



Biblios

E-ISSN: 1562-4730

editor@bibliosperu.com

Julio Santillán Aldana, ed.

Perú

de Sales, Rodrigo; Godoy Viera, Angel Freddy
Grupos e linhas de pesquisa sobre recuperação da informação no Brasil
Biblios, núm. 28, abril-junio, 2007, pp. 1-14
Julio Santillán Aldana, ed.
Lima, Perú

Disponível em: <http://www.redalyc.org/articulo.oa?id=16114070004>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica
Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

Grupos e linhas de pesquisa sobre recuperação da informação no Brasil

Rodrigo de Sales

Mestrando em Ciência da Informação.
Universidade Federal de Santa Catarina – Brasil
rodrigo.sales.s@gmail.com

Angel Freddy Godoy Viera

Professor do Departamento de Ciência da Informação
Universidade Federal de Santa Catarina – Brasil
godoy@cin.ufsc.br

Resumo

A recuperação da informação é assunto capital nos dias atuais devido à infinidade de informações produzidas e divulgadas em grande escala em nível Global, amplamente difundidas pelas novas tecnologias de informação e comunicação. Consiste em uma das linhas de pesquisa mais estudadas na área de Ciência da Informação, assim como em outras áreas como a Ciência da Computação, engenharias, entre outras. Este trabalho consiste em um mapeamento no Diretório dos Grupos de Pesquisa no Brasil do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) dos grupos brasileiros que estudam a recuperação de informação. As variáveis procuradas para a análise e interpretação foram a identificação dos grupos de pesquisa, quando foram criados, quais as instituições os abrigam, a distribuição geográfica dos grupos em território brasileiro, e quais as áreas do conhecimento dos grupos. Foi constatado que a maioria dos grupos de pesquisas estão ligadas à Instituições de Ensino Superior e foram formados a partir da década de 1990. Foi possível verificar também que mais da metade dos grupos estão localizados na região sudeste do Brasil, que é a região mais desenvolvida economicamente no país. As áreas de conhecimento que predominantemente investigam a recuperação da informação no Brasil são a Ciência da Computação e a Ciência da Informação. As linhas de pesquisa sobre questões tecnológicas predominam nos grupos da área de Ciência da Computação e, questões mais focadas para o usuário da informação são estudadas predominantemente nos grupos de Ciência da Informação. Observou-se que a pesquisa em recuperação da informação no Brasil está em ascensão e, embora esteja sendo desenvolvida predominantemente pelos grupos de Ciência da Computação e da Ciência da Informação, desperta interesse de pesquisadores dos mais variados domínios do conhecimento. A tendência da pesquisa brasileira aponta para as soluções inteligentes de recuperação da informação e banco de dados, confirmando a preocupação em fazer com que os sistemas de informação sejam sistemas que possibilitam a interação entre pessoas e informação.

Palavras-chave

Recuperação da Informação – Brasil. Grupos de Pesquisa. Linhas de Pesquisa. Diretório dos Grupos de Pesquisa do Brasil.

Abstract

Information retrieval is a key topic nowadays due to the amount of information created and published in large scale worldwide, highly spread by means of the new information and communication technologies. It is one of the most studied research areas in Information Science, as well as in others such as Computer Science, Engineering, among others. This study consists of a mapping in the *Diretório dos Grupos de Pesquisa no Brasil* from the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) of the Brazilian groups that study the retrieval information. The variables that were searched for this analysis and interpretation were the identification of the research groups, when they were created, which institutions they belong to, the geographical distribution of the groups around the Brazilian territory, and what the areas of knowledge of the groups are. It was found that the majority of the research groups are connected with higher education institutions and were formed between the years 1990 and 2006. It was possible to also verify that more than a half of the groups are located in the Southwest, which is the most developed region in terms of economy in the country. The areas of knowledge that mainly study information retrieval in Brazil are Computer Science and the Information Science. Research areas around technological issues are predominant in the Computer Science groups and issues focused with information users are studied mainly by the Information Science groups. It was possible to observe that research in information retrieval in Brazil is increasing, though it is being mainly done by Computer Science and Information Science groups, and it arises interest from researchers from the most different areas of knowledge. The trend in terms of the research conducted in Brazil points to intelligent information retrieval solutions and databases, confirming the preoccupation with information systems being systems which are able to provide interaction between people and information.

Keywords

Information Retrieval – Brazil. Research Groups. Research Areas. Research Groups of Brazil Directory

1. Introdução

A recuperação da informação é assunto fundamental na sociedade contemporânea que esta imersa em uma infinidade de informações. Com a evolução exponencial no desenvolvimento de tecnologias de informação e comunicação, sendo seus principais expoentes a rede mundial de computadores Internet e a Web, a produção e a divulgação de informação se tornaram mais numerosas e velozes. Com isso, tornou-se mais dificultosa a recuperação eficiente da informação. São várias as áreas do conhecimento que vêm focando seus esforços de pesquisa na recuperação da informação.

Em relação ao Brasil, alguns questionamentos que este trabalho visa responder são: quem está estudando os fundamentos da Recuperação da Informação no Brasil?, quais as áreas do conhecimento em que esses estudos são mais frequentes?, que especificamente está sendo estudado?.

Para responder essas perguntas, realizou-se um levantamento da situação atual da pesquisa referente à recuperação de informações no Brasil, para logo traçar algumas considerações sobre o assunto. Neste trabalho realizou-se um levantamento na Web no Diretório dos Grupos de Pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) do Brasil, sobre os grupos de pesquisa em Recuperação da Informação no Brasil, buscando mapear aquilo que vem sendo estudado referente o tema, e quem vem potencialmente fomentando tais estudos. Para tanto, identificou-se a evolução cronológica, a distribuição geográfica e institucional, além da distribuição temática das linhas de pesquisa e das áreas predominantemente hospedeiras dos respectivos grupos de estudo.

2. A Ciência da Informação e a Recuperação da Informação

A literatura da Ciência da Informação é concebida com divergências no que tange o surgimento dessa ciência. Existem duas linhas que sustentam diferentes pontos de vista histórico para o surgimento da Ciência da Informação, a primeira atribui aos estudos de Paul Otlet e Henri La Fontaine, referentes à bibliografia e à documentação, os elementos predecessores da Ciência da Informação; e a segunda atribui o aparecimento dessa ciência ao surgimento de tecnologias para o trato da informação científica e tecnológica no período Pós Segunda Guerra Mundial. Esta segunda impulsionada pelo artigo publicado em 1945 por Vannevar Bush, intitulado "As We May Think", que trás as tecnologias de informação como preocupação capital para o período Pós Guerra. Hayes (1999), traçando uma evolução histórica do desenvolvimento da Ciência da Informação no Estados Unidos, concorda com a idéia apresentada nesta segunda linha de pensamento afirmando que o primeiro período marcante da evolução da Ciência da Informação, compreendido no período de 1948 à 1964, foi marcado pela necessidade de investir em tecnologias que visem a automatização da informação e unidades de informação, iniciado pelo pensamento de Vannevar Bush.

Para Ingwersen (1992), o uso do termo Ciência da Informação data de 1958 quando se formou o Institute of Information Scientists (IIS) no Reino Unido com o intuito de diferenciar o cientista da informação do cientista de laboratório. Os membros do IIS foram cientistas de diversas áreas e estavam preocupados com o estudo da informação e dos processos envolvidos na comunicação científica. Lidavam por tanto, com o problema de organização, incremento e disseminação do conhecimento registrado, gerado antes da Segunda Guerra Mundial.

Mas, no que diz respeito aos aspectos conceituais da Ciência da Informação, definições amplamente divulgadas convergem em afirmar que a Ciência da Informação

é uma disciplina que investiga as propriedades e o comportamento da informação, as forças que regem o fluxo informacional e os meios de processamento da informação para a otimização do acesso e uso. Está interessada com esse corpo de conhecimento relacionado com a criação, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação (Borko 1968:3).

Le Coadic (1996), concorda com Borko, Shera e Cleveland em afirmar que a Ciência da Informação tem o objetivo de estudar as propriedades gerais da informação (tais como natureza, gênese, efeitos), e acrescenta afirmando que a Ciência da

Informação visa analisar os processos de construção, comunicação e uso da informação, e a concepção dos produtos e sistemas que tornam possíveis a construção, comunicação, armazenamento e uso. Segundo Le Coadic (1996), a Ciência da Informação é uma ciência social rigorosa baseada em tecnologias também rigorosas.

Saracevic (1999), afirma que a Ciência da informação possui três características gerais, elas são: Ela é interdisciplinar por natureza, esta conectada inexoravelmente à tecnologia e participa ativamente na evolução da sociedade da informação, tendo um forte componente social e humano, acima e mais além da tecnologia.

O autor afirma que a história da Ciência da Informação pode ser resumida pelos principais conceitos utilizados para lidar para o tratamento da informação. A primeira idéia original que emerge nos anos 1950 é a recuperação da informação para o processamento da informação baseada na lógica formal, a segunda surge pouco tempo depois é a relevância orientando e associando o processo com as necessidades de informação dos usuários e avaliação e terceiro, que chegou aproximadamente duas décadas depois a interação, que permite intercambio direto e retroalimentação entre sistemas e os usuários engajados no processo de recuperação da informação.

Marchiori definiu a Ciência da Informação como

a ciência que se ocupa do estudo da informação em si, isto é, a teoria e prática que envolve sua criação, identificação, coleta, validação, representação, recuperação e uso, tendo como princípio o fato de que existe um produtor/consumidor de informação que busca, nesta, um sentido e uma finalidade (Marchiori 2002: 61).

Em relação ao relacionamento intrínseco da Ciência da Informação com a recuperação da informação, Saracevic (1999) enfatiza que a Recuperação da Informação (RI) foi e continua sendo a principal área de pesquisa da Ciência da Informação.

Sobre pesquisa em recuperação da informação o autor afirma que nos anos 1990 surgem novas áreas de pesquisa entre elas estudos de interação, busca na Internet, recuperação de informação multimídia, recuperação de informação poliglota e bibliotecas digitais.

A recuperação da informação é assunto essencial na sociedade moderna submerso em uma infinidade de informações. Com a rápida evolução das tecnologias de informação e comunicação, a produção e a divulgação do conhecimento em forma de informação se tornaram mais numerosas e velozes, dificultando a recuperação eficiente das informações.

Em relação ao origem do termo Recuperação da Informação ela foi difundido no cenário internacional por Calvin Mooers na década de 1950. Segundo Ingwersen (1992), a recuperação da informação esta interessada nos processos que envolvem a representação, armazenamento, busca e descoberta da informação que é relevante para os requerimentos de informação desejados pelos usuários. Ele também afirma que a recuperação da informação é um dos campos de pesquisas principais da ciência da informação que tem por objetivo estudar e compreender os processo de recuperação da informação para projetar, construir e testar sistemas de recuperação que podem facilitar a efetiva comunicação das informações desejados entre o produtor e consumidor da informação.

Baeza-Yates e Ribeiro-Neto (1999), vão de encontro com a definição anterior e asseveram que a recuperação da informação lida com a representação, armazenamento, organização e acesso aos itens de informação. Eles fazem também uma distinção entre recuperação de dados e recuperação da informação, afirmam que recuperar dados consiste principalmente em identificar quais documentos de um

determinado acervo possuem como palavras-chave de representação de conteúdo o mesmo termo utilizado pelo usuário no momento da consulta.

Para Baeza-Yates e Ribeiro-Neto (1999), esse processo de recuperação não satisfaz a necessidade de informação do usuário, portanto, não passa de recuperação de dados. Segundo os mesmos, recuperação de informação consiste em recuperar informações a respeito de um assunto desejado, e não simplesmente recuperar documentos que satisfaçam sentenças de consulta. Com as atenções voltadas para a satisfação de necessidades de informação, os estudos e programas de recuperação da informação passaram a explorar a recuperação inteligente da informação.

Refletindo a respeito de como os sistemas de recuperação da informação podem agir de maneira inteligente, Belkin (2004) propõe uma abordagem para que os sistemas de RI estendam a funcionalidade de servir como meio de interação entre pessoas e informação. A abordagem de Belkin sugere que os sistemas de RI funcionem integrando o comportamento da busca por informação realizada pelo usuário com o texto, ancorados na observação do comportamento do usuário no momento da busca e em diversas situações. Segundo o autor, esse novo modo de enxergar os sistemas de RI pode alcançar uma melhor compreensão no momento de construí-los.

O campo de estudo da Recuperação da Informação (RI), para Marin Milanês e Torres Velásquez (2006), esteve, e continua estando, sob a influência de dois paradigmas ou tendências de suas práticas no decorrer de seu desenvolvimento: uma focada nos sistemas de informação e outra focada nos usuários de informação.

Ingwersen (1992) reforça esta afirmação indicando que as duas maiores abordagens que caracterizaram a pesquisa e desenvolvimento da recuperação da informação desde os anos 50 são a abordagem tradicional ou clássica e a abordagem baseados no usuário, afirma a mediados dos anos 80 se fundem, se transformando numa abordagem cognitiva.

Sobre os conceitos e problemas estudados na área da recuperação da informação, Ingwersen (1992) assevera que do ponto de vista da pesquisa são:

- A representação da informação, definida pelo autor como sobre que assunto um determinado documento, texto ou imagem trata e diferencia a representação do autor da representação do indexador. A representação do autor é definida pelo autor como sendo uma representação da informação com termos derivados do próprio documento, sendo a base para as teorias de indexação automática e técnicas de emparelhamento de texto. A representação do indexador refere-se a um indexador humano que realiza a indexação ou classificação do documento, resumindo o conteúdo da mensagem em cada documento. Os sistemas de classificação facetada, tesouros e indexação com vocabulário controlado são instrumentos utilizados neste caso para a representação da informação.
- A relevância, definida como a medida do grau de correspondência ou utilidade existente entre um texto ou documento e a consulta ou necessidades de informação determinados por uma pessoa e afirma que a relevância é um valor de natureza pragmática, associada com o espaço do problema e o estado de conhecimento do cada usuário.
- A avaliação, lida com a avaliação dos modelos de recuperação da informação.

Evidencia-se uma transformação no lidar com a recuperação da informação, acompanhando a evolução nos estudos de Ciência da Informação atrelados às tecnologias de informação e necessidades de informação.

Este trabalho apresenta um levantamento realizado com o intuito de mapear a situação atual das pesquisas voltadas à recuperação da informação no Brasil. Através da qual será possível exibir um panorama da diversidade nos estudos realizados no país nesta área temática.

3. Metodologia

O presente artigo é uma pesquisa documental realizada na web durante o mês de setembro de 2006. Trata-se de um levantamento da situação da pesquisa em Recuperação da Informação (RI) no Brasil. Para o referido levantamento, buscou-se primeiramente identificar as fontes de informação on-line disponíveis na web.

Por se tratar de uma investigação cujo foco é a pesquisa científica e tecnológica no Brasil, optou-se em examinar as páginas dos órgãos responsáveis pela pesquisa no país. O Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) são as principais agências nacionais responsáveis pelo fomento da pesquisa no Brasil. O CNPq (<http://www.cnpq.br>) é uma agência do Ministério da Ciência e Tecnologia (MCT) destinada ao fomento da pesquisa científica e tecnológica e à formação de recursos humanos para a pesquisa no país. Sua história está diretamente ligada ao desenvolvimento científico e tecnológico do Brasil contemporâneo. A CAPES (<http://www.capes.gov.br>) vem desempenhando papel fundamental na expansão e consolidação da pós-graduação stricto sensu (mestrado e doutorado) em todos os estados da Federação Brasileira, e suas maiores conquistas estão no campo da Educação e da Ciência & Tecnologia.

Após a verificação no sítio da CAPES, foi constatado que seria grande a dificuldade de identificar os grupos e as linhas de pesquisas, assim como suas características fundamentais, pois o referido sítio apenas fornece como subsídio (útil à esta pesquisa) os sítios dos inúmeros cursos de pós-graduação no Brasil. Tal fato exigiria que fosse realizada uma busca por sítio de cada curso de pós-graduação de diversas áreas do conhecimento, pois são várias as áreas que pesquisam sobre RI. Outra dificuldade em se adotar essa estratégia, é o fato de que nem todos os cursos consultados apresentam um sítio com ergonomia de fácil navegação. Tais dificuldades exigiriam um tempo maior de pesquisa.

Ao examinar o sítio do CNPq foi identificado uma base de dados que serviu de fonte para a presente investigação -o Diretório dos Grupos de Pesquisa no Brasil. O Diretório dos Grupos de Pesquisa no Brasil, projeto desenvolvido no CNPq desde 1992, constitui-se em bases de dados que contêm informações sobre os grupos de pesquisa em atividade no país.

O Diretório realiza censos bi-anual e mantém uma base de dados corrente, que pode ser atualizada continuamente. As informações contidas nessas bases dizem respeito aos recursos humanos constituintes dos grupos (pesquisadores, estudantes e técnicos), às linhas de pesquisa em andamento, às especialidades do conhecimento, aos setores de aplicação envolvidos, à produção científica e tecnológica e aos padrões de interação com o setor produtivo.

Além disso, cada grupo é situado no espaço e no tempo. Os grupos de pesquisa inventariados estão localizados em universidades, instituições isoladas de ensino

superior, institutos de pesquisa científica, institutos tecnológicos e laboratórios de pesquisa e desenvolvimento de empresas estatais ou ex-estatais. Os levantamentos não incluem os grupos localizados nas empresas do setor produtivo.

Cabe ressaltar que nem todos os grupos de pesquisa do Brasil estão vinculados ao CNPq, o que leva a consideração que as informações levantadas neste trabalho restringisse apenas aos grupos brasileiros que apresentam algum tipo de vínculo com o respectivo órgão do Ministério da Ciência e Tecnologia.

As informações contidas no Diretório do CNPq são fornecidas pelos líderes dos grupos cadastrados. Além das informações descritas acima, o Diretório disponibiliza ainda o currículo de cada membro dos grupos.

Como dispositivos de busca, o Diretório possibilita a busca por grupos, por líderes, por pesquisadores e por estudantes, além de possibilitar a filtragem por Estado da Federação, por instituição, e por área de conhecimento do grupo. A sentença de busca pode ser delimitada como "Todas as Palavras", "Qualquer Palavra", ou pela "Frase Exata".

A coleta dos dados foi realizada entre os dias 01 e 10 de setembro de 2006, adotando como estratégia de busca os dispositivos busca por grupo e frase exata, e a sentença de busca utilizada foi recuperação informação.

As variáveis procuradas para a análise e interpretação da situação da pesquisa em RI no Brasil foram: a identificação dos grupos de pesquisa, quando foram criados, quais as instituições os abrigam, a distribuição geográfica dos grupos em território brasileiro, e quais as áreas do conhecimento dos grupos.

A apresentação e análise dos resultados obtidos na pesquisa são apresentadas na seguinte seção.

4. Resultados e Análise das Informações

Como resultado da busca no Diretório dos Grupos de Pesquisa no Brasil identificou-se 37 grupos de pesquisa em áreas similares e específicas da Recuperação da Informação vinculadas ao CNPq.

Foi constatado que os grupos de pesquisa examinados predominantemente estão situados nas Instituições de Ensino Superior (IES), totalizando 86,48% dos 37 grupos identificados. Os 13,52% restante estão formados por grupos de pesquisa em empresas, laboratórios, organização nacional (Por exemplo: Comissão Nacional de Energia Nuclear - CNEN e Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA) e centro de ensino técnico.

Em relação ao ano da criação dos grupos de pesquisa, observa-se que está em ritmo ascendente, pois na década de 1970 foi formado 2,7% dos grupos pesquisados, 8,11% na década de 1980, 29,73% na década de 1990, e 59,46% foram formados do ano 2000 até o ano corrente. O Gráfico 1 mostra a evolução cronológica de criação dos grupos de pesquisa em recuperação da informação no Brasil.

Percentual de Grupos de Pesquisa do Brasil no Decorrer das Décadas

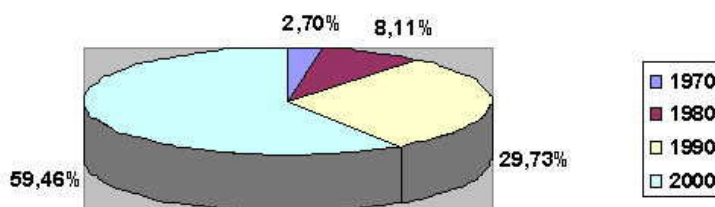


Gráfico 1 – Evolução Cronológica dos Grupos de Pesquisa no Brasil

Sobre a distribuição geográfica dos grupos de pesquisa encontrou-se que estão predominantemente situado na região Sudeste do Brasil (56,76%), Nordeste (16,22%), Sul (13,51%), Centro-Oeste (8,11), e Norte (5,4%). (**Região Sudeste:** São Paulo, Rio de Janeiro, Minas Gerais e Espírito Santo. **Região Nordeste:** Maranhão, Piauí, Rio Grande do Norte, Ceará, Pernambuco, Paraíba, Alagoas, Sergipe e Bahia. **Região Sul:** Paraná, Santa Catarina e Rio Grande do Sul. **Região Centro-Oeste:** Mato Grosso, Mato Grosso do Sul, Goiás e Distrito Federal. **Região Norte:** Amapá, Amazonas, Roraima, Pará, Rondônia e Acre.). Veja o Gráfico 2.

Quantidade Percentual de Grupos de Pesquisa por Região.

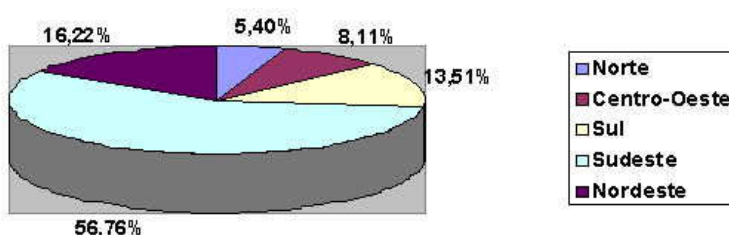


Gráfico 2 – Distribuição Geográfica dos Grupos de Pesquisa

As áreas de conhecimento que abrigam os respectivos grupos são em primeiro lugar a Ciência da Computação com 54,05% das ocorrências, seguida da área da Ciência da Informação (27,03%) e das Engenharias (8,11%). Outras áreas constatadas foram a Geociências (5,41% das ocorrências), a Física (2,7%) e a Comunicação (2,7%). O Gráfico 3 mostra os grupos de pesquisa por área de conhecimento.

Quantidade Percentual de Grupos por Área

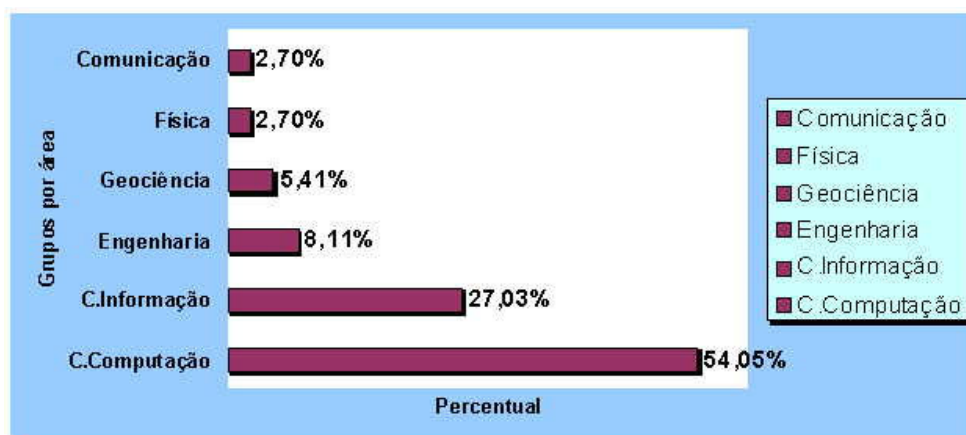


Gráfico 3 – Grupos de Pesquisa Distribuídos por Área de Conhecimento

Na Tabela 1 exposta abaixo, mostra a frequência de ocorrência das linhas de pesquisa achadas por área de conhecimento no Brasil. É importante ressaltar que algumas linhas de pesquisa expostas na Tabela 1 foram agregadas em linhas de temática semelhante (por exemplo: aprendizado de máquina foi agregada à inteligência artificial), ou então foram desconsideradas as linhas que não estão relacionadas com o tema recuperação da informação (por exemplo: jogos interativos multimídia).

Tabela 1 – Linhas de Pesquisa por Área de Conhecimento no Brasil.

LINHAS DE PESQUISA	FREQUÊNCIA DE LINHAS POR ÁREA DE CONHECIMENTO						TOTAL
	Ciência da Computação	Ciência da Informação	Engenharias	Geociências	Física	Comunicação	
Recuperação da Informação	7	5	2	0	0	0	14
Sistemas Especialistas	11	0	1	0	1	0	13
Banco de Dados	11	0	1	0	0	0	12
Comunicação e Disseminação da Informação	2	6	0	0	0	2	10
Inteligência Artificial	10	0	0	0	0	0	10
Indexação	4	3	1	0	0	0	8
Interoperabilidade	7	0	1	0	0	0	8
Sistemas de Informação Geográfica	4	0	1	3	0	0	8
Processamento e Análise de Imagens	4	0	2	0	0	0	6
Linguagem Documentária e Linguagem Natural	2	2		0	0	1	5
Sistemas Voltados à Saúde	2	0	2	0	0	0	4

Sistemas Multi-Agentes		0	0	0	0	0	4
Preservação e Conservação de Dados e Informação	0	3	0	0	0	0	3
Organização e Aquisição do Conhecimento	1	2	0	0	0	0	3
Sistemas de Informação Avançados	3	0	0	0	0	0	3
Mineração de Dados	3	0	0	0	0	0	3
Ontologias	3	0	0	0	0	0	3
Sistema de Apoio à Decisão	3	0	0	0	0	0	3
Métodos Gerenciais	3	0	0	0	0	0	3
DataWarehousing	3	0		0	0	0	3
Fontes de Informação	0	2	0	0	0	0	2
Sistemas Voltados à Educação	1	1	0	0	0	0	2
Biblioteca Digital	1	1	0	0	0	0	2
Tecnologias de Informação e Comunicação	1	1	0	0	0	0	2
Interfaces	1	1	0	0	0	0	2
Sistemas Multimídia	2	0	0	0	0	0	2
Arquitetura da Informação	0	1	0	0	0	0	1
Classificação e Extração de Dados	1	0	0	0	0	0	1
Websemântica	1	0	0	0	0	0	1
Caracterização e Armazenamento de Informação	1	0	0	0	0	0	1
Modelagem de Dados	1	0	0	0	0	0	1
TOTAL	97	28	11	3	1	3	143

Na Tabela 1 observa-se que a linha de pesquisa de recuperação da informação é a mais freqüente 9,79% das ocorrências, dos quais 50% são grupos da área de Ciência da Computação, 35,71% da área de Ciência da Informação e 14,28% das Engenharias. A linha de Sistemas Especialistas (9,09%), Banco de Dados (8,39%) e Inteligência Artificial (6,99%) são linhas estudadas quase exclusivamente pela área de Ciências da Computação e em grau muito menor pelas Engenharias.

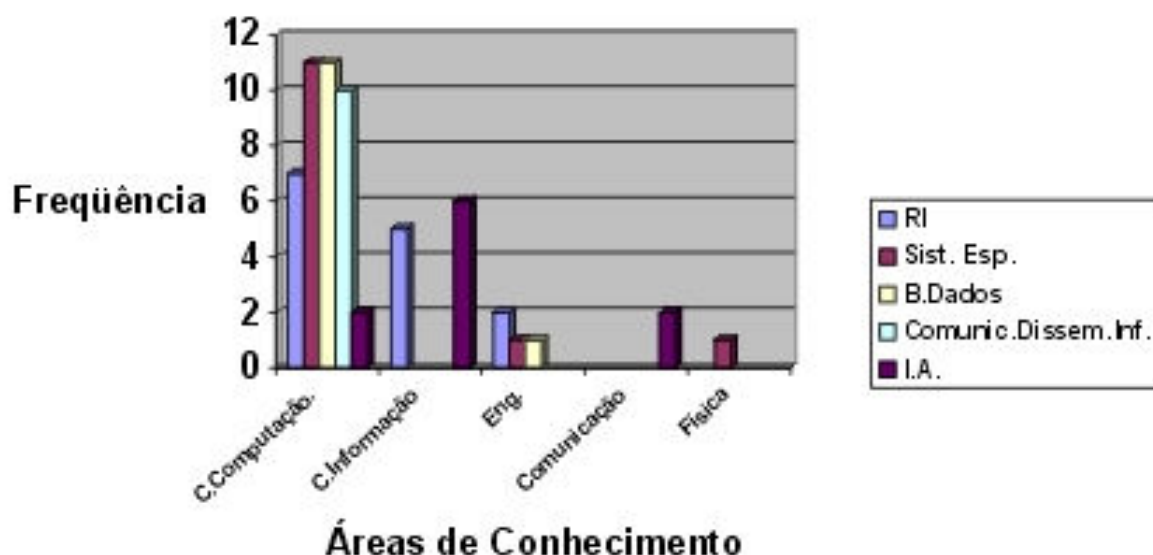
As linhas de pesquisa estudadas pelos grupos de pesquisa da área de Ciências da Informação, em ordem decrescente de ocorrência são: Comunicação e Disseminação da Informação (21,43%), Recuperação da Informação (17,86%), Indexação (10,72%), Preservação e Conservação de Dados e Informação (10,72%), Linguagem Documentária e Linguagem Natural (7,14%), Organização e Aquisição do Conhecimento (7,14%), Fontes de Informação (7,14%), Sistemas Voltados à Educação (3,57%), Biblioteca Digital

(3,57%), Tecnologias de Informação e Comunicação (3,57%), Interfaces (3,75%) e Arquitetura da Informação (3,57%).

Das linhas de pesquisa mencionadas a área de Ciência da Informação tem predomínio sobre as outras áreas nas linhas de Comunicação e Disseminação da Informação, Preservação e Conservação de Dados e Informação (linha esta não encontrada em nenhuma outra área de conhecimento), Organização e Aquisição do Conhecimento, Fontes de Informação e Arquitetura da informação.

As linhas de pesquisa com igual número de ocorrências na área de Ciência da Informação e a Ciência da Computação são: Linguagem Documentária e Linguagem Natural, Sistemas Voltados à Educação, Biblioteca Digital, Tecnologias de Informação e Comunicação e Interfaces.

Percebem-se que nos grupos brasileiros existe uma tendência em pesquisar a RI atrelada às tecnologias de informação e comunicação com necessidades de informação, pois as linhas que mais estão avançando são as de Recuperação da Informação, Sistemas Especialistas, Banco de Dados, Inteligência Artificial e Comunicação e Disseminação da Informação. Ver o Gráfico 4.



Explicação da Legenda:

RI = Recuperação da Informação; Sist. Esp. = Sistemas Especialistas; B. Dados = Banco de Dados; Comunic. Dissemin. Inf. = Comunicação e Disseminação da Informação; I.A. = Inteligência Artificial

Gráfico 4 - Ocorrências das Principais Linhas de Pesquisa por Área de Conhecimento no Brasil.

O ranking das linhas de pesquisa é liderado pela área do conhecimento que contempla maior número de grupos de estudo – a Ciência da Computação (CC) participa na maioria das linhas de pesquisa identificadas neste levantamento 67,83%, a segunda área do conhecimento em número de grupos de estudo, também é a segunda área em número de ocorrências das linhas de pesquisa – a Ciência da Informação (CI) 19,58%, terceiro está Engenharias 7,69%, empatadas na quarta posição Geociências e Comunicação com 2,10% e na última posição a Física com 0,70%. O Gráfico 5 mostra o ranking por área.

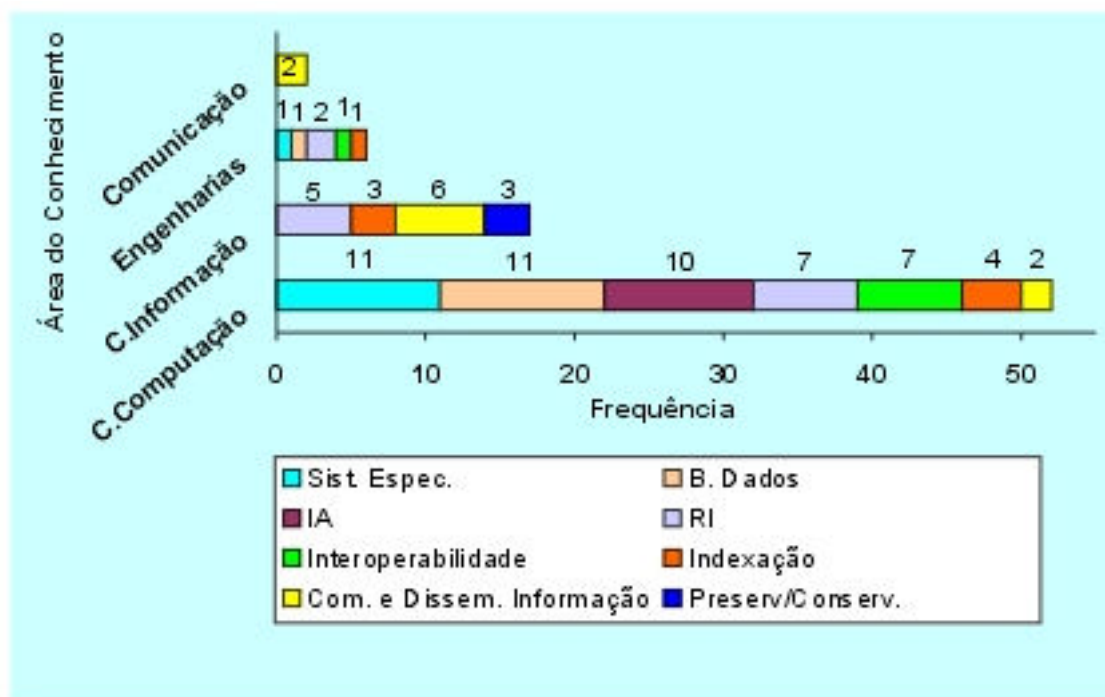


Gráfico 5 - Principais Linhas de Pesquisa por Área.

5. Considerações Finais

Com o exposto neste artigo é possível afirmar que a pesquisa brasileira, no que se refere à Recuperação da Informação (RI), volta suas atenções para as tecnologias de informação e comunicação atrelada às necessidades e demandas de informação, evidenciando que os sistemas de recuperação da informação atuais não são apenas meios de interação entre pessoas e informações, mas sim, uma meio de interação entre comportamentos de busca de informação e texto digital. Contudo é possível tecer algumas considerações referentes à situação da pesquisa em Recuperação da Informação no Brasil.

É evidente o aumento exponencial no número de grupos de pesquisa existentes no país entre as décadas de 1990 e 2000, já que nas décadas anteriores (1970 e 1980) é possível observar um crescimento tímido, e, que num país com as dimensões geográfica e populacional como as do Brasil, passa despercebido.

O mencionado crescimento exponencial do número de grupos de pesquisa no Brasil pode perfeitamente ser atribuído à disseminação da web, no país e no mundo na década de 1990, potencializando-se com muita rapidez na primeira década do século XXI. As Instituições de Ensino Superior são as principais fomentadoras de pesquisas de RI no território nacional.

Os grupos de pesquisa estão predominantemente concentrados na região mais desenvolvida economicamente do país, a região Sudeste, sobretudo nos estados de São Paulo, Minas Gerais e Rio de Janeiro. A Ciência da Computação e a Ciência da Informação são as áreas que mais desenvolvem pesquisas no tema RI no Brasil, e, conseqüentemente são as áreas com maior número de linhas de pesquisa, e com maior diversidade temática. Os Sistemas Especialistas, os Bancos de dados e a Inteligência

Artificial são áreas temáticas que mais despertam interesses entre os pesquisadores membros dos grupos investigados nesta pesquisa.

Chama a atenção que a preocupação com a preservação e conservação da informação se encontra exclusivamente na área da Ciência da Informação, o que mostra um certo descuido das outras áreas com este tema muito importante na atualidade, caracterizada pelas rápidas mudanças nas tecnologias e formatos existentes.

No decorrer deste artigo, sobretudo na seção dos resultados e análise dos resultados, foi possível observar que a pesquisa em recuperação da informação no Brasil está em ascensão e, embora esteja sendo desenvolvida predominantemente pelos grupos de Ciência da Computação e da Ciência da Informação, desperta interesse de pesquisadores dos mais variados domínios do conhecimento. A tendência da pesquisa brasileira aponta para as soluções inteligentes de recuperação da informação e banco de dados, confirmando a preocupação em fazer com que os sistemas de informação sejam sistemas que possibilitam a interação entre pessoas e as informações.

Referências

- Baeza-Yates, R. & Ribeiro-Neto, B. (1999). *Modern information retrieval*. New York, NY: Addison Wesley.
- Belkin, N. J. (2004). *Intelligent information retrieval: whose intelligence?*. Retrieved 18 April, 2006 from <http://mariner.rutgers.edu/tipster3/iirs.html>
- Borko, H. (1968). Information science: what is it?. *American Documentation*, 19(1), 3-5.
- Hayes, R.M. (1999). History review: The development of information science in the United States. Proceedings of the 1998 Conference on the History and Heritage of Science Information Systems, 223-236. Retrieved 15 December, 2006 from http://www.chemheritage.org/explore/ASIS_documents/ASIS98_Hayes.pdf
- Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq. Retrieved 01 September, 2006 from <http://www.cnpq.br>
- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES. Retrieved 02 September, 2006 from <http://www.capes.gov.br>
- Conselho Nacional de Desenvolvimento Científico e Tecnológico.(n.d.). *Diretório dos Grupos de Pesquisa no Brasil*. Retrieved 03 March, 2006 from <http://dgp.cnpq.br/buscaoperacional/>
- Ingwersen, P. (1992). *Information retrieval interaction*. London: Taylor Graham.
- Le Coadic, Y-F. (1996). *A ciência da informação*. Brasília, DF: Briquet de Lemos/Livros.
- Marchiori, P. Z. (2002). O campo de atividades, a ciência e o espaço da gestão da informação. In C.A.Castro (Ed.). *Ciência da informação e biblioteconomia: Múltiplos discursos* (pp. 51-71). São Luís, MA: EDUFMA.
- Marín Milanés, F. & Torres Veslásquez, A. (2006). Una mirada en torno al desarrollo de la Recuperación de información. *DataGramaZero: Revista*

- de Ciência da Informação*, 7 (4). Retrieved 10 September, 2006 from http://www.dgz.org.br/ago06/F_I_art.htm
- Saracevic, T. (1999). Information Science. *Journal Of The American Society For Information Science*, 50(12), 1051-1063.
- Robredo, J. (2003). *Da ciência da informação revisitada aos sistemas humanos de informação*. Brasília, DF: Thesaurus; SSRR Informações.
- Rowley, J. (2002). *A biblioteca eletrônica*. (A. A. B. de Lemos, Trad.). Brasília, DF: Briquet de Lemos/Livros.

Dados dos Autores

Rodrigo de Sales

Mestrando em Ciência da Informação (UFSC – Brasil)
Bacharel em Biblioteconomia (UFSC – Brasil)
Nascido em Florianópolis, Brasil; em 1980.
Bolsista de Desenvolvimento Tecnológico Industrial G | - CNPq

Angel Freddy Godoy Viera

Professor do Programa de Pós-Graduação em Ciência da Informação (UFSC – Brasil)
Professor do Departamento de Ciência da Informação (UFSC – Brasil)
Nascido em Assunção, Paraguay; em 1969.
Possui Doutorado em Engenharia de Produção pela Universidade Federal de Santa Catarina (2001) e Mestrado em Ciências da Computação pela Universidade Federal de Santa Catarina (1996). Atualmente é Professor Adjunto III da Universidade Federal de Santa Catarina. Tem experiência na área de Ciência da Informação, com ênfase em Recuperação Inteligente de Informação, atuando principalmente nos seguintes temas: recuperação da informação e automação de Bibliotecas.