

RECUPERAÇÃO DE INFORMAÇÃO NO BANCO DE DADOS DA EMBRAPA SOJA: ESTUDO DO AINFO

Thais Batista Zaninelli

Maria Elisabete Catarino

Resumo

Apresenta os resultados de uma pesquisa realizada com os utilizadores do sistema AINFO-Web que foi desenvolvido pela Embrapa para tornar disponível a informação existente na empresa. Para verificar se os usuários utilizam adequadamente os operadores booleanos e se estes efetivamente contribuem para suprir suas necessidades informacionais com qualidade, desenvolveu-se uma pesquisa descritiva. A coleta de dados junto aos pesquisadores, estagiários e técnicos da instituição, teve como instrumento um questionário que abordou a identificação dos sujeitos e pesquisa no AINFO-Web. Os resultados evidenciaram que a maioria dos usuários considera importante e utilizam os operadores booleanos ao formularem suas estratégias de busca, porém, alguns demonstraram possuir pouco conhecimento sobre seus mecanismos.

Palavras chave

Embrapa – AinfoWeb; Recuperação da informação; Pesquisa em base de dados.

1 INTRODUÇÃO

O avanço da ciência, aliado ao expressivo progresso das telecomunicações e das tecnologias de informação, impulsiona a produção de documentos pelas instituições, especialmente aquelas que têm

objetivos acadêmicos, voltados para pesquisa e conseqüentemente, produtoras do conhecimento.

Hoje, o meio eletrônico é uma forma rápida de armazenar, transmitir, compartilhar e recuperar documentos e informa-

ções. Para as bibliotecas e centros de documentação existem disponíveis várias tecnologias de informação que garantem a armazenagem, organização e recuperação de seus documentos.

Um recurso muito utilizado e eficaz para a disponibilização e troca de informações entre organizações públicas e privadas, é o banco de dados e, conseqüentemente, as suas bases de dados que ampliam significativamente a qualidade das buscas informacionais, visto que essas bases apresentam diversificados pontos de acesso à informação.

Uma base de dados permite consulta a milhões de itens de informações armazenados nas memórias de seus computadores, sendo que alguns deles estão conectados à Internet, como por exemplo, os catálogos em linha de acesso público, os *On-line Public Access Catalog (OPACS)*, que têm como uma de suas formas de divulgação a Internet.

Para o sucesso de uma busca de informação em bases de dados, o primeiro passo é escolher uma base especializada no assunto desejado. Em seguida, formular uma estratégia de busca que esteja em conformidade com os operadores booleanos e outros recursos de busca que ela oferece. Considera-se que uma das formas mais eficazes é o uso de operadores booleanos, que relacionam as informações

contidas em base de dados, de forma a possibilitar que o resultado da busca seja o mais pertinente e relevante possível para o usuário.

Operadores booleanos são recursos do *software* de recuperação de informação que podem ser usados para compor um argumento de pesquisa (ou estratégia de busca) mais completo a partir de termos individuais.

Com a quantidade excessiva de informações nas bases de dados, o uso de operadores booleanos é fundamental no momento da pesquisa, uma vez que são eles que definem o tipo de relacionamento entre os termos a serem pesquisados. Como a forma de expressão dos operadores booleanos varia de uma base para outra, o usuário pode confundir determinados símbolos no momento de sua pesquisa, que podem ser diferentes dependendo da base de dados utilizada.

A Embrapa Soja possui um Sistema de Informação nomeado de AINFOWeb, que disponibiliza várias bases de dados, dentre elas a base de dados documental da área agrícola, que oferece registros de informações sobre agricultura e ciências afins, com enfoque maior nas culturas de soja, trigo e girassol, objeto de estudo deste trabalho. Essa base de dados disponibiliza para os usuários vários recursos de busca para facilitar a pesquisa.

São os próprios usuários que formam suas estratégias, que escolhem o caminho que querem percorrer para chegar a um objetivo, no caso, um resultado de pesquisa que traga a resposta satisfatória para suas necessidades de informação.

Esta pesquisa pretendeu identificar as diferentes formas de realizar uma busca com qualidade e dentre elas, a que melhor recupera as informações desejadas. Um de seus principais objetivos foi verificar se os usuários da AINFOWeb utilizam os operadores booleanos para realizarem suas pesquisas e, caso afirmativo, verificou-se se estes recursos têm aumentado a qualidade do resultado final da busca.

Espera-se que os resultados obtidos nesta pesquisa venham contribuir para a comunidade interna da empresa no que se diz respeito à melhor qualidade na pesquisa no menor tempo possível, pois, se os usuários (no caso estagiários, pesquisadores e técnicos da Embrapa Soja) souberem de fato elaborar uma estratégia de busca utilizando-se dos operadores booleanos oferecidos pela base de dados, obterão qualidade na informação recuperada, fator fundamental para os pesquisadores, pois são estas informações que garantirão a segurança e veracidade de seus estudos. Conseqüentemente, os pesquisadores terão mais tempo para se dedicarem às suas pesquisas, trazendo assim benefícios para a própria Empresa.

É importante esclarecer que esta pesquisa, desenvolvida como trabalho de conclusão de curso de Biblioteconomia da Universidade Estadual de Londrina, trabalhou uma única base de dados e portanto, para se chegar a conclusões mais gerais há necessidade de estudar outras bases e *sites* na Internet para realizar comparações entre diversos *softwares* existentes e disponíveis no mercado.

2 O CONTEXTO

A Embrapa Soja atua como executora de pesquisa, como parceira em diversos empreendimentos de Pesquisa e Desenvolvimento (P&D) e como ponto de referência para a cultura de soja no Brasil, também o sendo para diversos outros países do Hemisfério Sul, marcadamente para os vizinhos do continente sul-americano. Também atua nas áreas de difusão de tecnologia e de concentração, processamento e repasse de informação sobre a cultura da soja, além de interfaces com o restante da cadeia produtiva, bem como na área de soja na alimentação. (EMBRAPA, 2000, p.21).

Visualiza-se claramente a tendência da transformação da soja em alimento funcional,¹ por suas propriedades terapêuticas

¹ Todo aquele alimento ou ingrediente que, além das funções nutricionais básicas, quando consumido como parte da dieta usual, produz efeitos metabólicos e/ou fisiológicos e/ou efeitos benéficos à

e preventivas para diversos distúrbios e enfermidades que afetam os seres humanos. Deverão ser viabilizadas soluções tecnológicas para o uso direto da soja na alimentação humana e para produção de soja orgânica. Em decorrência, também deverão ser conduzidos estudos para viabilizar soluções tecnológicas, a fim de melhorar a soja utilizada para a alimentação animal.

A empresa conta com uma rede de parcerias formada por instituições de pesquisa, de ensino, empresas privadas do setor produtivo e fundações de apoio à pesquisa. Essas parcerias ampliam e agilizam a capacidade de desenvolvimento, validação e transferência das tecnologias geradas, além estabelecer um canal permanente de prospecção de demandas junto à sociedade.

A Embrapa Soja torna disponível tecnologia, serviços e produtos capazes de viabilizar soluções para os agronegócios da soja e do girassol. Isso tem sido feito, principalmente, através dos inúmeros produtos e tecnologias gerados e desenvolvidos pela Unidade, artigos em revistas técnico-científicas nacionais e internacionais, palestras, dias de campo,² exposições a-

gropecuárias e divulgação de informação tecnológica por meio dos diferentes veículos de comunicação.

A Base de Dados da Pesquisa Agropecuária (BDPA), que tem por objetivo contribuir para o uso e a disseminação da informação técnico-científica produzida e armazenada pelas bibliotecas das unidades de pesquisa da Embrapa, é um sistema de informação desenvolvido pela Embrapa Informática Agropecuária e permite o gerenciamento, de forma integrada, de bases de dados bibliográficos e automação de processos bibliográficos, proporcionando rapidez e flexibilidade na recuperação de informações. Possui linguagem própria e permite a combinação entre operadores booleanos e outros recursos.

Para a recuperação de informação em bases de dados, o uso dos recursos como ferramentas de busca e operadores booleanos, são instrumentos indispensáveis no momento da pesquisa, tanto na pesquisa livre, como na estruturada, seja ela realizada em bases de dados ou na Internet. Para o usuário ter acesso às informações oferecidas pelas bases de dados, é necessário que ele faça estratégias de busca utilizando diferentes ferramentas e operadores, usados para relacionar termos ou palavras em uma expressão de pesquisa.

saúde, devendo ser seguro para consumo sem supervisão médica. (BRASIL, 1999).

² Demonstração de resultados de pesquisa e transferência de tecnologia, conhecimentos, inovações e imagens, através de demonstrações práticas, por meio de visitas aos campos experimentais da Em-

presa, áreas demonstrativas ou dia de campo na TV.

Os operadores booleanos *AND*, *NOT* e *OR* são normalmente utilizados por inúmeras bases de dados podendo ser substituídos por sinais, variando de uma base para outra. No AINFOWeb os operadores booleanos são representados por (e), (não) e (ou).

Nesta pesquisa foram estudadas as formas de recuperação de informação usadas pelos estagiários, pesquisadores e técnicos da Embrapa Soja. Junto com os recursos de busca disponíveis por esta base, é necessário destacar que os materiais nela contidos são: livros, periódicos, teses, *folders*, anais de eventos, folhetos, entre outros documentos que abordam especificamente os temas: soja, trigo e girassol.

3 AS TECNOLOGIAS DE INFORMAÇÃO

A sociedade industrial está se transformando em sociedade da informação por meio do avanço das telecomunicações e das novas tecnologias de informação. De repente, adaptamo-nos e passamos a viver nesta nova era em que a informação cresce em quantidade e numa velocidade que há bem pouco tempo não imaginávamos, modificando fundamentais valores econômicos e sociais. A velocidade com que esse avanço tecnológico acontece faz com que organizações, produtos, serviços e a sociedade se adaptem a essas mudanças

rapidamente, para que não se desqualifiquem.

Oliveira (apud ANTUNES; OLIVEIRA, 1998, p. 82), ressalta que “o ritmo acelerado de mudanças está impactando até os segmentos empresariais mais conservadores ou que nunca necessitaram de suporte tecnológico”.

Neste contexto, as tecnologias vêm transformando cenários e exercendo influências significativas no funcionamento de bibliotecas, centros de documentação e serviços de informação. A inserção das tecnologias de informação nas unidades de informação, resultou na eficiência de serviços em geral, beneficiando aqueles que delas se utilizam.

Atualmente, os *softwares* desenvolvidos especialmente para unidades de informação geralmente têm como objetivo gerar bases de dados que visam melhorar as atividades de processamento de informações, aquisição de materiais, catalogação, controle de empréstimos, manutenção do acervo bibliográfico, controle de periódicos, dentre outras atividades desenvolvidas em uma biblioteca.

4 BASES DE DADOS EM UNIDADES DE INFORMAÇÃO

Em unidades de informação, as bases de dados constituem ferramentas fun-

damentais aos serviços oferecidos, principalmente no que se refere à organização, estruturação e disponibilização da informação, para um atendimento ágil às necessidades do cliente.

Profissionais e pesquisadores utilizam bases de dados como fontes de informação porque facilitam e agilizam pesquisas, buscas bibliográficas e a localização de documentos, uma vez que as bases de dados são freqüentemente atualizadas. Para eles, as questões referentes à busca e uso de informações se tornam fundamentais, em consequência da relevância da informação para suas atividades.

Para Le Codiac (1996 apud CRESPO; CAREGNATO, 2003, p.273)

A informação é o sangue da ciência. Sem informação, a ciência não pode se desenvolver e viver [...].A atividade de pesquisa constitui, com efeito, a aplicação do raciocínio ao corpo de conhecimentos acumulados ao longo do tempo e armazenados nas bibliotecas e centros de documentação.

De acordo com a área de interesse e atuação profissional do pesquisador, no entanto, a maneira de buscar e de utilizar a informação será diferenciada, porém será contínua e permeia todas as etapas do processo de construção do conhecimento, desde o planejamento do projeto até a apresentação dos resultados das pesquisas.

O uso dos computadores tornou cada dia mais fácil a realização de buscas de informação. O número de bases de dados cresce exponencialmente e o que os usuários precisam é de orientação de especialistas em informação para identificar as bases de dados apropriadas do que ajuda na recuperação propriamente. Muitos preferem realizar suas próprias buscas, outros delegam a realização das buscas a estagiários, por falta de tempo ou porque acham importante o incentivo à pesquisa.

As bases de dados correspondem hoje em dia ao meio mais rápido e eficiente de obter informações [...] nas mais variadas áreas do conhecimento. Este fato, somando à perspectiva dos lucros financeiros advindos da produção e comercialização de base de dados, contribui para o surgimento e a crescente expansão de uma verdadeira “indústria” voltada a essas fontes computadorizadas de informação. (PAIVA, 2002, p.22).

As bases de dados podem ser constituídas por referências bibliográficas ou por textos na íntegra. Rowley (2002) classifica em bases de dados de referências e bases de dados de fontes. As de referências encaminham o usuário a outras fontes, para que ele tenha acesso a informações adicionais, ou o texto na íntegra de um determinado documento.

As bases de dados de fontes disponibilizam dados originais e “constituem um tipo de documento eletrônico”, sem neces-

idade de buscar a fonte original. “Os dados encontram-se disponíveis tanto em formato legível por computador, quanto em formato impresso.” (ROWLEY, 2002, p. 110).

4.1 Estratégias de Busca

Tanto as bases de dados referenciais, quanto as textuais exigem planejamento cuidadoso da estratégia de busca, visando à recuperação de informação de acordo com as necessidades do usuário. Portanto, para a realização de buscas nessas bases de dados exige-se treinamento e conhecimento de seus mecanismos de uso.

Rowley (1994, p. 82) argumenta que

a forma mais rudimentar de se fazer uma busca numa base de dados é percorrê-la, registro por registro, à procura do elemento de dado aproximado.

Normalmente as bibliotecas que disponibilizam bases de dados a seus usuários têm um bibliotecário habilitado para ajudá-los a selecionar a melhor base, de acordo com suas necessidades específicas, e cabe a esse profissional ensinar o uso dos diferentes mecanismos de busca, pois os comandos oferecidos pelas bases diferem de uma para outra.

Geralmente, as bases de dados são dotadas de um sistema de recuperação de informação com interfaces e comandos

que são baseados em menus, conforme o pacote de *software* utilizado, tornando os sistemas mais acessíveis.

Grogan (1995, p. 131) descreve que

as bases de dados constituem-se em sistemas de buscas baseadas em comandos ou menus, ou em uma mistura de ambas as partes ou palavras-chave, empregando se necessário, a lógica booleana [...].

Em se tratando de recuperação de informação, a estratégia de busca pode ser definida, segundo Lopes (2002, p.61), como

[...] uma técnica ou conjunto de regras para tornar possível o encontro entre uma pergunta formulada e a informação armazenada em uma base de dados. Isto significa que, a partir de um arquivo, um conjunto de itens que constituem a resposta de uma determinada pergunta será selecionado.

Sabe-se que alguns usuários são mais cuidadosos do que outros na formulação das estratégias de busca em suas pesquisas; porém, o objetivo de todos, geralmente, é recuperar registros relevantes e evitar a recuperação de registros excessivos, irrelevantes e insignificantes.

Complementando, concordamos com Bertholino, quando diz que:

A formulação da estratégia de busca é fundamental para refinar a busca e [...] obter resultados relevantes aos interesses do cliente. É importante, na definição das palavras chave, informar os termos sinônimos, correlacionados e equiva-

lentes, bem como suas respectivas definições no idioma adotado pela base [...] (1999, p.151).

A estratégia de busca é uma técnica usada para tornar possível o cruzamento entre uma pergunta e uma informação armazenada em uma base de dados que a responda.

Quanto ao nível, as ferramentas de busca oferecem mais de um, geralmente conhecidos como busca simples e avançada, dependendo do conhecimento e experiência do usuário. A busca avançada fornece maiores opções e recursos para elaboração das estratégias de busca.

Os bancos de dados, criados para indicar coleções organizadas de dados armazenados em computadores digitais oferecem múltiplas bases de dados para consulta, algumas com milhões de itens de informação, oferecendo também, recursos de busca que permitem a consulta textual por meio de palavras contidas nos títulos, resumos e no próprio texto. Bourne afirma que:

[...] a disponibilidade e uso de todas essas facilidades têm contribuído para o desenvolvimento de linguagens de indexação associadas com essas bases de dados. Observa, entretanto, que a quantidade de erros de ortografia, formas variantes de palavras e mesmo erros ortográficos dos termos de indexação afetam consideravelmente os resultados das buscas (apud LOPES, 2002, p.62).

Harter (1984 apud LOPES, 2002, p. 62), analisou os diferentes estilos de busca em linha, com ênfase no comportamento e atitude dos intermediários (os bibliotecários, por exemplo). O estudo revelou que existem grandes diferenças entre os intermediários que operacionalizam as buscas. Os intermediários de instituições acadêmicas diferem daqueles de outros tipos de biblioteca, principalmente das especializadas, no que se refere ao aprofundamento das estratégias de busca.

Pode-se relacionar esse estudo aos usuários da Embrapa Soja que possuem grau profissional mais elevado, e necessitam aprofundar seus conhecimentos, muitas vezes para solucionar problemas de pesquisa em andamento.

Bates (apud LOPES, 2002, p.63) conceituou a estratégia de busca como o "estudo da teoria, princípios e prática de planejar e executar táticas e estratégias de busca". Adaptando as definições ao ambiente de recuperação da informação, deu ênfase ao conceito de comportamento de busca, indagando: "o que as pessoas fazem, e como pode ser determinado o que elas pensam, quando estão executando uma busca de informação?"

[...] cabe uma análise semântica dos principais termos que compõem o objeto de estudo, lembrando que o planejamento e a execução de uma estratégia de busca é uma "arte de escolher onde, quan-

do e com que investigar cuidadosamente" a fonte de informação para alcançar os objetivos específicos do solicitante. A escolha de onde buscar, identificando qual a base de dados é a mais bem qualificada [...], depende de uma série de fatores relacionados com a abrangência de assunto e período da base, dos tipos de documentos indexados, da linguagem de indexação, dos campos de busca disponibilizados em determinado banco de dados e outros. (apud LOPES, 2002, p.63).

Ojala (1995 apud LOPES, 2002, p.65) analisou os diferentes recursos de busca oferecidos pelos sistemas em linha que podem alterar a decisão do usuário na escolha do melhor sistema para a realização de uma busca de informação em uma determinada base de dados. "As variações no uso de operadores booleanos, o uso de operadores de proximidade, a quantidade de frases da estratégia de busca que podem ser operacionalizadas", são alguns fatores mencionados por Ojala.

Na estratégia de busca são utilizadas palavras-chave, descritores, operadores booleanos e outros recursos que, organizados, formam uma expressão que recupera informações relevantes sobre uma temática, em sistemas automatizados, como as bases de dados.

A lógica booleana utilizada na maioria dos sistemas liga os termos e linguagens controladas e naturais. Os operadores lógicos e booleanos são *E*, *OU*, *NAO*. Com esses operadores, as estratégias de busca podem ser refinadas, utilizando-se mais do que um operador para se

chegar a um resultado satisfatório (BERTHOLINO, 1999, p.150).

Há sistemas de busca que somente reconhecem os operadores **AND**, **NOT** e **OR**, em inglês e com letras maiúsculas. Outros sistemas somente reconhecem os operadores em sua forma gráfica, os sinais gráficos. E há também sistemas que reconhecem os operadores em português. Para ter certeza da forma correta de realizar a pesquisa, o usuário deve consultar a página de ajuda do sistema de busca que estiver utilizando.

Oldroyd e Citroen (apud LOPES, 2002, p.70) identificam três etapas para decisão de estratégia de busca: "decisão sobre melhor base de dados, para um determinado tema; decisão referente à seleção dos termos de busca e sua adequação para a base a ser consultada; decisão sobre a formulação lógica da estratégia".

A utilização correta dos operadores e dos demais recursos interfere na qualidade das respostas obtidas quanto à pertinência e exaustividade, como uma busca por nome próprio que pode ser otimizada levando-se em conta a diferenciação que a ferramenta faz ou não para esta busca.

Como forma de facilitar a pesquisa, principalmente para o usuário leigo, algumas ferramentas de busca fornecem interface gráfica mais amigável, embutindo em opções de escolha os operadores que utili-

zam para busca. Assim, não é exigido do usuário o conhecimento prévio para elaborar a sintaxe da busca.

Além disso, permitem que o próprio usuário determine a exibição dos resultados dentre as opções existentes, tais como: quantas e quais informações devem ser exibidas para cada documento recuperado; quantos documentos devem ser exibidos; como devem estar ordenados os resultados dentre as opções: cronológica, alfabética ou relevância.

A formulação da estratégia de busca, bem como a escolha do banco/base de dados a ser utilizada para responder a uma informação específica, exige, no caso dos técnicos, estagiários e pesquisadores da Embrapa Soja, conhecimentos especializados tais como: estrutura da base, conhecimento da linguagem de busca do banco de dados, idiomas, lógica booleana, instrumentos de ajuda para identificação da terminologia a ser usada e outros recursos disponíveis para a obtenção dos resultados desejados pelo solicitante, requisitos que esta pesquisa também pretende identificar.

4.2 AINFOWeb

O AINFOWeb Versão 1.2, é um *software* de recuperação de informação utilizado para consulta *on-line* às bases de dados: Acervo Documental; Coleção de Periódicos; Cadastro de Recursos Huma-

nos e Cadastro de Instituições que são gerenciadas pelo AINFO 2001. Permite a combinação de operadores booleanos e recursos como truncagem, mascaramento, busca por proximidade e frase, dentre outras possibilidades na formulação de estratégias de busca.

O AINFOWeb foi desenvolvido pela Embrapa Informática Agropecuária, em atendimento à demanda das Bibliotecas da Embrapa, com o objetivo de viabilizar o acesso, via Web, às informações técnico-científicas contidas nas referidas bases de dados. Atualmente está em uso nas unidades da Embrapa, bem como em Empresas Estaduais de Pesquisa Agropecuária, Institutos de Pesquisa, Universidades, ONGs etc.

O **AINFO** é uma tecnologia de informação que permite o gerenciamento de bases de dados documentais e processos bibliográficos, possibilitando o armazenamento, atualização recuperação e disseminação de informações de forma simples e rápida (ANTUNES; LENK, 1997, p.7, grifo nosso).

As bases de dados do AINFOWeb podem ser assim apresentadas:

- **Acervo Documental:** permite o cadastramento e uso dos materiais bibliográficos das unidades da Embrapa: livros, teses, folhetos, *proceedings*, materiais, relatórios técnicos, dentre outros. A junção e compatibilização dos acervos das Unidades geram o Acervo Documental Embrapa.

- **Coleção de Periódicos:** Controla o recebimento de novos fascículos e novos títulos de periódicos, facilitando o controle das assinaturas. A junção e compatibilização das coleções de periódicos das Unidades da Embrapa geram o Catálogo Coletivo de Periódicos da Embrapa (CCPE), bem como a atualização do Catálogo Coletivo Nacional (CCN).

- **Recursos Humanos:** O principal objetivo desta base é, além de operar como cadastro de usuários da biblioteca, prover um banco de currículos, facilitando a prestação de vários serviços como o de orientação de produtores rurais.

- **Instituições:** Possibilita o cadastramento das Unidades da Embrapa, associações, cooperativas, agências de financiamentos, organizações governamentais, não-governamentais, universidades, institutos de pesquisas, empresas estaduais de pesquisa agropecuária, escritórios de extensão rural, rádios, jornais, entre outros veículos de comunicação (ANTUNES; LENK, 1997).

Os relacionamentos entre essas bases facilitam o processo do controle de aquisição de materiais, empréstimos, comunicação e disseminação da informação. Inicialmente foi elaborado com ferramentas disponíveis no Ambiente de *Software* NTIA, tecnologia gerada pela própria CNPTIA.

Atualmente, é utilizado o sistema gerenciador de banco de dados de domínio público chamado Firebird, que é derivado do Interbase, da Bordland. A junção e compatibilização dos arquivos das Unidades, permitem a geração das Bases de Dados da Pesquisa Agropecuária (BDPA), que se encontram disponíveis, via Internet no *site* <http://www.bdpa.cnptia.embrapa.br>.

A recuperação de informação no A-INFOWeb é realizada por meio da combinação lógica de palavras, números ou datas, existentes em qualquer um dos campos da base, utilizando-se de operadores booleanos e outros recursos como truncagem, mascaramento, proximidade, aspas, busca por frase, disponibilizado pelo módulo Recinfo.

O RECINFO é o módulo do Ambiente de *Software* NTIA que faz indexação e recuperação de informação. Possui linguagem própria de sintaxe simples e flexível e permite e combinação entre comandos, operadores booleanos e outros recursos como truncagem e adjacência, oferecendo dessa forma inúmeras possibilidades na formulação das estratégias de busca (ANTUNES; LENK, 1997, p.184).

5 METODOLOGIA

Esta pesquisa foi desenvolvida na Embrapa Soja (Londrina/PR), centro de referência para informações sobre soja, trigo e girassol.

Para cumprir os objetivos propostos

utilizou-se da pesquisa descritiva tendo como instrumento de coleta de dados, o questionário, que constituiu-se de duas partes: identificação do sujeito e pesquisa no AINFOWeb, onde se reproduziu exatamente a tela de recuperação do sistema e, em seguida, questões relacionadas aos operadores booleanos e aos registros recuperados, questões essas que buscaram atingir os objetivos específicos da pesquisa.

A população alvo deste estudo foi 115 sujeitos. Deste total, 73 pesquisadores,

sendo 15 mestres, 58 doutores (9 com pós doutorado) 30 estagiários de nível médio, graduação, bolsistas de mestrado e de doutorado, provenientes de diferentes cursos das faculdades e universidades de Londrina e 12 técnicos.

6 RESULTADOS

Identificou-se que grande parte dos usuários do AINFOWeb são estagiários (53,3%). Este dado se justifica pois os estagiários são os que mais utilizam do AINFOWeb.

Tabela 1 – Identificação dos sujeitos da pesquisa

Sujeitos	Nº	%
Pesquisadores	04	26,6
Estagiários	08	53,3
Técnicos	03	20
Total	15	100

Com relação aos pesquisadores, obteve-se 26,6% de respostas. Chegou-se a este resultado, pelo fato dos pesquisadores justificarem, nos questionários devolvidos, que normalmente não fazem suas próprias pesquisas, e sim solicitam aos seus estagiários ou aos bibliotecários da instituição.

O restante das respostas – 20% dos sujeitos – é de técnicos agrícolas, que corresponde à minoria, comparando-se ao número de pesquisadores e estagiários,

que é alto, e também pelo fato de que os técnicos agrícolas não trabalham diretamente com pesquisas.

O baixo índice de realização de buscas em bases de dados muitas vezes é atribuído ao desconhecimento de seus mecanismos de uso, sendo, portanto, o bibliotecário o profissional habilitado a prestar este tipo de serviço, de acordo com as necessidades específicas de cada usuário.

A fim de verificar a frequência de utilização do AINFOWeb, foram dadas aos sujeitos as seguintes alternativas: de 1 a 5 vezes, de 5 a 10 vezes e acima de 10 vezes por semana. De acordo com a Tabela 2, obteve-se as seguintes respostas: 100% dos pesquisadores utilizam o AINFOWeb de 1 a 5 vezes por semana. Dentre os es-

tagiários 87,5% responderam que utilizam a base de 1 a 5 vezes por semana e 12,5% assinalaram acima de 10 vezes. 100% dos técnicos assinalaram de 1 a 5 vezes por semana. Portanto, dentre as opções, verifica-se uma baixa frequência de uso do AINFOWeb em todas as categorias de usuários.

Tabela 2 – Frequência de utilização do AINFOWeb

Nº de vezes por semana	Pesquisadores		Estagiários		Técnicos	
	Nº	%	Nº	%	Nº	%
1 - 5	4	100	7	87,5	3	100
5-10	0	0	0	0	0	0
Acima de 10	0	0	1	12,5	0	0
Total	4	100	8	100	3	100

Foi solicitado ao sujeito, no momento de sua pesquisa, que escrevesse exatamente como digitou a expressão de busca, tendo como opção três oportunidades de pesquisa – uma vez que o instrumento de coleta de dados reproduziu a mesma tela do AINFOWeb – e que indicassem a quantidade de registros recuperados e destes, os relevantes para sua pesquisa. Esta questão teve como objetivo verificar se os usuários do AINFOWeb utilizam, ou não, os operadores booleanos em suas pesquisas.

Para se chegar a um resultado mais esclarecedor, foi analisado separadamente cada questionário, classificando os res-

pondentes por letras e estabelecendo-se uma tabela para cada sujeito.

Nas tabelas 3 e 4, a seguir, consta o resumo desta análise, mostrando, por categoria (pesquisadores, estagiários e técnicos), se foram ou não utilizados os operadores booleanos, a quantidade de registros recuperados e destes, quantos foram relevantes para suas pesquisas, mostrando também a porcentagem de aproveitamento do resultado da busca.

Constatou-se que a maioria das pesquisas realizadas tiveram o uso de operadores booleanos, porém, o índice de registros relevantes foi muito baixo. Com o exposto, identificou-se que os operadores

booleanos não estão sendo utilizados corretamente pelos usuários do AINFOWEb. Isso pode ser atribuído ao fato de não re-

ceberem treinamento adequado para a sua utilização.

Tabela 3 – Índice de relevância das pesquisas realizadas divididas por categoria de sujeitos com o **USO** de Operadores Booleanos.

Sujeitos	Nº pesquisa realizada	Nº pesquisa com uso de operadores booleanos		Registros recuperados	Registros Relevantes	% de aproveitamento
Pesquisador	12	8	66,7%	364	81	22%
Estagiário	24	17	70,8%	1.106	121	10,9%
Técnico	9	7	77,7%	1.158	190	16,4%

Observa-se na tabela 3 que na categoria pesquisador, foram realizadas 12 pesquisas sendo oito usando operadores booleanos (66,7%), resultando em 364 registros recuperados, sendo 81 relevantes para a pesquisa, o que equivale a 22% de aproveitamento. O operador mais utilizado pelos pesquisadores foi o “E” que tem como função juntar as palavras.

Na categoria estagiário foram realizadas 24 pesquisas. Destas, 17 (70,8%) com o uso de operadores booleanos, recuperando 1106 registros, sendo 121 relevantes, totalizando 10,9% de aproveitamento. Os operadores mais utilizados pelos estagiários foram “E” e a ferramenta de busca aspas (“ ”), que tem como função buscar palavras ou frases exatamente como foi digitado. Os técnicos realizaram nove pesquisas, sete (77,7%) utilizando-se de operadores; 1158 foram os registros recuperados nesta categoria, sendo 190 os relevan-

tes, o que equivale a 16,4% de aproveitamento para a sua pesquisa, sendo que o operador mais utilizado também foi o “E”.

Sendo assim, considera-se que o operador booleano “E”, é o mais utilizado pelos usuários do AINFOWEb. Em outras bases de dados o operador “E” pode ser substituído por “AND” ou sinais como: “+”, “&”.

Constatou-se também que outra ferramenta de busca mais utilizada pelos sujeitos foi a aspas (“ ”), pois esta ferramenta usada de forma correta busca expressões exatas; então, quando um usuário necessita de um autor, nome de livro, artigo, data, o operador facilita e limita bastante o número de resultados, recuperando documentos de relevância para a pesquisa. O recurso das aspas (“ ”) é também utilizado em diversas bases de dados.

Na Tabela 4 estão os resultados das pesquisas em que não foram utilizados os operadores booleanos.

Tabela 4 – Índice de relevância das pesquisas realizadas divididas por categoria de sujeitos SEM USO de Operadores Booleanos

Sujeitos	Nº pesquisa realizada	Nº pesquisa sem uso de operadores booleanos		Registros recuperados	Registros Relevantes	% de aproveitamento
Pesquisador	12	4	22,3%	13	9	69,2%
Estagiário	24	7	29,2%	400	24	6%
Técnico	9	2	22,3%	27	4	14,8%

Das 12 pesquisas realizadas por pesquisadores, quatro (22,3%) foram sem a utilização de operadores booleanos. Treze registros foram recuperados, sendo nove relevantes para suas pesquisas, totalizando 69,2% de aproveitamento do resultado.

Os estagiários realizaram um total de 24 pesquisas, sendo 7 (29,2%) sem operadores booleanos, recuperando 400 registros, e utilizando-se de 24, pois foram os relevantes para suas pesquisas. Sendo assim, obteve-se aproveitamento de 6%.

As pesquisas realizadas pelos técnicos foram nove, sendo duas (22,3%) com operadores booleanos. Recuperaram 27 registros, quatro relevantes, ou seja, o total de aproveitamento da pesquisa foi de 14,8%.

No caso de pesquisa sem o uso de operadores booleanos, a categoria pesquisador se diferenciou bastante no índice de relevância (69,2%), enquanto as demais

categorias – estagiários e técnicos – tiveram registros relevantes de 6% e 14,8% respectivamente.

Podemos deduzir que os pesquisadores tiveram um número maior de registros relevantes pelo fato de todos eles utilizarem, pelo menos em uma das pesquisas realizadas, um operador booleano.

Analisando as expressões de busca formuladas pelos pesquisadores nos questionários, nota-se que apesar do baixo índice de utilização direta do AINFOWEb, são eles os que mais utilizaram os operadores, embora não de forma correta.

Um dos objetivos dessa pesquisa foi identificar se os sujeitos usam os operadores booleanos de forma adequada, analisando suas pesquisas e registros recuperados. Percebe-se que apesar do alto índice de uso dos operadores, existe uma quantidade baixa de registros relevantes,

ou seja, de aproveitamento dos resultados das pesquisas.

Os usuários do AINFOWEb, não tiveram treinamento formal para utilizar os operadores e as ferramentas de pesquisa disponíveis na base. Quando um usuário necessita utilizar o AINFOWEb, recebe uma orientação superficial; conseqüentemente, o resultado de sua pesquisa não

terá um índice de relevância alto. Apesar do manual de instrução de uso do AINFOWEb estar disponível *on-line*, conclui-se que os usuários não o consultam antes de realizarem suas pesquisas.

O gráfico 1, foi elaborado para obter uma melhor visualização do grau de dificuldade na utilização dos operadores booleanos oferecidos pelo AINFOWEb.

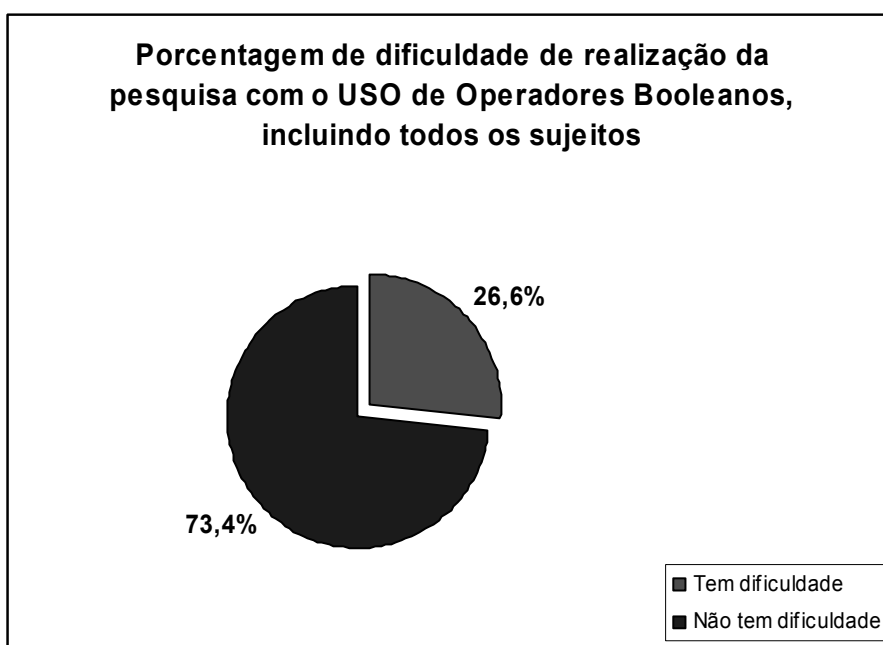


Gráfico 1 – Dificuldade na realização da pesquisa com USO de Operadores Booleanos

Analisando todos os sujeitos, obteve-se um resultado satisfatório, pois, como mostra o referido gráfico, 73,4% dos questionários respondidos assinalaram que utilizam os operadores booleanos sem dificuldades, e apenas 26,6% têm dificuldade de utilizar os operadores oferecidos pela base .

O gráfico 2, mostra em porcentagem geral dos sujeitos, que 80% deles acham importante a utilização de operadores booleanos, e 20% consideram que não seja importante utilizar operadores booleanos para fazer busca no AINFOWEb.

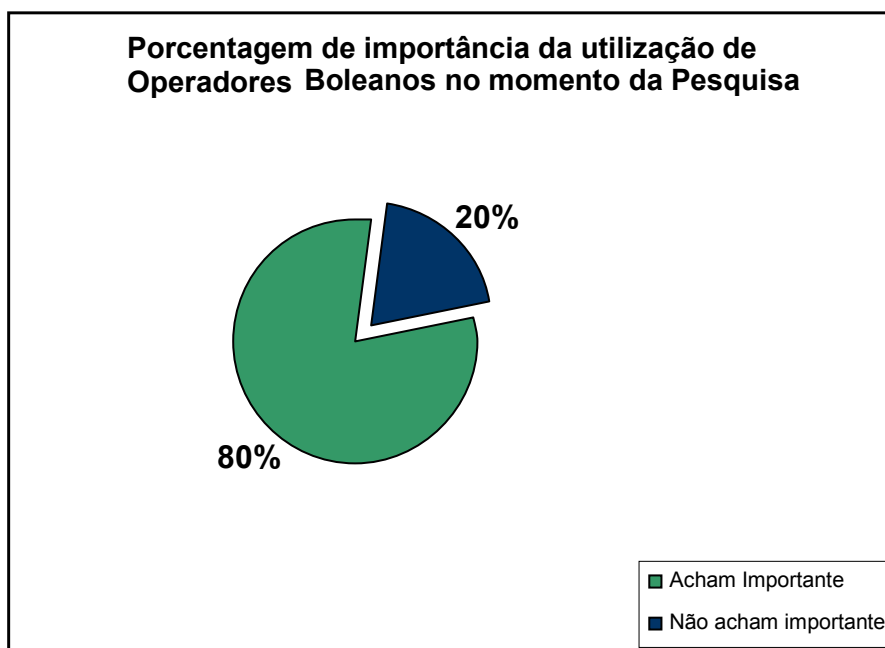


Gráfico 2 – Importância do uso de Operadores Booleanos.

A pesquisa também constatou que os usuários do AINFOWeb, utilizam-se de outras bases de dados para expandir suas pesquisas. As mais citadas foram; Capes, Agrícola, Agris2000, BDPA, CAB e Scielo.

7 CONSIDERAÇÕES FINAIS

O uso das tecnologias informacionais nas atividades de busca de informação realizadas por usuários, principalmente em bases de dados e recursos disponibilizados pela Internet, é atualmente uma realidade. Assim a pesquisa “Recuperação de informação no banco de dados da Embrapa Soja: estudo do AINFOWeb” teve como objetivos, verificar se os pesquisadores, estagiários e técnicos, formulam suas expressões de busca ao realizam suas pesquisas por meio dos operadores booleanos, bem como as ferramentas de busca

disponíveis no AINFOWeb.

A base de dados AINFOWeb é um sistema de recuperação de informação utilizado pelos usuários da Embrapa Soja, que procura dispor de informação de qualidade, fornecida em tempo real e com valor agregado para seus clientes.

De maneira geral, constatou-se que a maioria dos usuários – que realizam suas próprias pesquisas no AINFOWeb – utiliza os operadores booleanos ao formularem suas estratégias de busca, porém, alguns demonstraram possuir pouco conhecimento sobre seus mecanismos. Esse fato foi constatado, pois os usuários do AINFOWeb, não recebem nenhum treinamento formal, apenas no momento da pesquisa, quando solicitam ajudas dos bibliotecários.

A pesquisa demonstrou que a melhor

forma de pesquisa é a que utiliza os operadores booleanos, que filtram as informações, recuperam aquelas mais específicas, suprimindo assim as necessidades informacionais do usuário. Na literatura, comprovase que o uso de operadores booleanos é fundamental para uma pesquisa com qualidade e aproveitamento dos resultados.

Um aspecto investigado, e que consideramos importante, é que todos os sujeitos da pesquisa mencionaram utilizar outras bases de dados para complementar suas pesquisas. Assim, as bases de dados em CD-ROM e *on-line*, que atendam a cultura da soja e áreas relacionadas, constituem-se nas fontes secundárias mais utilizadas pelos pesquisadores, estagiários e técnicos da Embrapa Soja, na busca de informações.

Averiguou-se que os usuários do A-INFOWEB não têm total esclarecimento da função, utilização e importância do uso dos operadores booleanos para uma pesquisa com qualidade. Constatou-se também que os usuários não têm instrução suficiente para o uso correto dos operadores. Esse fato se concretiza quando se leva em consideração que os usuários não têm treinamento adequado e formalizado para utilizar a base; assim, eles optam por utilizar outras bases de dados com ferramentas de pesquisa, bem como a colocação dos operadores de uma forma mais simples.

Espera-se que esta pesquisa sirva de subsídio para a Embrapa Soja, bibliotecas e serviços de informação, de maneira a contribuir para melhorar cada vez mais o atendimento das necessidades de seus usuários e das especificidades de cada área.

REFERÊNCIAS

ANTUNES, J.F.G.; LENK, L.M. *AINFO*: versão 4.0: guia de referência. Campinas: EMBRAPA-CNPTIA, 1997. (EMBRAPA-CNPTIA. Manuais, 2).

_____.; OLIVEIRA, S.R.M. *Ainfo: a experiência da Embrapa na disponibilização e recuperação de informação. Ciência da Informação*, Brasília, v.27, n.1, p.82-89, 1998.

BERTHOLINO, M.L.F. Buscas em bases de dados. In: RAMOS, M.E.M. (Org.). *Tecnologia e novas formas de gestão em bibliotecas universitárias*. Ponta Grossa: UEPG, 1999.

BRASIL. Ministério da Saúde. Secretaria de Vigilância Sanitária. Portaria nº 398 de 30 de abril de 1999. Aprova o regulamento técnico que estabelece as diretrizes básicas para análise e comprovação de propriedades funcionais e ou de saúde alegadas em rotulagem de alimentos. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 03 maio 1999.

CRESPO, Isabel Merlo; CAREGNATO, Sonia Elisa. Comportamento de busca de informação: uma comparação entre dois modelos. *Em Questão*, Porto Alegre, v.9, n.2, jul./dez. 2003, p.271-281.

EMBRAPA Soja. *II Plano diretor*. Embrapa soja: 2000-2003. Londrina: Embrapa Soja, 2000. (Embrapa Soja. Documentos).

GROGAN, D. Buscas. In: _____. *A prática do serviço de referência*. Brasília: Briquet de Lemos/Livros, 1995. p. 109-152.

LOPES, Ilza Leite. Estratégias de busca na recuperação da informação: revisão de literatura. *Ciência da Informação*, Brasília, v.31, n.2, p.60-71, 2002.

PAIVA, Márcia Regina. *Base de dados e ferramentas de busca: seu uso no meio acadêmico*. Monografia (Especialização em Gerencia de Unidades de Informação) – Universidade Estadual de Londrina, 2002.

ROWLEY, J. *Informática para bibliotecas*. Brasília: Briquet de Lemos/Livros, 1994.

_____. *A biblioteca eletrônica*. Brasília: Briquet de Lemos/Livros, 2002.

Thais Batista Zaninelli

Mestranda em Gestão de Informação na
Fac. de Engenharia da Univ. do Porto
Graduada em Biblioteconomia pela UEL
tb_zaninelli@hotmail.com

Maria Elisabete Catarino

Docente do Deptº de Ciência da Informação da UEL
Doutoranda na Universidade do Minho
Mestre em Biblioteconomia pela
PUCCAMP
beteca@uel.br

tribute in a effective way to supply their information needs with quality, It developed a descriptive research. The data collection with the researchers, trainees and technicians of institution, had as instrument a questionnaire boarded users' identification and research in AINFOWeb. The results evidenced that most users consider important and use the operators booleanos to the formulate her search strategies, however, some demonstrated own little knowledge about their mechanisms.

Keywords

Embrapa – AinfoWeb; Information Recovery; Database research

Título

Recuperación de información en el banco de datos de Embrapa Soya: estudio de AINFO-Web

Resumen

Presenta los resultados de una investigación realizada con los utilizadores del sistema AINFOWeb que fue desarrollado por la Embrapa para tornar disponible la información existente en la empresa. Tuvo como objetivo principal averiguar si los usuarios utilizan adecuadamente los operadores booleanos y se estos efectivamente contribuyen para suprir sus necesidades informacionales con cualidad.

Palabras clave

Embrapa – AinfoWeb; Recuperación de la Información; Investigación en base de datos

Title

Information recovery of Soy Embrapa's Data Base: a study of AINFOweb system

Abstract

This paper presents the results of a research carried out with AINFOWeb system's users, a system that was developed by Embrapa institution to make available the existing information in the company. To verify if the users of this system make an adequate use of the boolean operators and if these boolean operators con-