíduo cada vez mais envolvido com tecnologia seja na vida profissional, estudantil ou pessoal – por meio de computadores, iPads, celulares, televisores digitais, iPods e PDAs. Outra razão para atenção especial ao letramento digital é a questão da inclusão social. Pessoas pouco letradas nas novas tecnologias ficam à margem dos avanços que vêm ocorrendo.

¹⁶ Buscas realizadas em 4 de jan. 2011.

4.2 Letramento digital

O termo letramento digital ganhou espaço a partir dos 1990 devido, especialmente, ao surgimento da web. Letramento digital é a habilidade de entender a informação e – mais importante – de avaliar e integrar a informação em múltiplos formatos fornecidos por um computador conectado a uma rede – sem levar em conta o formato, seja computador de mesa, notebook, netbook, celulares, smart phones etc. Entender informação encontrada na web é impossível sem avaliar as fontes ou colocá-la dentro de um contexto, como afirma Gilster (1997, p. 1), autor de *Digital Literacy*, em entrevista concedida em 1997 à revista *Educational Leadership*.

Para o autor, lidar com a informação encontrada na web é diferente de lidar com informações encontradas em livros, e por algumas razões. Primeiro, na web nem tudo é texto, pode haver vídeos, fotos, áudios. Segundo, a maneira como indivíduos encontram informação é diferente da maneira como nós usávamos um cartão de referência em uma biblioteca. “Um computador multimídia conectado à internet permite que as pessoas realmente ‘construam’ informação de todos os lugares do mundo” (GILSTER *apud* POOL 1997, p. 1). Terceiro, ser digitalmente letrado significa ser multidimensional e interativo.

Dois anos antes, Lanham (1995 *apud* BAWDEN, 2001, p. 246) tratou o termo como sinônimo de letramento multimídia (*multimedia literacy*). O autor argumenta que *literacy per se*, em uma era digital como a que se vive então, significa habilidade para entender informação independente de como ela se apresente. Letramento digital envolve a habilidade de decifrar imagens, sons etc., assim como textos.

Em 2010, Gilster retoma a definição anterior com pequenas alterações.

Letramento digital envolve ser capaz de entender um problema em particular o suficiente para criar uma série de questionamentos que servirão de base para buscar mais informações sobre o problema ao consultar a web. Mais do que consultar um conjunto estático de fontes, agora se trabalha com uma biblioteca hipertextual que está sempre se expandindo. Uma pessoa letrada digitalmente aprende com criar uma provisão de informações que são sempre atualizadas quando surgem novas fontes. A natureza desse material demanda um letramento que está focado em hábitos fortes de pensamento crítico. (GILSTER, 2010)

Gilster aponta também as diferenças entre ser letrado digitalmente e ser um usuário constante da web. Para o autor, usar um navegador, trabalhar com correio eletrônico, criar página na web ou navegar pelo *Facebook* são procedimentos. Ser digitalmente letrado não exige apenas esses conhecimentos de como usar a web, o diferencial é a análise e os julgamentos que essas pessoas são capazes de fazer para diferenciar quais informações são

relevantes e confiáveis em meio a todo o material que se pode encontrar em uma busca ou navegando pela web. “Gosto de falar que uma pessoa digitalmente letrada é alguém que questiona as fontes dos materiais recuperados e sabe como rastrear essas fontes em busca de provas e de outras fontes que comprovem o valor da informação recuperada.” (GILSTER, 2010). O autor afirma ainda que essas habilidades não são procedimentos possíveis de serem aprendidos em uma tarde, mas são hábitos que demandam o desenvolvimento da capacidade de julgamento e de avaliação – Menou já ressaltava em 1995 a importância de aprender a fazer julgamentos.

De acordo com o autor, pensamento crítico é a principal competência que um estudante deveria desenvolver ao procurar informações. Na web, existem fontes com dados editados e não editados. Em muitos casos, como afirma Gilster, verificar a autoria e o contexto de uma informação não é fácil. Para isso, aprende-se a utilizar os mecanismos de busca, assim como as habilidades comunicacionais, como e-mail ou outros métodos para descobrir o contexto do material encontrado. Competências de busca envolvem desde aprender a refinar as consultas até como usar mecanismos de busca para descobrir outras opiniões ou comentários sobre o assunto recuperado. Outra competência essencial é chamada pelo autor de *knowledge assembly*, em uma tradução livre “ajuntamento de conhecimento”, ou seja, a capacidade de extrair informações de diferentes fontes e formatos, seja um texto, um vídeo ou um tópico em uma lista de discussão, conseguindo avaliar o valor e as armadilhas de cada um.

Antes da consolidação da expressão letramento digital, quando computadores começaram a surgir, o termo em voga era letramento computacional (*computer literacy*). Nos anos 1960 e 1970, de acordo com Leahy e Dolan (2009, p. 149), era considerada letrada computacionalmente a pessoa que soubesse como programar um computador. Normalmente, isso se aplicava àqueles que trabalhavam na área de informática. Em artigos científicos, a expressão começou a aparecer nos 1970, quando os computadores passaram a ser usados em empresas. Para determinadas funções era exigido conhecimento sobre computadores. Nessa época, havia ainda a clara separação entre técnicos e usuários comuns. Esse cenário modificou-se nos anos 1980, quando os computadores pessoais popularizaram-se e os usuários finais começaram a desenvolver seus textos, tabelas e pequenos bancos de dados com o auxílio de computadores. Nos anos 1990, surgiu o termo letramento digital, cuja definição assemelha-se em diversos aspectos ao de letramento informacional, apesar de o

segundo termo ser mais amplo. Relacionada às duas expressões, apresenta-se também letramento ciber (*cyber literacy*).

Mesmo com variações, as definições tanto de letramento informacional quanto de letramento digital têm em comum o fato de envolver raciocínio, pensamento crítico e habilidades que tornem o indivíduo capaz de discernir informações confiáveis e fidedignas de materiais que devem ser evitados por não haver comprovação de fonte confiável, autoria ou veracidade. Dentro desse contexto, aos profissionais da informação pode-se atribuir o papel de intermediários (WERSIG, 1993, p. 4), de responsáveis por ajudar as pessoas confusas com a situação do uso do conhecimento a encontrar resultados apropriados para suas demandas informacionais.

4.3 Comportamento informacional

Em recente entrevista sobre recuperação da informação, Saracevic (2010, p. 162) cita a expansão de estudos sobre o comportamento informacional dos indivíduos ao buscar respostas que atendam as suas necessidades informacionais. De acordo com o pesquisador, estudos nessa área expandiram-se para cobrir ampla gama de processos empregados pelas pessoas engajadas com informação e relacionar estados e efeitos cognitivos a sociais. Saracevic destaca que além da Ciência da Informação, outras áreas estão interessadas no assunto, como Psicologia, Ciências Cognitivas, Neurociência, Comunicação, Sociologia, Filosofia.

Usuário da informação, de acordo com Costa, Silva e Ramalho, pode ser definido como aquele indivíduo ou coletivo que necessitando de informação a utiliza. Guinchat e Menou (1994, p. 482 *apud* COSTA, SILVA e RAMALHO, 2009, p. 4) afirmam que o usuário é um agente essencial na concepção, avaliação, enriquecimento, adaptação, estímulo e funcionamento de qualquer sistema de informação. Os dois autores arriscam-se a dividir os usuários em três grupos principais: 1) os usuários que ainda não estão na vida ativa profissional ou estudantes; 2) os usuários engajados na vida ativa, cujas necessidades de informação se originam da sua vida profissional; e 3) o cidadão cujas necessidades de informação são gerais e ligadas a sua vida social.

Costa, Silva e Ramalho (2009, p. 5-6) descrevem o trabalho de Bernal e Urquhart, apresentado em 1948 na Conferência de Informação Científica da Royal Society, como as primeiras reflexões sobre estudos orientados aos usuários, minuciosamente enfocando a

maneira como os cientistas e técnicos procediam para obter informação ou como indivíduos usavam a literatura técnica em suas respectivas áreas de atuação. No século XX, esses estudos passaram por diferentes estágios:

1. Inicialmente, no final da década de 1940, os estudos de usuário tinham como objetivo tornar mais rápidos e aperfeiçoar serviços e produtos prestados pelas bibliotecas. Estes estudos restringiram-se à área de Ciências Exatas.
2. Na década de 1950, intensificaram-se os estudos sobre o uso da informação entre grupos específicos de usuários, abrangendo já as Ciências Sociais Aplicadas.
3. Só nos anos 1960 é que se enfatizou o comportamento dos usuários, surgindo estudos de fluxo da informação, canais formais e informais. Os tecnólogos e educadores começam a ser pesquisados
4. Já na década de 1970, a preocupação maior passou a ser o usuário e a satisfação de suas necessidades de informação, atendendo a outras áreas do conhecimento como Humanidades, Ciências Sociais e Administrativas. Datam dessa década os primeiros trabalhos na literatura especializada sobre o tema.
5. A partir de 80, os estudos estão voltados à avaliação de satisfação e desempenho. (COSTA, SILVA e RAMALHO, 2009, p. 6)

Pinheiro (1982, p. 5), em ampla revisão sobre o tema, afirma que estudos sobre usuários da informação são importantes para o conhecimento do fluxo de informação científica e técnica, de sua demanda, da satisfação do usuário, dos resultados ou efeitos da informação sobre o conhecimento, do aperfeiçoamento, relações e distribuição de recursos de sistemas de informação e tantos outros aspectos relacionados à informação. Por serem tão abrangentes, esses estudos acabam compreendendo desde levantamentos de empréstimos em bibliotecas até pesquisas sobre o comportamento dos usuários, como o realizado nesta dissertação.

Como o que os usuários querem, independente da área de atuação ou do grau de conhecimento, é recuperar o máximo de informações relevantes gastando o mínimo de tempo possível, estudos de usuários em Ciência da Informação poderiam muito bem ater-se a experiências relacionadas à transmissão da informação e, conseqüentemente, à modificação do conhecimento do receptor.

Wilson (1999 *apud* GASQUE e COSTA, 2010, p. 22) sugere inserir o tema estudos de usuários no campo do comportamento humano e denominá-lo comportamento informacional, sendo que esse tipo de ação refere-se às atividades de busca, uso e transferência de informação nas quais a pessoa se engaja quando identificada as próprias necessidades de informação. O autor propõe quatro definições de comportamento informacional:

1. Comportamento informacional: a totalidade do comportamento humano em relação ao uso de fontes e canais de informação, incluindo a busca da informação passiva ou ativa;
2. Comportamento de busca da informação: a atividade ou ação de buscar informação em consequência da necessidade de atingir um objetivo;

3. Comportamento de pesquisa de informação: o nível micro do comportamento, no qual o indivíduo interage com sistemas de informação de todos os tipos;
4. Comportamento do uso da informação: constitui o conjunto dos atos físicos e mentais e envolve a incorporação da nova informação aos conhecimentos prévios do indivíduo. (WILSON, 1999 *apud* GASQUE e COSTA, 2010, p. 22)

Os estudos realizados no âmbito do comportamento informacional remetem não apenas à Ciência da Informação, mas a interseção com as Ciências Cognitivas.

4.4 Interseção entre Ciência da Informação e Ciências Cognitivas

O comportamento dos usuários de informação passou a ser estudado com um enfoque cognitivista a partir do começo da década de 1970. De Mey (1977) apresentou em 1977, durante o *International Workshop on the Cognitive Viewpoint*, a relação entre Ciência da Informação e as Ciências Cognitivas afirmando que “o ponto de vista cognitivo da Ciência da Informação implica que cada ato de processamento da informação, seja perceptivo ou simbólico, é mediado por um sistema de categorias e conceitos, os quais, para o mecanismo de processamento da informação, constituem um modelo de mundo”. (MEY, 1982, p. 4 *apud* LIMA, 2003, p.81).

Para Lima (2003, p. 82), a informação está associada ao contexto e à maneira de cada indivíduo ver o mundo e consiste no somatório de diferentes estruturas do conhecimento. Assim, todo estágio cognitivo implica contexto, que é organizado pelo sistema conceitual da informação. Para Varela (2010, p. 54), a base do conhecimento é a atividade mental que permite a construção/desconstrução/reconstrução de esquemas que tecem redes de significação. “Este processo estabelece relações entre esquemas de conhecimento que já possuímos, em direção a um significado próprio para objetos de conhecimento objetivamente existentes.”

Para Brookes (1980, p. 131 *apud* CAMPOS; VENANCIO, 2007, p. 111), “[...] a absorção da informação em uma estrutura de conhecimento pode causar não apenas acréscimo, mas algum ajuste na estrutura, bem como mudança nas relações entre dois ou mais conceitos já admitidos”. Varela (2010, p. 59) afirma que o processo cognitivo envolve atividades mentais como o pensamento, a imaginação, a recordação, a solução de problemas, a percepção, o julgamento, a aprendizagem da linguagem, entre outras, as quais ocorrem diferentemente em cada indivíduo, dependendo do grau de habilidade de cada um. Seguindo esse raciocínio, ao realizar uma busca na web, o indivíduo protagoniza processos cognitivos em diferentes momentos, da elaboração da consulta à escolha das fontes relevantes, da tomada

de decisão sobre qual mecanismo utilizar até como a informação recuperada será aproveitada para atender à necessidade que desencadeou todo o processo.

Outra abordagem é a do modelo proposto por Belkin, Oddoy e Brooks em 1982, uma das primeiras pesquisas sobre busca da informação sob o ponto de vista do usuário. Os autores consideram a natureza do estado do conhecimento do usuário como fenômeno básico para a compreensão do processo de recuperação da informação. Assim, “o usuário é considerado um sujeito cognoscente que utiliza os sistemas de recuperação ao reconhecer anomalias ou estados de conhecimento inadequados ou incoerentes com relação a algum assunto ou problema, denominado estado anômalo do conhecimento (*Anomalous State of Knowledge - ASK*)” (CAMPOS; VENÂNCIO, 2007, p. 112).

É preciso considerar o período em que o modelo foi desenvolvido, ou seja, ainda antes da criação da web. De qualquer forma, princípios estabelecidos no modelo de Belkin, Oddoy e Brooks continuam pertinentes. Como, por exemplo, o fato de um sentimento de inadequação em seu estado de conhecimento levar indivíduos em situações problemáticas a buscar por informação. Já naquela época, 1982, os autores consideravam relevante a interação do usuário com o sistema, em um processo dinâmico. O usuário interage com o sistema de informação e avança no processo construtivo de busca de informação. Cada ciclo do processo de recuperação da informação depende da avaliação realizada pelo usuário nos ciclos anteriores, tendo em vista que a anomalia e a percepção do indivíduo com relação ao problema são modificadas a cada interação com o sistema.

E ainda, dentro do contexto de interseção entre Ciência da Informação e Ciências Cognitivas existe a teoria do *sense-making* de Dervin e Nilan, desenvolvida em 1983. De acordo com Varela (2010, p. 59), essa teoria entende a informação como um dado incompleto, ao qual o indivíduo atribui sentido a partir da intervenção de seus esquemas mentais interiores. A busca da informação é orientada por uma lacuna ou falha na estrutura do conhecimento do usuário, que demanda mudanças nas estratégias de busca de respostas pelo usuário para o problema, ou seja, a necessidade informacional que motivou a busca.

Varela cita também os estudos realizados por Kuhlthau (1991), os quais dizem respeito a padrões do processo de busca e uso de informação, explorando a visão construtivista do aprendizado, defendendo que a estrutura cognitiva se reestrutura consoante à aprendizagem significativa, considerando que o processo de apropriação de conhecimento é

dinâmico e que as estruturas de conhecimento de um indivíduo estão em constante transformação.

Pela abordagem da significação de Kuhlthau, o processo de busca de informação caracteriza-se como uma atividade construtivista do usuário para dar significado e atribuir um novo significado às informações, a fim de ampliar seus conhecimentos sobre determinado problema ou assunto. A base teórica respalda-se na teoria do construto pessoal, a qual representa a experiência dos indivíduos envolvidos no processo de construção do significado, a partir da informação encontrada. (VARELA, 2010, p. 60)

A autora argumenta ainda que a busca de informação fundamenta-se em uma série de estratégias, procedimentos e conhecimentos, baseados nos quais se procura obter com maior rapidez, economia de tempo e pertinência a informação desejada. A partir desse contexto, Varela defende a importância de estudos cognitivos no campo da Ciência da Informação que ajudem a levar a buscas realizadas com autonomia e garantia de desenvolvimento pessoal.

Tomando como referência as teorias e revisões apresentadas até aqui, entende-se que cada grupo de usuários – independente do sistema a ser utilizado – tem características particulares quanto ao tratamento dispensado à informação, desde a elaboração da consulta até as decisões tomadas diante dos resultados recuperados. Presume-se também que o contexto em que os usuários estão inseridos desempenha influência significativa na performance dos integrantes de determinado grupo, assim como sugere maior ou menor letramento informacional e digital e maior ou menor habilidade ao protagonizar um processo de busca e recuperação da informação.

No planejamento do presente trabalho, apresentou-se como possibilidade de aprofundamento dos assuntos abordados no referencial teórico a realização de um estudo exploratório no âmbito da Ciência da Informação, tomando como população-alvo justamente um grupo que tem a informação como objeto de estudo. A partir dessa escolha, pretendeu-se investigar como profissionais da área comportam-se ao vivenciar uma experiência de busca na web, sem perder de vista questões sobre letramento informacional e processos cognitivos.

5. Investigação do comportamento e do letramento informacional

A metodologia de uma pesquisa científica determina o roteiro de ações programadas e coordenadas que leva à produção de algum conhecimento novo. Como afirma Gonzalez de Gómez (2000, p. 1), a metodologia designa, de maneira ampliada, o início e a orientação de um movimento de pensamento, cujo esforço e intenção direcionam-se à produção de um novo conhecimento. Na presente pesquisa, determinados os objetivos, esse movimento iniciou-se com um levantamento bibliográfico sobre os temas relacionados ao processo de busca e recuperação da informação na web, que teve como fim proporcionar maior familiaridade da pesquisadora com os elementos e os modelos conceituais presentes neste campo de pesquisa e permitir-lhe relacioná-los aos elementos específicos do cenário em que foi desenvolvida a investigação.

Alinhado à natureza do campo em que está inserido, o presente estudo tem um pano de fundo interdisciplinar, representado especialmente pela interseção entre a Ciência da Informação, campo do conhecimento que tem como objeto de estudo as propriedades gerais dos processos e sistemas de construção, comunicação e uso da informação (LE COADIC, 2004, p. 25), e as Ciências Cognitivas, campo voltado ao estudo do que é o conhecimento, como ele pode ser representado e manipulado nas suas diversas formas (MEY, 1982 *apud* LIMA, 2003, p. 3).

No campo da Ciência da Informação, como apresentado no referencial teórico, destaca-se a revisão bibliográfica sobre a noção de informação, a própria Ciência da Informação, a recuperação da informação, a recuperação da informação na web, o letramento informacional, o letramento digital e os avanços tecnológicos. Dentro do campo das Ciências Cognitivas, o foco da revisão concentra-se nos estudos sobre o comportamento informacional dos usuários, abordagem que se firma como substituta dos estudos de usuários. Também são

revisadas noções e informações sobre operadores booleanos e outros aspectos importantes para o desenvolvimento da pesquisa de campo.

Combinando teoria e prática, e em uma aproximação com a realidade, esta pesquisa exploratória¹⁷ englobou, posteriormente à revisão bibliográfica, pesquisa de campo de natureza qualitativa, com alguns elementos subsidiários para análises de natureza quantitativa. Para tal, foi realizada aplicação de questionário estruturado, complementado por entrevista semiestruturada. A pesquisa de campo teve início com a aplicação do questionário. Após a análise das respostas, a população-alvo foi novamente contatada e, por meio de entrevista, ocorreu elucidação de pontos específicos que não haviam ficado claros nas respostas apresentadas por escrito. Com isso, alcançou-se o nível desejado de aprofundamento.

A seguir, passa-se a descrever os diferentes tópicos relacionados à metodologia empregada na investigação, assim como características e considerações sobre o seu desenvolvimento e aplicação.

5.1 População-alvo

No sentido de atender aos requisitos do universo de estudo da presente pesquisa, de maneira que este possa ser caracterizado por meio de um conjunto de características observáveis, definiu-se como população-alvo um conjunto de pós-graduandos a que o pesquisador tivesse acesso direto e pudesse estabelecer processo de comunicação e vínculo facilitados. Com isso, procurou-se garantir-se a uniformidade enquanto unidade de investigação e assegurar ao máximo a adesão ao processo de coleta de informações pertinentes à pesquisa.

O conjunto de pós-graduandos escolhidos como população-alvo são alunos do Programa de Pós-Graduação em Ciência da Informação do convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro (PPGCI/IBICT-UFRJ), por meio da Faculdade de Administração e Ciências Contábeis (FACC). Durante a realização desta dissertação, o corpo discente do PPGCI/IBICT-UFRJ apresentava-se formado por mestrandos e doutorandos que ingressaram nas seleções para as turmas de 2009 e 2010.

¹⁷ Pesquisa exploratória: visa a proporcionar maior familiaridade com o problema com vistas a torná-lo explícito ou a construir hipóteses. Envolve levantamento bibliográfico; entrevistas com pessoas que tiveram experiências práticas com o problema pesquisado; análise de exemplos que estimulem a compreensão. (GIL, 1991)

A homogeneidade do grupo foi um dos fatores que estimularam a escolha desta população-alvo. Todos os integrantes são pós-graduandos, apesar de divididos em dois níveis de formação distintos, mestrado e doutorado. Todos têm a informação como objeto de estudo, mesmo que com pontos de vista e abordagens diferentes ou aplicações variadas. Todos são especialistas em suas respectivas áreas, o que colabora para o perfil interdisciplinar do PPGCI e confere abrangência maior ao trabalho.

Também caracteriza o grupo, especialmente no momento em que foi desenvolvida esta dissertação, o compartilhamento de uma necessidade comum. Ao trabalharem na produção de suas teses e dissertações, os pós-graduandos deparam-se com a necessidade de buscar informações. Apesar de diferente na semântica, a natureza da informação buscada é semelhante na forma, ou seja, mestrandos e doutorados buscam por artigos acadêmicos, publicações em periódicos científicos, livros, citações etc. Como enumera Miller (2002, p. 3), são quatro as principais necessidades informacionais de cientistas: informação atual para manter-se atualizado sobre o campo em que estão inseridos; informação específica relacionada às tarefas desempenhadas; informações para criar marcos referenciais que constituem a base de projetos de pesquisa; e revisões de áreas relacionadas ou periféricas à pesquisa.

Os tipos de fontes pesquisadas tendem à semelhança, uma vez que há grande probabilidade de as informações científicas buscadas estarem armazenadas, por exemplo, em bancos de dados acadêmicos, revistas científicas especializadas, repositórios de teses e dissertações.

Influenciou na escolha da população-alvo em questão o fato de a autora desta dissertação estar inserida no grupo. Inicialmente, considerou-se que esta particularidade facilitaria o acesso ao grupo. De fato, no desenvolvimento do projeto, constatou-se que essa característica reverteu-se efetivamente em alta participação do grupo pesquisado, assim como em demonstração de forte interesse em colaborar com a pesquisa empreendida.

Importante ressaltar que por se tratar de pesquisa eminentemente qualitativa, não foram considerados grupos estatísticos, mas, como define Gaskell (2008, p. 69), foi escolhido um “grupo natural”, composto por pessoas que partilham um projeto comum, como é o caso da população-alvo selecionada.

5.2 Elaboração do questionário

O questionário elaborado para esta pesquisa de campo é composto por duas partes distintas, tomando como base o modelo desenvolvido por Biolchini (1998). Na primeira parte, procura-se levantar informações sobre o perfil e a trajetória acadêmica dos pós-graduandos, assim como sobre os projetos individuais de pesquisa em desenvolvimento no PPGCI/IBICT-UFRJ pelos sujeitos estudados. Na segunda parte, obtém-se o relato detalhado de experiências vividas pelos participantes ao realizarem buscas na web direcionadas aos seus respectivos projetos. Cada participante relatou uma experiência.

Junto às duas partes do questionário, no material entregue à população-alvo incluiu-se um texto curto de apresentação do presente trabalho de investigação, com instruções básicas sobre o preenchimento e esclarecimentos sobre o projeto de pesquisa, o qual pode ser lido nos apêndices B, C, D e E ao final da dissertação. Também foi incluído ao material o termo de consentimento, apresentado no apêndice A.

A parte inicial do questionário, sobre a caracterização dos pós-graduandos, foi formulada em duas versões, uma destinada aos mestrandos e outra, aos doutorandos. Ambas contêm perguntas sobre a área de origem e o ano de graduação, assim como questões sobre o projeto desenvolvido no segundo semestre de 2010 no PPGCI/IBICT-UFRJ. Nos questionários destinados aos doutorandos foram inseridas três perguntas relativas à formação em nível de mestrado realizada pelos integrantes deste grupo. Os participantes informaram sobre a área do conhecimento tratada na dissertação, a instituição onde foi obtido o título de mestre e o ano de conclusão da pós-graduação.

A decisão de utilização de questionário, que na sequência foi complementado por entrevista, deu-se por se tratar de um instrumento de fácil aplicação, podendo ser entregue pessoalmente ou, se necessário, enviado por e-mail. Como o preenchimento não exige acompanhamento direto do pesquisador, foi possível nesta pesquisa de campo realizar a aplicação do questionário, por exemplo, a um grupo numeroso de pós-graduandos simultaneamente. Outro aspecto positivo diz respeito à possibilidade de incluir perguntas fechadas, as quais garantem maior uniformidade tanto para o registro quanto para a futura avaliação.

A segunda parte do questionário esteve focada no processo de busca e recuperação e foi desenvolvida a partir da técnica do incidente crítico – um dos métodos da área de estudos denominada Análise da Tarefa Cognitiva (CTA) –, detalhada mais adiante.

Para que o pós-graduando localizasse a situação que serviria de base para responder a todas as perguntas presentes na segunda parte do questionário foi-lhe sugerido na primeira questão que se lembrasse de uma busca relevante, e apenas uma, realizada recentemente e que estivesse associada ao desenvolvimento do projeto de dissertação ou tese. Solicitou-se também que especificasse há quanto tempo a busca havia sido efetuada, de acordo com unidades de tempo determinadas, ou seja, há quantas horas, dias ou semanas.

Para a elaboração das questões seguintes, além do objetivo geral e dos objetivos específicos, serviram de fundamento os elementos do processo de recuperação da informação sugeridos por Meadow (1967, p. 106), ou seja: a) formulação da consulta; b) submissão da consulta formulada às bases de dados existentes e a determinação da relevância da base de dados escolhida para a consulta realizada; c) a recuperação das informações; d) a avaliação do processo de recuperação; e e) o desenvolvimento de estratégias para a utilização dos recursos disponíveis para atingir o objetivo do pesquisador.

No encadeamento das perguntas procurou-se seguir a trajetória empregada em um processo de busca, como se vê na Figura 4. A disposição das perguntas no questionário obedeceu à sequência cronológica dos procedimentos inseridos no próprio processo de busca e recuperação da informação para que se tornasse mais fácil ao pós-graduado organizar seu relato com começo, meio e fim, como se objetiva que seja feito ao se utilizar a técnica do incidente crítico.

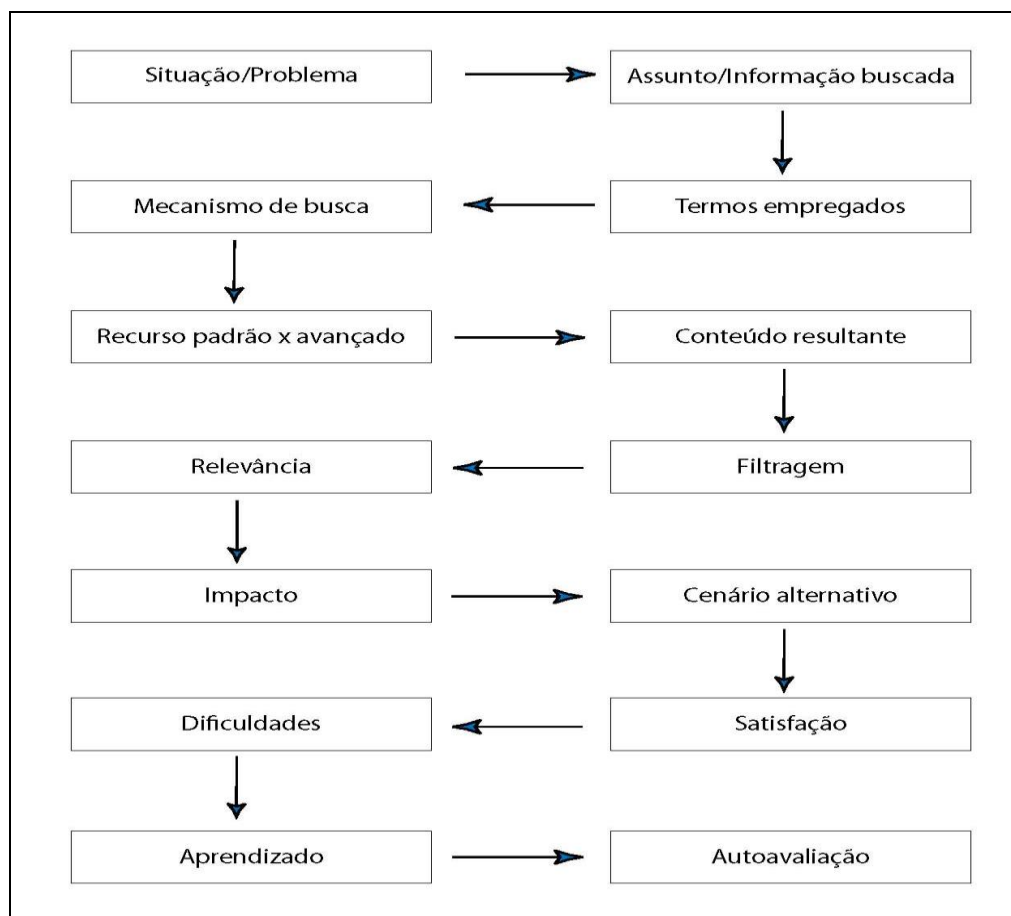


Figura 4 – Assuntos tratados nas perguntas inseridas no questionário, inspirados no processo de busca

Um problema a ser resolvido e/ou necessidade informacional a ser sanada dentro de situação específica suscitam um processo de busca. Após escolher qual busca seria relatada, o pós-graduando foi submetido a duas perguntas sobre o problema que a motivou:

2.2. Que situação o levou a realizar essa busca? Qual era o problema a ser solucionado?

2.3. Que informação específica você estava procurando ao recorrer à web? Isto é, qual era o assunto da busca? Qual era a questão que estava necessitando de uma resposta?

Após refletir sobre o problema a ser solucionado, o usuário passou à definição da consulta, como seria formatada, quais termos ou palavras-chave seriam incluídas, incentivado pela pergunta:

2.4. Quais termos/palavras foram inseridos no campo de busca?

Consulta, de acordo com Meadow (1967, p. 108), é a forma de comunicação por meio da qual os usuários “interrogam” uma biblioteca ou um estoque de informação – atualmente deve-se incluir nessa definição os repositórios digitais. Para Spink e Jansen (2005, p. 77), a

consulta na web consiste em um ou vários termos de busca inseridos em um campo determinado de uma página. Uma consulta é componente integral da busca na web e a expressão do problema informacional do usuário para o mecanismo de busca. “Ao usar um mecanismo na web, o usuário normalmente tem um problema informacional que o leva a usar esse recurso. O usuário deve expressar esse problema informacional para um mecanismo por meio de consultas. A busca na web tem de ser inserida em formato compreensível ao mecanismo escolhido.” (SPINK; JANSEN, 2005, p. 77)

O processo de busca, como recém-visto, tem continuidade com a decisão do indivíduo sobre o recurso informacional mais adequado, ou seja, qual base de dados ou mecanismo de busca melhor atenderá à necessidade em questão. Outro aspecto abordado no questionário refere-se à decisão de uso de recurso padrão ou avançado do mecanismo de busca ou repositório escolhido. Nesta etapa, foram utilizadas as seguintes perguntas:

2.5. Por meio de qual recurso informacional deu início à busca?

2.6. Que motivo o levou recorrer a essa fonte para buscar a informação de que necessitava?

2.7. Durante o processo de busca, utilizou recurso padrão oferecido pelo mecanismo escolhido ou recurso avançado? Se usou recurso avançado, especifique um pouco mais quais foram os campos utilizados e por quê.

Esperou-se com essas perguntas verificar também se profissionais da informação pertencentes à população-alvo utilizam outros subsídios, como, por exemplo, os operadores booleanos.

Na etapa seguinte, diante dos resultados foi contemplado o reconhecimento do tipo de conteúdo que mais atendeu ao que o pós-graduando estava procurando, por meio da pergunta:

2.8. Que tipo de conteúdo resultou da sua busca?

No processo de busca, após receber a lista de resultados, segue-se a tomada de decisões sobre como o material mais relevante para o projeto de pesquisa será filtrado:

2.9. Quais critérios foram utilizados para filtrar os resultados obtidos, como fez para separar o que seria útil para seu projeto?

Esta fase está relacionada à questão do desenvolvimento de estratégias para a utilização dos recursos disponíveis a fim de atingir o objetivo.

Na elaboração do questionário, foram considerados ainda aspectos relacionados à satisfação, ao aprendizado e às dificuldades enfrentadas. Em referência novamente a Meadow, o indivíduo foi convidado a refletir sobre o impacto causado por aquele processo de busca em seu projeto de pesquisa e sobre o quanto a busca em questão afetou as decisões futuras.

No contexto dos objetivos da presente pesquisa, o conceito de impacto refere-se ao efeito transformador da informação, podendo-se expressar sob diferentes aspectos, tais como: no sentido de uma contribuição relevante da mesma para o projeto do pesquisador investigado, da influência da informação recuperada sobre os rumos do desenvolvimento do projeto, de uma mudança de direcionamento do mesmo frente à presença ou à ausência da informação obtida no processo de busca e recuperação, do atendimento à necessidade informacional e à consolidação de uma etapa do projeto, assim como de uma mudança cognitiva resultante de aprendizagem possibilitada pela própria experiência de busca e recuperação da informação.

Essa parte é representada no questionário por uma série de perguntas:

2.10. Em que aspectos específicos o resultado obtido foi relevante para a sua pesquisa de mestrado? No que ele contribuiu para o seu projeto?

2.11. Qual foi o impacto da informação recuperada para o seu projeto? Como o resultado obtido afetou os passos seguintes no desenvolvimento do seu projeto?

2.12. O que poderia ter ocorrido se o resultado não tivesse sido obtido?

2.13. Esta busca atendeu satisfatoriamente a sua necessidade informacional e deu por encerrado o processo? Ou foi necessário buscar mais informações?

2.14. Qual foi o seu grau de satisfação em relação à busca, em uma escala de 0 a 10, sendo 0 muito baixo e 10 muito alto?

2.15. Qual foi a maior dificuldade encontrada no processo em relação aos conteúdos obtidos? O que você fez para resolvê-la?

2.16. Qual foi o maior aprendizado proporcionado por esta experiência?

O questionário continha ainda duas perguntas relacionadas às habilidades e à experiência do usuário em relação a busca e recuperação da informação na web com o objetivo de se obter informações sobre como os integrantes da população-alvo avaliam suas próprias competências. Em dois momentos, no início e ao final do questionário, foram inseridas duas questões de autoavaliação. Com a primeira pergunta, pretendia-se observar

como o indivíduo avalia sua capacidade de buscar informações científicas na web em geral. Com a segunda questão, após ter relatado uma experiência completa de busca e recuperação de informações, pretendia-se observar o quanto e como a avaliação se mantinha e também para estimular a metacognição. As oscilações registradas nas respostas sugerem a reflexão por parte dos pós-graduandos sobre sua própria atuação.

Entendeu-se que com o estudo dessa situação, ao acompanhar as etapas pelas quais os participantes passam ao realizar busca de informação na web, seu processo de raciocínio e a condução do processo como um todo, torna-se possível atingir os objetivos propostos nesta dissertação.

5.3 A técnica do incidente crítico

Como afirmado anteriormente, o questionário utilizado neste estudo foi elaborado a partir da técnica do incidente crítico, um conjunto de procedimentos para a coleta de observações diretas do comportamento humano, como a define Flanagan (1954, p. 327). Denominadas incidentes críticos, essas observações baseiam-se em situações reais definidas – e não teóricas ou hipotéticas –, que representem contribuição significativa para a atividade pesquisada, e que apresentem um grau de dificuldade acima do habitual. Por meio do uso da técnica, foram coletadas amostras representativas de informação sobre as questões em estudo.

A técnica do incidente crítico, desenvolvida na área da psicologia do trabalho, inicialmente para um programa de psicologia da aviação na década de 1940, mostrou-se eficaz e adequada ao caso por permitir que os indivíduos contribuíssem com informações baseadas em suas próprias experiências. Como afirma Flanagan (1954, p. 328), com a utilização da técnica do incidente crítico, é possível obter o registro de comportamentos específicos relacionados a exigências críticas para o sucesso de um indivíduo em uma tarefa definida.

Flanagan ressalta ainda que um incidente é crítico se contribuir significativamente – positiva ou negativamente – para o propósito geral da atividade. A técnica do incidente crítico faz parte, na atualidade, de um campo metodológico de investigação interdisciplinar, denominado Análise da Tarefa Cognitiva (*Cognitive Task Analysis - CTA*), cuja dimensão informacional, como afirmam Crandall, Klein e Hoffman (2006, p. 20), encontra-se no cerne de suas preocupações e de seus objetivos.

Os autores ressaltam ainda que o propósito da CTA é capturar a maneira como a mente dos indivíduos funciona, capturar a cognição. Um pesquisador, ao empregar a análise da tarefa cognitiva, está tentando entender e descrever como os indivíduos veem a ação que estão praticando e como esta faz sentido. Se esses indivíduos estão conseguindo efetivamente cumprir o que se propõem e dão conta de circunstâncias complexas, a CTA deveria descrever a base de sua performance habilidosa. Se estão cometendo erros, então a análise deveria explicar o que está levando a isso. “Estudos sobre a análise da tarefa cognitiva tentam capturar o que as pessoas estão pensando, no que elas estão prestando atenção, as estratégias que usam para tomar decisões ou para detectar problemas, o que estão tentando realizar e o que sabem acerca de como o processo funciona.” (CRANDALL; KLEIN; HOFFMAN, 2006, p. 9)

Investigações no campo da CTA apresentam três aspectos primários: elicitación do conhecimento (*knowledge elicitation*), referente à coleta de dados, ao conjunto de métodos empregados para obter informações sobre o que as pessoas sabem e como elas o sabem, ou seja, estratégias, habilidades e julgamentos; análise de dados (*data analysis*), referente à avaliação realizada a partir dos dados coletados; e representação do conhecimento (*knowledge representation*), que trata sobre a divulgação e comunicação dos resultados obtidos, para que a comunidade científica e de usuários interessados tome conhecimento.

Dentro dessa perspectiva, o método do incidente crítico, enquanto CTA, está inserido no grupo dos métodos mais poderosos de elicitación do conhecimento. Em métodos de análise da tarefa cognitiva, segundo Crandall, Klein e Hoffman (2006, p. 13), as pessoas falam sobre todos os tipos de detalhes, desafios, sinais sutis, influências e estratégias que nunca viriam à tona com a utilização de outros métodos.

No presente estudo, a elicitación do conhecimento deu-se por meio da aplicação do questionário e, posteriormente, da realização da entrevista. Crandall, Klein e Hoffman destacam que entrevistas são eficientes, mas observam que a obtenção de bons dados sempre depende de os participantes serem capazes e desejarem refletir de maneira aprofundada sobre seu desempenho. Nesta pesquisa de campo, os pós-graduandos foram convidados a relatar uma experiência vivida no passado, o que pode resultar em algumas desvantagens, como esquecimento ou distorção de detalhes. Por essa razão, optou-se pela aplicação de questionário conjugada à complementação do recurso de aprofundamento por meio de entrevista semiestruturada, para esclarecer aspectos obtidos através do primeiro instrumento de coleta de informações.

Com o material obtido, a partir da utilização sequencial dos dois instrumentos de eliciação do conhecimento, foi efetuado o cruzamento dos dados e a análise dos mesmos. A representação do conhecimento dá-se agora com a apresentação desta dissertação.

5.4 Elaboração da entrevista

Como planejado, após a aplicação do questionário, foi realizada a rodada de entrevistas. Gaskell (2008, p. 73) ressalta que com a entrevista “o pesquisador pode obter esclarecimentos e acréscimos em pontos importantes com sondagens apropriadas e questionamentos específicos”. A entrevista foi composta por número menor de perguntas que o questionário, as quais foram adaptadas e aplicadas de acordo com as necessidades identificadas em cada caso. As perguntas da entrevista funcionaram como complementação de algumas questões correspondentes que já haviam sido respondidas no questionário. Dessa forma, os participantes nem sempre responderam às mesmas perguntas durante o processo da entrevista.

Para Gaskell (2008, p. 43), toda pesquisa é uma forma de interação, uma troca de ideias e significados, na qual várias realidades e percepções são exploradas e desenvolvidas. “Com respeito a isso, tanto o entrevistado quanto o entrevistador estão, de maneiras diferentes, envolvidos na produção do conhecimento.” (GASKELL, 2008, p. 43) Na entrevista conduzida através da presente pesquisa aprofundou-se especialmente aspectos relacionados aos pontos críticos do processo de busca e recuperação, tais como dificuldades e erros, por serem estes os aspectos que revelaram maiores dificuldades de serem relatados espontaneamente por meio da interação do respondente com o questionário.

A elaboração da entrevista ocorreu paralelamente à elaboração do questionário. Para cada pergunta inserida no questionário foram aprioristicamente pensadas opções que pudessem vir a esclarecer ou complementar possíveis pontos que poderiam não ficar claros somente pela aplicação do questionário. Durante o processo de elaboração dos instrumentos, perguntas inicialmente programadas para a entrevista terminaram por ser incluídas no questionário, como foi o caso da questão sobre o aprendizado que o processo de busca e recuperação trouxe ao pós-graduando.

Após o recebimento dos questionários respondidos, foi realizada uma avaliação individual dos resultados obtidos e feita então a seleção das perguntas que seriam então

aplicadas de modo específico e individual àqueles respondentes que deixaram respostas incompletas ou passíveis de dúvidas.

A inclusão de entrevista ao processo de elicitação deu-se em função da premissa de que haveria necessidade de complementação de dados, dada a complexidade do processo investigado e a diversidade de informações a serem obtidas relativamente ao universo de estudo. Uma das vantagens da entrevista é precisamente a oportunidade que a mesma fornece como instrumento de elicitação para obtenção de informações que não se encontram em fontes documentais, e são fortemente determinadas pela vivência humana de situações, com importante componente subjetivo, além de relevantes e significativas.

5.5 Teste do questionário

Como planejado no desenvolvimento cronogramático da pesquisa, ao ser concluída a primeira versão do questionário, em outubro de 2010, realizou-se então o teste-de-campo, por meio de um estudo-piloto com a participação de um conjunto de mestres e doutorandos em Ciência da Informação. Sete voluntários, sendo seis pós-graduandos em Ciência da Informação pelo convênio IBICT-UFF e um pela Universidade Federal de Santa Catarina (UFSC), responderam às perguntas, mensuraram a duração do preenchimento das respostas, realizaram comentários sobre o formato e sugeriram melhorias e alterações.

A aplicação do teste permitiu um aperfeiçoamento das perguntas e eventual troca ou alteração daquelas que não estavam suficientemente claras em sua compreensão ou com sentido vago. Também foi a partir do teste que se tomou a decisão de retirar a terceira parte do questionário original, apontada como repetitiva por todos os participantes do teste, e não respondida pela maior parte dos voluntários por estes não a considerarem relevante. Ao final desta dissertação, o questionário-piloto elaborado para os mestrandos encontra-se como apêndice B e o questionário-piloto elaborado para ser respondido por doutorandos, como apêndice D.

Inicialmente, o questionário foi elaborado contendo três etapas. A primeira, na qual era feita a caracterização dos participantes, permaneceu inalterada. Nas duas etapas seguintes, havia uma que tratava da experiência positiva ao buscar informações na web e outra na qual os participantes relatavam uma experiência negativa. Após o teste, a terceira parte foi suprimida, em função de seu custo-benefício ser considerado insuficiente, aportando pouca informação útil e relevante e estendendo em demasia a duração da aplicação do questionário.

Foi possível concluir, por meio desses parâmetros de avaliação provenientes do estudo-piloto, que a exclusão desta terceira parte da versão original não afetaria os resultados almejados pela pesquisa de campo.

Conforme mencionado anteriormente, ao responderem às perguntas, os voluntários mediram o tempo de realização do questionário. A média de preenchimento ficou em 25 minutos, considerada satisfatória para uma pesquisa dessa natureza. Este dado foi usado posteriormente nas instruções aos respondentes com o objetivo de dar aos pós-graduandos a noção de quanto tempo gastariam em média para responder às perguntas, contribuindo para a sua motivação em participar da pesquisa como voluntários e para o agendamento e a programação da sua participação efetiva.

5.6 Aplicação da pesquisa

O contato inicial com a população-alvo foi realizado via e-mail e de forma individual. Dos 65 pós-graduandos que fazem parte do PPGCI/IBICT-UFRJ, integrando as turmas que ingressaram nos anos de 2009 e 2010, 63 mostraram-se dispostos a colaborar, o equivalente a 97% da população-alvo.

O agendamento da aplicação presencial do questionário foi realizado com 40 colegas. Outros 12, por incompatibilidade de horários, dispuseram-se a imprimir o questionário e preenchê-lo, deixando-o na secretaria do PPGCI/IBICT-UFRJ. No total, 52 pós-graduandos aderiram como respondentes e participaram da aplicação do questionário, perfazendo 80% do total da população-alvo escolhida. O percentual de participação é representativo dentro do grupo pesquisado. As entrevistas foram realizadas com 32 pós-graduandos, representando 61% do grupo de respondentes, por telefone e e-mail. Conforme mencionado anteriormente, sua realização se deu nos casos em que foi identificada a necessidade de complementação das informações obtidas pela aplicação do questionário.

5.7 Análise dos dados

Meadow (1967, p. 85) afirma que o objetivo amplo da análise de dados é procurar sentidos e compreensão. “É vital garantir que toda interpretação esteja enraizada nas próprias entrevistas, de tal modo que, quando a análise é feita, o *corpus* pode ser trazido para justificar as conclusões.” (MEADOW, 1967, p. 86) No caso da pesquisa de campo empreendida para

esta dissertação, o *corpus* analisado foi formado pelas respostas dos pós-graduandos às perguntas do questionário e da entrevista.

Em um primeiro momento da apresentação dos resultados desta pesquisa foi realizado um resumo sobre a participação dos pós-graduandos. Em seguida, organizou-se o conteúdo das respostas em tabelas, separadas tanto por pós-graduando quanto por pergunta, para permitir uma análise comparativa. Respostas às questões como campos de origem, universidades em que os pós-graduandos concluíram a graduação, ano de formação e linha de pesquisa atual foram agrupadas já em uma primeira fase e transformadas em gráficos. Os campos estudados pelos pós-graduandos, além da Ciência da Informação, foram agrupados de acordo com a Tabela de Áreas do Conhecimento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e transformados em uma nuvem de etiquetas¹⁸ (*tag cloud*), como poderá ser conferido logo adiante.

As respostas oferecidas pelos doutorandos sobre as áreas do conhecimento envolvidas em seus projetos tanto de mestrado quanto de doutorado serviram de base para análise sobre a evolução ou mudança de rumo na formação dos acadêmicos.

No estágio inicial da análise foram trabalhados dados sobre as formas de aquisição de habilidades de busca e recuperação de informação e sobre as auto-avaliações executadas pelos participantes. No que pode ser caracterizado por um estágio mais complexo foram executadas análises longitudinais de todas as respostas dadas pelos pós-graduandos na segunda parte do questionário, na qual foram detalhadas experiências vivenciadas pelos participantes ao realizar uma busca na web.

¹⁸ Nuvem de etiquetas é uma tradução livre da expressão em inglês *tag cloud*. Também é conhecida pela expressão inglesa *weighted list*, traduzível como uma lista ponderada, ou seja, uma relação de conceitos onde o peso de cada um deles define seu lugar dentro do arranjo visual. Na maioria das *tag clouds* ainda existe a opção de visualização por ordem alfabética ou por peso de conceitos. (GUEDES; SOUZA, 2008, p. 7)

6. Resultados da pesquisa de campo

Mestrandos e doutorandos que ingressaram em 2009 e 2010 no Programa de Pós-Graduação em Ciência da Informação do Instituto Brasileiro de Informação em Ciência e Tecnologia em convênio com a Universidade Federal do Rio de Janeiro, como apresentado anteriormente, compõem a população-alvo desta pesquisa de campo.

Dos 65 integrantes deste grupo¹⁹, 39 (60% do total) são mestrandos e 26 (40%), doutorandos. Em divisão por ano de ingresso, 28 (43%) pós-graduandos pertencem à turma de 2009 (sendo 18 mestrandos e 10 doutorandos, respectivamente 27,6% e 15,4% do total) e 37 (57%), à turma de 2010 (sendo 21 mestrandos e 16 doutorandos, respectivamente 32,4% e 24,6% do total), como apresentado na Figura 5.

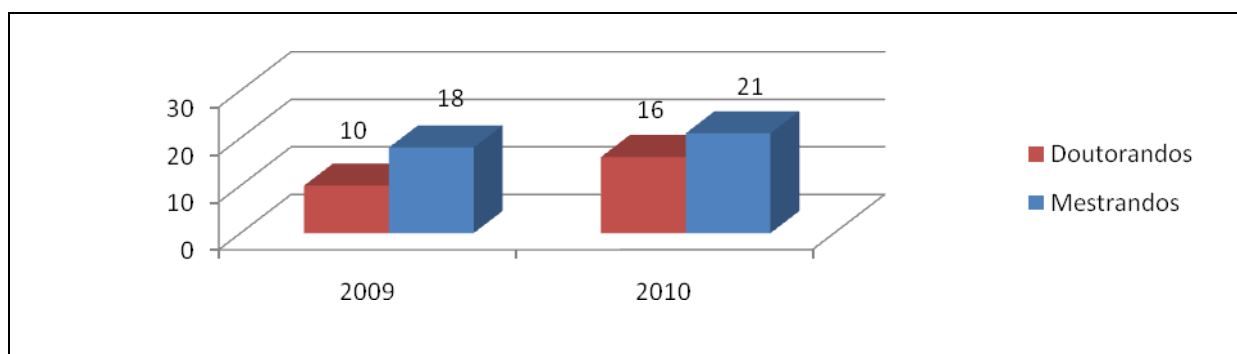


Figura 5 – Número de pós-graduandos do PPGCI/IBICT-UFRJ ingressados em 2009 e 2010

Como apresentado na metodologia, foi realizada uma abordagem inicial, através do envio de e-mails individuais, aos 65 pós-graduandos, dos quais 63 manifestaram interesse e disposição em participar da pesquisa, o equivalente a 97%. Durante o período de 5 a 26 de novembro de 2010, 40 pós-graduandos (63,5% do total de 63 participantes) responderam ao questionário em encontros presenciais. Doze (19%), por incompatibilidade de agenda ou dificuldade de deslocamento, receberam o questionário por e-mail. Após imprimi-lo e respondê-lo, essa parcela da população-alvo entregou o conjunto de respostas, junto com o

¹⁹ Ao se incluir a autora desta dissertação chega-se ao número de 66 pós-graduandos ingressados no PPGCI/IBICT-UFRJ nos anos de 2009 e 2010, mas nesta dissertação considerou-se 65 como número total

termo de consentimento para participação na pesquisa, assinado, na secretaria do PPGCI/IBICT-UFRJ.

Durante a aplicação do questionário foi novamente medido o tempo gasto para completar as perguntas. A média, tanto entre mestrandos quanto entre doutorandos, foi de 23 minutos, dois minutos a menos do que o tempo obtido no estudo-piloto. A pequena diferença de tempo constatada demonstra a pertinência das informações coletadas no teste.

Ao final da aplicação do questionário, registrou-se a participação de 52 pós-graduandos, o equivalente a 80% do grupo correspondente à população-alvo.

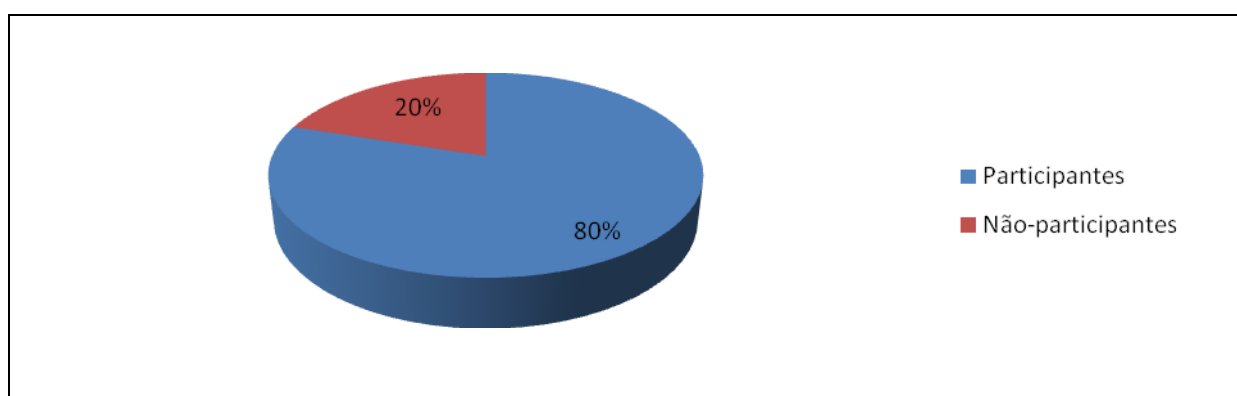


Figura 6 – Proporção de participantes da pesquisa de campo entre pós-graduandos do PPGCI/IBICT-UFRJ

Do total de respondentes ao questionário, 77% o fizeram presencialmente, enquanto que 23% por e-mail.

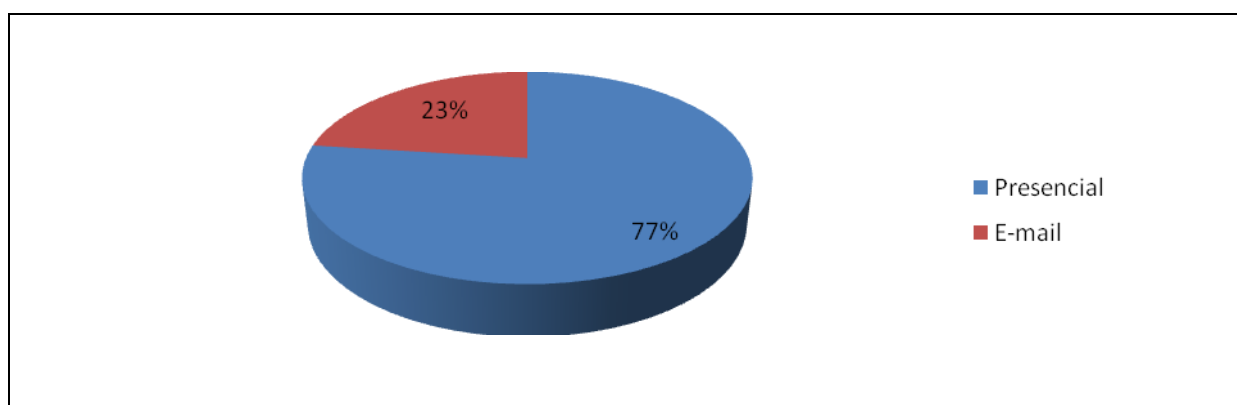


Figura 7 – Proporção de participantes que responderam ao questionário de forma presencial x por e-mail

Desses, 34 são alunos do mestrado – sendo 18 (53% do subgrupo de mestrandos estudados) da turma de 2009 (no qual se obteve 100% de participação) e 16 (47% dos mestrados estudados) da turma de 2010 (com 76% de participação) – e 18 de doutorado, sendo oito (44% do subgrupo de doutorandos estudados) da turma de 2009 (com 80% de

participação) e 10 (56% dos doutorandos estudados) da turma de 2010 (com 62,5% de participação).

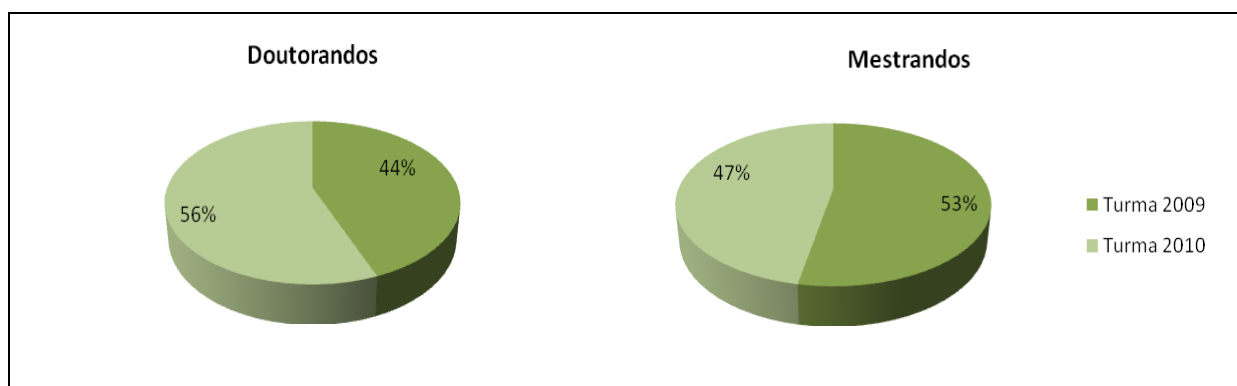


Figura 8 – Porcentagem de mestrandos e doutorandos estudados, em relação ao ano de ingresso

Depois de apresentar na Figura 8 a porcentagem de mestrandos/doutorandos da turma de 2009 participantes em relação aos participantes da turma de 2010, na Figura 9 observa-se a porcentagem de participação de mestrandos/doutorandos por ano de ingresso.

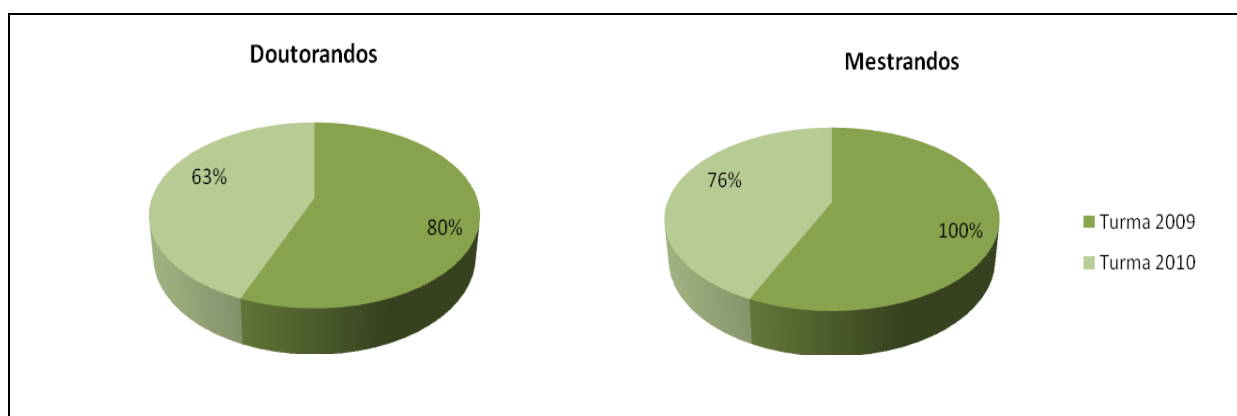


Figura 9 – Porcentagem de participação de doutorandos e mestrandos por ano de ingresso

Os dados apontam a adesão de 87% (34 de 39) dos mestrandos e de 62% (18 de 26) dos doutorandos à pesquisa. A partir desses percentuais percebe-se que houve maior participação de pós-graduandos que ingressaram em 2009 do que em 2010, correspondente a 32% a mais para os mestrandos (100% / 76%) e a 28% a mais para os doutorandos (80% / 62,5%). Acredita-se que esse resultado pode ter se dado por se tratar da turma à qual pertence a autora desta dissertação e ser, consequentemente, o grupo com o qual houve maior troca de informações sobre a presente pesquisa.

Após análise dos questionários, em relação ao seu grau de completude das informações, realizada entre 27 de novembro e 9 de dezembro de 2010, foram identificados aqueles respondentes para os quais seria necessário completar o processo de elicitação através

das entrevistas. Estas foram então aplicadas, no período de 10 a 18 de dezembro de 2010, a 23 dos 34 mestrandos participantes (67%), e a nove dos 18 doutorandos (50%) envolvidos.

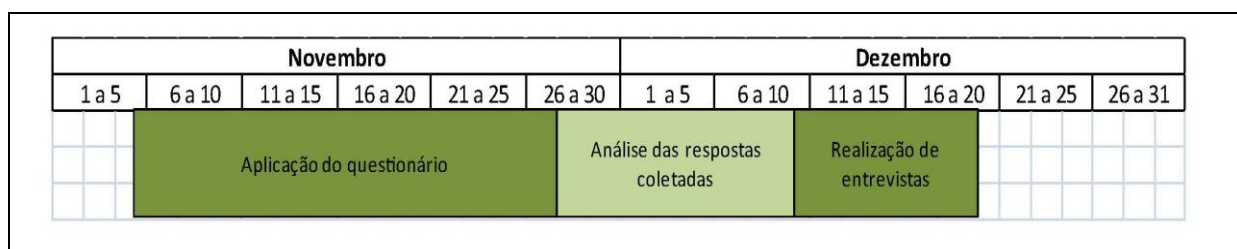


Figura 10 – Cronograma de coleta de dados para pesquisa de campo

A realização de número relativamente menor de entrevistas em relação ao de questionários aplicados deve-se aos resultados satisfatórios obtidos na primeira etapa da coleta de dados com uma parcela dos respondentes. Os percentuais de aplicação das entrevistas corroboram a necessidade da conjugação de dois instrumentos de coleta de informações, no sentido de obter tanto a abrangência quanto a profundidade das respostas almejadas de acordo com os objetivos da investigação.

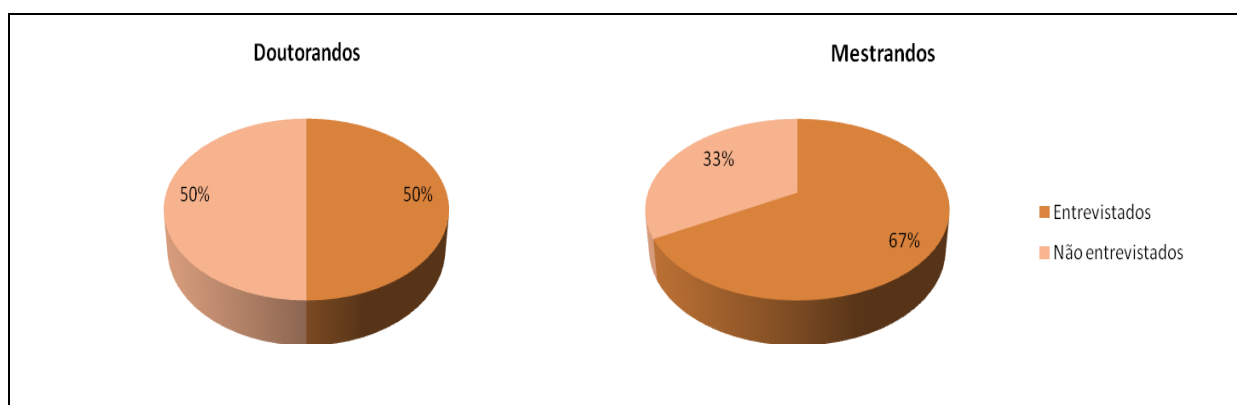


Figura 11 – Porcentagem de pós-graduandos entrevistados, divisão por mestrandos e doutorandos

Os esclarecimentos mais frequentes obtidos através da realização da entrevista disseram respeito especialmente a três aspectos: a motivação dos pós-graduandos ao escolher o mecanismo de busca utilizado, pois as justificativas relatadas no questionário não haviam sido apresentadas de maneira clara; a forma exata como os termos foram inseridos no campo de busca, não descritos com precisão no questionário; e a alteração da nota para mais ou para menos na segunda autoavaliação, a fim de promover um entendimento maior sobre as razões que levaram a mudanças nos valores atribuídos nas duas avaliações.

6.1 Formação dos pós-graduandos participantes

Ciência da Informação é, como define Foskett (1980, p.64) e já foi apontado no referencial teórico, uma disciplina que surge de “fertilização cruzada” de ideias que incluem a velha arte/ciência da biblioteconomia, a nova arte/ciência da computação, as artes dos novos meios de comunicação e ciências como psicologia e linguística que, em suas formas contemporâneas, têm a ver diretamente com todos os problemas da comunicação – a transferência do conhecimento organizado. O levantamento dos campos do conhecimento de origem dos pós-graduandos endossa a ideia de interdisciplinaridade apresentada por Foskett, o que pode ser constatado tanto entre os mestrandos, como se vê na Figura 12, quanto entre os doutorandos, como apresentado na Figura 13.

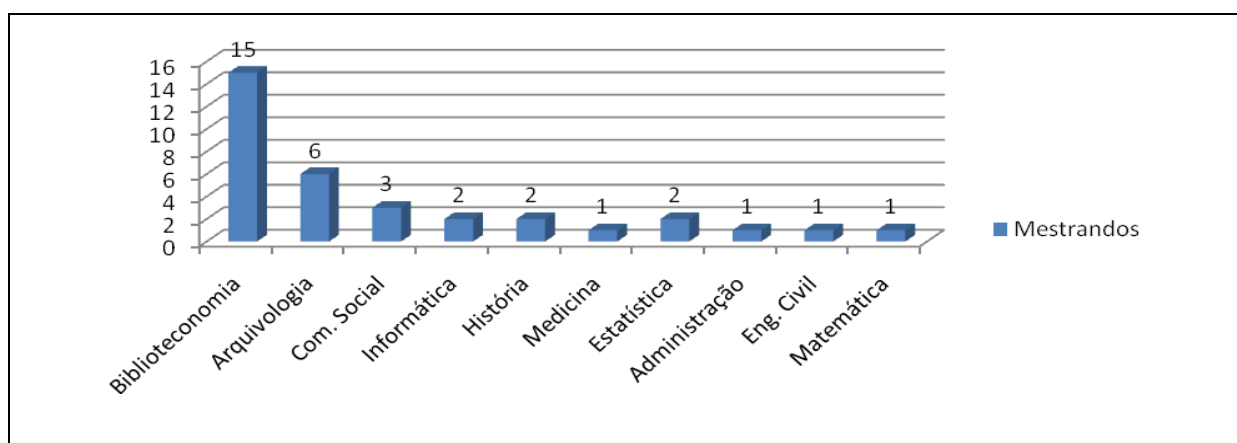


Figura 12 – Campos de origem dos mestrandos participantes da pesquisa de campo

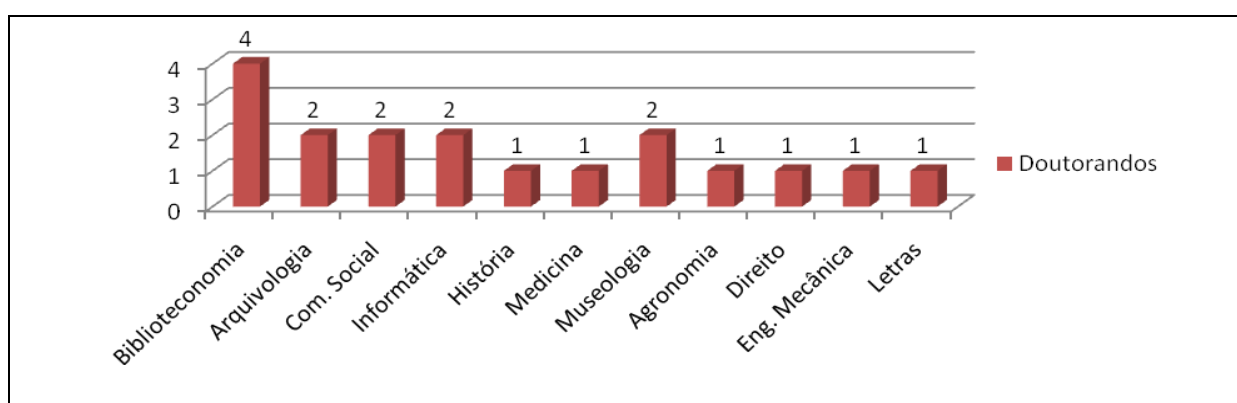


Figura 13 – Campos de origem dos doutorandos participantes da pesquisa de campo

Seis áreas do conhecimento coincidentes estão presentes tanto no segmento de mestrado quanto no de doutorado do PPGCI/IBICT-UFRJ: Biblioteconomia em primeiro lugar, seguida por Arquivologia, Comunicação Social, Informática, História e Medicina. Entre

os mestrandos contam-se 10 diferentes áreas de graduação, enquanto que entre os doutorandos, nota-se a ocorrência de 11 diferentes cursos, como exposto na Figura 14.

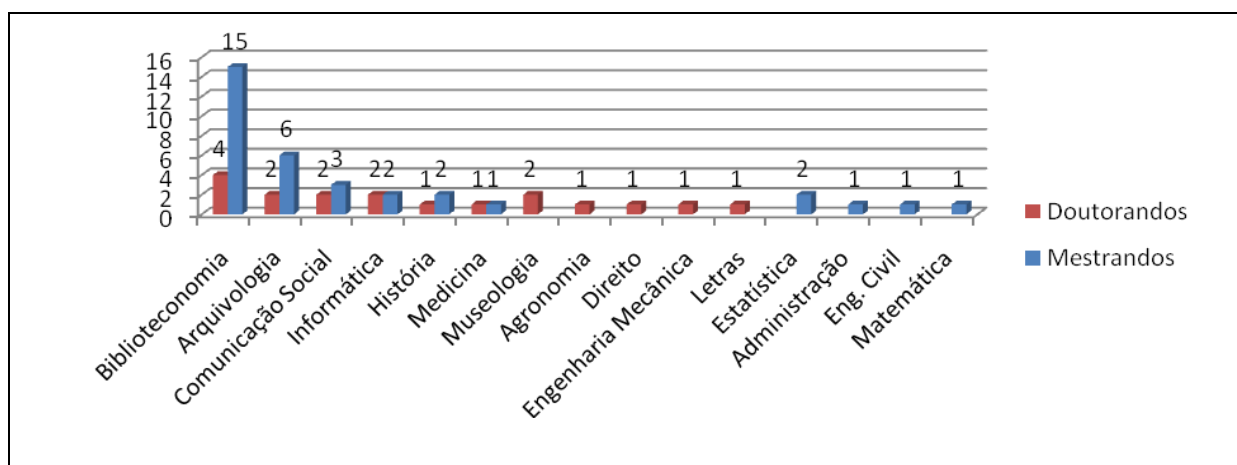


Figura 14 – Campos de origem dos pós-graduandos participantes da pesquisa de campo

Apesar de o questionário ter sido testado por sete mestres e doutores em Ciência da Informação, notou-se durante a aplicação à população-alvo que ainda outra alteração teria sido adequada. Mais da metade dos pós-graduandos não se sentiram seguros ao ler a questão que pedia para ser apontada a área de origem. Pergunta recorrente foi se tal questão referia-se à graduação.

Mais de três quartos dos participantes concluíram a graduação em instituições públicas de ensino superior (IES) – 85% dos mestrandos e 77% dos doutorandos. Noventa por cento estudaram em universidades no estado do Rio de Janeiro. A Figura 15 apresenta a distribuição das universidades por estado. No detalhe, aparecem representadas as instituições fluminenses frequentadas pelos pós-graduandos.

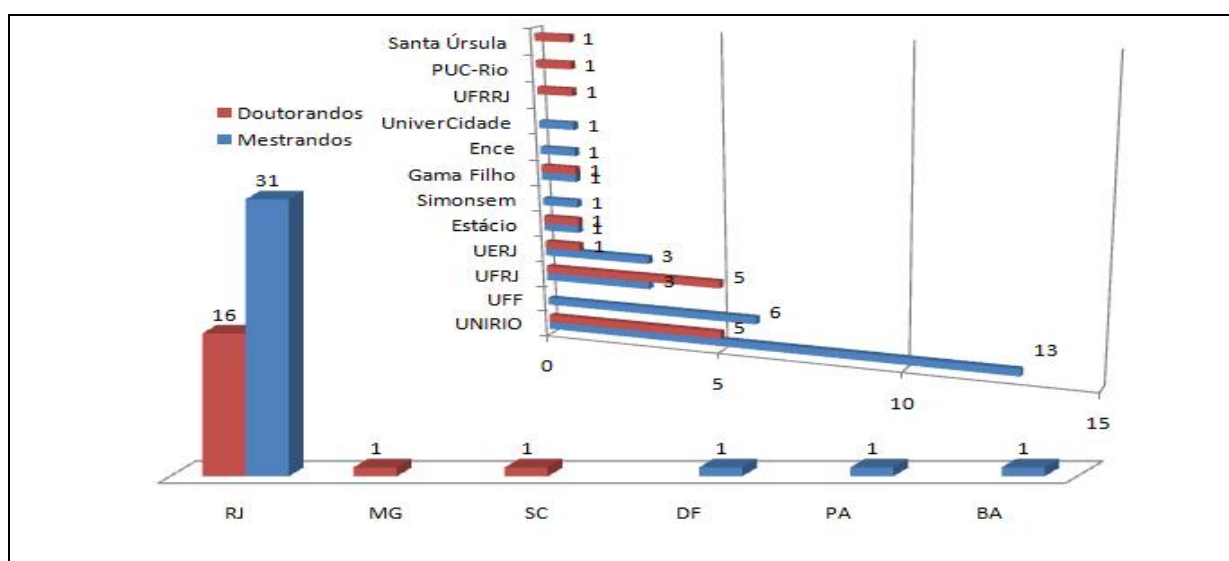


Figura 15 – IES, por estado, frequentadas na graduação; no detalhe, IES do estado do Rio de Janeiro

Como mencionado na metodologia, aos doutorandos foram realizadas duas perguntas relacionadas ao mestrado. Com exceção de apenas um doutorando, que obteve título de mestre em Santa Catarina, todos frequentaram instituições federais localizadas na região Sudeste, especialmente na região metropolitana do Rio de Janeiro.

Pôde-se observar também que apenas um estudante realizou o mestrado em um estado diferente daquele em que havia concluído a graduação. Entre os 18 doutorandos participantes, como apresentado na Figura 16, seis (um terço deles) são egressos do programa de pós-graduação do IBICT, sendo cinco destes do período em que o instituto mantinha convênio com a Universidade Federal Fluminense (UFF) e um do convênio anterior com a Escola de Comunicação da Universidade Federal do Rio de Janeiro (UFRJ).

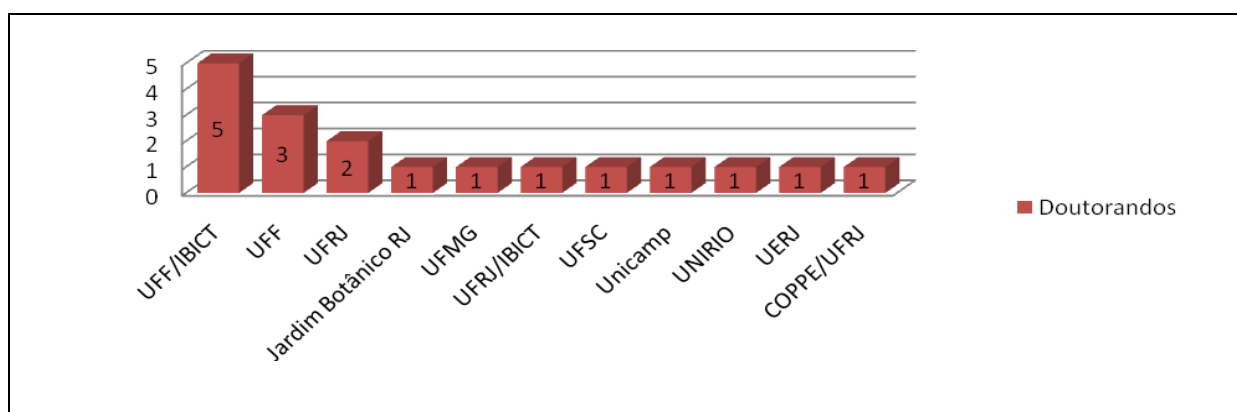


Figura 16 – Instituições nas quais doutorandos das turmas 2009/2010 concluíram seus mestrados

Os mestrandos que contribuíram com esta pesquisa graduaram-se em seus respectivos cursos entre 1972 e 2007. A maior concentração ocorreu nos anos 2000, com 19 graduandos (36% dos participantes), seguida pelos anos 1980, com oito (15%), pela década de 1990, com cinco (10%), e pelos anos 1970, com dois (4%). Entre os doutorandos, o período de graduação estende-se de 1969 a 2006, sendo que um graduou-se na década de 1960 (2%), um na década de 1970 (2%), sete nos anos 1980 (13%), quatro nos anos 1990 (8%) e cinco nos anos 2000 (10%).

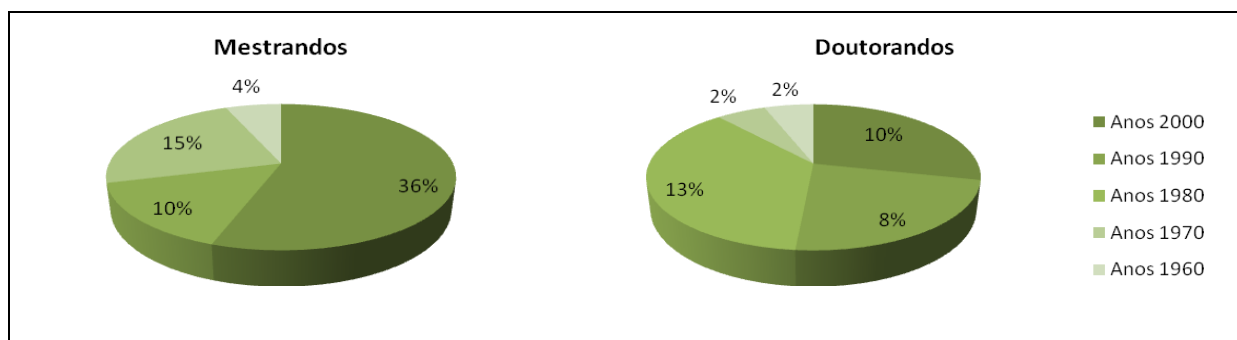


Figura 17 – Ano de graduação dos pós-graduandos participantes desta pesquisa de campo

A Figura 18 representa a trajetória dos estudantes de doutorado. Para a criação do gráfico foram usados o ano de graduação, o ano de conclusão do mestrado, assim como o ano de ingresso no PPGCI/IBICT-UFRJ.

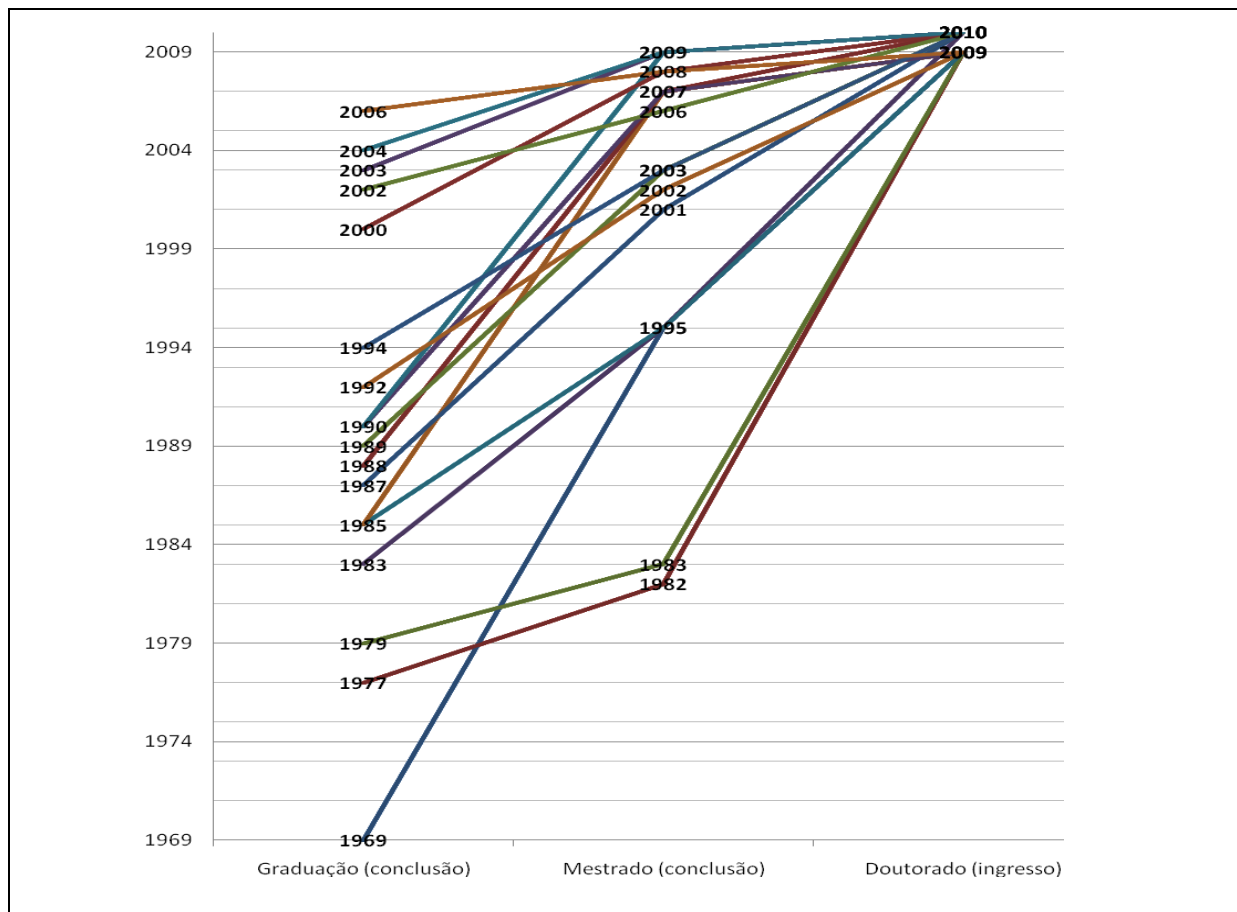


Figura 18 – Trajetória dos doutorandos, de acordo com ano de graduação, de conclusão do mestrado e de ingresso no PPGCI/IBICT-UFRJ

Os dados expostos no gráfico permitem demonstrar duas constatações:

A primeira é que a proporção de doutorandos é maior para aqueles cujo intervalo (em número de anos) entre as conclusões da graduação e do mestrado supera o intervalo entre a conclusão do mestrado e o ingresso no doutorado. Ou seja, a quantidade de doutorandos que realizou seu mestrado após anos de prática profissional é maior que a quantidade daqueles que o fizeram pouco tempo após a graduação, indicando uma maior proporção de mestres com maior experiência profissional anterior em suas respectivas áreas.

A segunda constatação é que o número de doutorandos que realizou seu mestrado recentemente é equivalente ao daqueles que o fizeram há um número maior de anos, demonstrando assim dois perfis idênticos de formação acadêmica.

6.2 Campos do conhecimento dos pós-graduandos

Ao apresentar campos e áreas do conhecimento sobre os quais os pós-graduandos desenvolvem suas pesquisas de dissertação e tese, fez-se necessário recorrer a Amaral (1990 *apud* PINHEIRO, 2006, p.115) em busca de definições apropriadas para campo, área e linha de pesquisa. De acordo com o autor,

[...] campo designa o território total cuja investigação o curso pretende operar. Medicina, Filosofia e Comunicação são campos. Área é uma subdivisão do campo, um corte introduzido artificialmente por motivos de estratégia exploratória. Teorias da Comunicação e da Cultura e Tecnologias da Imagem são áreas, como Filosofia Brasileira ou Cirurgia. [...] Linha de pesquisa existirá cada vez que, dentro de uma área, certos temas aglutinadores deem forma à cooperação entre pesquisadores. Estes então se reúnem para, trabalhando em conjunto em torno desses temas, aprofundarem a área e desenvolverem o campo. (AMARAL, 1990 *apud* PINHEIRO, 2006, p. 115)

Apesar de as linhas de pesquisa do PPGCI/IBICT-UFRJ terem passado por atualizações e adequações em 2010, aplicadas formalmente para o processo de seleção de mestrands e doutorandos para 2011, por uma questão de coerência nesta pesquisa foram consideradas as definições utilizadas no momento das seleções de 2009 e 2010, correspondentes aos pós-graduandos estudados. Sendo assim, a Linha 1 relaciona-se a *Epistemologia e Interdisciplinaridade na Ciência da Informação*; a Linha 2, a *Organização, Estrutura e Fluxos da Informação*; e a Linha 3, a *Informação, Sociedade e Gestão Estratégica*.

Entre os doutorandos pesquisados, a maior concentração ocorre na Linha 1 (38,89%), com sete pós-graduandos, sendo seguida pela Linha 3, com seis (33,33%), e pela Linha 2, com cinco (27,28%). Entre os mestrands, há mais projetos na Linha 3 (38,24%), somando 13. Os resultados, no entanto, mostram equilíbrio, como se vê na Figura 19, uma vez que 11 (32,35%) pesquisam sobre a Linha 1 e 10 (29,41%) sobre a Linha 2.

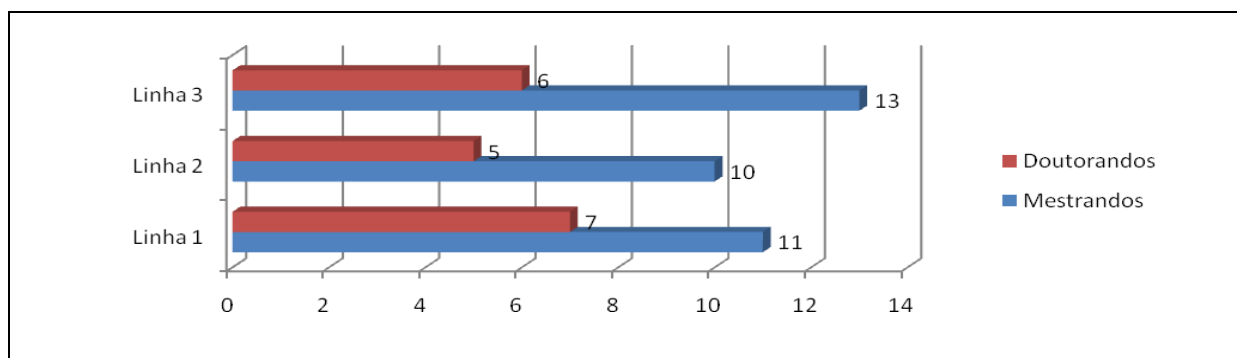


Figura 19 – Comparativo da divisão, por linha de pesquisa, de mestrands e doutorandos

No total, dezenove (36,5%) dos pós-graduandos na Linha 1, dezoito deles (34,6%) na Linha 3 e quinze (28,9%) na Linha 2.

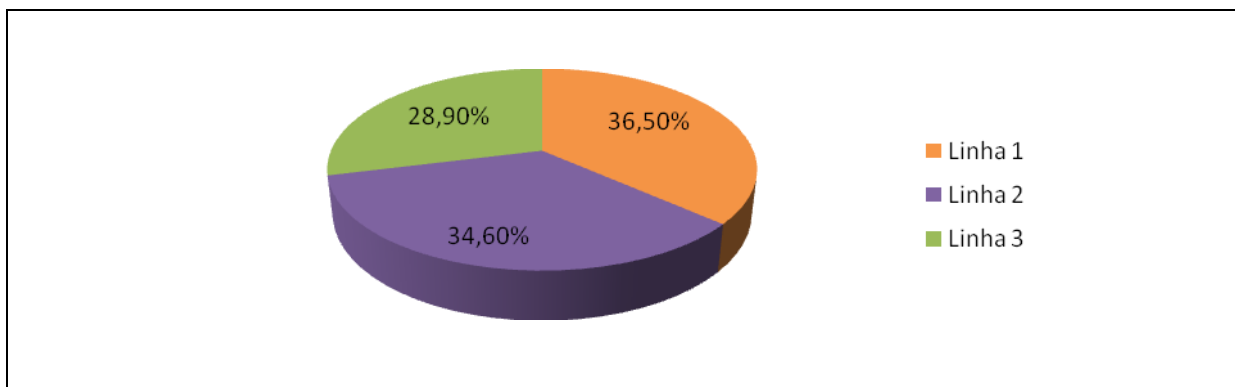


Figura 20 – Divisão dos pós-graduandos por linha de pesquisa

Assim como no mapeamento dos campos de origem dos pós-graduandos observava-se o caráter interdisciplinar da Ciência da Informação, na análise dos campos e áreas presentes nos projetos de dissertação e tese reforça-se a ideia da interdisciplinaridade. De acordo com Japiassu e Marcondes (1991 *apud* PINHEIRO, 2006, p. 113), interdisciplinaridade é o método de pesquisa e de ensino suscetível de fazer com que duas ou mais disciplinas interajam entre si. Esta interação pode ir da simples comunicação das ideias até a integração mútua dos conceitos, da epistemologia, da terminologia, da metodologia, dos procedimentos, dos dados e da organização da pesquisa.

No questionário, os participantes apontaram em quais campos e áreas do conhecimento, além da Ciência da Informação, estão inseridos seus respectivos projetos. As respostas foram classificadas de acordo com a Tabela de Áreas do Conhecimento do CNPq. Ciências Humanas e Ciências Sociais Aplicadas são as grandes áreas que incluem as mais citadas, mas, como se observa na Figura 10, estão presentes também algumas das Ciências Exatas e da Terra, das Ciências Biológicas, das Ciências da Saúde e de Linguística, Letras e Artes.

Na Figura 21, em formato de nuvem de etiquetas, estão representados os campos e áreas estudados pelos mestrands. O tamanho da fonte está diretamente ligado à frequência em que o termo foi citado. Percebe-se que Sociologia, Biblioteconomia e Filosofia foram as mais mencionadas pelos mestrands.



Figura 21 – Nuvem de etiquetas com campos e áreas do conhecimento inseridas nos projetos de mestrado das turmas de 2009 e 2010 do PPGCI/IBICT-UFRJ

Na Figura 22 é mostrada a representação dos campos e áreas apontados pelos doutorandos como integrantes de seus projetos. Em destaque apenas Filosofia, mas a nuvem de etiquetas revela a diversidade de focos de interesse dos doutorandos. Há desde estudos clássicos de Biblioteconomia até estudos sobre Judaísmo, Etnomusicologia e Física.



Figura 22 – Nuvem de etiquetas com campos e áreas do conhecimento inseridas nos projetos de doutorado das turmas de 2009 e 2010 do PPGCI/IBICT-UFRJ

Como afirma Pinheiro (2006, p. 113), as disciplinas têm enfoques específicos e o real significado de cada uma é sempre reduzido ao ângulo de visão particular dos seus especialistas, que se ampliaria na medida das interconexões com outras disciplinas. A autora cita Japiassu (1976 *apud* PINHEIRO, 2006, p. 113), que explica: “a pesquisa interdisciplinar se faz das aproximações, das interações e dos métodos comuns às diversas especialidades”.

Na Figura 23 foram combinadas as duas nuvens de etiquetas. Ao observar o resultado apresentado nas Figuras 21 e 22, bem como nas agregação exposta na Figura 23, conclui-se que, para ambos os grupos, mestrandos e doutorandos, as Ciências Humanas e as Ciências

Sociais Aplicadas são as áreas mais representadas. Por estarem relacionadas a disciplinas classicamente vinculadas à Ciência da Informação, nota-se presença destacada em ambos os níveis de pós-graduação, de campos como a Biblioteconomia, a Comunicação e a Ciência da Computação. Observa-se também que tanto nos projetos de mestrado quanto nos de doutorado, o campo da saúde é representado, assim como os da estatística e da linguística.



Figura 23 – Nuvem de tags com campos e áreas do conhecimento inseridas nos projetos de dissertação e tese dos pós-graduandos das turmas de 2009 e 2010 do PPGCI IBICT-UFRJ

Com os dados obtidos no questionário respondido pelos doutorandos, também foi possível notar se houve mudança no foco em relação aos campos envolvidos anteriormente na elaboração da respectiva dissertação. A Figura 24 indica que pelo menos a metade dos doutorandos deu continuidade aos estudos dentro do mesmo campo do conhecimento ou em um campo afim. Entre os demais, constatam-se mudanças radicais no foco, como no caso do doutorando que da área de Botânica transferiu seus estudos para uma combinação entre Filosofia, Sociologia, Psicologia e Linguística ou o doutorando que das questões sobre resíduos sólidos estudadas no mestrado passou a dedicar-se à tese relacionada a um entrecruzamento entre sobre física, judaísmo e neurociência.

Mestrado	Doutorado
Saúde Coletiva	Saúde Coletiva
Fotografia/Filosofia	Fotografia/Filosofia/Arte
Filosofia	Filosofia/Filologia
Administração	Administração/Filosofia
Sociologia/Administração	Sociologia/Antropologia
Comunicação/Cultura	Comunicação/Linguística
Biblioteconomia/Arquivologia/Museologia	Biblioteconomia/Arquivologia
Tecnológica	Ciência da Computação/Design
Linguística Aplicada	Educação/Administração
Política Científica e Tecnológica	Linguística e Comunicação
Engenharia de Produção	Ciência da Computação
Eng. de Prod./Gestão/Estratégia/Finanças Emp.	Planejamento/Avaliação/Des.Organizacional/Indicadores
Memória Social	Antrop./Música/Economia/História/Etnomusicologia
História/História das Ideias	Direito/Epistemologia da CI
Administração	Comunicação Social/Jornalismo Científico
Botânica	Filosofia/Sociologia/Psicologia/Linguística
Ciências Sociais	Estatística/Sociologia da Ciência
Eng. Sanitária/Resíduos Sólidos	Física/Judaísmo/Neurociência

Figura 24 – Comparação entre as áreas estudadas pelos atuais doutorandos para elaboração da dissertação e áreas inseridas na elaboração da tese

Saracevic (1999 *apud* PINHEIRO, 2006, p. 120) apresenta duas justificativas para a interdisciplinaridade da Ciência da Informação. A primeira, e mais importante, é que seus problemas não podem ser resolvidos por abordagens ou construtos de uma única disciplina. A segunda justificativa é que a interdisciplinaridade foi introduzida na área, e continua sendo até hoje, pelas muitas diferenças de formação (*background*) e áreas de origem das pessoas que tratam dos problemas da Ciência da Informação.

Ambas podem ser atestadas neste estudo, à medida que as pesquisas, tanto de mestrado quanto de doutorado, são baseadas em diferentes disciplinas, e que as áreas de origem dos pós-graduandos são diversificadas, como já apresentado na Figura 1.

6.3 Habilidades na busca e recuperação da informação

Ainda na primeira seção do questionário, sobre a caracterização do pós-graduando, foram inseridas duas perguntas com pelo menos dois objetivos. O primeiro deles foi obter informações sobre como o pós-graduando adquiriu conhecimento e/ou aperfeiçoou suas habilidades na busca e recuperação da informação na web. O segundo objetivo foi estimular a avaliação do participante sobre a atividade, a qual se inseriu de tal maneira no cotidiano das

peessoas atualmente, que comumente é executada sem muita reflexão sobre seus aspectos constitutivos.

Na primeira pergunta – *Por meio de qual forma você adquiriu conhecimento sobre como realizar o processo de busca e recuperação da informação na web?* – são oferecidas seis opções fechadas e uma aberta, na qual poderia ser descrita uma opção não citada anteriormente. Dessa lista, o participante podia escolher apenas duas modalidades, sendo que uma deveria ser apontada como a mais importante e outra, como a segunda mais importante. As opções apresentadas foram as seguintes:

- a) Experiência própria, aprendendo na prática;
- b) Com colegas, amigos, parentes, conhecidos;
- c) Treinamento ou curso especializado sobre o assunto;
- d) Leitura de livros e/ou artigos especializados sobre o assunto;
- e) Leitura de conteúdos sobre o assunto em sites na Internet;
- f) Curso a distância, via internet;
- g) Outra modalidade. Qual?

Entre as respostas dos mestrados sobre como adquiriram conhecimento e desenvolveram a habilidade no processo de busca, sobressaem-se a alternativa a) *Experiência própria, aprendendo na prática* como primeira opção e b) *Com colegas, amigos, parentes, conhecidos* como segunda, como se vê na Figura 25.

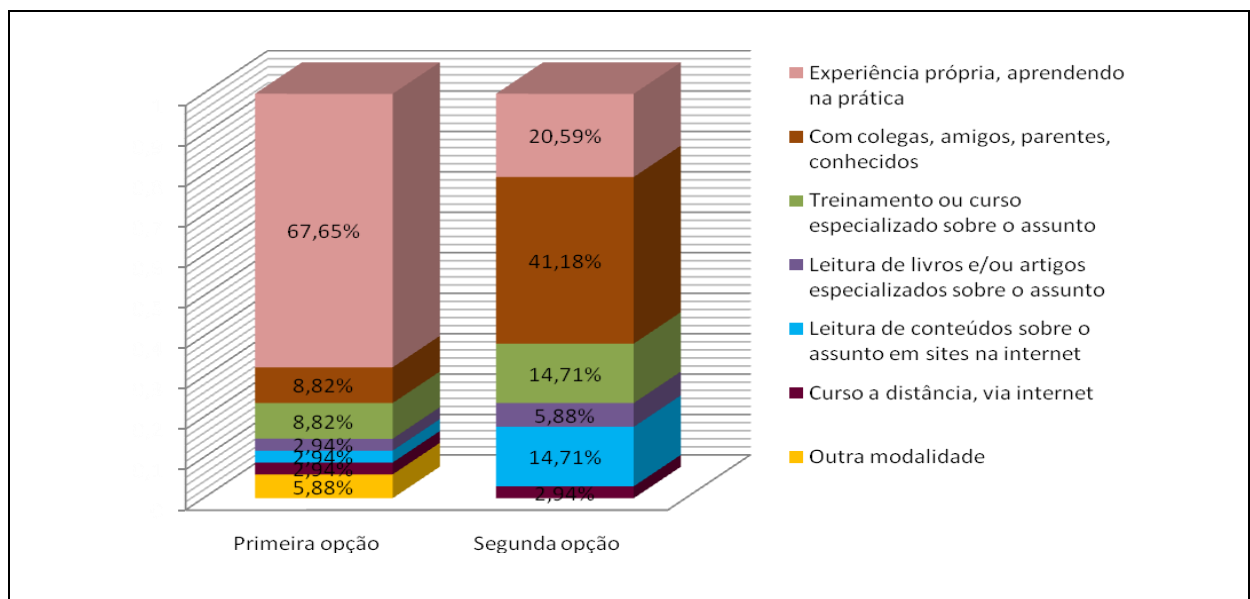


Figura 25 – Primeira e segunda maneiras pelas quais mestrados adquiriram conhecimento sobre busca na web

Entre os doutorandos, a escolha mais frequente como primeira opção mais importante também foi *a*). Quanto à segunda opção mais importante, os doutorandos assinalaram principalmente a alternativa *b*). As respostas compõem a Figura 26.

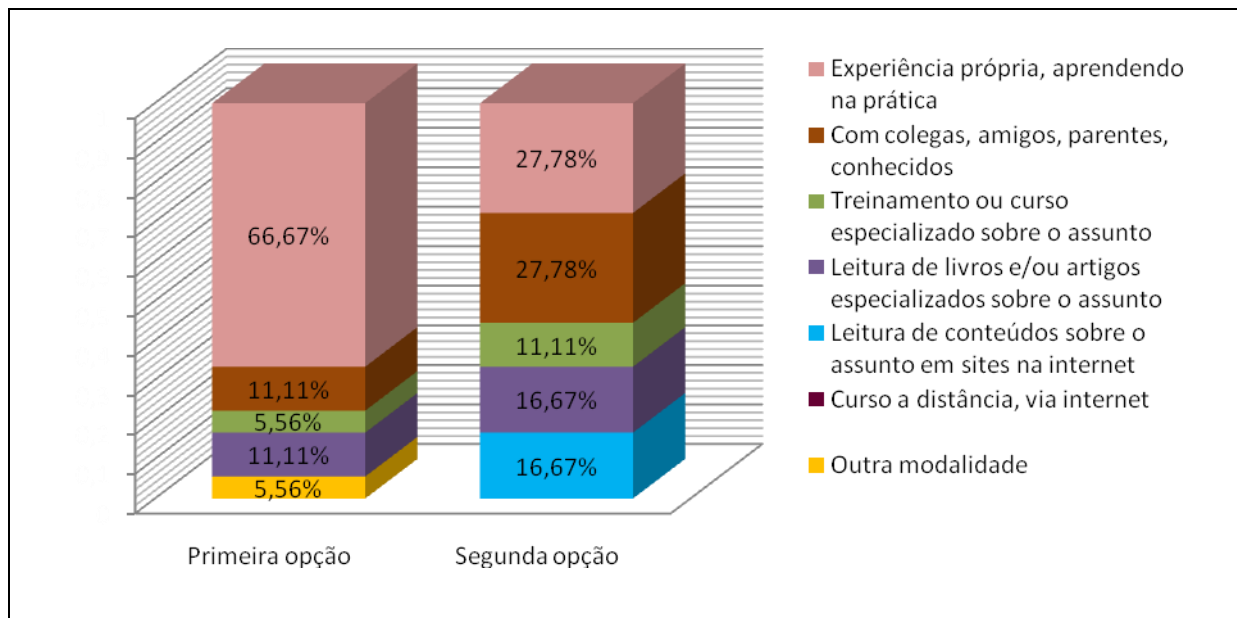


Figura 26 – Primeira e segunda maneiras pelas quais doutorandos adquiriram conhecimento sobre como realizar buscas na web

Em seguida, foram reunidas na Figura 27 as opções mais selecionadas por mestrandos e doutorandos. O gráfico conjunto assemelha-se mais à representação das opções dos mestrados.

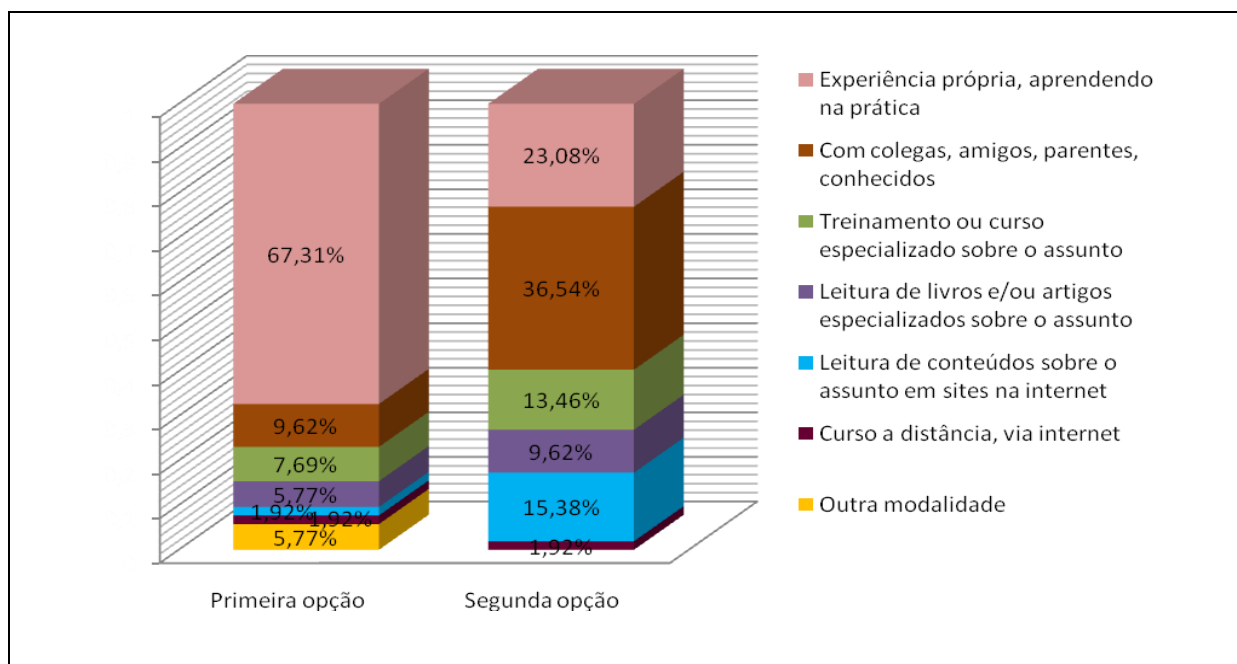


Figura 27 – Primeira e segunda opções por meio das quais pós-graduandos adquiriram conhecimento sobre como realizar busca e recuperação na web

A alternativa *a) Experiência própria, aprendendo na prática* desponta como a maneira mais frequente por meio da qual os participantes adquiriram conhecimento sobre como buscar e recuperar informações na web – o equivalente a 67% da população-alvo. Combinando os resultados da opção *a)* como a primeira mais importante e como a segunda mais importante, chegamos a 90% de ocorrência. A segunda opção mais importante, destacada por 36% dos participantes, foi a *b) Com colegas, amigos, parentes, conhecidos*.

De acordo com as respostas dos pós-graduandos, as combinações mais frequentes foram: *a) Experiência própria, aprendendo na prática + b) Com colegas, amigos, parentes, conhecidos* (com 19 ocorrências) seguida por *a) Experiência própria, aprendendo na prática + e) Leitura de conteúdos sobre o assunto em sites na internet* (com sete ocorrências) e *a) Experiência própria, aprendendo na prática e d) Leitura de livros e/ou artigos especializados sobre o assunto* (com cinco ocorrências). A Figura 28 representa as combinações registradas.

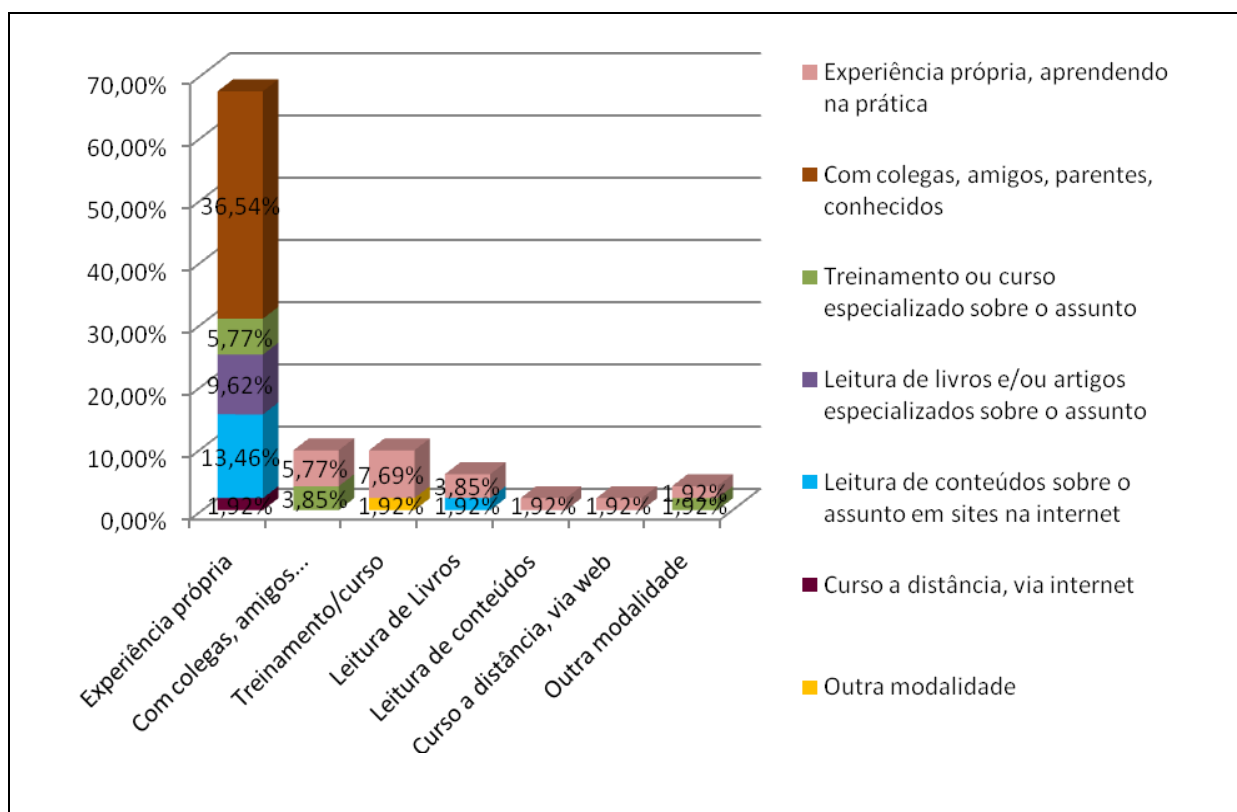


Figura 28 – Frequência com que as opções foram combinadas

Em apenas três das 52 respostas (6%), ocorreram combinações não incluindo as alternativas *a) Experiência própria, aprendendo na prática* e *b) Com colegas, amigos, parentes, conhecido*. A partir desse resultado, presume-se que neste grupo o aprendizado ocorre especialmente de forma autodidata e que pessoas próximas aos pós-graduandos

exercem papel importante na troca de informações e de experiências que levam ao desenvolvimento das habilidades na busca e recuperação da informação na web.

Após essa visão geral sobre os resultados obtidos nesta questão, partiu-se para a análise considerando os campos de origem dos pós-graduandos. Esta análise foi iniciada pela formação em Biblioteconomia, por se tratar da formação mais frequente entre os participantes, como também por ser uma das primeiras disciplinas a atuar no campo da informação – Biblioteconomia, Ciência da Computação, Ciências Cognitivas e Comunicação –, as quais, segundo Saracevic (2010, p. 2), “continuam tendo forte interação com a Ciência da Informação”.

Apesar de as alternativas *a) Experiência própria, aprendendo na prática + b) Com colegas, amigos, parentes, conhecidos* serem a combinação mais frequente entre as respostas da população-alvo, entre os profissionais oriundos da Biblioteconomia ocorre variação no resultado. Dos 19 bibliotecários envolvidos na pesquisa, apenas dois (10,5%) escolheram a combinação citada. Nos demais casos, tanto a alternativa *a)* quanto a alternativa *b)* estão presentes em 16 respostas (84,2%), mas sempre junto com alternativas relacionadas a cursos, treinamentos, disciplinas da graduação ou leitura de conteúdos sobre o tema. Presume-se então que bibliotecários, também por estarem inseridos em uma área que tradicionalmente trabalha com busca e recuperação da informação, vão além da formação autodidata ou do aprendizado realizado de maneira informal com amigos, conhecidos ou colegas.

No cruzamento dos resultados com o ano de graduação dos participantes, não foram observados nas respostas padrões que indicassem tendências de comportamento em decorrência do tempo de formado.

6.4 Capacidade de buscar e recuperar informação na web

A primeira parte do questionário foi encerrada com autoavaliação da capacidade de buscar e recuperar informações na web. A pergunta antecede a etapa em que o participante relata em detalhes uma experiência de busca na web. Ao inserir a pergunta nesta etapa do questionário pretendeu-se estimular o pós-graduando à rápida reflexão sobre o quão preparado se sentia ao realizar buscas na web. Pergunta semelhante foi realizada depois que o usuário completou o relato sobre a vivência.

Nesta primeira autoavaliação as notas conferidas pelos pós-graduandos às suas competências variaram de 5 a 10, sendo que a nota 8 foi a mais frequente, tanto entre mestrandos quanto doutorandos, como pode ser observado na Figura 29. Entre os doutorandos, a nota 7 ocorreu mais vezes, ocupando o segundo lugar na autoavaliação, seguida da nota 9, indicando uma maior concentração na faixa situada entre 7 e 9, em comparação com os mestrandos, cuja distribuição é mais diluída entre os diferentes graus autoatribuídos, à exceção da nota 8. Coincidentemente, a soma dos resultados para as notas 8, 9 e 10, conferidas tanto por mestrandos quanto por doutorandos, resulta em 61%. A maior parte dos pós-graduandos concedeu-se nota superior a 7. No grupo dos mestrandos, a média de autoavaliação corresponde a 7,53, enquanto que para os doutorandos equivale a 7,72.

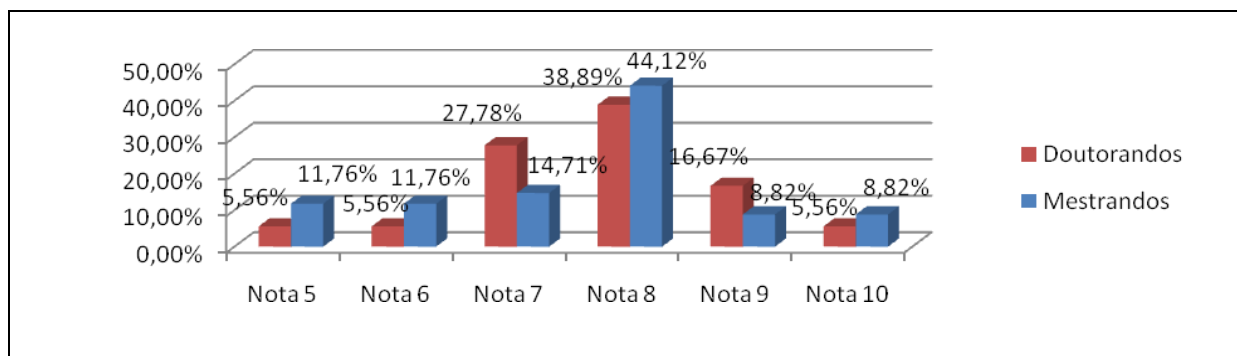


Figura 29 – Autoavaliação da capacidade de buscar e recuperar informações na web - primeira etapa

A segunda autoavaliação, após o relato de uma experiência de busca, teve como objetivo estimular a reflexão sobre o processo recém-realizado e incentivar o exercício de metacognição. Para Weinert (1987 *apud* RIBEIRO, p. 110), as metacognições podem ser consideradas cognições de segunda ordem: pensamentos sobre pensamentos, conhecimentos sobre conhecimentos, reflexões sobre ações. Matlin (2003 *apud* ANDRETTA et al., 2010, p. 8) refere-se à metacognição como o conhecimento que a pessoa tem de seus próprios processos cognitivos, isto é, o conhecimento, a consciência e o controle dos processos cognitivos. De acordo com Ribeiro (2003, p. 110), a metacognição diz respeito, entre outras coisas, ao conhecimento do próprio conhecimento, à avaliação, à regulação e à organização dos próprios processos cognitivos.

Cotidianamente, a metacognição pode auxiliar na execução de muitas tarefas, inclusive na seleção de estratégias de memória que sejam mais adequadas para determinadas situações, o que possibilita maior economia de tempo e melhor aprimoramento do conhecimento adquirido. Andretta et al. recorrem a Jou e Sperb (2006) para declarar que a compreensão que as pessoas têm do próprio processamento cognitivo envolve a percepção

dos pensamentos e das ações realizadas. O sistema cognitivo monitora, planeja e regula seus processos, e estes se tornam mais conscientes com o passar do tempo. Sendo assim, o indivíduo adquire a capacidade de monitoramento, autorregulação e elaboração de estratégias para aumentar sua cognição.

Peixoto (2007 *apud* ANDRETTA et al., 2010, p. 9) afirma que o conhecimento metacognitivo permite decidir sobre eventos, tais como prosseguir ou não no ritmo atual de estudo, intensificar esforços, reduzir o empenho ou abandonar a tarefa. A partir dessa noção, o autor afirma que a metacognição desenvolve-se a partir da capacidade do homem de refletir sobre seu processo de conhecimento, durante a realização de tarefas, sobre os processos mentais que facilitam essa realização e sobre as estratégias que utiliza para a resolução de problemas.

As respostas dos pós-graduandos para a segunda avaliação estão representadas na Figura 30.

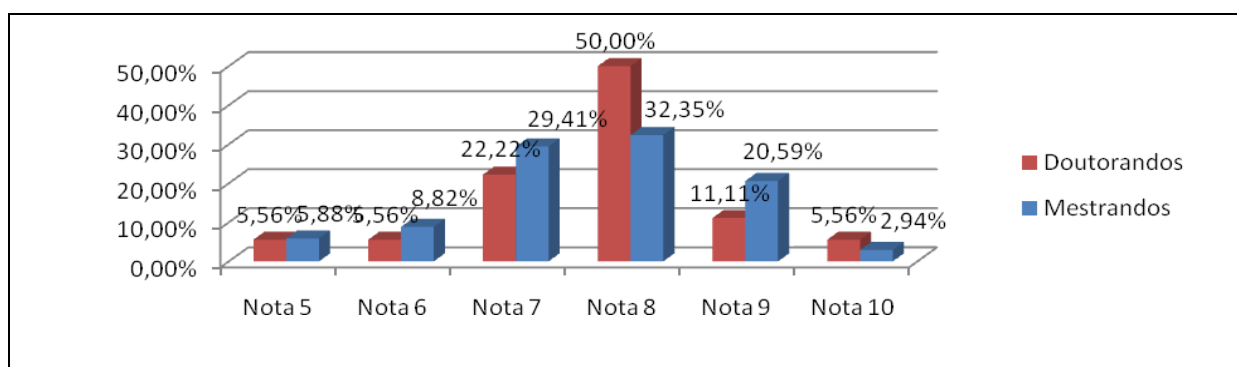


Figura 30 – Autoavaliação da capacidade de buscar e recuperar informações na web - segunda etapa

Os dados demonstram que as autoavaliações permanecem no intervalo compreendido entre as notas 5 e 10, para ambos os grupos, mestrandos e doutorandos.

Constatou-se que dos doutorandos, 11 (61,11%) não alteraram a avaliação. Entre os mestrandos, o mesmo ocorreu com 58,82%, ou seja, 20 deles decidiram permanecer com a mesma avaliação. Enquanto que entre os doutorandos, 4 (22,22%) aumentaram a nota, contra 3 (16,67%) que a diminuíram, entre os mestrandos, 7 (20,58%) cresceram-na e 7 (20,58%) a reduziram em um ponto. Diferente do que ocorreu na primeira avaliação, quando tanto mestrandos quanto doutorandos apresentavam notas superiores a 7, empatando em 61% na totalização das avaliações compreendidas entre 8 e 10, nesta segunda avaliação mais doutorandos apareceram com nota igual ou superior a 8 (66,67%).

No entanto, como o número dos doutorandos que reduziram o grau de sua autoavaliação foi equivalente nos subgrupos de nota 7 e 9, migrando para a nota 8, a média global permaneceu inalterada, equivalendo aos mesmos 7,72 da primeira autoavaliação. Entre os mestrandos, verificou-se queda, sobrando 55,88% do grupo com nota igual ou acima de 8. Tal resultado se refletiu também na média global do grupo dos mestrandos, a qual reduziu para 7,09 na segunda autoavaliação.

Durante as entrevistas foram apuradas algumas das razões que levaram os pós-graduandos a alterar a nota. Uma mestranda, que havia inicialmente se autoavaliado com nota 8 e modificou a avaliação para 7, justifica dizendo que a alteração ocorreu após a análise do desempenho ao responder ao questionário. Em rápida reflexão, a pós-graduanda avaliou que não encontra muito do que procura na web e demora tempo demais quando consegue recuperar a informação desejada. Outros participantes baixaram a nota por acreditar que, apesar de dominar as técnicas de busca, sempre há algo a ser melhorado. Já um mestrando que aumentou em um ponto sua nota na segunda avaliação justificou que ao longo dos processos de busca que desenvolve descobriu macetes que melhoraram sua atuação.

Retomando a questão da metacognição, Brown, ao tratar sobre o comportamento de jovens leitores, conclui que a dificuldade na compreensão de uma tarefa, ou tornar-se consciente de que não se compreendeu algo, é uma habilidade que pode distinguir os bons dos maus leitores. Os primeiros sabem avaliar suas dificuldades e/ou ausências de conhecimento, o que lhes permite, nomeadamente, superá-las, recorrendo, muitas vezes, a inferências feitas a partir daquilo que sabem. (BROWN, 1978 *apud* RIBEIRO, 2003, p. 110).

Andretta et al. (2010, p. 12) concluem que o indivíduo experimenta diversas estratégias de aprendizagem desde criança, tendendo sempre a repetir aquela com a qual obteve maior eficácia. Assim, cada um tem uma estratégia que funciona melhor para si; o processo metacognitivo é diferente para cada um. Quanto melhor a capacidade metacognitiva, melhores as habilidades de aprendizagem, sendo possível aperfeiçoar o tempo e a qualidade do estudo. A mesma regra vale para o desempenho no processo de busca e recuperação da informação.

6.5 A experiência de busca na web

A primeira decisão tomada pelos pós-graduandos ao responder a segunda parte do questionário foi escolher um exemplo recente de experiência de busca na web que houvesse

resultado em conteúdo relevante para o projeto de mestrado ou doutorado. A pergunta utilizada foi a seguinte:

2.1. Baseando-se em um exemplo recente de uma busca realizada na web que tenha resultado na recuperação de conteúdo relevante para o seu projeto de mestrado, responda? (artigos em periódicos, páginas da web, arquivos em formato pdf, doc ou ppt, teses e dissertações, sites, blogs, livros, citações, páginas não encontradas, etc.) Há quanto tempo ocorreu a busca? Há () horas / Há () dias / Há () semanas.

De acordo com o declarado pelos participantes, quase 40% dos pós-graduandos relataram experiências vivenciadas nos seis dias anteriores à aplicação do questionário. Onze por cento realizaram a busca no dia anterior. Houve casos de relatos referentes a experiências mais antigas, como nota-se no Figura 31. Observou-se também que 19% não definiram o número de horas, dias ou semanas.

No artigo de Flanagan (1954) sobre a técnica do incidente crítico, o autor apresenta estudo realizado em 1949 no qual profissionais de uma mesma área coletam experiências semelhantes em momentos diferentes. O primeiro grupo registrou os incidentes no mesmo dia em que ocorreram; o segundo, ao final de uma semana; o terceiro, depois de duas semanas. Os primeiros relataram 315 incidentes; os segundos, 155; e os terceiros, 63.

Com isso, pôde-se inferir que relatos mais recentes tendem a ser mais completos e ricos em detalhes do que aqueles relacionados a situações ocorridas há mais tempo.

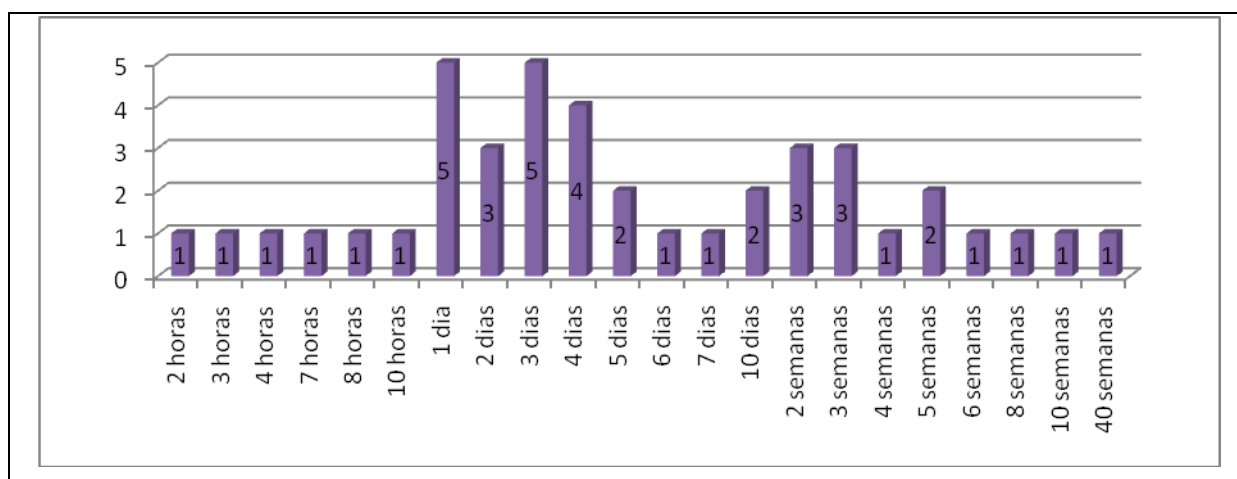


Figura 31 – Quantidade de tempo transcorrida desde a realização da experiência relatada

Seguindo orientação prévia, as buscas descritas pelos pós-graduandos deveriam ter relação com os respectivos projetos de dissertação ou tese. Porém, sete dos 52 participantes, optaram por relatar buscas recentes relacionadas a trabalhos de aula, conteúdo que

posteriormente poderá ou não ser incluído em suas teses e dissertações. Em comum, todos os sete integravam a turma de pós-graduandos ingressados em 2010 e, no momento da pesquisa, cumpriam parte da carga horária estipulada no regulamento do PPGCI. Estes pós-graduandos, em especial, ainda encontravam-se em fase de coleta de artigos para delimitação de seus projetos ou início da redação da dissertação.

Dois participantes, apesar de buscarem por informações no âmbito da Ciência da Informação, não as utilizariam em seus projetos, mas sim em evento científico da área e para elaboração de uma aula para o curso de Biblioteconomia. Houve ainda o caso de um participante que relatou busca relacionada à área de formação e ao tema do projeto desenvolvido, mas que não se tratava de tema acadêmico.

6.6 Identificação do problema

Na primeira questão, foram abordados aspectos do problema, assim como da situação, que levou à busca, a partir da seguinte pergunta: 2.2. *Que situação o levou a realizar essa busca? Qual era o problema a ser solucionado?* Dois pós-graduandos relataram não ter um problema específico a resolver, mas que teriam sido movidos um por curiosidade sobre determinada combinação de palavras inserida em sua tese; outro, pela indicação de uma amiga.

As demais respostas foram então agrupadas por assuntos relacionados. Entre os demais 42 participantes, os problemas apontados puderam ser encaixados dentro de quatro categorias de assuntos buscados.

Quase metade dos pós-graduandos, 25 deles (48%), procuravam por literatura relacionada a assuntos inseridos em seus projetos, como se constata nas seguintes respostas: “Levantamento de bibliografia sobre avaliação de sistemas de informação para uso na tese”; “Necessidade de buscar informações específicas sobre o movimento de arquivos abertos e direito de autor. O problema era procurar entender como a CI está vendo a questão dos direitos autorais”; “Preparar material para leitura – levantamento bibliográfico de trabalhos acadêmicos sobre o uso de novas tecnologias no fazer historiográfico”.

Na sequência, 10 participantes (19%) sabiam exatamente o material específico a ser localizado, como, por exemplo, nas respostas: “Encontrar a definição de termos categoria e

categorização”; “Saber o quanto o livro de *Comunicação Científica* de Meadows é citado em artigos da área de Ciência da Informação, especialmente no que diz respeito a acesso livre”.

Em três casos (6%), os pós-graduandos pretendiam localizar conteúdos de determinado autor: “Identificar a produção de determinado autor na *Web of Science*”; “Buscar textos de um autor indicado pela banca de qualificação”; e “Interesse em buscar informações sobre um autor citado em sala de aula”.

Os dois últimos participantes (4%) buscavam por conteúdos semelhantes ao já inseridos em seus projetos: “Verificar se existe conteúdo semelhante ao meu projeto em artigos e blogs”; “Precisava saber se já havia classificação específica para a ‘cultura’ de forma abrangente, para saber se era possível utilizar na elaboração de estatísticas”.

6.7 Google, mecanismo de busca preferido

Após expor o problema e/ou situação que o levou à busca relatada, cada pós-graduando indicou a informação específica procurada e os termos inseridos no campo de busca. No entanto, antes de entrar na análise dessas informações, torna-se necessário apresentar os mecanismos de busca escolhidos e apontar se foi utilizado recurso padrão ou avançado. Após essas considerações será dada continuidade à avaliação do comportamento dos participantes.

O *Google*, que conquistou a posição de mecanismo de busca mais popular dos anos 2000, aparece em primeiro lugar na preferência dos mestrandos e empata com o *Google Acadêmico* na escolha dos doutorandos, como mostra a resposta da pergunta seguinte: 2.5. *Por meio de qual recurso informacional deu início à busca? (ex.: site, ferramenta de busca, mecanismo de busca, base de dados).* Os exemplos inseridos nesta pergunta com os objetivos de facilitar a compreensão e mostrar sinônimos para mecanismo de busca acabou gerando dúvidas em diversos participantes, que perguntaram sobre as diferenças entre as expressões. Alguns participantes responderam à questão dizendo ter usado ferramenta de busca, mas sem nomear qual foi especificamente. Essa informação foi recuperada durante a entrevista.

Dezenove dos 34 mestrandos participantes da pesquisa, o equivalente a 55,82%, recorreram ao *Google*, como pode ser visto na Figura 32. Cinco (14,7%) optaram por bases especializadas, especialmente duas mestrandas que basearam seu projeto no material recuperado de apenas dois sites, uma no *Web of Science*, outra no *Science Direct*.

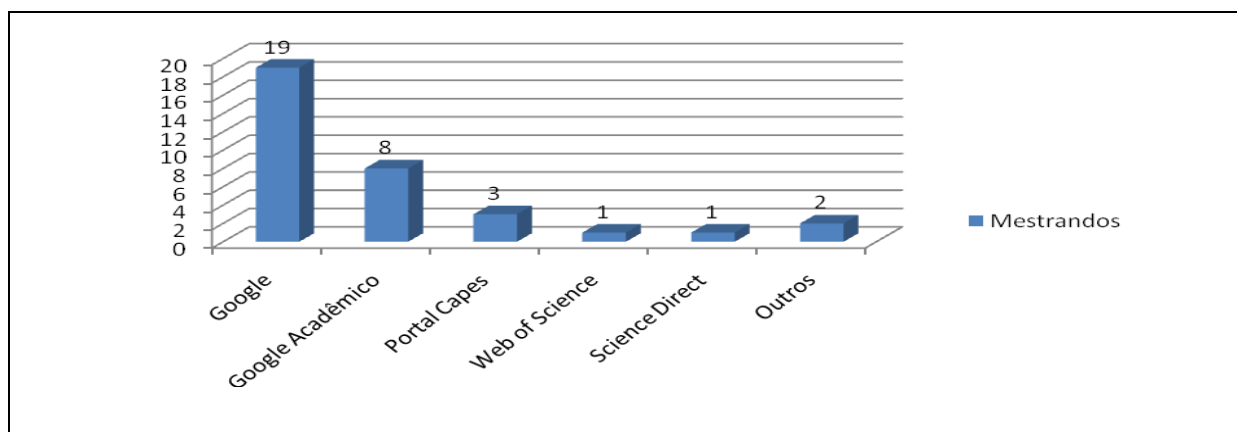


Figura 32 – Mecanismos de busca mais utilizados pelos mestrandos

Entre os doutorandos, *Google* e *Google Acadêmico* foram mencionados, cada um por seis (33,33%) pós-graduandos.

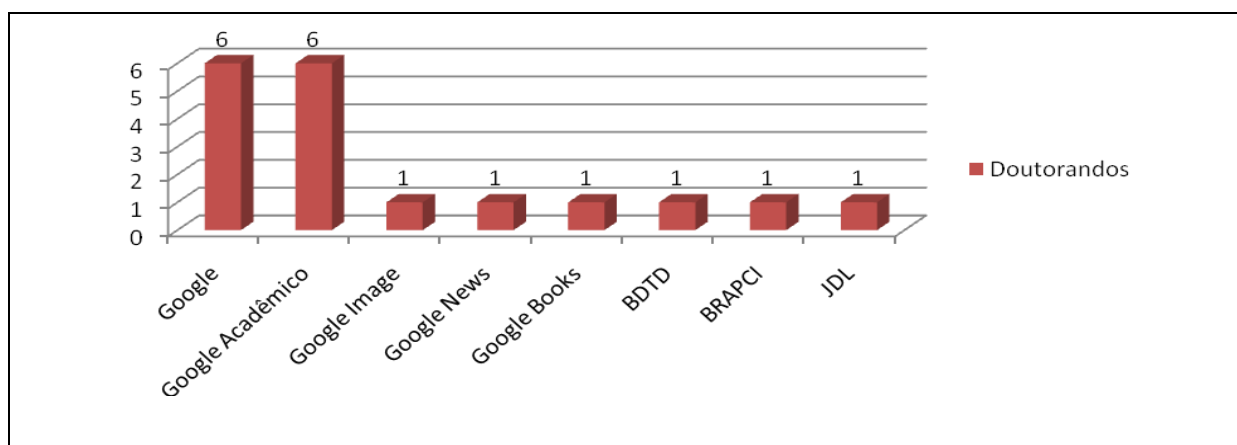


Figura 33 – Mecanismos de busca mais utilizados pelos doutorandos

Ao reunir as respostas de ambos os grupos em apenas um gráfico, fica evidenciada a preferência pelo *Google* e pelo *Google Acadêmico* tanto entre mestrandos quanto entre doutorandos, como se vê na Figura 34.

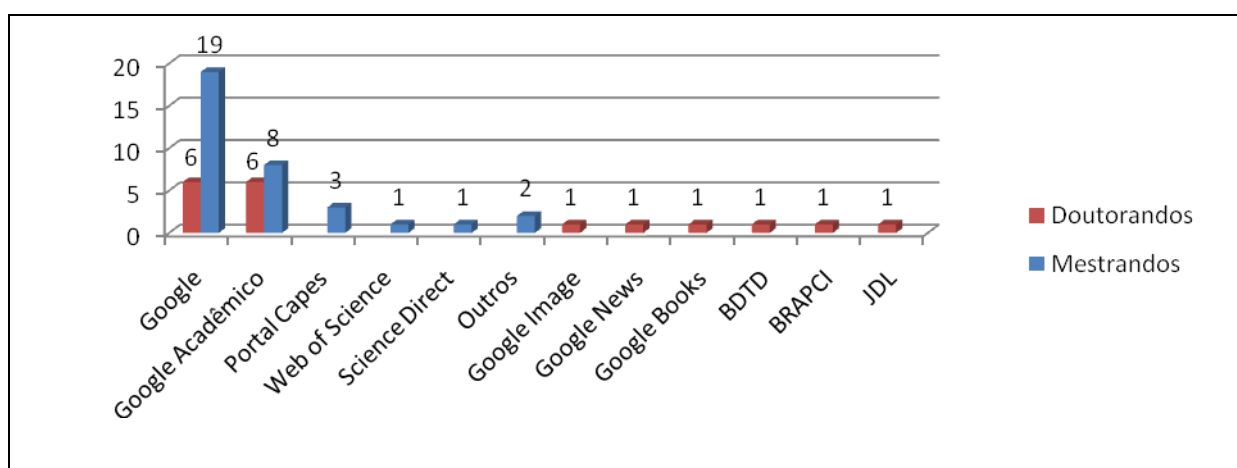


Figura 34 – Mecanismos de busca mais utilizados pelos pós-graduandos

Após indicar o mecanismo de busca empregado, os pós-graduandos foram questionados sobre as razões que levaram à escolha, na pergunta:

2.6. *Que motivo o levou a recorrer a essa fonte para buscar a informação de que necessitava?*

Rápido, fácil e prático são os adjetivos mais comuns atribuídos ao *Google* pelos mestrandos. A partir dos resultados abrangentes apresentados pelo mecanismo é então realizada nova busca para refinar as respostas obtidas. “O *Google* se mostra como um primeiro caminho, que já apontou referências relevantes em outras buscas, opção genérica, a partir da qual depois se pode realizar buscas em sites mais específicos”, afirma um mestrando. Outra característica apresentada pelos pós-graduandos é o fato de o mecanismo reunir informações de vários sites, de bases de dados e repositórios e ter a possibilidade de recuperar rapidamente conteúdos desses diferentes lugares, “poupando o tempo do pesquisador”. Outro mestrando é categórico: “É o site de busca que sempre utilizo e com o qual sinto mais satisfação.”

Durante a análise observou-se que muitos participantes declararam ser o *Google* o local inicial da busca, que tem continuidade em outras bases de informação, algumas vezes com consultas mais elaboradas e complexas. Uma dúvida surgida a partir dessa declaração refere-se ao fato de não ter sido relatada no questionário essa busca seguinte, já em outro mecanismo e ocorrida depois de maior reflexão.

No caso dos doutorandos, como recém-atestado, 33,33% preferem o *Google*, mas mesmo percentual utiliza o *Google Acadêmico*. Assim como para os mestrandos, o *Google* serve como primeira fonte para os doutorandos, aquele recurso que resultará no maior volume de resultados, que mais tarde será refinado. Intimidade com o mecanismo, agilidade, praticidade, abrangência e a “certeza de encontrar uma resposta” são alguns dos aspectos ressaltados, tanto para o *Google* quanto para o *Google Acadêmico*. “Escolhi o *Google* por ser o principal portal agregador de notícias”, justifica um doutorando. “Sempre começo pelo *Google*”, comenta outro.

A “família” *Google* aparece bem representada entre os doutorandos. Além da versão tradicional, são utilizados também o *Google News* – “por ser o principal agregador de notícias” –, o *Google Image* – “quando se quer estudar fotografia ou fotografos, ferramentas de busca por imagem são mais interessantes” – e o *Google Books* – “porque os livros consolidam artigos, facilitando a compreensão de contextos”.

Em terceira posição figura o *Portal Capes* (Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES), nesta pesquisa apontado apenas por mestrandos. Ao Portal Capes são atribuídas qualidades como confiabilidade e abrangência no âmbito da Ciência da Informação.

Dos 52 participantes, apenas 10 (13%) utilizaram bancos de dados especializados, vide Figura 35, como a Biblioteca de Teses e Dissertações (BDTD) do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) ou a Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI), da Universidade Federal do Paraná (UFPR).

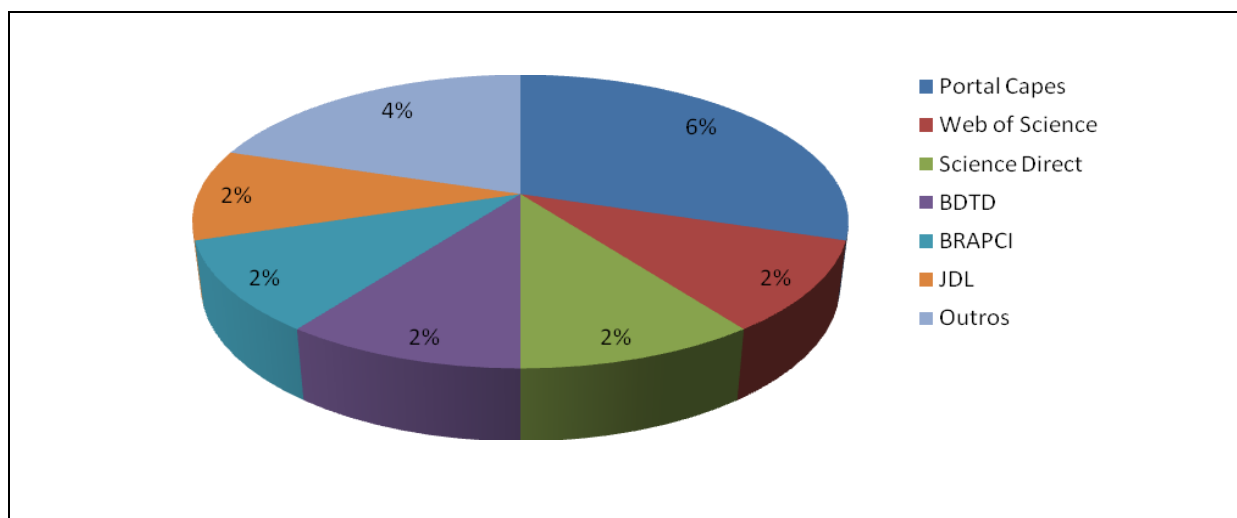


Figura 35 – Bases especializadas usadas pelos pós-graduandos

Doze (23% dos participantes) dos 27 pós-graduandos que empregaram mecanismos que não o *Google – Google Acadêmico*, Portal Capes, BDTD – têm formação em Biblioteconomia.

6.8 Recurso padrão x recurso avançado

A conduta dos indivíduos ao buscar informações na web, como constatado em pesquisa semelhante no campo da saúde (IVANITSKAYA; CHES; CASEY, 2006), aponta que baixa porcentagem de consumidores de informação usa recursos avançados em suas buscas, especifica precisamente as palavras-chave ou limita suas buscas de alguma maneira.

Isso se comprovou também nesta pesquisa com as respostas à pergunta 2.7. *Durante este processo de busca, você utilizou recurso padrão ou recurso avançado oferecido pelo mecanismo escolhido? (ex.: busca avançada, pesquisa avançada) Se usou recurso avançado, especifique um pouco mais quais foram os campos utilizados e por quê.*

Ao relatarem a experiência de busca, 26 (76,4%) dos 34 mestrandos confirmaram utilizar o recurso padrão do mecanismo escolhido ao invés da opção avançada.

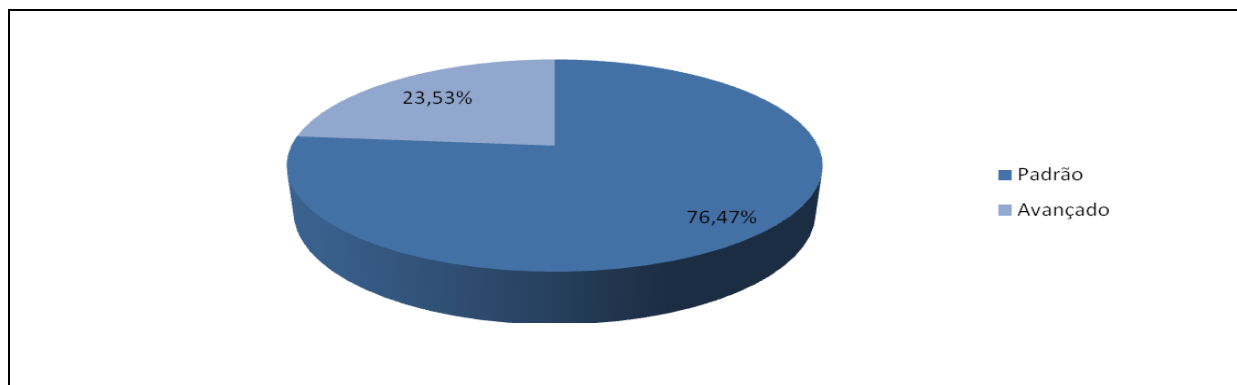


Figura 36 – Representação da escolha dos mestrandos por recurso padrão ou avançado

Entre os doutorandos, o recurso padrão foi utilizado por 16 pós-graduandos (89%), um percentual ainda maior.

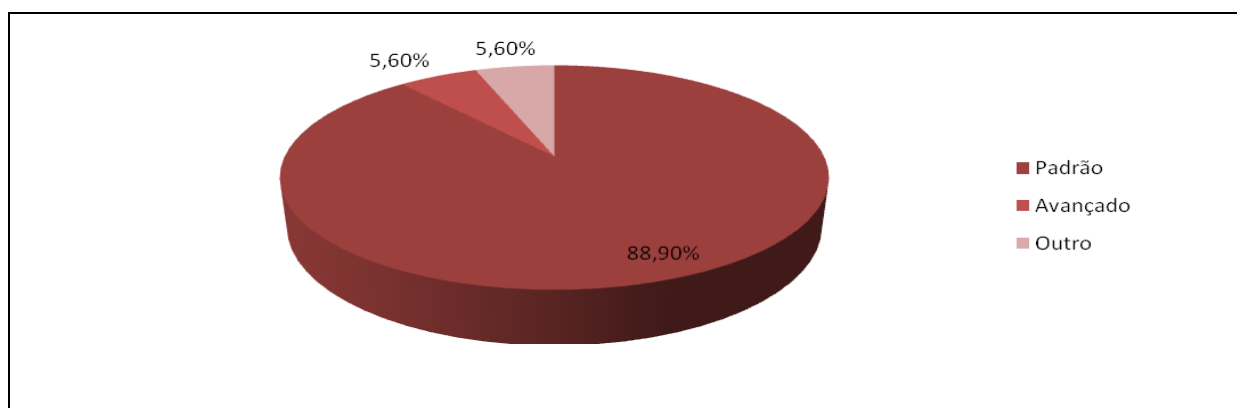


Figura 37 – Representação da escolha dos mestrandos por recurso padrão ou avançado

Assim como ocorreu na escolha do mecanismo de busca que será utilizado no início de uma busca por informações na web, parte pequena dos participantes afirmou utilizar o recurso avançado ao começar nova busca. “Já tentei usar o avançado, mas eu não gosto muito dos filtros, a não ser que seja algo muito específico, o que não foi o caso nesta pesquisa relatada, pois primeiro queria vasculhar a web, ter uma visão mais abrangente.”

Doze por cento afirmam começar a busca sempre pelo modo mais simples, ou seja, o recurso padrão, para depois partir para um refinamento, que é feito pesquisando nos resultados obtidos ou fazendo novos cruzamentos de palavras-chave. “No início, utilizei padrão, mas como queria um artigo na íntegra e com o uso de recurso padrão não o encontrava, recorri ao critério avançado, restringindo a busca a artigos acadêmicos do autor procurado.”

A utilização de recurso avançado, na opinião dos pós-graduandos, aumenta a precisão e filtra de maneira mais efetiva o conteúdo desejado. Além dessa forma de obter resultados melhores, 12% da população-alvo afirmam utilizar operadores booleanos ao realizarem suas buscas.

Os operadores booleanos surgiram no século XIX, criados pelo matemático George Boole. A chamada álgebra booleana é um sistema binário no qual existem somente dois valores possíveis para qualquer símbolo algébrico: 1 ou 0, verdadeiro ou falso. Como revisa Ferneda (2003, p. 22), no modelo booleano um documento é representado por um conjunto de termos de indexação que podem ser definidos de forma intelectual (manual) por profissionais especializados ou automaticamente, por meio da utilização de algum tipo de algoritmo computacional. As buscas são formuladas por meio de expressão booleana composta por termos ligados através dos operadores lógicos AND, OR e NOT (E, OU e NÃO) e apresentam como resultado os documentos cuja representação satisfaz às restrições lógicas da expressão de busca.

O autor resume o funcionamento dos três operadores em uma busca:

A expressão conjuntiva de enunciado t_1 AND t_2 recuperará documentos indexados por ambos os termos (t_1 e t_2). Esta operação equivale à interseção do conjunto dos documentos indexados pelo t_1 com o conjunto dos documentos indexados pelo termo t_2 , representado pela área cinza na Figura 38.

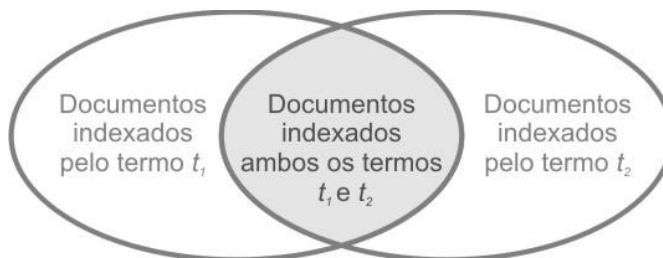


Figura 38 - Representação do resultado de expressão booleana conjuntiva (AND)

A expressão disjuntiva t_1 OR t_2 recuperará o conjunto dos documentos indexados pelo termo t_1 ou pelo termo t_2 . Essa operação equivale à união entre os conjuntos dos documentos indexados pelo termo t_1 e o conjunto dos documentos indexados pelo termo t_2 , como ilustrado na Figura 39.

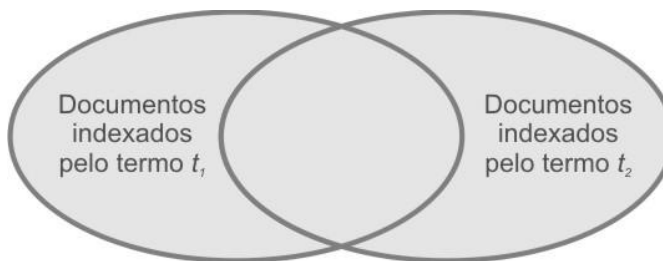


Figura 39 - Representação do resultado de uma busca booleana disjuntiva (OR)

A expressão que utiliza apenas um termo t_1 terá como resultado o conjunto de documentos indexados por t_1 . A expressão NOT recuperará os documentos que não são indexados pelo termo t_1 , representados pela área cinza da Figura 40.

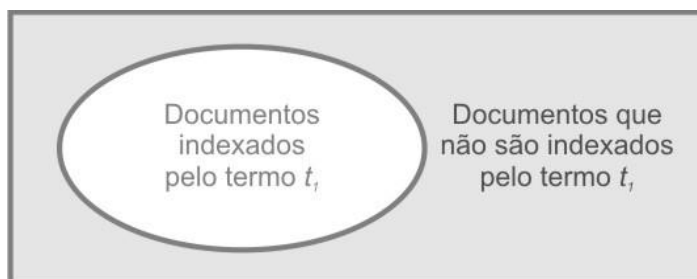


Figura 40 - Resultado de busca negativa (NOT)

As expressões t_1 NOT t_2 ou t_1 AND NOT t_2 terão o mesmo resultado: o conjunto dos documentos indexados por t_1 e que não são indexados por t_2 (Figura 41). Neste caso, o operador NOT pode ser visto como um operador da diferença entre conjuntos. Assim a área cinza da figura 8 representa o conjunto dos documentos indexados pelo termo t_1 menos (subtraído de) o conjunto dos documentos indexados por t_2 .

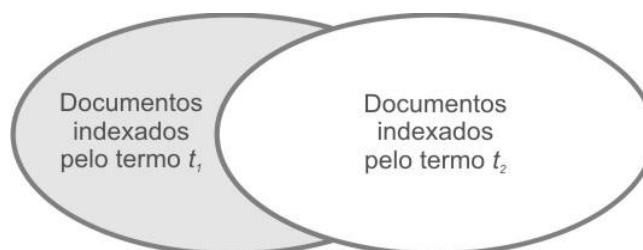
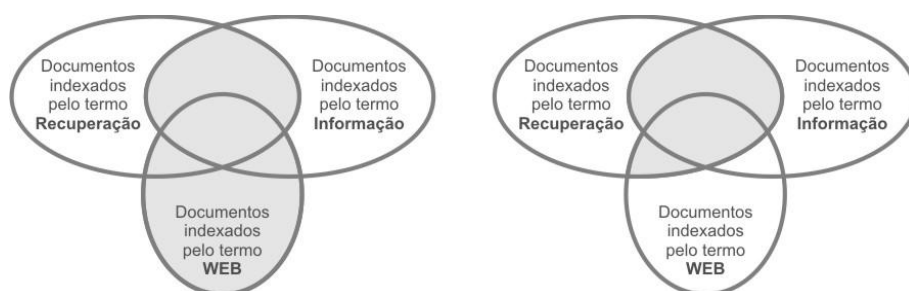


Figura 41 - Resultado de busca booleana com o operador NOT

Termos e operadores booleanos podem ser combinados para especificar buscas mais detalhadas ou restritivas. Como a ordem de execução das operações lógicas de uma expressão influencia no resultado da busca, muitas vezes é necessário explicitar essa ordem delimitando partes da expressão através de parênteses. Na ausência de parênteses, a expressão booleana será interpretada de acordo com o padrão utilizado pelo sistema, que pode ser a execução da expressão da esquerda para a direita ou em uma ordem pré-estabelecida, geralmente NOT – AND – OR.



a) (Recuperação AND Informação) web b) Recuperação AND (Informação OR web)

Figura 42 - Resultado de busca booleana utilizando parênteses

As áreas cinza da Figura 42 representam o resultado de duas expressões de busca que utilizam os mesmos termos e os mesmos operadores, mas diferem na ordem de execução. Na primeira expressão (a) inicialmente é executada a operação AND entre os termos “recuperação” e “informação”. Com o resultado obtido, é executada a operação OR com o termo “web”. A segunda expressão (b) executa a operação OR entre os termos “informação” e “web” e com o resultado é efetuada a operação AND com o termo “recuperação”. (FERNEDA, 2003, p. 22-23)

Ao se analisar como os termos foram inseridos no campo de busca pelos pós-graduandos, observa-se que na prática a utilização dos operadores booleanos mostra-se mais significativa. Ao relatarem suas experiências, especialmente na questão sobre os termos usados no campo de busca, 15 (28,8%) participantes, valeram-se da utilização do “+”. Um (1,9%) pós-graduando utilizou a forma clássica AND. Outro recurso adotado por 12 (23,1%) integrantes da população-alvo foi aspas, eficiente para definir uma cadeia de expressões e manter a ordem das mesmas.

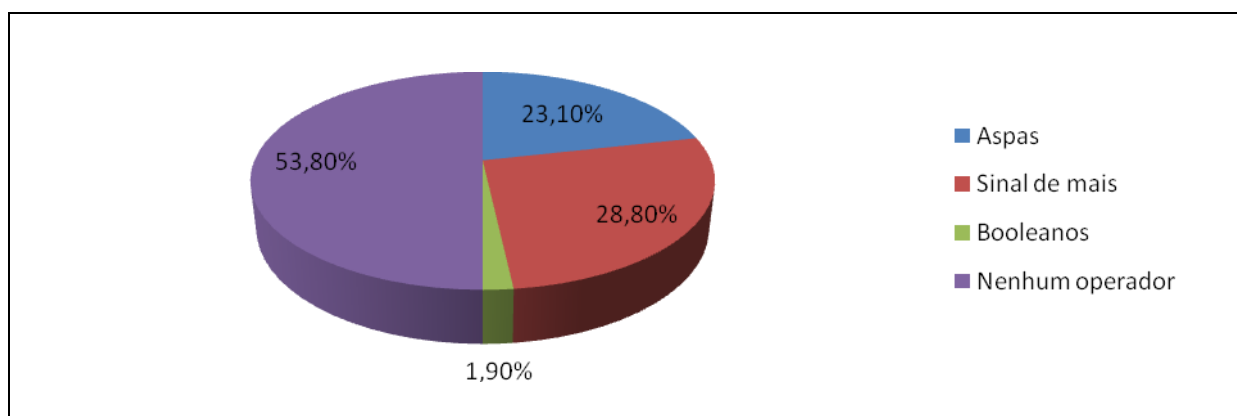


Figura 43 – Utilização de operadores para refinamento da busca

6.9 Da informação específica à elaboração dos termos da consulta

A fim de identificar coerências e incoerências nas buscas empreendidas pelos pós-graduandos, assim como estabelecer a correlação entre as diferentes etapas do processo de busca e recuperação realizado por cada um dos participantes da pesquisa, foi conduzida uma análise longitudinal das respostas.

Inicialmente, foram consideradas as respostas das três primeiras questões, nas quais o pós-graduando explicou o problema (2.2. *Que situação o levou a realizar essa busca? Qual era o problema a ser solucionado?*), descreveu o que definiu como informação específica (2.3. *Que informação específica você estava procurando ao recorrer à web? Isto é, qual era o assunto da busca? Qual era a questão que estava necessitando de uma resposta?*) e os termos escolhidos (2.4. *Quais termos/palavras foram inseridos no campo de busca?*). Em seguida, a análise foi complementada por respostas a outras perguntas, quando se fez necessário e pertinente.

A revisão das 52 experiências revelou pontos positivos e negativos comuns a diversos casos, relacionados a, por exemplo, delimitação da consulta; formato da consulta; escolha da

base de dados ou do mecanismo de busca; e uso de operadores booleanos e outros recursos que aumentam a precisão.

6.9.1 Delimitação da consulta

Determinadas experiências demonstram desatenção a detalhes que poderiam tornar mais eficiente o processo de busca e facilitar a avaliação e o aproveitamento dos resultados obtidos. Alguns exemplos em particular permitem ilustrar esses aspectos. Ao receber a orientação para buscar sobre “information seeking behavior”, o pós-graduando tomou a decisão acertada de procurar por artigos no Portal Capes. Porém, ao inserir no campo de busca *information seeking behavior*, não delimitou os termos com aspas e nos resultados obteve textos além do assunto desejado, como diversos artigos que tratavam separadamente sobre “information”, “seeking” e “comportamento”. Apesar da quantidade de resultados, a maior dificuldade apontada esteve relacionada à falta de domínio do inglês, idioma presente em praticamente todos os artigos recuperados. Mesmo assim, o grau de satisfação foi 8. A experiência foi relatada três semanas após ter ocorrido.

Em outro caso, o pós-graduando procurava por informações sobre iniciativas sustentáveis em diferentes organizações. Na experiência, descrita no questionário um dia após ter ocorrido como incidente crítico, o usuário diz ter inserido no campo de busca *sustentabilidade + programas*. O problema detectado foi não citar que procurava por programas de sustentabilidade em organizações, o que acabou resultando na dificuldade apontada: “grande volume de resultados”. A filtragem foi realizada selecionando entre os resultados aqueles que remetiam a sites de empresas. Por se tratar de uma busca inicial, optou pelo *Google*, por ser mais abrangente. “Tenho a impressão de que usei a ferramenta adequada, já que o que eu procurava estava disperso na internet. Com uma base de dados, talvez eu não tivesse recuperado tantos resultados e com o acesso direto para os links das organizações.” Grau de satisfação: 9.

Caso semelhante ocorreu com o pós-graduando que precisava encontrar material teórico a respeito dos conceitos de difusão e disseminação da inovação, especificamente conceitos e/ou definições que ajudassem a diferenciar os termos e como são utilizados pelos teóricos da área. A experiência foi relatada três dias após ter ocorrido. No momento de elaborar a consulta, porém, o usuário optou pelos termos “difusão da inovação” + disseminação, o que gerou resultados sobre temas que iam além do escopo desejado. Todavia, o grau de

satisfação foi 8, deduz-se que pela escolha acertada do *Google Acadêmico*, que gerou como resultado artigos em periódicos, teses, dissertações, citações e livros. O material foi filtrado de acordo com a adequação do título, o que levou à leitura superficial dos resumos. Enfrentou dificuldades com a falta de objetividade dos resumos, mas conseguiu encontrar conteúdo interessante para incluir na tese.

No caso seguinte, o pós-graduando realizou a busca baseado na necessidade de informações sobre determinado fotógrafo citado em uma palestra, o qual trabalha com fotos de museus. A experiência de busca ocorreu três dias antes da aplicação do questionário. O mecanismo escolhido foi o *Google Image*, “por parecer a ferramenta de busca de imagem mais interessante”. Pela resposta ao questionário, os termos escolhidos foram: Photography museum. Curiosamente observa-se que o usuário não inseriu o nome do fotógrafo desejado nesta busca, o que pode ter comprometido consideravelmente o resultado, composto por arquivos em formato *jpg*. As imagens recuperadas foram filtradas de acordo com a avaliação do usuário sobre a idoneidade do site que disponibilizou a fotografia. “A imagem fotográfica conduz a informações que podem levar a textos acadêmicos relevantes para a pesquisa”. Como a busca resultou em muita redundância, o usuário sentiu a necessidade de mudar os parâmetros e fazer nova busca. Tudo isso motivou nota 5 ao grau de satisfação.

6.9.2 Comportamento dos usuários na formulação da consulta

Jansen et al. (*apud* YANG, 2005, p. 38), em estudo de 1998, registraram que usuários costumam utilizar consultas curtas e empregam o menor esforço possível ao avaliar e refinar suas buscas. Na pesquisa empreendida pelos pesquisadores, mais de 30% dos usuários utilizaram um único termo na consulta, mas a média de palavras foi de 2,35. Em estudo no mesmo ano, Silverstein et al. (1998 *apud* YANG, 2005, p. 38) encontraram a mesma média de 2,35, mas o número de consultas com uma única palavra subiu para 77%. O baixo número de termos, como descrevem os pesquisadores, evidencia dedicação pequena ao processo.

Resultado semelhante foi apontado por Wolfram et al. (2001) ao analisarem mais de 1 milhão de consultas realizadas por 200 mil usuários do site Excite. “Eles descobriram que buscadores na web tendem a colocar pouco esforço tanto na formulação de consultas quanto na avaliação dos resultados, o que confirma padrões reportados em estudos anteriores.” (YANG, 2005, p. 39)

Hölscher e Strube (2000), também citados por Yang, aos investigarem o comportamento de usuários novatos e de usuários experientes, constataram que os segundos exibiram comportamentos muito mais complexos do que os observados na média dos usuários da web. Como exemplo, os autores citam o engajamento no aprimoramento das técnicas de busca, como reformulação da consulta e uso de recursos avançados. Em geral, usuários experientes parecem preferir percorrer um trecho a mais para satisfazer suas necessidades informacionais, o que pode ser percebido com consultas mais longas que o habitual, com média de 3,64 palavras, e a revisão de maior número de resultados.

Estudos realizados por Spink e Jansen (2005, p. 77) a partir das consultas submetidas a três grandes mecanismos de busca, mostraram algumas tendências no começo dos anos 2000:

a) o tamanho da consultas está aumentando lentamente, quase chegando à média de três termos; 2) o uso de operadores está presente em quase 10% de todas as consultas; 3) Usuários geralmente olham entre dois e três documentos resultantes por consulta; 4) o uso de mecanismo de buscas como forma de navegação rápida está crescendo; 5) alguns usuários estão realizando buscas longitudinais; 6) usuários submetem geralmente consultas curtas; 7) mecanismos de busca estão sendo usados para localizar amplos conjuntos de informação; 8) tem havido pequeno crescimento no uso de linguagem natural nas consultas; e 9) existem padrões de consultas identificadas em mecanismos de busca. (SPINK; JANSEN, 2005, p. 77)

Na pesquisa de campo realizada para esta dissertação, a maior ocorrência foi de consultas formadas por dois termos (42,3% do total de participantes), sendo que a média de palavras por consulta ficou em 2,7. Este resultado é coincidente com a afirmação a) do estudo de Spink e Jansen (2005) citada acima. Apenas 17,3% utilizam mais de quatro termos no campo de busca.

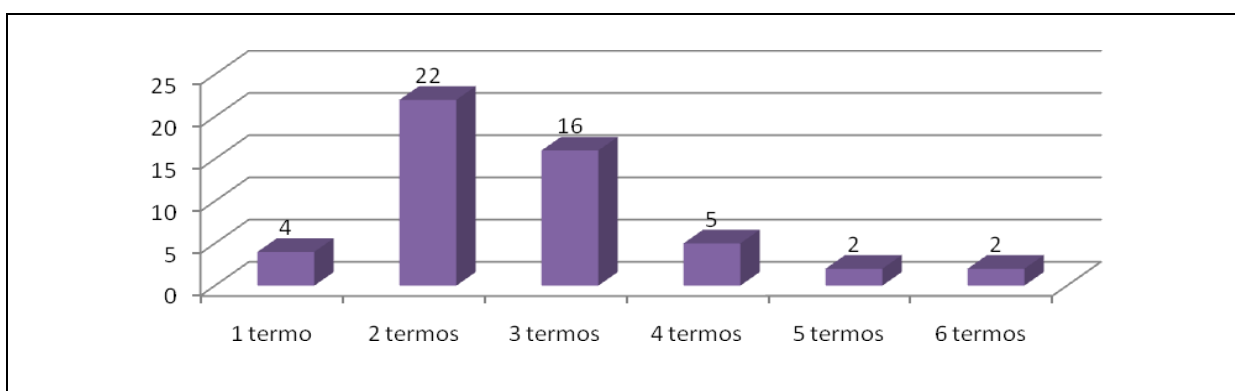


Figura 44 – Número de termos utilizados nas consultas relatadas pelos pós-graduandos

No intuito de investigar se havia relação entre a capacidade de busca na web e o número de termos utilizados, fez-se um cruzamento desses dois dados. Observou-se que entre os 11 pós-graduandos que autoavaliaram sua capacidade de busca na web com as notas 9 e 10, podendo ser tratados *a priori* como os mais capacitados, apenas três usaram mais de 4

palavras na formulação da consulta. Os demais, não passaram de três palavras. Assim, pode-se deduzir que não houve relação direta entre usuários que melhor se autoavaliaram e o número de termos empregado na busca.

6.9.3 Escolha da base de dados ou do mecanismo de busca

O sucesso do processo de busca está diretamente relacionado à formulação da consulta, mas tão importante quanto é a escolha da base onde será aplicada essa consulta. Em seguida, serão analisadas decisões tomadas por pós-graduandos em relação ao mecanismo mais apropriado para cada busca.

Pelo menos três experiências coletadas na pesquisa de campo demonstram como a escolha da melhor base de dados remete a resultados concretos. No primeiro caso, o pós-graduando selecionou a *Web of Science* como ponto de partida para a pesquisa que embasaria seu projeto de mestrado. A informação específica buscada era a produção indexada e o índice-H de diferentes autores brasileiros. A partir desta pré-definição, a consulta pôde ser constituída apenas pelo nome do autor. Apesar de bem sucedida quanto aos resultados, a busca revelou os problemas de indexação existentes na *Web of Science*. Os nomes de autores brasileiros podem aparecer grafados de diferentes formas, o que compromete o levantamento. O grau de satisfação foi 8.

No segundo caso analisado, o pós-graduando queria mapear as últimas teses defendidas no Brasil sobre epistemologia construtivista, mas somente aquelas que utilizaram a teoria ação construtivista de Habermas e o construtivismo de Piaget. Para realizar tal busca, foi escolhida a Biblioteca Digital de Teses e Dissertações do IBICT. No campo de busca, foram inseridos os termos Epistemologia Habermas. O resultado, de acordo com o pesquisador, foram teses e dissertações de autores brasileiros que tratam sobre o tema. Com isso, foi possível mapear a produção recente.

No terceiro exemplo selecionado, o pós-graduando procurava por embasamento sobre determinado protocolo de comunicação. Por se tratar de conteúdo para a dissertação, o material retornado deveria ser artigo científico. Essa razão levou ao *Google Acadêmico*, no qual foi inserida consulta delimitada por aspas: “sru/srw”. A busca resultou em artigos sobre o assunto, mas em grande volume. Apesar da dificuldade de filtrar manualmente o material resultante, o grau de satisfação foi 10 e o material recuperado “agilizou o desenvolvimento do projeto na medida em que permitiu que se esgotasse aquele tema e passasse para outro”.

Assim como foram encontrados casos de escolhas acertadas do mecanismo de busca mais apropriado, também observaram-se casos em que essa seleção apresentava incoerências. Neste caso, por exemplo, a busca foi motivada pela necessidade de saber o quanto o livro *Comunicação Científica* de Meadows era citado em artigos da área de Ciência da Informação, especialmente no tocante a acesso livre. A informação específica buscada pelo usuário dizia respeito a artigos científicos sobre acesso livre à informação científica. Em primeira análise, nota-se falta de linearidade entre as duas respostas, pois parece tratar-se de necessidades diferentes. A consulta formulada foi *Comunicação Científica* e Meadows, a qual foi inserida no *Google*, sem utilização de operadores ou recursos que aumentassem a precisão. A escolha do mecanismo foi justificada pelo pós-graduando como “hábito”. O resultado trouxe todo tipo de conteúdo – artigos em periódicos, artigos em formatos variados, teses, dissertações, citações e livros.

O grande volume de resultados foi apontado como dificuldade, pois como havia excesso de informações, mais tempo foi utilizado para separar o que realmente era importante e interessante. A maneira encontrada para filtrar os milhares de resultados obtidos foi realizando um refinamento da busca com a inclusão da expressão “acesso livre” à consulta. Ao final, a experiência foi considerada relevante pelo usuário por dar-lhe a certeza de que “o assunto vem merecendo muito destaque e importância”, levando-o a classificar seu grau de satisfação como 10. Como aprendizado, ele escreve que “a Internet e suas ferramentas de busca por informação são mecanismos muito úteis para a pesquisa científica, mas no final é o conhecimento humano que decide qual a informação que realmente interessa e esse processo toma tempo (que nem sempre temos)”.

Na presente análise, conclui-se que o tempo poderia ter sido usado de maneira mais eficiente se o usuário tivesse optado por mecanismo mais adequado ao tipo de busca realizada, assim como recursos de formulação da expressão, como as aspas, que ajudariam a aumentar a precisão da busca. A troca do *Google* pelo *Google Acadêmico* já significaria ganho de tempo, uma vez que este segundo mecanismo daria ideia do número de citações já no primeiro resultado. As decisões tomadas neste caso apontam para baixa reflexão no momento da busca, especialmente, como apontado anteriormente, sobre as ferramentas disponíveis na web.

Neste outro caso, a busca teria sido realizada de maneira mais ágil caso tivesse sido escolhido outro mecanismo. O assunto procurado era Folksonomia, especificamente sobre como os autores definem o conceito. Foi utilizado o *Google*, mas como no caso anterior,

haveria um refinamento natural caso tivesse se optado pelo *Google Acadêmico*. O resultado, apesar de volumoso, trouxe poucas fontes. Para selecionar algo relevante, o usuário teve de encontrar entre o material recuperado aqueles que fossem artigos acadêmicos. Apesar do resultado pobre, o grau de satisfação foi avaliado em 9.

6.10 Impacto provocado pela busca

Na questão seguinte, solicitou-se aos participantes que avaliassem o impacto provocado pela busca na elaboração do projeto de mestrado ou doutorado (2.11. *Qual foi o grau do impacto da informação recuperada para o seu projeto, considerando uma, e somente uma, das opções abaixo*). Foram oferecidas cinco graduações, de nenhum impacto até impacto fundamental.

Entre os mestrandos, 13 dos 34 (38,2% do total de participantes) avaliaram como fundamental a informação obtida.

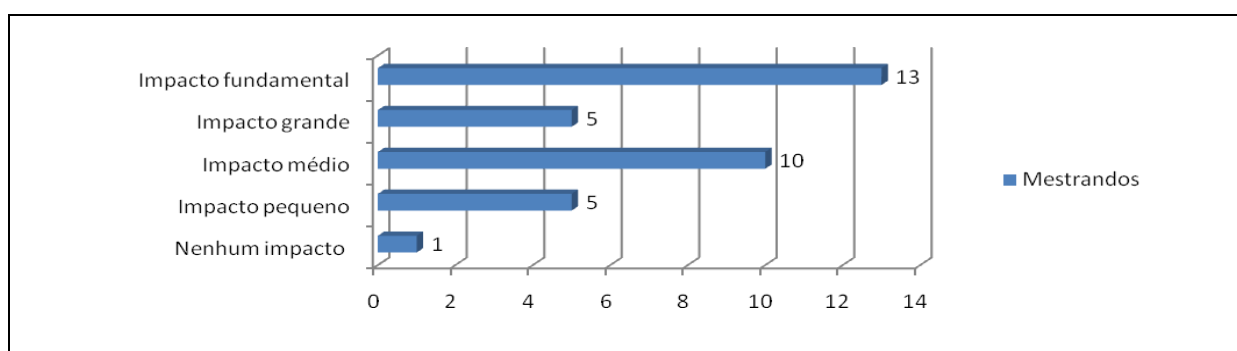


Figura 45 – Avaliação dos mestrandos sobre o impacto causado pela informação recuperada

Como mostra a Figura 45, entre os doutorandos, a opção impacto médio foi a mais frequente, com oito adesões, o equivalente a 44,4% do total.

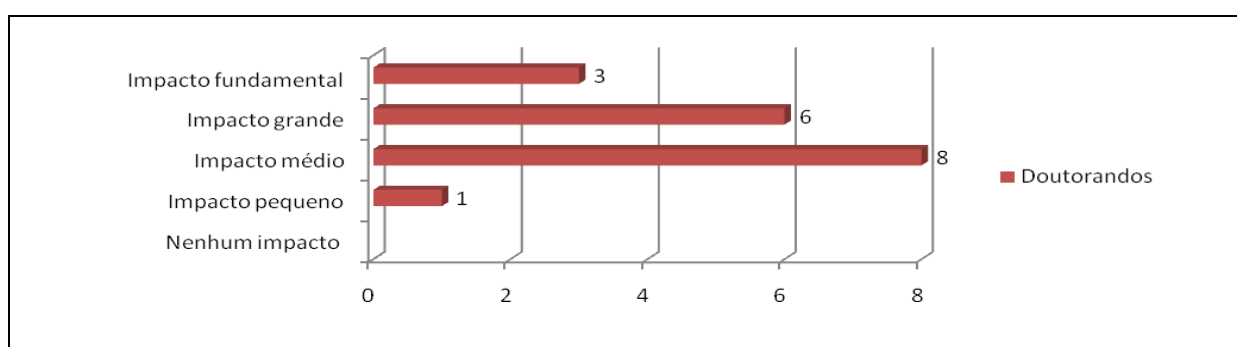


Figura 46 – Avaliação dos doutorandos sobre o impacto causado pela informação recuperada

Ao se comparar os resultados dos dois grupos, especialmente as opções impacto fundamental e impacto grande, percebe-se apenas pequena diferença percentual. Enquanto

53% dos mestrandos assinalaram essas alternativas, metade (50%) dos doutorandos também o fizeram. Já ao se comparar os resultados de impacto pequeno e nenhum impacto, nota-se que mais mestrandos identificaram-se com essas alternativas – 17,6% dos mestrandos contra 5,6% dos doutorandos. De modo geral, pode-se afirmar que o impacto das buscas realizadas foi relativamente maior entre os doutorandos, como demonstra a Figura 46.

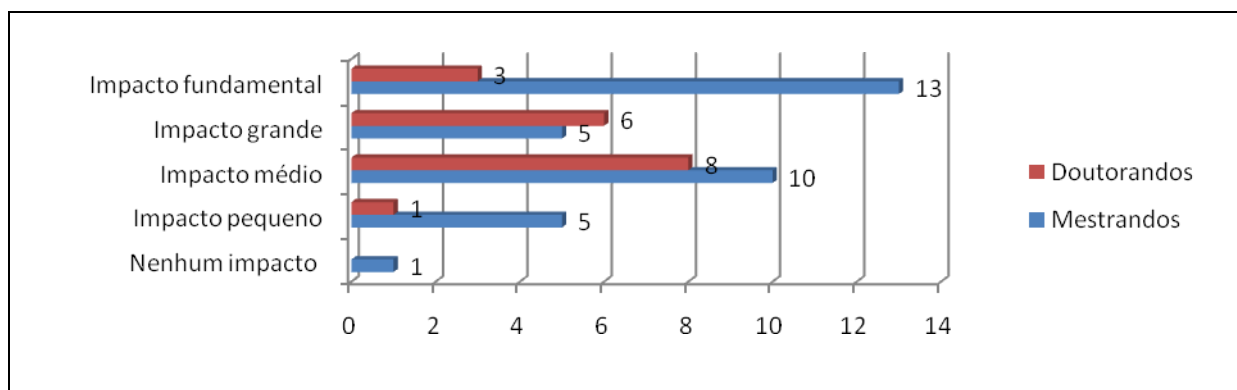


Figura 47 – Avaliação dos pós-graduandos sobre o impacto causado pela informação recuperada

6.11 Influência nos passos seguintes

Os pós-graduandos também foram questionados sobre o quanto a busca relatada afetou os passos seguintes no desenvolvimento da tese ou dissertação (2.12. *Como o resultado afetou os passos seguintes no desenvolvimento do seu projeto?*). As respostas remetem a diferentes tipos de influências que uma busca pode gerar. Entre as respostas, há desde aquelas dizendo que não houve influência alguma – “Ainda não afetou. Meu pensamento é ziguezagueante, não linear, vou e volto, viro à esquerda e à direita, descarto ou prossigo. Ainda não estou pensando em ‘próximos passos’”; até aquelas em que a influência foi baixa, mas existiu – “Ajudou a contextualizar alguns pontos de reflexão, mas não modificou o desenvolvimento do projeto”. Houve um caso em que a busca afetou de tal forma que o usuário desistiu da web e resolveu apelar para os recursos mais tradicionais para poder dar continuidade ao trabalho – “Fui a uma biblioteca pesquisar fisicamente”.

Houve ainda quem estivesse apenas reunindo conteúdos para mais tarde refinar como será estruturado o referencial teórico da tese, situação exemplificada pelas seguintes afirmações. “O resultado da pesquisa me ajudou a reorientar os objetivos do projeto.” “Acrescentou mais informação no grande processo de interpretação que é confirmar ou refutar hipóteses.”

Um terço dos participantes observou influências positivas, como nos seguintes exemplos: “Consegui conceituar um dos principais assuntos do meu trabalho, o que me rendeu um subcapítulo quase inteiro”; “Afetou positivamente, na medida em que apontou outras questões relacionadas a um processo de classificação colaborativa na web”; “Ajudou-me a entender melhor o tema pesquisado, ajudando a sistematizá-lo”; “Ratificou minha hipótese inicial e serviu como base de análise bibliográfica no assunto e como dado empírico”.

As buscas empreendidas também contribuíram para a organização dos materiais a serem lidos: “O resultado me levou a aprofundar mais as leituras acerca do autor e as obras de autores outros que possam sustentar a hipótese de que o desequilíbrio decorre de uma visão parcial do mundo”; “O resultado reorientou a ordem de relevância das leituras e ampliou os potenciais de interpretação do objeto proposto”; “Foi decisiva na seleção do material para a pesquisa”.

6.12 Cenário alternativo

E se a informação desejada não tivesse sido encontrada? Na pergunta 2.13 (*O que poderia ter ocorrido se o resultado não tivesse sido obtido?*), o pós-graduando foi convidado a refletir sobre essa hipótese.

As respostas variaram em grau de dramaticidade e senso prático. Enquanto alguns participantes acreditam que ficariam sem os conteúdos necessários para confirmar suspeita ou embasar tema abordado no projeto, outros acreditam que teriam de mudar o rumo teórico da pesquisa ou até desistir de colocar o assunto no projeto. “Acredito que haveria uma lacuna conceitual na pesquisa e afetaria os resultados no que se refere a aportes teóricos.” “Seria uma situação difícil, pois não há literatura direta sobre o assunto, só em livros importados (e em outro idioma).” “Teria optado por outro autor que tratasse do mesmo assunto, mas não daria o “peso” necessário à argumentação.”

Os mais práticos fariam nova busca, mudando o conteúdo da consulta, refinando mais os termos utilizados ou procurando mecanismo ou bases mais adequados – “Buscaria em base científica tipo *Pubmed* ou *Portal Capes* (no qual tenho muita dificuldade de recuperar o que preciso)”. Buscar em bibliotecas ou recorrer a colegas e professores também são alternativas apresentadas – porém, há quem responda que bibliotecas seriam a última opção: “Teria

recorrido aos "museus" (bibliotecas) da informação e, com sorte, talvez tivesse encontrado alguma informação.”

A questão do tempo também foi abordada. “Teria de continuar buscando e perderia muito tempo com isso.” Buscas malsucedidas implicam em perda de tempo, necessidade de dedicar mais horas à mesma tarefa e até em atraso de entrega do material.

6.13 Estratégias para filtrar o conteúdo

De acordo com as experiências reladas, as buscas resultaram principalmente em arquivos em formato *pdf*, documentos do *Word*, slides em *Power Point* e tabelas produzidas no *Excel*. Na sequência, foram recuperados em maior volume artigos em periódicos. Em terceiro lugar, aparecem páginas na web, que podem ser tanto artigos em publicações científicas quanto materiais sem fonte determinada. Entre os 52 participantes, 22 (42,3%) recuperaram teses e dissertações, assim como livros e citações, ambos encontrados por 18 (34,6%) participantes.

Para filtrar os conteúdos recuperados, os pós-graduandos adotaram diferentes estratégias. As mais repetidas foram a seleção pelos títulos, pelo tipo de arquivo e por meio da leitura dos resumos dos textos obtidos, como se vê na Figura 47.

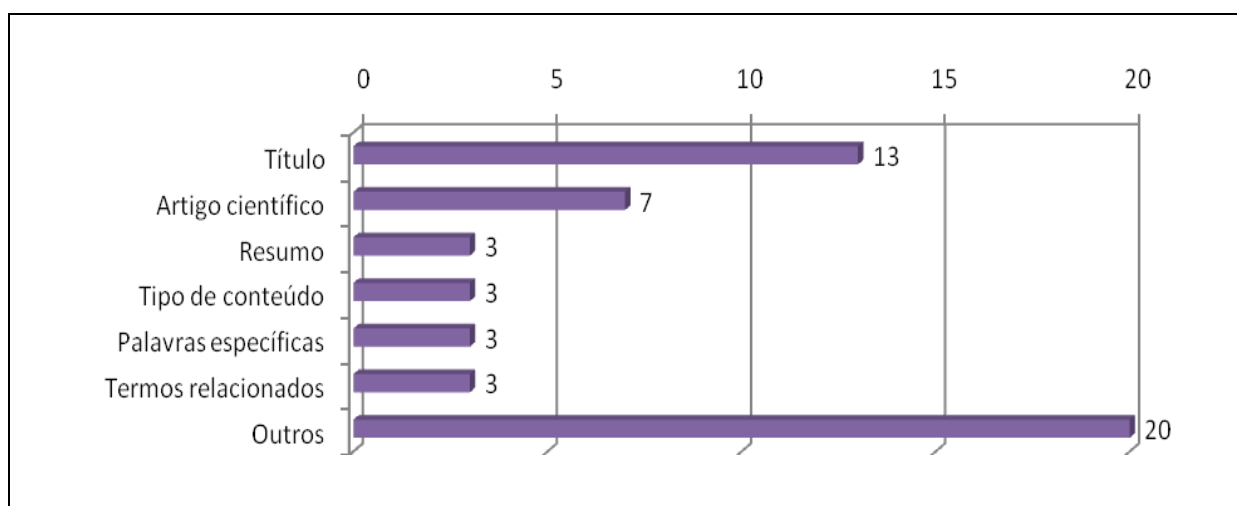


Figura 48 – Estratégias mais usadas para a filtragem do material recuperado

A escolha de como seria realizada esta etapa variou de acordo com o mecanismo utilizado e com o tipo de busca empreendido, mas análise de título do material recuperado foi a maneira mais frequente de filtragem, independente da base escolhida. “Abri os documentos que, pelo título, me pareciam úteis e li os resumos dos mesmos para separá-los.” “Li primeiro

os que, no título ou resumo, falavam da esfera pública na internet.” “Delimitar dentre os resultados aqueles que contemplassem em seu título o tipo de usuário diagnosticado pelo artigo. Ex.: ‘pesquisadores’.”

O tipo de artigo recuperado também foi decisivo. “Dou preferência a livros, teses e artigos científicos por uma questão de confiabilidade (ou sensação de confiabilidade).” “Os artigos mais acadêmicos.” “Procurei na lista de resultados os artigos de periódicos.” Apesar de estarem procurando por conteúdos científicos e com valor acadêmico, entre os pós-graduandos que usaram essa estratégia de filtragem, apenas um utilizou o *Google Acadêmico*. Os demais usaram o *Google*, mecanismo abrangente que recupera todos os tipos de conteúdos.

A leitura dos resumos foi outra estratégia adotada para selecionar conteúdos de interesse. “O filtro foi realizado através da leitura dos resumos, no caso de teses, dissertações e artigos, além da análise de dados pré e pós-textuais, como sumário e índices.” Houve também quem repetiu a busca acrescentando novas palavras à consulta, assim como quem optou por abrir todos os resultados das primeiras páginas de resultados.

6.14 Satisfação dos usuários

Na avaliação do grau de satisfação com a experiência relatada, 11 (32,35%) dos mestrands disseram-se completamente satisfeitos com o processo realizado, correspondente ao grau 10 de satisfação.

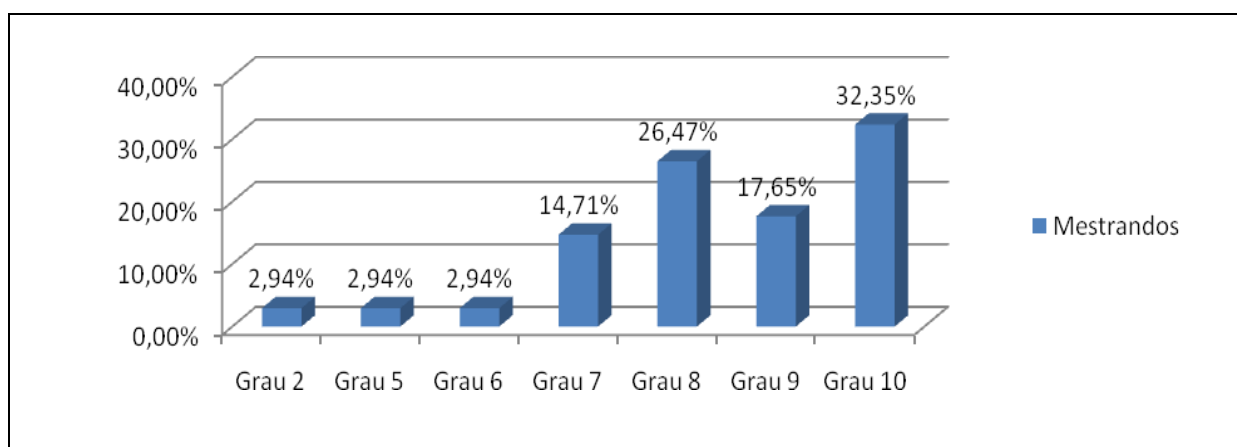


Figura 49 – Grau de satisfação dos mestrands em relação à experiência de busca relatada

Entre os doutorandos, apenas três (17%) conferiram nota máxima ao grau de satisfação. No resultado geral, porém, os doutorandos mostraram-se mais satisfeitos do que os mestrandos, como pode-se ver na Figura 49.

Considerando-se o somatório dos graus 8, 9 e 10, os mestrandos demonstraram 76,5% de satisfação, enquanto que os doutorandos somaram 83,3%.

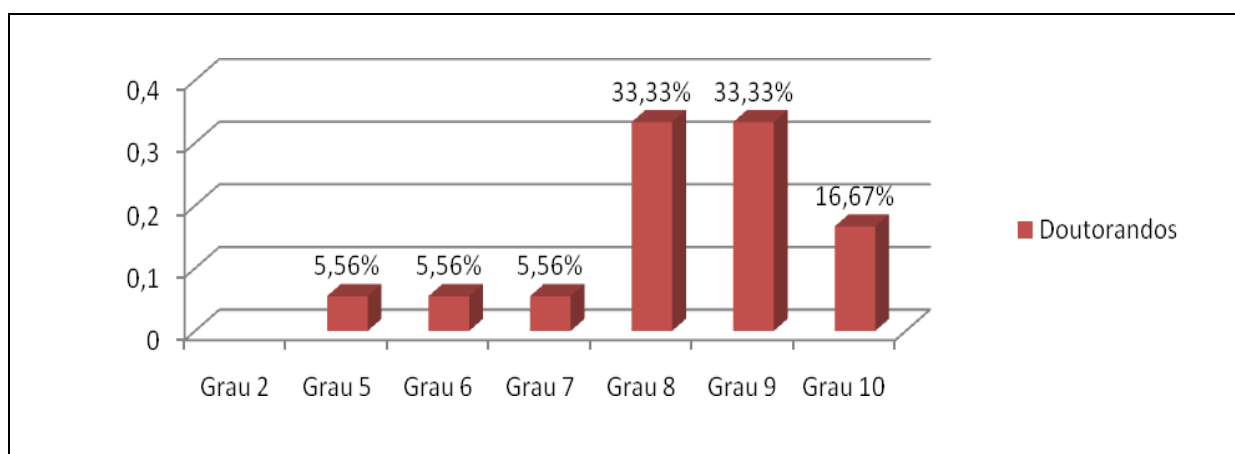


Figura 50 – Grau de satisfação dos doutorandos em relação à experiência de busca relatada

A Figura 50 demonstra que tanto entre mestrandos quanto entre doutorandos foi obtido alto índice de satisfação, superior a 75% nos dois grupos em relação aos graus mais elevados da escala, compreendendo o intervalo entre 8 e 10.

Se considerarmos o ponto de corte um grau abaixo, no entanto, compreendendo o intervalo entre os valores 7 e 10, os resultados de ambos os grupos se aproximam bastante, correspondendo a 91% entre os mestrandos e a 89% entre os doutorandos.

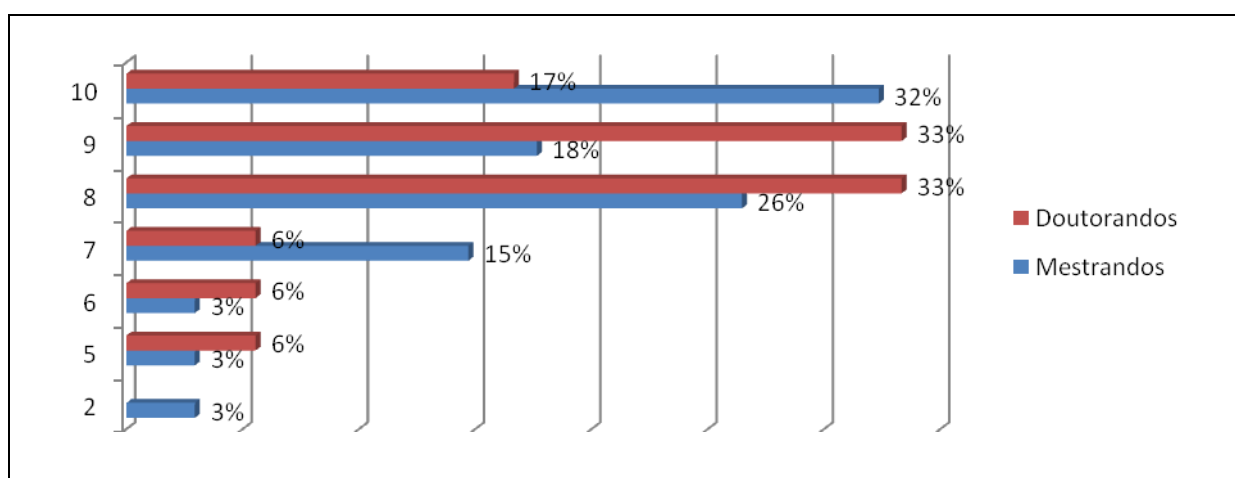


Figura 51 – Grau de satisfação dos pós-graduandos em relação à experiência de busca relatada

Satisfazer o usuário é uma das principais atribuições de um sistema de recuperação da informação – se não a mais importante. Em mais de um artigo, Lancaster (1979; 1981 *apud*

AL-MASKARI; SANDERSON, 2010, p. 2) sugere que um sistema de recuperação da informação pode ser avaliado de acordo com três critérios: a) a adequação em termos das tarefas específicas de recuperação da informação para as quais será utilizado; b) a eficiência na execução de uma tarefa; e c) o quanto o sistema satisfaz às necessidades informacionais dos usuários. Sendo assim, a satisfação dos indivíduos é um importante aspecto de avaliação de um sistema.

Porém, como alertaram Belkin e Vickery (1985) ainda mesmo depois de Lancaster, existem complicações inerentes à tentativa de medir a satisfação de um usuário: a própria definição ambígua da palavra satisfação e a dificuldade em medi-la. Em artigo de 2001, Hildreth questiona a confiabilidade das medidas de satisfação, devido à falta de independência de outros fatores que influenciam no processo. “Usuários finais de um sistema muitas vezes expressam satisfação tanto com os resultados obtidos quando com o desempenho global do sistema, mesmo quando uma análise mais detalhada comprove que os resultados obtidos são bastante pobres.” (HILDRETH, 2001 *apud* AL-MASKARI; SANDERSON, 2010, p. 3) Comportamento semelhante foi observado nesta pesquisa de campo.

Al-Maskari e Sanderson recorrem a Griffiths et al. (2007) para definir o que se entende por satisfação dos usuários: não é um construto único sobre o qual deve se basear a avaliação de eficácia de um sistema, mas algo influenciado por diversos fatores, como saída de dados do sistema, expectativa e atitude do usuário, percepção de facilidade de uso e utilidade do sistema, tipo de sistema e dificuldades de executar uma tarefa.

Em pesquisa relacionada ao tema, os autores investigam a relação entre satisfação do usuário e outros quatro fatores, que são efetividade do sistema, efetividade do usuário, esforço empreendido pelo usuário e as características do usuário.

A efetividade do sistema mede o quanto ele alcança seus objetivos. A efetividade do usuário é definida como a exatidão e plenitude com as quais os objetivos são atingidos. Após consultar diferentes autores, Al-Maskari e Sanderson especificam três critérios que levam à medida da efetividade do usuário: a) o número de tarefas completadas com sucesso; o número de documentos relevantes obtidos e c) o tempo gasto para realizar as tarefas.

Para os autores, a definição do esforço do usuário é similar à do comportamento na busca de informação (*information searching behavior*), ou seja, é o comportamento do usuário ao interagir com um sistema de recuperação da informação à procura de informação relevante. O esforço do usuário pode ser medido pelo número de cliques, número de

consultas, número de reformulações dessa consulta e número de resultados acessados para obter a informação relevante desejada.

O comprometimento do usuário é decisivo tanto para a recuperação de informações relevantes quanto para a satisfação com o processo como um todo. Outros fatores que influenciam no processo, como destacam os autores, são características do usuário que está realizando a busca. Al-Maskari e Sanderson citam Egan (1988) e Nielsen (1993) ao observar que fatores humanos e diferenças individuais são reconhecidos como peças-chave para se entender o comportamento dos usuários ao realizarem buscas. Alguns desses fatores são: familiaridade com o tópico buscado, isto é, conhecimento sobre a área tratada; motivação; e experiência em computação, biblioteconomia e busca de informação.

Os autores citam a afirmação de Marchionini (1995) de que todo indivíduo tem um conjunto único de habilidades para recuperar informação. Essas habilidades consistem em três componentes: a) expertise na área; b) expertise sobre o sistema utilizado; e c) expertise na busca de informação.

Os resultados obtidos por Al-Maskari e Sanderson apontam relação entre satisfação do usuário e efetividade do sistema, mas não de forma incisiva. Em relação à efetividade do usuário, notou-se que ela tem impacto direto sobre o grau de satisfação do usuário. Quanto maior o número de documentos relevantes obtidos, maior o grau de satisfação do usuário. Também foi constatado que quanto menor o tempo gasto na busca, maior a satisfação do usuário.

Já quando se trata do esforço empreendido em relação à satisfação, o impacto é o inverso. Quanto mais esforço, menor a satisfação do usuário. Na última análise, entre características do usuário e satisfação, Al-Maskari e Sanderson esperaram que usuários familiarizados com determinado assunto buscado tivessem maior satisfação, mas isso não se comprovou na pesquisa. Os pesquisadores avaliaram também a experiência de busca *on-line*. Os resultados apontaram que tanto usuários experientes ou familiarizados quanto inexperientes mostraram-se igualmente satisfeitos com os resultados obtidos.

6.15 Comparação entre autoavaliação e grau de satisfação

Na presente pesquisa, observou-se que dos 31 pós-graduandos (60%) que se autoavaliaram com notas entre 8 e 10, 28 deles (90%) conferiram também notas entre 8 e 10

ao grau de satisfação. A partir desse resultado se poderia inferir que, para este grupo, o grau de satisfação está associado à habilidade de buscar e recuperar informação na web. No entanto, ao se comparar os resultados das autoavaliações com notas iguais ou inferiores a 7 dos demais 21 pós-graduandos (40%), nota-se que apenas 8 deles (38%) tiveram grau de satisfação igual ou inferior a 7. A maioria, 13 dos 21 (62%), deu notas entre 8 e 10 ao grau de satisfação. Sendo assim, presume-se que a autoavaliação das habilidades possa não influenciar diretamente no grau de satisfação. Outra possibilidade a ser considerada é a de que possa existir um viés positivo de influência, para autoavaliações altas sobre o grau de satisfação, sem que haja necessariamente um viés negativo de interferência, para autoavaliações mais baixas sobre a satisfação, indicando apenas uma redução da positividade de alta de um possível viés.

6.16 Comparação entre impacto e grau de satisfação

Foram analisadas conjuntamente as respostas sobre o impacto provocado pela informação recuperada e as respostas sobre o grau de satisfação do usuário. Percebeu-se que o grupo pesquisado não estabeleceu relação direta entre impacto e satisfação, pois mesmo o pós-graduando que concluiu que os resultados não causaram nenhum impacto ao projeto, conferiu nota 8 ao seu grau de satisfação.

Observa-se, por exemplo, que o mesmo número de participantes conferiu nota 10 ao grau de satisfação tanto quando o impacto foi fundamental quanto quando foi pequeno. Independente do grau de impacto apontado, em todos é possível encontrar pós-graduandos completamente satisfeitos com o processo de busca empreendido, como pode ser visto na Figura 51.

As conclusões que se pode tirar dessa análise é que parece não haver reflexão suficiente sobre a relação entre os dois elementos (satisfação-impacto) ou, ainda, que a atribuição do grau de satisfação, para este grupo, parece ser independente do impacto que determinada informação possa causar nos processos de busca e recuperação realizados.

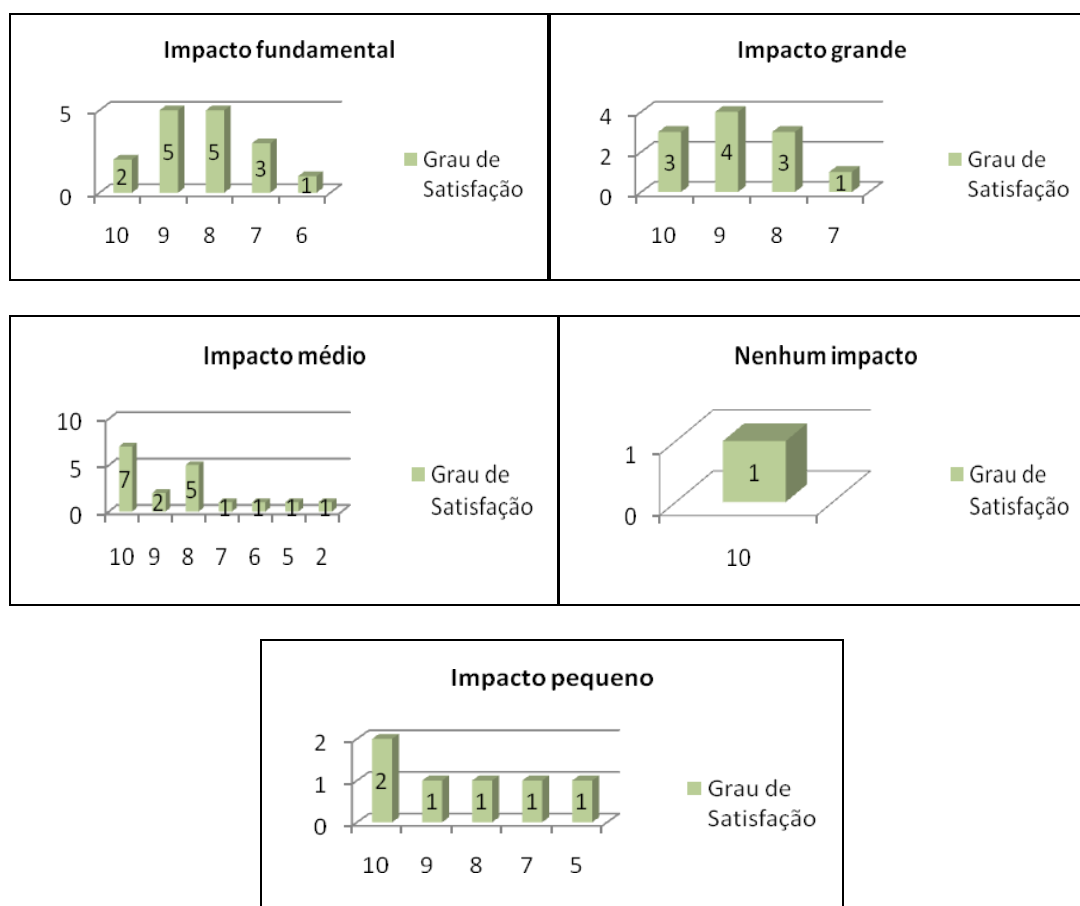


Figura 52 – Cruzamento de resultados do impacto causado pela informação obtida com o grau de satisfação dos pós-graduandos com a busca empreendida

6.17 Dificuldades enfrentadas no processo de busca

Na etapa seguinte da análise, percebeu-se relação entre dificuldades relatadas e o comportamento dos usuários ao longo da busca. Excessos de respostas, resultados que não atendem à demanda informacional, redundância e conteúdos em línguas estrangeiras não compreendidas pelos usuários são algumas das dificuldades identificadas.

Problemas detectados na análise da formulação da consulta e na escolha do mecanismo de busca utilizado ficam ainda mais aparentes ao se avaliar as dificuldades identificadas pelos participantes. A definição imprecisa dos termos utilizados, assim como a não utilização de caracteres auxiliares à linguagem natural, tais como aspas ou operadores booleanos, demonstrando a utilização de habilidades metalinguísticas, mostram-se como possíveis responsáveis pelo excesso de resultados, o que dificulta não apenas a filtragem, mas a obtenção de respostas em tempo razoável.

“Dificuldade encontrada sempre é o alto número de fontes indicadas. O primeiro passo que dou para a seleção é pelo título indicado.” “Excesso de informação. Necessidade de mais tempo para separar o que realmente importa e interessa. Vou voltar à mesma pesquisa em outra oportunidade.” Esses são relatos que exemplificam o alto volume de respostas. A fim de melhorar os resultados, uma solução é refinar a busca com uma consulta mais completa. Esse tipo de comportamento também ocorre quando o usuário identifica muita redundância nas respostas. “Recupero, nas primeiras buscas, conteúdo demais. Nesses casos, melho as palavras usadas na busca inicial.” “Quantidade de itens recuperados e baixa relevância. Tive de ler grande quantidade de resumos/introduções/referências para checar se o documento me interessava.”

Em determinados casos, os assuntos buscados são mais frequentes em línguas que não o português. O conhecimento insuficiente de idiomas estrangeiros aparece como umas das dificuldades enfrentadas. “A dificuldade maior foi a falta de domínio de uma língua estrangeira dominante nos resultados, especificamente o inglês.” “Dificuldade maior foi encontrar arquivos em alemão. Como não domino o idioma, descartei o arquivo.” “A questão de textos importantes em francês. Já combinei com a minha orientadora e terei de traduzi-los.” Quando o material mais importante sobre o assunto buscado existe em outros idiomas, o problema vai além do processo de busca, envolve outras habilidades. Modificar os parâmetros, para recuperar apenas conteúdos em português, não é exatamente a solução mais adequada para este tipo de problema.

Entre os 52 participantes, 13 (25%) afirmaram não ter enfrentado dificuldades, como pode-se identificar na Figura 52.

Vale ressaltar o baixo índice atribuído à acessibilidade do conteúdo, correspondendo apenas a 4% dos relatos.

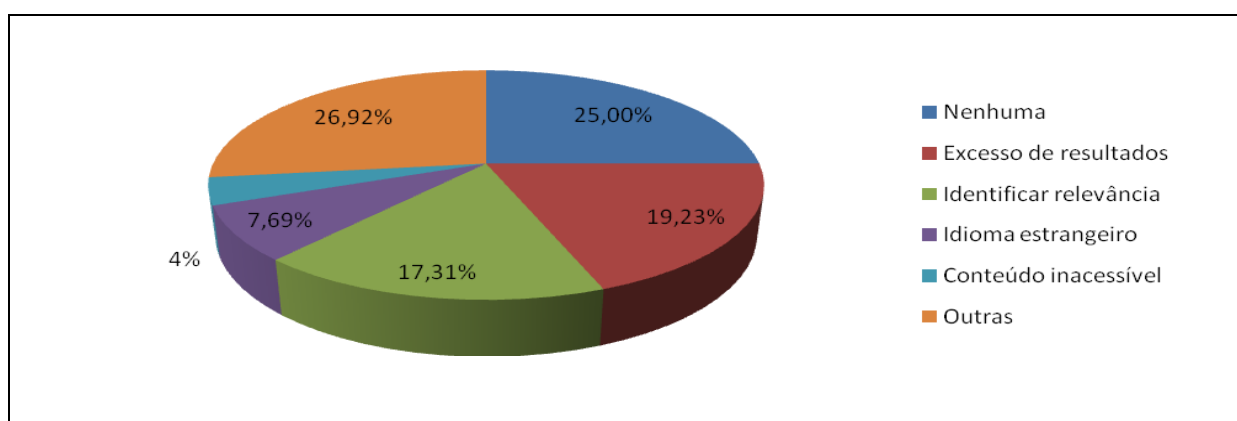


Figura 53 – Principais dificuldades identificadas pelos pós-graduandos

6.18 Aprendizado

Além da autoavaliação sobre a capacidade de buscar informações na web, os pós-graduandos refletiram sobre o aprendizado obtido com a experiência.

Alguns dos participantes fizeram um *mea culpa* sobre o esforço concedido à tarefa de buscar informação na web: “A web permite economizar tempo e dinheiro para encontrar a informação desejada de maneira eficiente, mas preciso me qualificar mais para pesquisar na rede”; “Com relação à utilização de ferramentas de busca, penso que ainda necessito melhorar minha capacidade de realização destas”; “Preciso aprender a fazer pesquisas mais profundas e selecionadas”.

Houve quem apontasse necessidades, como conhecer novas fontes na área de interesse, tais como outros sites, artigos ou blogs, ou adquirir novas habilidades, como ser mais específico ou aprender a refinar melhor a busca, para assim obter resultados mais relevantes. “Os critérios de busca devem ser os mais precisos possíveis para evitar resultados que não atendem à necessidade da pesquisa.” “É necessário ser criterioso na pesquisa, mais especificamente nos filtros de pesquisa para se obter um resultado que atenda à demanda.” Ou ainda “Quando eu não encontrar o que quero de imediato, devo olhar as referências bibliográficas de outros artigos relevantes com mais critérios”.

As respostas mostraram reflexões sobre o momento em que se vive. “O potencial de construção de conhecimento que temos nas bases on-line.” “O Google ainda é a fonte inicial relevante.” “A Internet e suas ferramentas de busca por informação são mecanismos muito úteis para a pesquisa científica, mas, no final, é o conhecimento humano que decide qual a informação que realmente interessa e esse processo toma tempo – o qual nem sempre temos.” A questão do tempo também foi abordada: “A internet é boa fonte de informações, porém a relevância dos documentos encontrados em mecanismos de busca comuns é ainda pouco eficiente, demandando tempo para filtragem dos documentos relevantes”.

Apesar da existência da preocupação com o processo, presente, por exemplo, na resposta “é preciso certa cautela frente aos resultados obtidos na busca on-line no sentido de diferenciar fontes relevantes de irrelevantes”, a avaliação de um dos pós-graduandos serve para ilustrar a maneira – por vezes displicente – como a busca na web pode ser encarada: “Sei lá. É uma busca como outra qualquer que já fiz e pretendo fazer. É algo quase automático. Não sei mensurar isso!”

7. Conclusões e perspectivas

A investigação realizada corrobora resultados de estudos semelhantes que demonstram que a natureza e a qualidade dos processos de planejamento e reflexão implicam em diferenças nos resultados obtidos pelos usuários ao realizarem busca e recuperação de informação na web. Raciocinar e ponderar sobre a estratégia mais adequada a cada etapa do processo que visa a recuperar a informação necessitada conduz a resultados mais apropriados, efetivos e eficientes, levando a um melhor aproveitamento do tempo investigado.

A necessidade de se garantir um aprimoramento contínuo da qualidade em um universo profissional é uma marca da contemporaneidade. Esta dimensão se torna ainda mais premente em ambientes de atividades que se fundamentam em planejamento investigativo dos procedimentos e capacitação no tratamento da informação, como na esfera da investigação científica. Particularmente, no âmbito de áreas intensivas em conhecimento, o aprimoramento desses processos de tratamento da informação se torna ainda mais relevante e urgente.

Ainda que a presente pesquisa se trate de um estudo exploratório, cujo potencial de generalização apresenta limites, o comportamento observado, por meio dos exemplos coletados, demonstra, de modo geral, pouca atividade de planejamento e de reflexão, tanto prévia quanto ao longo dos processos de busca e recuperação da informação científica relatados pela população-alvo.

A partir dos dados coletados na primeira parte do questionário que integrou a pesquisa de campo foi possível traçar um perfil da população-alvo estudada, composta por pós-graduandos na área da Ciência da Informação. De modo mais específico, a pesquisa de campo foi realizada no âmbito de um programa de pós-graduação, visando a garantir a homogeneidade da população em estudo. Realizar atividades de investigação de natureza recursiva, sobre a própria área que fornece os instrumentos conceituais e metodológicos de estudo, traz desafios adicionais aos normalmente encontrados em qualquer atividade de

investigação científica. Dentre eles, podemos citar a delimitação do campo empírico e de suas variáveis de observação, a garantia de adesão dos pares enquanto ‘objeto de estudo’, a aplicação de instrumentos investigativos em sujeitos treinados a se situar como observadores e não observados, a redução de vieses avaliativos, bem como o tratamento adequado, científica e eticamente, de resultados que possam ser interpretados como negativos.

Face a esses desafios, a autora da presente investigação considera ter obtido resultados satisfatórios na condução da pesquisa, de forma a permitir dar uma parcela de contribuição para o desenvolvimento das questões de estudo, assim como da própria área da Ciência da Informação.

De modo mais específico, em relação aos resultados obtidos por meio da pesquisa de campo, podem ser feitas algumas considerações.

Aproximadamente 70% dos pós-graduandos estudados são oriundos da Biblioteconomia (36,5%), Arquivologia (15,3%), Comunicação Social (9,6%) e Informática (7,7%), três delas citadas em 1996 por Saracevic como as áreas com maior interação com a Ciência da Informação. Em entrevista de 2010, o autor reafirma o forte intercâmbio da Biblioteconomia, da Ciência da Computação e da Comunicação, além das Ciências Cognitivas, com a Ciência da Informação, apesar de ser esse um campo do conhecimento procurado por profissionais com diferentes *backgrounds*.

Noventa por cento dos participantes graduaram-se em universidades federais, assim como 100% dos doutorandos concluíram o mestrado em instituições públicas – sendo seis egressos do PPGCI/IBICT. A partir da presente pesquisa de campo não é possível apontar a causa desta característica, mas em uma pesquisa futura e mais detalhada poderiam ser levantados dados, por exemplo, referentes aos tipos de faculdades existentes – se públicas ou particulares – no período em que os pós-graduandos frequentaram a universidade ou obtiveram título de mestre. Sabe-se, a este respeito, que houve um aumento no número de faculdades privadas nos últimos 15 anos.

Relacionado a esse assunto, observa-se pelos dados coletados que entre os mestrandos o período de conclusão da graduação estende-se de 1972 a 2007, com predominância nos anos 2000. Entre os doutorandos, a janela de tempo vai de 1969 a 2006, com maior número de graduações nos anos 1980. Em análises cruzadas com outros dados coletados, o ano de formação não pareceu ter influenciado no nível das dificuldades relatadas ou nos resultados obtidos pelos participantes. Pode ser pensando que, a despeito da Internet em geral e a web

em particular terem se desenvolvido nos últimos quinze anos, a formação profissional em nível de graduação não determina uma influência sobre o grau de performance ou expertise em sua utilização, ainda que sua existência venha tendo uma importância crescentemente marcante na esfera acadêmica, científica e profissional. Na medida em que a maioria dos participantes da pesquisa adquiriu conhecimento sobre busca e recuperação de informação científica na web a partir de sua própria experiência, aprendendo na prática, qual o papel e a influência formadora que a graduação profissional vem exercendo em relação a tais competências?

Ainda na parte de caracterização dos participantes, observou-se equilíbrio na distribuição dos pós-graduandos entre as linhas de pesquisa existentes nos anos de 2009 e 2010 – Linha 1: *Epistemologia e Interdisciplinaridade na Ciência da Informação*; Linha 2: *Organização, Estrutura e Fluxos da Informação*; e Linha 3: *Informação, Sociedade e Gestão Estratégica*. Quase 30% dos projetos desenvolvidos pelos mestrados encaixam-se na Linha 2. As linhas 1 e 3 têm, respectivamente, 32% e 38%. No doutorado esse equilíbrio persiste, estando quase 39% inseridos na Linha 1, 33% na Linha 2 e 28% na Linha 3.

A análise dos dados sobre como a habilidade de buscar e recuperar informação na web foi adquirida, estimulada por pergunta específica, revelou a formação autodidata dos pós-graduandos. Essa inferência foi feita a partir do predomínio das alternativas a) *Experiência própria, aprendendo na prática* e b) *Com colegas, amigos, parentes, conhecidos* como respostas mais marcadas tanto como primeira quanto como segunda opção – a combinação a/b foi escolhida por 19 participantes e a combinação b/a por três, totalizando 42,3% do total de participantes. Notou-se que as alternativas relacionadas a cursos, treinamentos e leitura de conteúdos foram assinaladas principalmente por bibliotecários, donde estima-se que por esse tipo de atividade fazer parte das atribuições profissionais específicas dessa formação.

Tanto a pergunta sobre como foram adquiridos os conhecimentos para realizar uma busca na web quanto as duas autoavaliações realizadas pelos pós-graduandos apresentam conexão com o terceiro objetivo específico, o qual trata sobre o grau de conhecimento dos pós-graduandos em relação a diferentes competências e habilidades na condução do processo de busca e recuperação. Ao responder como aprenderam a encontrar informações na web, os participantes encaixaram-se em pelo menos dois perfis. O primeiro deles remete ao aprendizado autodidata e ao método de tentativa e erro, ambos ocasionando demanda maior de tempo, e, conseqüente baixa eficiência de aprendizagem; mas, ao mesmo tempo, por se tratar de um ambiente virtual, a possibilidade de testar ampla gama de possibilidades poderia

lhes trazer alguma vantagem na aquisição de conhecimento tácito, ainda que com maior dispêndio de tempo e de esforço. No segundo perfil estão inseridos aqueles que se submeteram a algum tipo de treinamento, seja imposto pela formação escolhida, por mera curiosidade ou por interesse pessoal em aperfeiçoar o conhecimento na área em questão.

Nas duas autoavaliações os pós-graduandos pontuaram sua capacidade de realizar buscar na web, antes e depois de relatar uma de suas experiências ao executar uma tarefa dessa natureza. Na primeira avaliação, notou-se pelos resultados que mais da metade tanto de mestrandos quanto de doutorandos autoavaliou-se com notas iguais ou superiores a 8. Ao planejar a segunda autoavaliação, esperou-se estimular uma reflexão sobre a ação recém-realizada, o relato de uma experiência de busca e recuperação na web. Sessenta e um por cento não alteraram nota, o que não significa necessariamente a inexistência de metacognição.

Entre os 40% que alteraram as notas, 21,15% a aumentaram contra 19,2% que a diminuíram. Enquanto a média de notas iguais ou superiores a 8 aumentou entre os doutorandos, na média global ela se manteve, enquanto ela baixou um pouco entre os mestrandos. As razões tanto para aumentar quanto para diminuir as pontuações assemelham-se, apesar de irem em direções contrárias. Quem baixou a nota teria se julgado não merecedor da nota autoconferida inicialmente, após constatar os eventos do próprio incidente crítico?; Quem a aumentou teria observado que seu desempenho foi melhor do que o imaginado inicialmente?

Inserido em determinado contexto, cada grupo de usuários tem particularidades, as quais levam ao estabelecimento de expectativas e hipóteses. Diferente da análise de um grupo envolvido em uma pesquisa do dia-a-dia, ou busca-breve como a nomeou Marchionini (2006), quando as expectativas são baixas, a observação da experiência de busca de pós-graduandos em Ciência da Informação induz à perspectiva de comportamentos mais complexos e elaborados. Dentro do esquema apresentado por Marchionini, a população-alvo da pesquisa de campo realizada para esta dissertação está envolvida em uma busca exploratória, ou seja, aquela que abarca interpretação dos resultados e processo cognitivo. Em especial quando se trata de ampliação de um conhecimento, como no caso estudado, espera-se do grupo envolvido julgamentos qualitativos, análise efetiva do material recuperado e que os resultados sejam avaliados criticamente antes de serem incorporados.

Nesse contexto, há a expectativa de que a busca empreendida pelos pós-graduandos em Ciência da Informação para recuperar informação científica afaste-se das buscas cotidianas de conteúdos não acadêmicos.

Os participantes da pesquisa de campo escolheram experiências ocorridas de 2 horas até 40 semanas anteriores ao preenchimento do questionário. Percebem-se duas principais janelas de tempo privilegiadas pelos pós-graduandos: uma entre um e cinco dias e outra entre uma e cinco semanas. Apesar desta constatação não foi encontrada uma explicação que justificasse a ocorrência. De todo modo, estas duas janelas de tempo de maior ocorrência podem indicar que, em futuros estudos desta natureza, voltados a ampliar o conhecimento de alguns aspectos do campo de fenômenos estudados, poderiam servir como parâmetros delimitadores da elicitación de incidentes críticos por parte dos respondentes. Desta forma, poder-se-ia garantir uma maior qualidade e quantidade do material obtido pelos instrumentos de coleta de informações na pesquisa de campo.

Coerentes com a especificação de que deveriam relatar uma experiência relacionada ao desenvolvimento do projeto de mestrado ou de doutorado, mais da metade dos pós-graduandos informaram estar à procura de literatura relacionada aos assuntos de suas futuras dissertações e teses. Vinte e cinco dos 52 participantes, ou seja, 48%, estavam empreendendo a busca com o objetivo de realizar um levantamento bibliográfico.

Pela natureza desta busca, como visto com Marchionini, ela difere-se das tantas realizadas cotidianamente. Sabe-se que realizar buscas na web transformou-se em ato corriqueiro. É chegado então o momento em que o usuário deve dissociar a prática mais simplificada do dia-a-dia desta busca, por assim dizer, especializada. Ao decidir procurar por informações na web e não em outro repositório, seja uma biblioteca ou uma base de dados *off-line*, o usuário precisa refletir sobre como a consulta deverá ser construída, quais os termos mais apropriados, o uso de recursos para refinamento da busca e onde, em que sistemas de informação e bases de dados, ela será melhor realizada, para ampliar sua efetividade e eficiência.

Após tabulação dos dados fornecidos pelos pós-graduandos, chegou-se aos buscadores/mecanismos de busca na web mais utilizados. *Google* foi o preferido de 55,88% dos mestrados e 33,33% dos doutorandos, sendo o mecanismo mais utilizado. Em seguida, despontou o *Google Acadêmico*, com adesão de 23,5% dos mestrados e 33,33% dos

doutorandos. Em terceiro lugar, mas somente entre os mestrandos (8,8% deles), revela-se o *Portal Capes*.

Inferre-se que o hábito de abrir o buscador mais popular para procurar qualquer tipo de conteúdo não seja deixado de lado, para uma boa parcela dos usuários, nem quando o assunto buscado exige refinamento nas respostas. Apesar de a experiência relatada pelos pós-graduandos caracterizar-se pela busca de informações científicas que seriam usadas em nível acadêmico, muito mais rigoroso que outros meios, notou-se uma relativa baixa preocupação na escolha dos locais onde fontes confiáveis poderiam ser localizadas – em pelo menos um caso ficou aparente o desconhecimento de que em outros meios o processo de busca poderia alcançar resultados mais efetivos.

Tratando-se de busca por material para fundamentar projetos de mestrado e doutorado, a escolha da base de dados mais adequada mostra-se como essencial para a recuperação de documentos que tenham valor científico e possam ser usados nesse tipo de trabalho. Há casos, certamente, em que a utilização de um mecanismo de busca mais abrangente se justifique. No entanto, o ponto a ser ressaltado é precisamente que seja justificável em função de sua adequação aos propósitos almejados, e não por sua prevalência ou popularidade de uso. Desta forma, a pesquisa contribui para apontar que em determinadas experiências relatadas pode-se constatar julgamentos equivocados na escolha da base utilizada, o que contribuiu para levar a resultados de menor qualidade e/ou que determinaram esforço adicional para sua satisfatória completude.

A questão principal, porém, não é o mecanismo de busca escolhido, mas sim a falta de julgamentos qualitativos sobre o processo como um todo. A decisão pela utilização do recurso padrão ao invés do recurso avançado de um mecanismo de busca é exemplo disso. Mais de 75% dos participantes julgaram o recurso padrão como mais apropriado. Também nota-se o baixo uso de operadores que podem facilitar e ampliar a qualidade dos resultados, como os booleanos ou mesmo as aspas.

O comportamento da população-alvo desta pesquisa assemelha-se a conduta detectada em estudos de diferentes áreas. A pesquisa de Ivanitskaya, O'Boyle e Casey (2006) sobre como estudantes de medicina recuperam informações sobre saúde na web aponta que poucos consumidores de informação usam recursos avançados de busca, especificam precisamente as palavras-chave utilizadas ou limitam suas buscas de alguma maneira. No estudo empreendido por Jansen, Spink e Saracevic (2000), observou-se o desconforto dos usuários com operadores

booleanos, além de o costume de não passar da primeira página de resultados. Ou ainda, na pesquisa de 2001, desenvolvida por Spink et al., a qual mostrou consultas curtas, pouco modificadas e com estrutura simples.

Essas observações também ajudam a se perceber que nas buscas empreendidas por grande parte dos pós-graduandos parece preponderar a lei do menor esforço, em detrimento de outros princípios que garantam maior efetividade e eficiência – paradoxalmente levando a um maior esforço posterior ao exercício da busca e, conseqüentemente, ao processo de busca e recuperação como um todo.

Se no passado, o mais prático era recorrer às bibliotecas mais próximas de casa, com a web o comportamento não denota ser muito diferente. Acaba-se utilizando o caminho mais conhecido, não necessariamente o melhor, como ilustrado nos resultados pelo comportamento de escolher a mesma base de dados para encontrar tanto informações do dia-a-dia quanto para efetuar uma pesquisa acadêmica – em muitos casos independentemente dos conteúdos informacionais, das fontes prováveis, dos tipos de documentos, da precisão almejada, da confiabilidade informacional e da complexidade ou da especificidade do assunto esperados em uma busca mais complexa.

Diante da situação desenhada, percebe-se uma realidade que é dissonante do entendimento apresentado por diferentes pesquisadores em relação aos requisitos de qualidade de um processo de busca – não apenas nesta pesquisa, mas também naquelas citadas no referencial teórico, como as de Jansen, Spink e Saracevic (2000). Para os cientistas, comportamento reflexivo está intrinsecamente ligado aos procedimentos desenvolvidos em uma busca por informação. Tomando-se Menou como exemplo, depois de citar o processo material como uma das atividades da ação de transformação e transporte da informação, o autor pondera que haverá um processo intelectual antes da transmissão, utilização e mesmo da assimilação do que foi encontrado.

Participantes que demonstraram maior empenho ao longo das fases de busca obtiveram resultados mais consistentes não apenas ao final, mas durante o processo. Observou-se esse comprometimento em diferentes etapas do processo, tal como na tomada de decisão sobre a base mais apropriada, na elaboração cuidadosa da consulta, inclusive com a utilização de recursos que contribuem para maior precisão, e na escolha dos parâmetros para a filtragem dos resultados.

Todavia, na maior parte das experiências relatadas foram detectados problemas nas diferentes fases do processo, os quais levaram a dificuldades comuns a diversos participantes. O excesso de resultados e a dificuldade em encontrar material relevante entre os conteúdos recuperados denotam, novamente, a baixa reflexão ao longo do processo. Outro aspecto interessante apontado nos resultados foi a forma como a capacidade de busca na web é aperfeiçoada. Dentro de um grupo determinado, o de bibliotecários, notou-se que a capacidade foi ampliada de maneira formal, por meio de cursos e treinamentos, indicando a necessidade de ampliação de tais iniciativas educacionais em diferentes esferas da formação acadêmica em geral, independentemente da área de atuação profissional, como recurso capaz de permitir a aquisição de maior competência e letramento informacional no contexto contemporâneo da sociedade do conhecimento.

Observou-se ainda no estudo que o grau de satisfação dos participantes é alto, mesmo quando são citados percalços ao longo do processo. Entre os mestrandos, 76% dos participantes pontuaram o grau de satisfação com nota igual ou superior a 8. Entre os doutorandos, esse comportamento chegou a 83% dos participantes. Acrescentando-se aqui as dificuldades relatadas, dados como esses podem levar à hipótese nada alentadora de que independentemente do que for recuperado, o usuário estará satisfeito. Essa constatação induz a pensar que uma parcela da população-alvo pode ter níveis abaixo do esperado em termos de letramento informacional. Entre os aspectos que identificam uma pessoa informacionalmente letrada estão as habilidades para desenvolver estratégias de pesquisa bem-sucedidas e a utilização da informação recuperada de forma crítica. Outra possibilidade a ser considerada refere-se a poder desmembrar, em futuros estudos similares, a avaliação do grau de satisfação em componentes e fatores que influenciam sua determinação.

Pensamento crítico, aliás, é competência essencial no âmbito da recuperação da informação, tanto na desenvoltura para refinar as consultas quanto como subsídio para uso mais eficiente dos mecanismos de busca. Como afirma Dudziak, indivíduos letrados informacionalmente estão aptos a avaliar a informação segundo critérios de relevância, objetividade, pertinência, lógica e ética, incorporando as informações selecionadas ao seu próprio sistema de valores e conhecimentos.

No entanto, para confirmar, ou não, a hipótese do baixo letramento informacional seria necessário seguir nesta linha de investigação com um maior aprofundamento, por meio de uma pesquisa na qual fossem analisados não apenas relatos de uma única experiência, mas o comportamento ao longo de uma série de buscas derivada de uma mesma necessidade

informativa. Uma futura pesquisa de campo, por exemplo, poderia ser empreendida tanto com uma população-alvo com características semelhantes quanto com usuários de áreas diversas, mas com o mesmo objetivo de recuperar informações acadêmicas, para analisar como os usuários agem e elaboram os refinamentos necessários, a partir dos resultados de uma primeira tentativa, para chegar a respostas melhores.

Observou-se nos relatos coletados comentários sobre como a partir da primeira busca – e, na maioria dos casos, a busca relatada no questionário –, os usuários conseguiram aumentar o refinamento em uma segunda ou terceira tentativas. Curiosamente, ao escolher qual seria a experiência relatada, os participantes optaram por descrever a primeira tentativa, avaliada por eles mesmos como a mais simples. Essa constatação do pesquisador ao analisar os dados poderá servir, em uma futura pesquisa, para aprimorar as instruções dadas aos participantes antes da aplicação do questionário, ajudando-os a escolher o melhor exemplo. Nem mesmo a questão do tempo transcorrido entre a realização da busca e a aplicação do questionário justifica a escolha, pois as tentativas posteriores seriam mais recentes.

Com base nas pesquisas referenciadas, o comportamento da população-alvo da presente pesquisa poderia ser considerado próximo do de outros indivíduos pesquisados. Porém, deve-se não perder de vista, como dito anteriormente, o grupo ao qual pertencem os participantes desta pesquisa de campo. Apesar de se tratar de profissionais oriundos de diferentes campos do conhecimento, ao estarem todos inseridos no campo da Ciência da Informação, considerados assim profissionais da informação, aumenta-se a expectativa em relação à postura desses usuários no que diz respeito ao tratamento da informação. A baixa reflexão apontada nesta pesquisa não seria problema fundamental caso a população-alvo estivesse tentando suprir outros tipos de necessidades informacionais.

A realização desta dissertação, especialmente no que diz respeito à pesquisa de campo, pode ser caracterizada como um constante desafio pessoal para a autora. Em primeiro lugar, desafio em inteirar-se sobre uma nova área, na qual informação tem significado bastante diverso daquele empregado desde a graduação em Jornalismo. Em segundo, um desafio sugerido pelo tempo reduzido para a realização de uma pesquisa de campo envolvendo a aplicação de instrumentos conjugados, ou seja, tanto aplicação de questionário quanto realização de entrevistas. Além disso, como um terceiro desafio, o fato de pertencer à própria população-alvo. Ao mesmo tempo em que essa característica foi bastante útil no desenvolvimento da pesquisa, especialmente pela disponibilidade e boa vontade encontrada junto aos colegas, ela poderia ter sido um fator complicador na fase da análise dos resultados.

Avaliar o comportamento de pessoas tão próximas na esfera profissional e acadêmica poderia ter gerado sensação de insegurança e desconforto, para ambos os lados da interação entre os atores sociais envolvidos. Porém, a solução para tais preocupações revelou-se na própria execução da pesquisa, a qual pôde ser realizada de maneira transparente, contando com a receptividade e colaboração dos colegas como participantes respondentes, e baseando as análises e interpretações na garantia empírica pela obtenção de dados concretos fornecidos pela população-alvo.

Espera-se com esta investigação ter conseguido entender e descrever como os indivíduos veem a ação de busca e recuperação da informação na web, ressaltando do material coletado tanto as performances mais habilidosas quanto aquelas que poderiam ser submetidas a mudanças no sentido de seu potencial aprimoramento. Na certeza de que a reflexão científica e ética sobre as lacunas do conhecimento constitui-se num grande impulsionador do seu desenvolvimento, pode-se afirmar que os pós-graduandos que participaram da pesquisa foram partícipes imprescindíveis na contribuição de um maior conhecimento humano sobre as complexas ciência e arte dos processos de busca e recuperação da informação, tanto através de seus acertos quanto de seus erros e dificuldades.

O usuário é agente essencial na concepção, avaliação, enriquecimento, adaptação, estímulo e funcionamento de qualquer sistema de informação. Por isso, assim como para benefícios próprios, a reflexão deve ser insistentemente estimulada e realizada. Estudos de usuários são importantes para o conhecimento do fluxo de informação científica e técnica, de sua demanda, da satisfação do usuário, dos resultados ou efeitos da informação sobre o conhecimento, do aperfeiçoamento dos processos, das relações com o conhecimento, da distribuição de recursos de sistemas de informação e de tantos outros aspectos relacionados à informação. Somente assim, tanto os resultados recuperados com maior qualidade poderão trazer maior satisfação e benefícios aos usuários quanto permitirão estimular melhorias significativas nos próprios sistemas existentes. Dentro desse contexto, com a presente dissertação espera-se estimular a reflexão não apenas em relação ao grupo populacional estudado, mas a todos aqueles usuários que buscam o aperfeiçoamento constante do seu conhecimento, da sua prática e da sua vida na sociedade.

8. Referências

- AL-MASKARI, A.; SANDERSON, M. A review of factors influencing user satisfaction in information retrieval. *Journal of the American Society for Information (JASIS)*, v. 61, n. 5, p. 859-868, 2010.
- ALPERT, J.; HAJAJ, N. We knew web was big... The official Google Blog, jul. 2008. Disponível em: <http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html>. Acesso em: 25 fev. 2011.
- ANDRETTA, I; SILVA, J. G.; SUSIN, N.; FREIRE, S. D. Metacognição e aprendizagem: como se relacionam? *Psico*, Porto Alegre, PUCRS, v. 41, n. 1, p. 7-13, jan./mar. 2010.
- BARRETO, A. Poucas notas para a elaboração de seu projeto, dissertação ou tese. ExtraLibris. Disponível em: <<http://extralibris.org/ensaio/como-elaborar-a-dissertacao/>>. Acesso em: 15 nov. 2010.
- BARRETO, A. As estruturas de suporte da informação no processo do conhecimento: o papel da fluência digital. *DataGramaZero - Revista de Ciência da Informação*, v.7, n.4, ago. 2006.
- BAEZA-YATES, R; RIBEIRO-NETO, B. Modern Information Retrieval. New York: ACM Press., 1999.
- BAWDEN, D. Information and digital literacies: a review of concepts. *Journal of Documentation*, v. 57, n. 2, mar. 2001, p. 218-259.
- BELKIN, N.; ROBERTSON, S. E. Information Science and the phenomena of information. *Journal of the American Society for Information Science (JASIS)*, v. 27, n. 4, p. 197-204, jul/ago 1976.
- BIOLCHINI, J. C. A. *Da Experiência ao Conhecimento: a informação clínica em homeopatia*. 1998. Dissertação (Mestrado em Ciência da Informação) – IBICT-UFRJ/ECO, Rio de Janeiro, 1998. Orientadora: Rosali Fernandez de Souza.
- BORKO, H. Information science: What is it? *American Documentation*, v. 19, n. 1, p. 3-5, jan. 1968.
- BRAGA, G.. Informação, ciência da informação: breves reflexões em três tempos. *Ciência da Informação*, Brasília, DF, Brasil, v. 24, n.1, abr. 1995. Disponível em: <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/534>. Acesso em: 29 Mar. 2011.
- BUSH, Vannevar. As we may think. *Atlantic Monthly*, v. 176, n. 1, p. 101-108, jul. 1945.
- CAMPELLO, B. S. Letramento Informacional no Brasil: Práticas educativas de bibliotecários em escolas de ensino básico. 2009. 208 f. Tese (Doutorado em Ciência da Informação) – UFMG, Belo Horizonte, 2009.
- CAMPELLO, B. S. O movimento da competência informacional: uma perspectiva para o letramento informacional. *Ciência da Informação*, Brasília, v. 32, n. 3, p. 28-37, set./dez. 2003.
- CAMPOS, M. L. A. *As cinco leis da biblioteconomia e o exercício profissional*. Disponível em <http://www.conexaorio.com/bitl/mluiza/index.htm>. 1999. Acesso em: 8 jan. 2011.
- CAPURRO, R; HJØRLAND, B. O conceito de informação. *Perspectivas em Ciência da Informação*, v. 12, n. 1, 2007. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/54>. Acesso em: 1 fev. 2011.
- CHOO, C. W. *The knowing organization: how organizations use information to construct meaning*, 2. ed. New York, Oxford: Oxford University Press, 2006.

CRANDALL, B.; KLEIN, G.; HOFFMAN, R.R. Overview of Cognitive Task Analysis Methods. In: CRANDALL, B.; KLEIN, G.; HOFFMAN, R.R. *Working Minds: A practitioner's guide to cognitive Task Analysis*. Cambridge, MA: The MIT Press, 2006. Cap. 2, p. 9-26.

COSTA, F. L. da; SILVA, A. C.; RAMALHO, F. A. (Re)visitando os estudos de usuário: entre a “tradição” e o “alternativo”. *DataGramaZero – Revista de Ciência da Informação*, v. 10, n. 4, ago. 2009. Disponível em: <http://www.datagramazero.org.br/ago09/Art_03.htm>. Acesso em: 23 jan. 2011.

Digital Literacy: A Neglected Skill? - Interview with Paul Gilster. In: *ESL-Globe*, v. 7, n. 1, 2010. Disponível em: <http://www.ncsu.edu/eslglobe/gilster.html>. Acesso em: 3 jan. 2011.

DUDZIAK, E. A. Information literacy: princípios, filosofia e prática. *Ciência da Informação*, Brasília, v. 32, n. 1, mai. 2003. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/123>>. Acesso em: 20 mai. 2010.

Entrevista Tefko Saracevic. In: *InCID: Revista Ciência da Informação e Documentação*, Ribeirão Preto, v. 1, n. 2, p. 161-167, jul./dez. 2010.

FERNEDA, E. *Recuperação de informação: Análise sobre a contribuição da Ciência da Computação para a Ciência da Informação*. 2003. 147 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes da USP, São Paulo, 2003. Orientadora: Johanna Wilhelma Smit

FLANAGAN, J. C. The critical incident technique. *Psychological Bulletin*, v. 51, n. 4, p. 327-358, 1954. Disponível em: <<http://www.apa.org/pubs/databases/psycinfo/cit-article.pdf>>. Acesso em: 20 mai. 2010.

FLICK, U. Entrevista episódica. In: BAUER, M. W.; GASKELL, G. *Pesquisa qualitativa com texto, imagem e som: Um manual prático*. 7. ed. Petrópolis, RJ: Vozes, 2008. Cap. 5, p. 114-136.

FOSKETT, D. J. Informática. In: *Ciência da Informação ou Informática?* Org. Hagar Espanha Gomes. Rio de Janeiro, Calunga, 1980. p. 9-51.

GASQUE, K. C. G. D.; COSTA, S. M. de S. Evolução teórico-metodológica dos estudos de comportamento informacional de usuários. *Revista Ciência da Informação*, v.39, n.1, p. 21-32, 2010.

GASKELL, G. Entrevistas individuais e grupais. In: BAUER, M. W.; GASKELL, G. *Pesquisa qualitativa com texto, imagem e som: Um manual prático*. 7. ed. Petrópolis, RJ: Vozes, 2008. Cap. 3, p. 64-89.

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

GONZALEZ DE GÓMEZ, M. N. Metodologia de pesquisa no campo da Ciência da Informação. *Datagramazero*, v. 1, n. 6, dez. 2000. Disponível em: <http://www.dgz.org.br/dez00/Art_03.htm>. Acesso em: 07 jun. 2010.

GUEDES, R. de M; SOUZA, R. S. Navegando entre nuvens de etiquetas: uma proposta de utilização da tag cloud em catálogos eletrônicos de bibliotecas. *PontodeAcesso*, Salvador, v. 2, n. 3, p. 2-13, dez. 2008. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/viewFile/3186/2334>>. Acesso em: 21 fev. 2011.

HATSCHBACH, M. H. de L. *Information Literacy: aspectos conceituais e iniciativas em ambiente digital para o estudante de nível superior*. 2002. 109 f. Dissertação (Mestrado em Ciência da Informação – IBICT-UFRJ/ECO), Rio de Janeiro, 2002. Orientadora: Gilda Olinto.

INGWERSEN, P.; JÄRVELIN, K. *The Turn: Integration of information seeking and retrieval in context*. Holanda: Springer, 2005.

IVANITSKAYA, L.; O'BOYLE, I.; CASEY, A. M. Health Information Literacy and Competencies of Information Age Students: Results from the Interactive Online Research Readiness Self-Assessment (RRSA). *Journal of Medical Internet Research*, v. 8, n. 2, p. 1-17, abr./jun. 2006.

JANSEN, B; SPINK, A; SARACEVIC, T. Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing and Management*, v. 36, n. 2, p. 207-227, 2000.

KRONBERGER, N.; WAGNER, W. Palavras-chave em contexto: análise estatística de textos. In: BAUER, M. W.; GASKELL, G. *Pesquisa qualitativa com texto, imagem e som: Um manual prático*. 7. ed. Petrópolis, RJ: Vozes, 2008. Cap. 17, p. 416-441.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos/Livros, 1993.

LAWRENCE, S.; GILES, L. Searching the World Wide Web. *Science*, v. 280, abr. 1998.

LE COADIC, Y. F. *A ciência da informação*. 2. ed. Brasília: Briquet de Lemos, 2004.

LEAHY, D.; DOLAN, D. Digital Literacy – Is It Necessary for eInclusion? In: *HCI and Usability for i-enclusion: Lecture Notes in Computer Sciences*. Berlin-Heidelberg: Springer-Verlag, 2009. p. 149-158.

LESK, M. The seven ages of information retrieval. Conference for the 50th anniversary of *As We May Think*, 12-14 October 1995, MIT, Cambridge, Mass.

LÉVY, P. *Cibercultura*. 2. ed. São Paulo: Editora 34, 2000.

LIMA, G. A. B. Interfaces entre a ciência da informação e a ciência cognitiva. *Ciência da Informação*, Brasília, v. 32, n. 1, p. 77-87, jan./abr. 2003.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. *Introduction to information retrieval*. Cambridge UP, 2008. Disponível em: <<http://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>>. Acesso em: 20 Jun. 2010.

MARCHIONINI, G. Exploratory search: from finding to understanding. *Communications of the ACM*, v. 49, n. 4, p. 41-46, abr. 2006.

MEADOW, C. T. *The analysis of information systems – a programmer's introduction to information retrieval*. EUA: John Wiley & Sons Inc., 1967.

MENOU, M. J. Trends in. a critical review. The impact of information -II. Concepts of information and its value. *Information Processing & Management*, v.31, n. 4, p. 479-490, 1995.

MORVILLE, P. e ROSENFELD, L. *Information Architecture for the World Wide Web*. 3. ed. EUA: O'Reilly Media, 2006.

MORVILLE, P. *Ambient Findability*. Califórnia: O'Reilly Media, 2005.

PEREIRA, M. N. F. Bibliotecas virtuais: realidade, possibilidade ou alvo de sonho. *Ciência da Informação*. Brasília, v. 24, n. 1, 1995.

PINHEIRO, L.V. R.; LOUREIRO, J.M.M. Traçados e limites da Ciência da Informação. *Ciência da Informação*, Brasília, v. 24, n. 1, p. 42-53, jan./abr. 1995.

PINHEIRO, L. V. R. Informação – Esse obscuro objeto da Ciência da Informação. *Revista Morpheus*, v. 2, n. 4, 2004. Disponível em: <<http://www.unirio.br/morpheusonline/Numero04-2004/lpinheiro.htm>>. Acesso em: 29 nov. 2010.

PINHEIRO, L. V. R. Ciência da Informação: desdobramentos disciplinares, interdisciplinaridade e transdisciplinaridade. In: GONZÁLEZ DE GÓMEZ, M. N.; ORRICO, E. G. D. (Org.) *Políticas de memória e informação: reflexos na organização do conhecimento*. Natal: Editora Universitária da UFRN/EDUFRN, 2006, p. 111-141.

PINHEIRO, L. V. R. *Usuário – Informação: o contexto da ciência da tecnologia*. Rio de Janeiro: IBICT, 1982.

POOL, C. A new digital literacy: a conversation with Paul Gilster. In: *Educational Leadership*, v. 55, n. 3, p. 6-11, nov-1997.

RIBEIRO, C. Metacognição: Um Apoio ao Processo de Aprendizagem. *Psicologia: Reflexão e Crítica*, v.16, n. 1, p. 109-116, 2003.

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.

SARACEVIC, T. Information Science. *Journal of The American Society for Information Science*, v. 50, n. 12, p. 1051-1063, 1999.

SANTIAGO, M. C. C. *Metadados para recuperação da informação em ambiente virtual*. 2004. 122 f. Dissertação (Mestrado em Ciência da Informação) – IBICT-UFRJ/ECO, Rio de Janeiro, 2004. Orientadora: Lena Vania Ribeiro Pinheiro.

SPINK, A.; WOLFRAM, D.; JANSEN, J.; SARACEVIC, T. Searching the web: the public and their queries. *Journal of The American Society for Information Science*, v. 52, n. 3, p. 226-234, 2001.

SPINK, A.; JANSEN, B. *Web search: public searching of the web*. Dordrecht, Holanda: Kluwer Academic Publishers, 2005.

The multiple faces of literacy: Reading, Knowing, Doing. Disponível em: <http://www.iasl-online.org/events/conf/2006/conference2006-gen.html>. Acesso em: 4 jan. 2011.

VARELA, A. Dimensão das relações entre a ciência da informação e as ciências cognitivas: caminhos percorridos e a percorrer. *Informação & Sociedade:Estudos*, João Pessoa, v. 20, n. 2, p. 51-64, maio/ago. 2010. Disponível em: <<http://www.brapci.ufpr.br/documento.php?dd0=0000009100&dd1=6f27f>>. Acesso em: 10 fev. 2010.

VIZINE-GOETZ, D. Using library classification schemes for internet resources. *OCLC Internet Cataloguing Project Colloquium*. Position Paper. Disponível em: <<http://staff.oclc.org/~vizine/InterCat/vizine-goetz.htm>>. Acesso em: 19 mar. 2010.

WERSIG, G. Information Science: the study of postmodern Knowledge usage. *Information Processing & Management*, v. 29, n. 2, p. 229-239, 1993.

YANG, K. Information Retrieval on the Web. *Annual Review of Information Science and Technology*, v. 39, p. 33-80, 2005. Disponível em: <http://elvis.slis.indiana.edu/kiyang/pubs/Webir_arist.pdf>. Acesso em: 26 jun. 2010.

ZURKOWSKI, P. G. Information services environment relationships and priorities. Washington D. C.: National Commission on Libraries, 1974. Disponível em: <<http://www.eric.ed.gov/PDFS/ED100391.pdf>> Acesso em: 1 fev. 2011.

9. Apêndices

APÊNDICE A – TERMO DE CONSENTIMENTO

Título do Projeto: *Busca e recuperação da informação na web: a experiência de profissionais da informação*

Pesquisadora responsável: Rafaela Boeira Giordano – (21) 9912-0816

Instituição: Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), em convênio com Universidade Federal do Rio de Janeiro (UFRJ)

Termo de consentimento

Colega, você está sendo convidado(a) para participar da pesquisa de campo do projeto de pesquisa *Busca e recuperação da informação na web: a experiência de profissionais da informação*, da mestranda Rafaela Giordano, orientada pelo prof. Dr. Jorge Biolchini.

Este projeto tem como objetivo estudar o comportamento de pós-graduandos em Ciência da Informação ao buscar e recuperar informações na web para a realização de suas dissertações ou teses, a fim de elaborar indicações que contribuam para a obtenção de resultados mais efetivos na recuperação da informação na web para fins acadêmicos.

Para o desenvolvimento do projeto será realizado um estudo exploratório e qualitativo, executado por meio de revisão bibliográfica e pesquisa de campo. Nesta segunda etapa, estão inseridos um questionário e uma entrevista, cuja elaboração baseou-se na técnica do incidente crítico.

Essa técnica reúne um conjunto de procedimentos metodológicos para elaboração da coleta de informações sobre o comportamento humano em uma situação real vivenciada pelo respondente, específica e definida, e que tenha apresentado relevância para o mesmo.

A população-alvo escolhida para esta pesquisa está composta por alunos do Programa de Pós-Graduação em Ciência da Informação (PPGCI) do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), em convênio com a Universidade Federal do Rio de Janeiro (UFRJ).

Você está sendo convidado(a) para responder ao questionário e à entrevista. As duas atividades serão realizadas, cada uma, em dois momentos distintos, previamente agendados, sendo que o questionário precede a entrevista. O questionário será respondido à mão, sem interferência do pesquisador. Após a análise do questionário respondido, será realizada a entrevista por telefone. O tempo estimado para responder o questionário é de 25-30 minutos, baseado no teste-piloto já efetuado em uma população semelhante.

Espera-se, com este projeto, entender melhor como pós-graduandos na área de Ciência da Informação realizam suas buscas na web para finalidades acadêmicas e, por meio desse levantamento, elaborar e oferecer a profissionais e cientistas da informação indicações que colaborem para a obtenção de melhores resultados na busca e recuperação da informação na web.

Muito obrigada.

Rafaela Giordano, mestranda em Ciência da Informação no PPGCI – IBICT/UFRJ

Eu, _____, RG nº _____, declaro ter sido informado(a) sobre a natureza e os objetivos do presente estudo e concordo em participar, como voluntário(a), do projeto de pesquisa acima descrito.

APÊNDICE B – QUESTIONÁRIO MESTRANDOS – TESTE

Caro mestrando,

este questionário é parte integrante do estudo de campo da dissertação de mestrado de Rafaela Boeira Giordano, orientada pelo prof. Dr. Jorge Biolchini, no programa de pós-graduação em Ciência da Informação (PPGCI) do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), em convênio com a Universidade Federal do Rio de Janeiro (UFRJ).

A finalidade deste questionário é entender o processo de busca e recuperação de informações na web utilizado por cientistas da informação em fase de desenvolvimento de sua pós-graduação. Para tanto, a população-alvo caracteriza-se pelos pós-graduandos em Ciência da Informação das turmas 2009 e 2010 do PPGCI - IBICT/UFRJ.

Sua participação é muito importante uma vez que as informações obtidas neste questionário serão o núcleo da investigação e permitirão conhecer as características, os resultados, as dificuldades, entre outros aspectos, do processo realizado por cientistas da informação ao buscar e recuperar informações na web.

Ao final desta dissertação de mestrado, espera-se entender melhor como pós-graduandos na área de CI realizam buscas e recuperam informações na web. Além disso, as ações e os padrões executados pela população-alvo no processo de busca poderão ser úteis para profissionais e cientistas da informação ao realizarem buscas para seus trabalhos científicos.

O tempo estimado para o preenchimento do questionário é de **xx-xx** minutos.

Muito obrigada pela sua valiosa contribuição.

Estudo sobre busca e recuperação da informação na web sob a perspectiva de pós-graduandos em Ciência da Informação

1. Caracterização do pós-graduando

- 1.1. Qual a sua área do conhecimento de origem?
- 1.2. Em qual unidade acadêmica e instituição universitária foi concluída a sua graduação? (ex.: *Faculdade de Biblioteconomia, Universidade Federal do Rio de Janeiro*)
- 1.3. Em que ano foi concluída a sua graduação?
- 1.4. Neste mestrado do PPGCI - IBICT/UFRJ, qual a sua linha de pesquisa?
- 1.5. Além da Ciência da Informação, que áreas do conhecimento estão inseridas em sua dissertação?
- 1.6. Qual o tema de sua dissertação?
- 1.7. Em uma escala de 0 a 10, em que 0 é fraca e 10 muito alta, como você avalia a sua capacidade de buscar e recuperar informações – sobre artigos, fontes, materiais científicos – na web?
- 1.8. Por meio de qual forma você adquiriu conhecimento sobre como realizar o processo de busca e recuperação da informação? Escolha somente as duas principais. O grau de importância de cada deve ser apontado com a utilização de (1) para a mais importante de todas e (2) para a segunda mais importante de todas.
 - () a. Experiência própria, aprendendo na prática
 - () b. Com colegas, amigos, parentes, conhecidos
 - () c. Treinamento ou curso especializado sobre o assunto
 - () d. Leitura de livros e/ou artigos especializados sobre o assunto

() e. Leitura de conteúdos sobre o assunto em sites na Internet

() f. Curso a distância via internet

() g. Outra modalidade? Qual?

2. Processo de busca e recuperação da informação com resultado positivo

2.1. Você se lembra de **exemplo recente** de uma busca realizada na web que tenha resultado na recuperação de artigo/conteúdo **relevante** para o seu projeto de mestrado? Há quanto tempo ocorreu?

Há () horas

Há () dias

Há () semanas

2.2. Que **situação** o levou a realizar essa busca? Qual era o **problema** a ser solucionado?

2.3. Que **informação específica** você estava procurando ao recorrer à web? Isto é, qual era o **assunto** da busca? Qual era a **questão** que estava necessitando de uma resposta?

2.4. Que **termos** foram inseridos no campo de busca? Com quais **palavras** a busca foi realizada?

2.5. Por meio de qual **recurso** informacional deu início à busca? (*ex.: site, ferramenta de busca, mecanismo de busca*)

2.6. Que **motivo** o levou recorrer a essa fonte para buscar a informação de que necessitava?

2.7. Durante esse processo de busca, utilizou **recurso padrão** oferecido pelo mecanismo escolhido ou **recurso avançado**? Se usou recurso avançado, especifique um pouco mais quais foram os campos utilizados e por quê. (*ex.: busca avançada, pesquisa avançada*)

2.8. Que tipo de **conteúdo** resultou da sua busca?

2.9. Quais foram os critérios utilizados para **filtrar** os resultados obtidos? Como foram extraídos os **mais relevantes** para sua busca?

2.10. Em que aspectos específicos o **resultado** obtido foi **relevante** para a sua pesquisa de mestrado? No que ele **contribuiu** para o seu projeto?

2.11. Qual foi o **impacto** da informação recuperada para o seu projeto? Como o resultado obtido **afetou** os passos seguintes no desenvolvimento do seu projeto?

2.12. O que poderia ter ocorrido se o resultado não tivesse sido obtido?

2.13. Esta busca atendeu **satisfatoriamente** a sua necessidade informacional e deu por encerrado o processo? Ou foi necessário buscar mais informações?

2.14. Durante o processo de busca, passou por alguma **dificuldade**? Qual? Caso tenha passado, o que poderia ser utilizado no sentido de **facilitar** a obtenção da informação buscada?

3. Processo de busca e recuperação da informação com resultado negativo

3.1. Você se lembra de **exemplo recente** de uma busca realizada na web que **não** resultou na recuperação de artigo/conteúdo relevante para o seu projeto de mestrado ou que tenha sido **insatisfatória**? Há quanto tempo ocorreu?

Há () horas

Há () dias

Há () semanas

3.2. Que **situação** o levou a realizar essa busca? Qual era o **problema** a ser solucionado?

- 3.3. Que **informação específica** você estava procurando ao recorrer à web? Isto é, qual era o **assunto** da busca? Qual era a **questão** que estava necessitando de uma resposta?
- 3.4. Que **termos** foram inseridos no campo de busca? Com quais **palavras** a busca foi realizada?
- 3.5. Por meio de qual **recurso** na web deu início à busca? (ex.: site, ferramenta de busca, mecanismo de busca)
- 3.6. Que **motivo** o levou recorrer a essa fonte para buscar a informação de que necessitava?
- 3.7. Durante esse processo de busca, utilizou **recurso padrão** oferecido pelo mecanismo escolhido ou **recurso avançado**? Se usou recurso avançado, especifique um pouco mais quais foram os campos utilizados e por quê. (ex.: busca avançada, pesquisa avançada)
- 3.8. Durante o processo de busca, passou por alguma **dificuldade**? Qual? Caso tenha passado, o que poderia ser utilizado no sentido de **facilitar** a obtenção da informação buscada?
- 3.9. Que tipo de **resultado** teria sido mais útil para seu projeto de pesquisa?
- 3.10. Qual **impacto** a não recuperação da informação trouxe para o seu projeto? Como a não obtenção desse resultado **afetou** os passos seguintes no desenvolvimento do seu projeto?
- 3.11. O que poderia ter acontecido caso você tivesse obtido algum resultado?
- 3.12. Como você reavaliaria a sua capacidade de buscar e recuperar informações – sobre artigos, fontes, materiais científicos – na web? Em uma escala de 0 a 10, em que 0 é fraca e 10 muito alta, qual seria a nova avaliação?

APÊNDICE C – QUESTIONÁRIO MESTRANDOS – VERSÃO FINAL

Caro mestrando

Este questionário é parte integrante do estudo de campo da dissertação de mestrado de Rafaela Boeira Giordano, orientada pelo prof. Dr. Jorge Biolchini, no programa de pós-graduação em Ciência da Informação (PPGCI) do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), em convênio com a Universidade Federal do Rio de Janeiro (UFRJ).

A finalidade deste questionário é entender o processo de busca e recuperação de informações na web empregado por cientistas da informação em fase de desenvolvimento de sua pós-graduação. Para tanto, a população-alvo caracteriza-se pelos pós-graduandos em Ciência da Informação das turmas 2009 e 2010 do PPGCI - IBICT/UFRJ.

Sua participação é muito importante uma vez que as informações obtidas neste questionário serão o núcleo da investigação e permitirão conhecer as características, os resultados, as dificuldades, entre outros aspectos, do processo realizado por cientistas da informação ao buscar e recuperar informações na web.

Ao final desta dissertação de mestrado, espera-se entender melhor como pós-graduandos na área de CI realizam buscas e recuperam informações na web. Além disso, as ações e os padrões executados pela população-alvo no processo de busca poderão ser úteis para profissionais e cientistas da informação ao realizarem buscas para seus trabalhos científicos.

O tempo estimado para o preenchimento do questionário é de **25-30 minutos**.

Muito obrigada pela sua valiosa contribuição.

Estudo sobre busca e recuperação da informação na web sob a perspectiva de pós-graduandos em Ciência da Informação

1. Caracterização do pós-graduando

1.1. Qual a sua área do conhecimento de origem?

1.2. Em qual unidade acadêmica e instituição universitária foi concluída a sua graduação? (ex.: *Faculdade de Biblioteconomia, Universidade Federal do Rio de Janeiro*)

1.3. Em que ano foi concluída a sua graduação?

1.4. Neste mestrado do PPGCI - IBICT/UFRJ, qual a sua linha de pesquisa?

1.5. Além da Ciência da Informação, que áreas do conhecimento estão inseridas em sua dissertação?

1.6. Qual o tema de sua dissertação?

1.7. Em uma escala de 0 a 10, em que 0 é muito baixa e 10 é muito alta, como você avalia a sua capacidade de buscar e recuperar informações – sobre artigos, fontes, materiais científicos – na web?

1.8. Por meio de qual forma você adquiriu conhecimento sobre como realizar o processo de busca e recuperação da informação? Escolha somente as duas principais. O grau de importância de cada deve ser apontado com a utilização de (1) para a mais importante de todas e (2) para a segunda mais importante de todas.

() a. Experiência própria, aprendendo na prática

() b. Com colegas, amigos, parentes, conhecidos

() c. Treinamento ou curso especializado sobre o assunto

() d. Leitura de livros e/ou artigos especializados sobre o assunto

- () e. Leitura de conteúdos sobre o assunto em sites na Internet
- () f. Curso a distância via internet
- () g. Outra modalidade? Qual? _____

2. Processo de busca e recuperação da informação com resultado positivo

2.1. Baseando-se em um **exemplo recente** de uma busca realizada na web que tenha resultado na recuperação de conteúdo **relevante** para o seu projeto de mestrado, responda? (artigos em periódicos, páginas da web, arquivos em formato pdf, doc ou ppt, teses e dissertações, sites, blogs, livros, citações, páginas não encontradas, etc.)

Há quanto tempo ocorreu a busca?

Há () horas

Há () dias

Há () semanas

2.2. Que **situação** o levou a realizar essa busca, qual era o **problema** a ser solucionado?

2.3. Em relação ao **assunto** buscado, qual **informação específica** você estava procurando ao recorrer à web?

2.4. Quais **termos/palavras** foram inseridos no campo de busca?

2.5. Por meio de qual **recurso** informacional deu início à busca?
(*ex.: site, ferramenta de busca, mecanismo de busca, base de dados*)

2.6. Que **motivo** o levou a recorrer a essa fonte para buscar a informação de que necessitava?

2.7. Durante este processo de busca, você utilizou **recurso padrão** ou **recurso avançado** oferecido pelo mecanismo escolhido? (*ex.: busca avançada, pesquisa avançada*) Se usou recurso avançado, especifique um pouco mais quais foram os campos utilizados e por quê.

2.8. Que tipos de **conteúdos** resultaram da sua busca?

- () Artigos em periódicos;
- () Páginas da web;
- () Arquivos em formato pdf, doc ou ppt;
- () Teses e dissertações
- () Citações
- () Livros
- () Outro. Qual?

2.9. Quais critérios foram utilizados para **filtrar** os resultados obtidos, como fez para separar o que seria útil para seu projeto?

2.10. Em que aspectos específicos o **resultado** obtido foi **relevante** para a sua pesquisa de mestrado?

2.11. Qual foi o grau do **impacto** da informação recuperada para o seu projeto, considerando uma, e somente uma, das opções abaixo:

- () Nenhum impacto;
- () Impacto pequeno;
- () Impacto médio;
- () Impacto grande;

() Impacto fundamental.

2.12. Como o resultado **afetou** os passos seguintes no desenvolvimento do seu projeto?

2.13. O que poderia ter ocorrido se o resultado **não** tivesse sido obtido?

2.14. Qual foi o seu grau de satisfação em relação à busca em uma escala de 0 a 10, sendo 0 muito baixo e 10, muito alto?

2.15. Qual foi a maior **dificuldade** encontrada no processo em relação aos conteúdos obtidos? O que você fez para resolvê-la?

2.16. Qual foi o maior **aprendizado** proporcionado por esta experiência?

2.17. Como você **reavaliaria** a sua capacidade de buscar e recuperar informações – sobre artigos, fontes, materiais científicos – na web? Em uma escala de 0 a 10, em que 0 é muito baixa e 10 é muito alta, qual seria a nova auto-avaliação?

APÊNDICE D – QUESTIONÁRIO DOUTORANDOS – TESTE

Caro doutorando,

este questionário é parte integrante do estudo de campo da dissertação de mestrado de Rafaela Boeira Giordano, orientada pelo prof. Dr. Jorge Biolchini, no programa de pós-graduação em Ciência da Informação (PPGCI) do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), em convênio com a Universidade Federal do Rio de Janeiro (UFRJ).

A finalidade deste questionário é entender o processo de busca e recuperação de informações na web utilizado por cientistas da informação em fase de desenvolvimento de sua pós-graduação. Para tanto, a população-alvo caracteriza-se pelos pós-graduandos em Ciência da Informação das turmas 2009 e 2010 do PPGCI - IBICT/UFRJ.

Sua participação é muito importante uma vez que as informações obtidas neste questionário serão o núcleo da investigação e permitirão conhecer as características, os resultados, as dificuldades, entre outros aspectos, do processo realizado por cientistas da informação ao buscar e recuperar informações na web.

Ao final desta dissertação de mestrado, espera-se entender melhor como pós-graduandos na área de CI realizam buscas e recuperam informações na web. Além disso, as ações e os padrões executados pela população-alvo no processo de busca poderão ser úteis para profissionais e cientistas da informação ao realizarem buscas para seus trabalhos científicos.

O tempo estimado para o preenchimento do questionário é de **XX** minutos.

Muito obrigada pela sua valiosa contribuição.

Estudo sobre busca e recuperação da informação na web sob a perspectiva de pós-graduandos em Ciência da Informação

1. Caracterização do pós-graduando

- 1.1. Qual a sua área do conhecimento de origem?
- 1.2. Em qual unidade acadêmica e instituição universitária foi concluída a sua graduação? (ex.: *Faculdade de Biblioteconomia, Universidade Federal do Rio de Janeiro*)
- 1.3. Em que ano foi concluída a sua graduação?
- 1.4. Em que área do conhecimento estava inserido o seu mestrado?
- 1.5. Em qual unidade acadêmica e instituição universitária foi concluído o seu mestrado? (ex.: *Faculdade de Biblioteconomia, Universidade Federal do Rio de Janeiro*)
- 1.6. Em que ano foi concluído o seu mestrado?
- 1.7. Neste doutorado do PPGCI - IBICT/UFRJ, qual a sua linha de pesquisa?
- 1.8. Além da Ciência da Informação, que áreas do conhecimento estão inseridas em sua tese?
- 1.9. Qual o tema de sua tese?
- 1.10. Em uma escala de 0 a 10, em que 0 é fraca e 10 muito alta, como você avalia a sua capacidade de buscar e recuperar informações – sobre artigos, fontes, materiais científicos – na web?
- 1.11. Por meio de qual forma você adquiriu conhecimento sobre como realizar o processo de busca e recuperação da informação? Escolha somente as duas principais. O grau de importância de cada deve ser apontado com a utilização de (1) para a mais importante de todas e (2) para a segunda mais importante de todas.

- () a. Experiência própria, aprendendo na prática
- () b. Com colegas, amigos, parentes, conhecidos
- () c. Treinamento ou curso especializado sobre o assunto
- () d. Leitura de livros e/ou artigos especializados sobre o assunto
- () e. Leitura de conteúdos sobre o assunto em sites na Internet
- () f. Curso a distância via internet
- () g. Outra modalidade? Qual?

2. Processo de busca e recuperação da informação com resultado positivo

2.1. Você se lembra de **exemplo recente** de uma busca realizada na web que tenha resultado na recuperação de artigo/conteúdo **relevante** para o seu projeto de doutorado? Há quanto tempo ocorreu?

Há () horas

Há () dias

Há () semanas

2.2. Que **situação** o levou a realizar essa busca? Qual era o **problema** a ser solucionado?

2.3. Que **informação específica** você estava procurando ao recorrer à web? Isto é, qual era o **assunto** da busca? Qual era a **questão** que estava necessitando de uma resposta?

2.4. Que **termos** foram inseridos no campo de busca? Com quais **palavras** a busca foi realizada?

2.5. Por meio de qual **recurso** informacional deu início à busca? (*ex.: site, ferramenta de busca, mecanismo de busca*)

2.6. Que **motivo** o levou recorrer a essa fonte para buscar a informação de que necessitava?

2.7. Durante esse processo de busca, utilizou **recurso padrão** oferecido pelo mecanismo escolhido ou **recurso avançado**? Se usou recurso avançado, especifique um pouco mais quais foram os campos utilizados e por quê. (*ex.: busca avançada, pesquisa avançada*)

2.8. Que tipo de **conteúdo** resultou da sua busca?

2.9. Quais foram os critérios utilizados para **filtrar** os resultados obtidos? Como foram extraídos os **mais relevantes** para sua busca?

2.10. Em que aspectos específicos o **resultado** obtido foi **relevante** para a sua pesquisa de doutorado? No que ele **contribuiu** para o seu projeto?

2.11. Qual foi o **impacto** da informação recuperada para o seu projeto? Como o resultado obtido **afetou** os passos seguintes no desenvolvimento do seu projeto?

2.12. O que poderia ter ocorrido se o resultado não tivesse sido obtido?

2.13. Esta busca atendeu **satisfatoriamente** a sua necessidade informacional e deu por encerrado o processo? Ou foi necessário buscar mais informações?

2.14. Durante o processo de busca, passou por alguma **dificuldade**? Qual? Caso tenha passado, o que poderia ser utilizado no sentido de **facilitar** a obtenção da informação buscada?

3. Processo de busca e recuperação da informação com resultado negativo

3.1. Você se lembra de **exemplo recente** de uma busca realizada na web que **não** resultou na recuperação de artigo/conteúdo relevante para o seu projeto de doutorado ou que tenha sido **insatisfatória**? Há quanto tempo ocorreu?

Há () horas

Há () dias

Há () semanas

3.2. Que **situação** o levou a realizar essa busca? Qual era o **problema** a ser solucionado?

3.3. Que **informação específica** você estava procurando ao recorrer à web? Isto é, qual era o **assunto** da busca? Qual era a **questão** que estava necessitando de uma resposta?

3.4. Que **termos** foram inseridos no campo de busca? Com quais **palavras** a busca foi realizada?

3.5. Por meio de qual **recurso** na web deu início à busca? (ex.: site, ferramenta de busca, mecanismo de busca)

3.6. Que **motivo** o levou recorrer a essa fonte para buscar a informação de que necessitava?

3.7. Durante esse processo de busca, utilizou **recurso padrão** oferecido pelo mecanismo escolhido ou **recurso avançado**? Se usou recurso avançado, especifique um pouco mais quais foram os campos utilizados e por quê. (ex.: busca avançada, pesquisa avançada)

3.8. Durante o processo de busca, passou por alguma **dificuldade**? Qual? Caso tenha passado, o que poderia ser utilizado no sentido de **facilitar** a obtenção da informação buscada?

3.9. Que tipo de **resultado** teria sido mais útil para seu projeto de pesquisa?

3.10. Qual **impacto** a não recuperação da informação trouxe para o seu projeto? Como a não obtenção desse resultado **afetou** os passos seguintes no desenvolvimento do seu projeto?

3.11. O que poderia ter acontecido caso você tivesse obtido algum resultado?

3.12. Como você reavaliaria a sua capacidade de buscar e recuperar informações – sobre artigos, fontes, materiais científicos – na web? Em uma escala de 0 a 10, em que 0 é fraca e 10 muito alta, qual seria a nova avaliação?

APÊNDICE E – QUESTIONÁRIO DOUTORANDOS – VERSÃO FINAL

Caro doutorando

Este questionário é parte integrante do estudo de campo da dissertação de mestrado de Rafaela Boeira Giordano, orientada pelo prof. Dr. Jorge Biolchini, no programa de pós-graduação em Ciência da Informação (PPGCI) do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), em convênio com a Universidade Federal do Rio de Janeiro (UFRJ).

A finalidade deste questionário é entender o processo de busca e recuperação de informações na web empregado por cientistas da informação em fase de desenvolvimento de sua pós-graduação. Para tanto, a população-alvo caracteriza-se pelos pós-graduandos em Ciência da Informação das turmas 2009 e 2010 do PPGCI - IBICT/UFRJ.

Sua participação é muito importante uma vez que as informações obtidas neste questionário serão o núcleo da investigação e permitirão conhecer as características, os resultados, as dificuldades, entre outros aspectos, do processo realizado por cientistas da informação ao buscar e recuperar informações na web.

Ao final desta dissertação de mestrado, espera-se entender melhor como pós-graduandos na área de CI realizam buscas e recuperam informações na web. Além disso, as ações e os padrões executados pela população-alvo no processo de busca poderão ser úteis para profissionais e cientistas da informação ao realizarem buscas para seus trabalhos científicos.

O tempo estimado para o preenchimento do questionário é de **25-30** minutos.

Muito obrigada pela sua valiosa contribuição.

Estudo sobre busca e recuperação da informação na web sob a perspectiva de pós-graduandos em Ciência da Informação

1. Caracterização do pós-graduando

1.1. Qual a sua área do conhecimento de origem?

1.2. Em qual unidade acadêmica e instituição universitária foi concluída a sua graduação?_ (ex.: *Faculdade de Biblioteconomia, Universidade Federal do Rio de Janeiro*)

1.3. Em que ano foi concluída a sua graduação?

1.4. Além da Ciência da Informação, em que área do conhecimento estava inserido o seu mestrado?

1.5. Em qual unidade acadêmica e instituição universitária foi concluído o seu mestrado? _ (ex.: *Faculdade de Biblioteconomia, Universidade Federal do Rio de Janeiro*)

1.6. Em que ano foi concluído o seu mestrado?

1.7. Neste doutorado do PPGCI - IBICT/UFRJ, qual a sua linha de pesquisa?

1.8. Além da Ciência da Informação, que áreas do conhecimento estão inseridas em sua tese?

1.9. Qual o tema de sua tese?

1.10. Em uma escala de 0 a 10, em que 0 é muito baixa e 10 é muito alta, como você avalia a sua capacidade de buscar e recuperar informações – sobre artigos, fontes, materiais científicos – na web?

1.11. Por meio de qual forma você adquiriu conhecimento sobre como realizar o processo de busca e recuperação da informação? Escolha somente as duas principais. O grau de importância de cada deve ser apontado com a utilização de (1) para a mais importante de todas e (2) para a segunda mais importante de todas.

- () a. Experiência própria, aprendendo na prática
- () b. Com colegas, amigos, parentes, conhecidos
- () c. Treinamento ou curso especializado sobre o assunto
- () d. Leitura de livros e/ou artigos especializados sobre o assunto
- () e. Leitura de conteúdos sobre o assunto em sites na Internet
- () f. Curso a distância via internet
- () g. Outra modalidade? Qual?

2. Processo de busca e recuperação da informação com resultado positivo

2.1. Baseando-se em um **exemplo recente** de uma busca realizada na web que tenha resultado na recuperação de conteúdo **relevante** para o seu projeto de doutorado, responda: (artigos em periódicos, páginas da web, arquivos em formato pdf, doc ou ppt, teses e dissertações, sites, blogs, livros, citações, páginas não encontradas, etc.)

Há quanto tempo ocorreu a busca?

Há () horas; ou

Há () dias; ou

Há () semanas

2.2. Que **situação** o levou a realizar essa busca, qual era o **problema** a ser solucionado?

2.3. Em relação ao **assunto** buscado, qual **informação específica** você estava procurando ao recorrer à web?

2.4. Quais **termos/palavras** foram inseridos no campo de busca?

2.5. Por meio de qual **recurso** informacional deu início à busca?
(*ex.: site, ferramenta de busca, mecanismo de busca, base de dados*)

2.6. Que **motivo** o levou a recorrer a essa fonte para buscar a informação de que necessitava?

2.7. Durante este processo de busca, você utilizou **recurso padrão** ou **recurso avançado** oferecido pelo mecanismo escolhido? (*ex.: busca avançada, pesquisa avançada*) Se usou recurso avançado, especifique um pouco mais quais foram os campos utilizados e por quê.

2.8. Que tipos de **conteúdos** resultaram da sua busca?

- () Artigos em periódicos;
- () Páginas da web;
- () Arquivos em formato pdf, doc ou ppt;
- () Teses e dissertações
- () Citações
- () Livros
- () Outro. Qual?

2.9. Quais critérios foram utilizados para **filtrar** os resultados obtidos, como fez para separar o que seria útil para seu projeto?

2.10. Em que aspectos específicos o **resultado** obtido foi **relevante** para a sua pesquisa de doutorado?

2.11. Qual foi o grau do **impacto** da informação recuperada para o seu projeto, considerando uma, e somente uma, das opções abaixo:

- () Nenhum impacto;
- () Impacto pequeno;
- () Impacto médio;
- () Impacto grande;
- () Impacto fundamental.

2.12. Como o resultado **afetou** os passos seguintes no desenvolvimento do seu projeto?

2.13. O que poderia ter ocorrido se o resultado **não** tivesse sido obtido?

2.14. Qual foi o seu grau de **satisfação** em relação à busca, em uma escala de 0 a 10, sendo 0 muito baixo e 10 muito alto?

2.15. Qual foi a maior **dificuldade** encontrada no processo em relação aos conteúdos obtidos? O que você fez para resolvê-la?

2.16. Qual foi o maior **aprendizado** proporcionado por esta experiência?

2.17. Como você **reavaliaria** a sua capacidade de buscar e recuperar informações – sobre artigos, fontes, materiais científicos – na web? Em uma escala de 0 a 10, em que 0 é muito baixa e 10 é muito alta, qual seria a nova auto-avaliação?