

Avaliação de repertórios brasileiros em agricultura, ciência da informação e direito: uma análise de conteúdo

José Antônio Moreiro González
Denise Gomes Pereira de Melo
Joana Coeli Ribeiro Garcia
Emeide Nóbrega Duarte
Maria Elizabeth B. C. Albuquerque
Maria de Lourdes de Arruda Melo
Dulce Amélia de Brito Neves

FATORES RELACIONADOS AO DESENVOLVIMENTO DA PESQUISA SOBRE AVALIAÇÃO DOS REPERTÓRIOS

A necessidade de dispor de mecanismos que permitam reconhecer a existência dos repertórios documentais, garantir sua qualidade e a eficácia dos produtos que oferecem são alguns elementos que justificam o trabalho.

Desde os anos 70, o interesse em avaliar tem-se evidenciado em todas as atividades documentais, sobretudo as que se referem à recuperação da informação e, em especial quando se realizam mediante arquivos automatizados. Esse interesse da documentação por pesquisas que avaliem seus sistemas e produtos se deve sobretudo: a) à aparição de novas formas ou maneiras de armazenar, tratar, recuperar e difundir a informação; b) ao aumento da oferta de serviços e produtos de informação; c) a crescente utilização da informação para ajudar na tomada de decisões.

Assim, propusemo-nos a avaliar os produtos mais comuns elaborados pelos sistemas de informação: os repertórios. Nos referimos aos repertórios como documentos secundários que reúnem registros analíticos dos documentos primários ou originais, e estes tanto em sua edição em papel (bibliografias analíticas, boletins de resumos, sumários analíticos de periódicos), como em formato eletrônico (bases de dados *on-line*, em CD-ROM, e as que são difundidas pela Internet).

Preocupados em identificar, analisar e avaliar alguns repertórios brasileiros, apareceu como primeira possibilidade de dispersão a existência de incontáveis bases de dados, com a conseqüente ausência de regularização em seus formatos, inclusive oferecendo produtos similares, porém tratados de forma totalmente distinta (Marcaccio, 1992). O projeto objetivou identificar a presença de critérios explicitados nas diferentes fases de confecção dos repertórios: seleção das publicações e artigos a incluir, sua análise documental e sua apresentação. Tratamos de fixar a cobertura de cada um dos repertórios e, em conseqüência, também as repetições e lacunas que apresentam sobre os originais das áreas que pretendem cobrir. No entanto, a finalidade primordial foi a de diagnosticar esses repertórios. Destacamos a relevância que uma indexação correta alcança na hora de se avaliar qualquer repertório enquanto um instrumento referencial dos conteúdos específicos de cada área do repertório.

O objeto do estudo apresentou de início algumas hipóteses. Já em 1988, duas professoras da Escola de Biblioteconomia da UFMG ao referirem as características e utilização das fontes de informação especializadas reconheceram que a "produção de bibliografias especializadas brasileiras tem sofrido influências que dificultam o seu desenvolvimento" (Campello, Campos, 1988, p. 144). Entre essas influências, situavam: a) utilização maciça da informação estrangeira; b) falta de demanda pela informação nacional (como conseqüência do item anterior); c) escassa estruturação comercial dos serviços

Resumo

Avaliação da análise de conteúdo que apresentam repertórios brasileiros em Agricultura (Agrobase e Base de Dados da Pesquisa Agropecuária), Ciência da Informação (Lici), e Direito (Bibliografia Brasileira de Direito). Após identificar os antecedentes, natureza, usuários potenciais e tipo de produto oferecido por cada um dos repertórios, mediu-se a percentagem da exaustividade de sua cobertura. As tarefas relativas à avaliação da indexação e do resumo se fizeram a partir da seleção de 20 registros de cada um deles, que foram comparados mediante inspeção direta com os correspondentes originais. Utilizou-se também um questionário para determinar a linguagem documental e os critérios utilizados para fazer a análise. Atendeu-se especialmente ao exame da indexação e do resumo feitos no original, assim como sua profundidade e pertinência, medindo-se inclusive o índice de consistência quando possível. Os resumos foram estudados de acordo com o traslado da superestrutura do original e a sua qualidade técnica, medindo-se os seus fatores de legibilidade. Os resultados foram apresentados por meio de dados comparativos.

Palavras-chave

Brasil-Repertórios documentais; Repertórios documentais-agricultura, ciência da informação, direito; Avaliação da qualidade; Análise de conteúdo.

bibliográficos; d) dificuldades de difusão das bibliografias subsidiadas pelo governo.

Para superar essas dificuldades de socialização de acesso dos recursos de informação (bases de dados, textos, bibliografias etc.) o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) vem coordenando os sistemas setoriais que integram o Sistema Brasileiro de Informação Científica e Tecnológica, convertendo-se, assim, na instituição que atualmente impulsiona a produção de instrumentos bibliográficos no Brasil (IBICT, s.d.). Situação esta que valoriza e oportuniza esta pesquisa, de acordo com um contexto que foi assim definido: "a produção das bibliografias especializadas, iniciada pelo IBICT em 1954 como uma de suas três linhas prioritárias, foi transferida aos setores específicos em 1986, sob forma de bases de dados setoriais. Alguns desses setores não deram continuidade ao trabalho, não se tendo notícia do estágio atual das bases de dados transferidas aos mesmos" (Vieira, 1994).

PROCESSO DA COLETA DE DADOS

Obtenção de informação sobre os repertórios avaliados:

a) **Fixação** das áreas. O registro bibliográfico nacional brasileiro apresenta vários campos de interesse, nos quais fizemos a primeira proposta de cobertura do trabalho: Agricultura; Ciência da Informação; Comunicação; Direito; Economia.

b) **Identificação** dos repertórios brasileiros existentes nessas áreas (incluindo bibliografias e bases de dados – *on-line*, em CD-ROM, ou pela Internet –). A identificação visou a obter conhecimento sobre os seguintes aspectos do repertório:

Antecedentes

Natureza da Bibliografia ou da Base de Dados

Usuários potenciais

Tipo de **produto** concreto oferecido, considerando: conteúdo, formato e periodicidade.

Forma de **obtenção da informação** desses produtos.

Medição da percentagem da exaustividade da **cobertura do repertório** (CR) (construção de uma lista de publicações periódicas brasileiras da área estudada, comparando-a com as que o repertório recolhe). Sob a fórmula:

$$CR = C/i \times 100,$$

onde **C** é o número de publicações periódicas da área cobertas pelo repertório e **i** o número total de publicações periódicas da área no Brasil. Deve-se considerar se as publicações periódicas recolhidas são depois analisadas integral ou seletivamente.

Avaliação da análise de conteúdo (indexação e resumo quando houver)

Requeriam-se a seleção de 20 registros de cada base de dados ou bibliografia atendida e a comparação posterior com os artigos originais a que esses registros se referem. A avaliação se fez mediante inspeção direta da seleção de registros, considerada representativa de cada repertório. Este método mostra muita lentidão na sua aplicação, mas assegura que o processo seja crítico e dedutivo.

a) *Tipo de linguagem utilizada na indexação*

Linguagem livre a partir do título, do resumo e considerando o texto completo.

Linguagem controlada mediante tesouro, lista de palavras-chave, glossário de descritores e outros.

b) *Crêterios utilizados na indexação*

Formação específica dos indexadores: experiência, conhecimento da terminologia e das regras de uso da linguagem utilizada. Existência de normas ou diretrizes: estabilidade dos critérios, supervisão, controle periódico da qualidade. Respeito às normas existentes.

Reutilização da indexação no original, corrigida ou não.

Profundidade (até que nível é representada a superestrutura) e pertinência (relação entre os conteúdos do original e os termos da indexação) (Lancaster, 1991. p. 54).

Índice de consistência quando possível (só nas áreas com vários repertórios):

$$IC (\%): \frac{100A}{A+M+N}$$

onde A são os termos comuns aos repertórios M e N. (ABAD, 1997. p. 130).

c) *Crêterios observados na elaboração dos resumos*

Reutilização do resumo feito no original (pelo autor). Considerações sobre seu melhoramento técnico, em busca da coerência e normalização desejáveis (Moreiro, 1986. p. 44).

Fatores de homogeneidade mediante a existência de normas ou diretrizes: estabilidade dos critérios, supervisão, controle periódico da qualidade.

Translado da superestrutura do original (de maneira informativa, seletiva ou indicativa) (Moreiro, 1993. p. 145-157):

Os resumos *informativos* devem transladar a Macroestrutura global do texto (Mg), além da Superestrutura (Sp), quer dizer, Macroestruturas parciais (Mp) ordenadas, e ter uma expressão explicativa.

Os resumos *seletivos* devem transladar a Macroestrutura global do texto (Mg), além da Superestrutura (Sp) quer dizer Macroestruturas parciais (Mp) ordenadas, e apresentam uma expressão indicativa.

Os resumos *indicativos* limitam-se a transladar a Macroestrutura global do texto (Mg) e, às vezes, alguma das macroestruturas parciais (Mp), sempre com expressões indicativas.

Qualidade técnica dos resumos: presença de redundâncias (frases inúteis, repetição do título, presença de expressões como: *O autor analisa ... Este artigo ... O documento que ...*); ausência de dados sobre a natureza documental do original e sobre o enfoque específico com que o autor trata o assunto; presença de erros ortográficos ou sintáticos (Moreiro, 1993, p. 125-139).

Quantidade de palavras dos resumos: parece excelente que a extensão dos resumos se situe entre 100 e 250 palavras (Tenopir, 1993, p.202).

Fatores da legibilidade dos resumos (Tenopir, 1993, p. 50):

A dificultam

- Presença da voz passiva.
- Subordinação de orações dentro de uma frase

A facilitam

- Presença de enlaces sintáticos.
 - Orações curtas:
 - Orações curtas: <15 palavras
 - Orações médias: 15-20 palavras
 - Orações compridas: >30 palavras
 - (Se obterá a percentagem por repertório)

d) *Apresentação comparativa de alguns dos resultados obtidos*

e) *Aplicação do questionário*

Para obter dados, além da informação oferecida pelos próprios repertórios, foi elaborado um questionário, que enviamos aos responsáveis pelos sistemas estudados (Anexo).

DIFICULDADES DE ACESSO

Embora as informações estejam disponíveis nos vários tipos de suportes existentes (papel, disquete, CD-ROM, bases de dados etc.), o acesso é dificultado, quando a informação encontra-se "disponibilizada" em suportes eletrônicos, principalmente nas bases de dados nacionais, como é o caso dos repertórios de comunicação e economia, áreas estas pesquisadas no início. As bases consultadas apresentam apenas o registro bibliográfico, dificultando a pesquisa e o acesso ao texto. Da mesma forma, as bibliotecas universitárias não têm recebido repertórios impres-

sos atualizados. Como uma dificuldade a mais, algumas delas não estão ligadas a redes. Diante destas barreiras, a pesquisa ficou prejudicada nas áreas referidas.

A difusão mal planejada dos produtos bibliográficos referenciados obrigou a que tivéssemos de fazer a análise exclusivamente nas áreas de Agricultura, Ciência da Informação e Direito, ainda que a intenção inicial tivesse sido mais ampla.

DESCRIÇÃO E ANÁLISE DOS REPERTÓRIOS PESQUISADOS

Agrobase e base de dados da pesquisa agropecuária

Na área agrícola, destacam-se duas instituições que mantêm relacionamento com sistemas e redes informacionais da área documental agrícola: a Empresa Brasileira de Pesquisa Agropecuária (Embrapa) e a Coordenação Geral de Informação Documental Agrícola (Cenagri).

A Embrapa, ligada ao Ministério da Agricultura e Abastecimento, sediada em Brasília, recebe as fitas magnéticas pelo Commonwealth Agricultural Bureaux (CAB). Oferece serviços de disseminação seletiva da informação e busca retrospectiva a partir deles. O CAB é financiado pelo governo inglês. Armazena literatura agrícola mundial e está representado no Brasil pela Embrapa, que, por sua vez, acha-se subdividida em 40 centros de pesquisas especializadas nas áreas da agricultura e pecuária espalhados por todo o país.

O Cenagri, órgão da Secretaria de Desenvolvimento Rural do Ministério da Agricultura, do Abastecimento e da Reforma Agrária, é a unidade central do Sistema Nacional de Informação e Documentação (Snida), que tem como meta básica, por meio da operacionalização do Sistema, organizar, orientar, coordenar e difundir a nível nacional a informação documental. O Cenagri mantém relacionamento com o Sistema Internacional de Informação para a Ciência e Tecnologia Agrícola (Agris) administrado pela FAO e a Rede Internacional de Bibliotecas Agrícola (Aglinet). No âmbito nacional, é um

dos centros distribuidores da Rede de Serviços de Informação em Ciência e Tecnologia – Antares/MCT/CNPq/IBICT) – que torna possível, para acesso em linha, suas bases de dados.

Agrobase

O Cenagri criou várias bases de dados nacionais: Agrobase, Base de Dados sobre Tecnologias Apropriadas ao Meio Rural (BDTA), Base de Dados de Endereços de Instituições Agrícolas (Agroinst), Base de Dados de Cadastramento de Especialistas em Agropecuária (Agroest).

A Base de Dados Agrícola Nacional (Agrobase) é constituída pela bibliografia Agropecuária Brasileira, incluindo Piscicultura e Silvicultura. Foi criada e é mantida pelo Cenagri. É produzida no Brasil com a finalidade de disseminar e preservar a memória agrícola nacional para atender aos profissionais da área que se ocupam tanto do ensino, quanto da pesquisa e da extensão.

A partir dessa base de dados, são produzidos e encontram-se disponíveis no Cenagri, a Bibliografia Brasileira de Agricultura, as Bibliografias Estaduais, Bibliografias Especializadas por assunto e edições de vários órgãos do Ministério da Agricultura e do Abastecimento.

Contém atualmente cerca de 155 mil referências bibliográficas correntes e retrospectivas, cobrindo de 1870 até o momento. O Serviço Automatizado de Informações Bibliográficas (Saiba) fornece listagens atendendo a solicitações específicas. O acervo está disponível na Rede Antares e no site da Embrapa, no endereço: www.cnptia.embrapa.br/bdpa.

Na Base de Dados do Acervo da Embrapa, encontram-se os seguintes registros para os artigos de periódicos: tipo de publicação, autoria, título, ano da publicação, fonte (título do periódico, volume, número, páginas e ano), idioma e conteúdo (resumo), palavras-chave e unidade onde se localiza o documento.

A Agrobases informa sobre identificação do documento, categoria de assunto, nível bibliográfico, autoria pessoal, título, idioma, notas, título do seriado, ISSN, data da publicação, colação do seriado e Thesagro (palavras-chave).

A PRÁTICA NA ANÁLISE DE CONTEÚDO

O responsável pela base de dados informou, nas respostas ao questionário, que adota as normas do Agris/FAO adaptadas à realidade nacional, para análise e indexação dos documentos e justifica a não-elaboração de resumos por possuir equipe de pessoal pequena para o grande acervo (cerca de 600 a 700 mil volumes incluindo a incorporação de novas obras) a ser tratado. Utiliza linguagem controlada tendo como fonte o Thesagro (Thesaurus Agrícola Nacional).

Informa ainda que exerce controle de qualidade sobre os trabalhos dos analistas que possuem formação nas áreas de Biblioteconomia e Documentação e Ciências Agrárias. Esses últimos profissionais são requisitados sempre que se tratar de informações muito específicas da área.

Sobre a indexação, na maioria das vezes, torna-se claro que os termos não são controlados pelo Thesagro. Quando os artigos estão indexados originalmente, alguns termos são complementados para lhes dar maior significação e se tornarem mais pertinentes. Algumas vezes os descritores referem-se às metodologias utilizadas e aos lugares geográficos onde as experiências/experimentos são realizados. Embora sejam termos pertinentes ao artigo, a exaustividade leva a uma precisão baixa, pois, enquanto permite que se recupere tudo que está indexado sob aquele termo geral, deixa ao usuário o trabalho de ter de selecionar dentre a grande quantidade de referências aquela específica que lhe interessa. Ou seja, a recuperação se realiza com certa dificuldade.

É importante ressaltar que essa base deveria ser o mais completa possível em termos de pertinência, exaustividade e precisão da indexação para superar a ausência do resumo.

Bases de dados da Pesquisa Agropecuária

O acervo documental da Embrapa contém 241 685 referências, abrangendo a produção técnico-científica dos pesquisadores da empresa, teses e a literatura nacional e internacional adquirida: Catálogo Coletivo de Periódicos da Embrapa com 7 440 títulos e suas respectivas coleções e base de instituições do Sistema Nacional de Pesquisa Agropecuária, com endereços, serviços prestados e áreas de atuação de 177 instituições, reunidos em um CD-ROM. Contém ainda todo o acervo bibliográfico das 40 bibliotecas da empresa espalhadas pelo Brasil, além de toda produção de teses e *papers* científicos de autoria de seus 2 000 cientistas. Para os artigos de periódicos são registrados: tipo de publicação, autoria, filiação, título, ano de publicação, fonte, idioma, conteúdo, palavras-chave e unidades que dispõem do documento. Inclui a Agrobases com 178 239 registros, Bracarís com 1 565 registros e a Base de Dados sobre o Cerrado, com 8 633 registros.

Quanto à **natureza**, a base bibliográfica é analítica, contendo resumos. Seus usuários potenciais são constituídos por técnicos da extensão, alunos de graduação e pós-graduação, professores e pesquisadores da área agropecuária e demais áreas correlatas.

Para obtenção dos produtos da base, os técnicos da Embrapa têm acesso por comutação da própria instituição sem ônus para os mesmos. Aos usuários externos, o acesso se dá por meio do programa de comutação do IBICT, por intermédio das bibliotecas conveniadas, mediante o pagamento antecipado por cupons.

De um total de 104 títulos de periódicos correntes nacionais da área agrícola, 74 possuem artigos técnicos e/ou científicos, portanto são aqueles que podem ser indexados em base de dados. A percentagem da exaustividade da Cobertura deste Repertório (CR) em referência a essas publicações é:

$$CR: 74/104 \times 100 = 71,15$$

mostrando, portanto, alta capacidade para representar a produção científica da área agrícola. Porém, a cobertura deste repertório tem uma percentagem de exaustividade de **100**, se consideramos que indexa todos os artigos técnicos e/ou científicos das revistas da área.

Nas respostas ao questionário, a pessoa responsável pela base de dados informou que, na análise dos conteúdos para indexação e elaboração de resumos, os analistas não usam normas ou diretrizes preestabelecidas. Utilizam linguagem natural com controle da qualidade do trabalho dos analistas. Tanto a gerente do CNTIA como os analistas têm como formação básica, a graduação em biblioteconomia e documentação.

Sobre a indexação, percebemos que, na maioria dos casos do estudo, os descritores são reutilizados dos originais, inclusive usam termos em inglês quando os autores também elaboram *abstract*. Muitas vezes, há acréscimo de outros termos secundários que não têm significância para a recuperação exata da informação que se deseja. Por outro lado, como reutilizam a indexação feita pelos autores, os problemas técnicos são simplesmente repassados.

ANÁLISE DA ELABORAÇÃO DOS RESUMOS

Foi possível perceber que os resumos são reutilizados dos originais elaborados pelos autores. Assim, os problemas de ordem técnica, ou seja, de desconhecimento das normas de elaboração de resumos, são repassados para a base de dados. Os resumos pretendem ser informativos, porém, em alguns casos, deixam de contemplar elementos que compõem a macroestrutura do texto. Verificou-se a existência de dois casos em que os resumos não apresentavam a conclusão do artigo. Um artigo cujo resumo estava em inglês, na base não constam nem a indexação a partir do título, nem o resumo traduzido.

Em dois casos identificamos linhas suprimidas no resumo da base, bem como palavras e/ou fórmulas químicas truncadas em uma demonstração clara de que não há controle de qualidade do trabalho realizado pelos indexadores. Se, por um lado, há problemas técnicos quando os autores elaboram os resumos; ao passar para a base de dados, esses problemas são acrescidos de outros de responsabilidade dos indexadores. Por essas razões, a legibilidade dos resumos deve ser considerada de nível médio. Como os resumos são elaborados pelos autores, as formas são muito variadas em estilo, tamanho do resumo, tamanho das frases que o compõem, variando de 15 a 20 palavras a mais de 30. Em três casos, houve presença de frases na voz passiva.

ÍNDICE DE CONSISTÊNCIA ENTRE OS TERMOS DE INDEXAÇÃO NAS DUAS BASES

Recuperamos 10 artigos que se encontravam registrados concomitantemente nas duas bases trabalhadas. A Agrobases apresentou 41 descritores para os artigos enquanto a Base da Embrapa, para os mesmos artigos, apresentou 78 termos. Destes, apenas 13 termos coincidiram estar tal e qual nas duas bases. Aplicando-se a fórmula para cálculo do índice de consistência chegou-se a **9,34** o que é um índice muito baixo. Em dois artigos não apareceu sequer uma coincidência de termos.

Lici (Ciência da Informação)

É produzida pelo IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia), órgão subordinado ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) do Ministério da Ciência e Tecnologia (MCT). Este centro tem como missão promover o desenvolvimento do setor de informação, por meio da proposição de políticas, execução de pesquisas e difusão de inovações, capazes de contribuir para o avanço da ciência e para a competitividade da tecnologia brasileira.

1. No âmbito nacional, realiza estudos e pesquisas na área de sua competência, coordena redes de informação, oferece e desenvolve produtos e serviços e proporciona, ainda, a formação e capacitação de profissionais em ciência da informação. Em âmbito internacional, atua em projetos de informação científica e tecnológica do Mercosul, participa do Programa Geral de Informação da Unesco e da Federação Internacional de Informação e Documentação (FID), participando, ainda, como membro da rede ISSN.

Antecedentes

Bibliografia Brasileira de Documentação. Rio de Janeiro: IBBD, 1960...
Compilador Edson Nery de Fonseca.

v. 1: 1911-1960. v. 2: 1960-1970

Só tinha descrição catalográfica. Com Índice KWIC.

Bibliografia Brasileira de Ciência da Informação. Brasília: IBICT.

v. 3/4: 1971-1977. Com Índice KWOC.

v. 5: 1978-1980. Com Índice KWOC.

v. 6: 1980-1983. Software SRB (Sistema de Registro Bibliográfico, feito pelo IBICT).

Continha lista dos periódicos analisados. Os registros eram feitos com resumo, e índices de autores pessoais, corporativos e assuntos (sob indexação coordenada com o Tesouro em ciência da informação do IBICT).

Natureza da Base de Dados Lici: Bibliográfica (Referencial).

Usuários potenciais: alunos de graduação e pós-graduação, professores, pesquisadores na área de ciência da informação e correlatas.

Produto oferecido: Conteúdos Biblioteconomia, Ciência da Informação, Arquivologia, Museologia e Informática. Formato: CALCO. Os elementos com que se registram os artigos de periódicos são: Título; Autor; Resumo; Publicação; Descritores e Idioma. A informação desses produtos é obtida pela Internet: <http://www.ibict.br/bibvirtu>. A interface é via Telnet.

Cobertura: Analisa integralmente todas as publicações recolhidas, entre as quais se destacam os anais de congressos, dissertações e teses dos campos analisados, assim como todos os artigos das revistas:

– *Revista da Escola de Biblioteconomia da Universidade Federal de Minas Gerais.* UFMG.

– *Revista do Núcleo da Universidade Federal Fluminense.* UFF.

– *Revista de Biblioteconomia de Brasília.* ABDF. UNB.

– *Revista Brasileira de Biblioteconomia e Comunicação.*

– *Revista Brasileira de Biblioteconomia e Documentação.*

– *Ciência da Informação.* IBICT.

– *Transinformação.* PUCCAMP.

– *Cadernos de Biblioteconomia.* UFPE.

– *Informare.* UFRJ. IBICT.

– *Biblos.*

– *Informação & Sociedade: Estudos.* UFPB.

– *Acervo: Revista do Arquivo Nacional.*

A percentagem da exaustividade da **cobertura deste repertório** (CR) em referência às publicações periódicas é:

$$\text{CR: } 12/14 \times 100 = \mathbf{81},$$

que deve ser considerado de notável capacidade para representar a produção científica das áreas a que atende.

Análise da indexação prática

No questionário, os responsáveis pela base de dados manifestaram que a indexação é feita, quando não se utiliza a que foi realizada pelos próprios autores, de forma controlada usando o *Tesaurus Spines* (da UNESCO) e o *Tesaurus Ciência da Informação* (TCI), construído e mantido pelo próprio IBICT.

Os analistas são graduados em Biblioteconomia e Documentação. A análise é feita de acordo com as normas para indexação e resumo da ISO, Associação Brasileira de Normas Técnicas (ABNT), e do IBICT. A este respeito, foi manifestado no questionário que as normas são atualizadas continuamente. Aparentemente, existe controle de qualidade do trabalho feito pelos analistas.

Sobre a prática analítica, observou-se que é reutilizada a indexação na original geralmente sem modificações. Quando a indexação não existe no original, é feita a partir dos títulos, raramente a partir dos resumos. Por isso, é uma indexação muito genérica que descreve só a informação global do documento correspondente à sua macroestrutura, inclusive, excluindo os descritores de lugar e os onomásticos. Porém, a pertinência pode ser considerada boa. A exaustividade de termos atribuídos proporciona uma precisão baixa ao utilizar às vezes termos muito amplos, o que resulta um ambíguo comportamento das representações de conteúdo na hora da recuperação. Nos 68 descritores dos registros analisados, foram encontradas, em quatro ocasiões, redundâncias, ocasionando alta revocação e uma escassa precisão na busca de informação.

ANÁLISE DA ELABORAÇÃO DOS RESUMOS

Reutilizam-se, quando existem, os resumos elaborados pelos autores dos originais em toda sua essência. Como não se introduz nenhuma modificação, conservam, por isso, as deficiências técnicas com que foram feitos: localizamos expressões inúteis, erros sintáticos, orações incompletas, ausência de legibilidade nas idéias apresentadas, fazendo, assim, que a preocupação com a qualidade dos resumos seja, nesses casos, escassa.

Os registros mais antigos, anteriores a 1980, carecem de resumos em 70% dos casos. Quando existem, são de modelos indicativos. A maioria dos trabalhos publicados em anais de congressos também não apresenta resumo. Os registros posteriores a 1980 transladam a superestrutura do original de maneira informativa em 30% dos casos, seletiva em 55% e, finalmente, indicativa nos 15% restante.

A qualidade técnica dos resumos fica dependente daqueles que os elaboram, que, muitas vezes, desconhecem as normas e as exigências da redação dos resumos. Por isso, a legibilidade deve ser considerada de nível médio. Os resumos de estilos e formas muito variadas dependeram só da vontade e do estilo do autor que fez o original. Onze dos resumos apresentavam todas as orações com menos de 20 palavras de extensão. Três apresentaram orações com mais de 30 palavras, alcançando, em um caso, uma média de 50 palavras, com grande subordinação de orações dentro de suas frases. Os restantes misturavam orações de menos de 30 palavras com outras maiores de 40. Em quatro casos, constatou-se a presença da voz passiva.

As informações contidas na base de dados dão a idéia de que não são submetidas a nenhum processo de controle na linguagem de indexação e nem de revisão na estrutura e apresentação dos resumos.

OUTRAS OBSERVAÇÕES

Houve dificuldades de leitura na tela de uma referência na íntegra: um texto com resumo longo apareceu incompleto, portanto a tela é pequena para tanta informação.

Os operadores booleanos *and* e *or* usados em uma só busca não foram compreendidos pelo sistema. Um dos registros apareceu como resultado de uma busca em que se utilizou a combinação dos descritores “**indexação and automática**”, quando este último conceito não aparecia em nenhum dos termos de indexação e também não era palavra que formasse parte do título ou do resumo.

Bibliografia Brasileira de Direito

Antecedentes: Bibliografia Brasileira de Direito

v.1 ao v.9 - Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Publicou até o ano de 1979.

Bibliografia Brasileira de Direito; Nova Série - Senado Federal.

v. 1 – período de 1980 a 1981 (setembro 87);

v. 2 – período de 1982 e 1983 (outubro 88);

v. 3 – período de 1984 e 1985 (1986);

v. 4 – ano de 1986 (outubro 87)

v. 5 – ano de 1987 (1988)...

Objetivo: colocar à disposição dos usuários as informações armazenadas pelas bibliotecas da Rede.

Arranjo: as referências bibliográficas são apresentadas segundo grupos de assuntos e, dentro de cada grupo, em ordem alfabética autor-título.

Natureza: Bibliográfica (Referencial). Bibliografia Impressa e Base de Dados acessível pela Internet. Usuários potenciais: alunos de graduação e pós-graduação, professores, advogados, magistrados e pesquisadores da área de Direito.

Produto Oferecido: Conteúdo: Direito. Reúne monografias, artigos de periódicos e artigos de jornais. Tratamento Técnico: as monografias e os artigos de periódicos receberam tratamento de forma descentralizada em esquema cooperativo pelas bibliotecas que alimentam os Bancos de Dados BIRD (monografias) e Peri (artigos de periódicos). Desde 1994, passou a incluir a Jorn (artigos de jornais), alimentada exclusivamente pela Seção de Materiais Especiais da Biblioteca do Senado Federal.

Formato: MARC (Machine Readable Cataloguing). Utiliza as regras de catalogação do Código de Catalogação Anglo-Americano (AACR2), segundo nível, respeitando como fator de qualidade, o que prescreve a norma ISO 2.709. Acesso: Por meio das bibliografias impressas ou pela Home Page do Senado Federal. Periodicidade: Anual

A publicação impressa apresenta índices que remetem para o número sequencial do item e, ainda, a relação dos periódicos analisados, com a indicação do endereço dos editores e a relação das siglas e endereços das bibliotecas da Rede. A Base de Dados eletrônica fornece indicações sobre o tipo de publicação, autor, título, notas tipográficas, indexação e especificam as bibliotecas que possuem os documentos.

Cobertura: Não é exaustiva em relação ao período coberto, registra publicações recebidas pelas bibliotecas da Rede, no campo do Direito, produzidas no Brasil. Fazem parte da Rede Sabi (Subsistema de Administração de Bibliotecas), gerenciada pela Biblioteca Acadêmica Luís Viana Filho, do Senado Federal. São 15 bibliotecas, a saber:

Senado Federal; Centro de Informática e Processamento de Dados do Senado Federal (Prodasen); Câmara dos Deputados; Estado Maior da Armada do Ministério da Marinha; Supremo Tribunal Federal; Ministério do Trabalho; Ministério da Justiça; Superior Tribunal de Justiça; Tribunal Superior do Trabalho; Advocacia Geral da União; Tribunal de Contas do Distrito Federal; Procuradoria Geral da República; Superior Tribunal Militar; Câmara Legislativa do Distrito Federal; Tribunal de Justiça do Distrito Federal e dos Territórios.

Análise da prática na indexação e na elaboração dos resumos

A indexação é feita de acordo com o Vocabulário Controlado Básico do Senado Federal (VCBS), que é atualizado sempre que necessário. Os analistas responsáveis pela indexação são graduados na área específica da Base de Dados – Direito – e utilizam manuais próprios para executar a atividade.

Os artigos originais não apresentaram indexação ou resumo, apesar de os analistas, segundo resposta ao questionário, elaborarem resumos quando julgam que a indexação não cobriu totalmente o assunto e utilizarem termos não-autorizados no vocabulário, quando consideram importante para a recuperação da informação. No entanto, não foi encontrado, na Base de Dados, nenhum registro com resumo.

Os descritores apresentados na Base de Dados correspondem, na sua maioria, à temática central do artigo. Os termos utilizados na indexação permitem recuperar a informação com alta precisão.

DADOS COMPARATIVOS DOS RESULTADOS OBTIDOS

Percentagem da exaustividade da cobertura

BD Embrapa: **71,15** (100 dos artigos)
LICI: **81**

Índice de Consistência entre os Termos de Indexação

Agrobase/Embrapa: **9,34**

Percentagem de modelos dos resumos apresentados na BD Embrapa

Informativos: 100%
Seletivo: 0
Indicativo: 0

Percentagem de modelos dos resumos apresentados na BD LICI

Informativos: 30%
Seletivos: 55%
Indicativos: 15%

Extensão média dos resumos em palavras

BD Embrapa: 200
BD LICI: 118

Percentagem da extensão das orações

	Embrapa	LICI
Curtas (<15)	11,11	7,4
Médias (15-30)	62,96	63,1
Longas (>30)	25,92	29,5

CONCLUSÕES E RECOMENDAÇÕES

Todas as bases de dados pesquisadas apresentaram falhas técnicas na elaboração dos resumos, quando estes existem, e nos termos de indexação. Esses problemas são decorrentes de falta de controle de qualidade no processo de indexação, na transcrição dos resumos e, em alguns casos, do uso da linguagem natural, afetando a recuperação da informação.

Manifestou-se, nas respostas aos questionários, que se exercia controle de qualidade do trabalho feito pelos analistas. Porém, constatou-se uma multiplicidade de uso tanto na confecção dos resumos como na representação dos conceitos na indexação. Recomenda-se maior homogeneidade na aplicação de critérios, observação das normas e um efetivo controle da qualidade.

A ausência de resumos e/ou descritores significativos dificultam sobremaneira o trabalho do pesquisador, por não permitir o aprofundamento no levantamento bibliográfico de seu interesse. Vale salientar que se faz necessária, por parte dos Conselhos Editoriais dos Periódicos, a exigência de descritores e resumos, o que traria contribuições tanto para os usuários/pesquisadores, quanto auxiliaria os indexadores das bases.

A inexistência de outros repertórios em algumas áreas impediu a análise comparativa, o que permitiria o estabelecimento do índice de consistência entre as bases e um estudo mais aprofundado do produto oferecido.

É importante ressaltar que, quando as bases de dados não apresentam resumos, sua indexação deveria ser a mais completa possível em termos de pertinência, exaustividade e precisão para superar a ausência da informação que deriva da ausência do resumo.

A reutilização dos resumos do autor permite a transferência rápida e menos onerosa da produção científica, mas as deficiências técnicas devem ser corrigidas nos serviços de análise quando se pretende alcançar produtos informativos de qualidade. A melhoria na qualidade dos resumos obriga a supervisionar a sua elaboração e mostra a necessidade da presença de um corpo de indexadores especializados e de normas de indexação e apresentação de resumos que sejam realmente aplicadas.

Observou-se a necessidade de facilitar o acesso pela Internet, de tornar mais amigáveis as operações de consulta e de envio dos registros recuperados, assim como de melhorar a exatidão da linguagem de busca.

Evaluation of the Brazilian repertories in agriculture, information science and law: the quality of content analysis

Abstract

Evaluation of content analysis which the Brazilian repertories in Agriculture (Agrobases and the Data Base of Agricultural and Animal Husbandry Research), Information Science (LIC), and Law (Brazilian Bibliography of Law) present. After identifying the antecedents, nature, potential users and type of product offered by each of these repertories, a percentage measure of the exhaustive breadth of their coverage was carried out. The tasks related to the evaluation of the indexing and abstracts were made based on a selection of twenty cases for each, which were compared by means of direct inspection with the corresponding originals. A questionnaire was also used to determine the documentary language and the criteria employed to carry out the analysis. Special attention was given to an examination of the indexing and abstract presented in the original, as well as their depth and relevance, measuring where possible the index of consistency. The abstracts were studied in accordance with the transcription of the superstructure of the original and its technical quality, measuring the factors of legibility. The results were presented by means of comparative tables.

Keywords

Brazil; Documentary repertories; Agriculture; Information science; Law; Evaluation of quality; Content analysis.

Os problemas detectados nas bases nacionais são transferidos para aquelas por elas alimentadas. O ideal seria verificar até que ponto os periódicos de resumo (LISA e ISA pelo menos) que indexam os artigos da área da Ciência da Informação aproveitam os resumos e os termos de indexação apresentados no documento original.

Em favor de maior difusão da produção científica brasileira, seria conveniente a apresentação das palavras-chave e dos resumos também em inglês (como língua utilizada majoritariamente pela comunidade científica) e em espanhol, pelas conotações contextuais da situação econômica e política brasileira.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ABAD GARCIA, Francisca. *Investigación evaluativa en documentación: aplicación a la documentación médica*. Valencia: Universidad de València, 1997.
2. BLAIR, D. *Language and representation in information retrieval*. Amsterdam: Elsevier, 1990.
3. CAMPELO, Bernadete Santos, CAMPOS, Carlita Maria. *Fontes de Informação especializada: características e utilização*. Belo Horizonte: Ed. UFMG, 1988.
4. EISENBERG, M. Measuring relevance judgments. *Information Processing and Management*. n.24, p.373-389, 1988
5. FOLSTER, M. B. A study of the use of information sources by social science researchers. *Journal of the Academic Librarianship*, n.1, p.7-11, 1989.
6. IBICT. *Bases de dados nacionais*. Rio de Janeiro, [s.d.].
7. LANCASTER, F. W. *El control de vocabulario en la recuperación de la información*. Valência: Universidad de Valência, 1990.
8. _____. *Indexing and abstracting in theory and practice*. Champaign: University of Illinois, 1991.
9. _____, WARNER, A. *Information retrieval today*. Arlington: Information Resources, 1993.
10. MARCACCIO, K. Y. (Ed.). *Directory of Online Databases*. Detroit, v.13, n.2, 1992.
11. MOREIRO GONZÁLEZ, José Antônio. *Aplicación de las ciencias del texto al resumen documental*. *Boletín Oficial del Estado*. Madrid: Universidad Carlos III, 1993.

12. _____. Los resúmenes de autor: su aplicación y su lugar en la comunicación científica. *Cuadernos de Documentación de Cajas de Ahorro*, n.78, p.33-56, 1986.
13. ROBREDO, Jaime, CUNHA, Murilo B. da. *Documentação de hoje e de amanhã; uma abordagem informatizada da biblioteconomia e dos sistemas de informação*. 2. ed. São Paulo: Global, 1994.
14. TENOPIR, Carol, JACSÓ, Péter. Quality of abstracts. *Online Review*. n.5, p. 44-55, 1993.
15. VIEIRA, Anna da Soledade. *Redes de ICT e a participação brasileira*. Brasília: IBICT-CNPq- SEBRAE, 1994.

José Antonio Moreiro González

Doutor em Geografia e História - Universidad Carlos III - Madrid - Espanha

jose@openline.com.br

Denise Gomes Pereira de Melo

Mestre em Biblioteconomia - UFPB

Joana Coeli Ribeiro Garcia

Mestre em Biblioteconomia - UFPB

Emeide Nóbrega Duarte

Mestre em Biblioteconomia - UFPB

Maria Elizabeth B.C.de Albuquerque

Mestre em Biblioteconomia - UFPB

Maria de Lourdes de Arruda Melo

Mestre em Ciência da Informação - IBICT

Dulce Amélia de Brito Neves

Mestre em Biblioteconomia - UFPB

ANEXO

QUESTIONÁRIO

01) Dispõe de normatização ou diretrizes para realizar a análise de conteúdo (indexação e resumo)?

Sim [] Não []

02) Caso a resposta tenha sido afirmativa, as normas utilizadas são:

ABNT [] **IBICT** [] **ISO** [] **Da própria Editora** []

Outras (especificar) _____

03) Foi necessário revisar as normas adotadas?

Sim [] Não [] Por quê? _____

04) Qual a formação dos analistas?

Graduado em Biblioteconomia e Documentação []

Graduado na área da Base de Dados []

Outra graduação (Especificar _____)

Sem graduação e/ou treinados na editora []

05) Utiliza linguagem natural [] ou linguagem controlada []?

06) Se utilizam linguagem controlada, qual é? _____

07) Utiliza normas para a elaboração de resumos?

Sim [] Não []

08) Em caso de não-elaboração de resumos, cite as razões porque não o fazem:

09) Existe algum controle de qualidade do trabalho dos analistas?

Sim [] Não []