

A indexação de clippings de coordenadorias de comunicação social em ambiente universitário: um estudo comparativo entre a prática manual e a automática

Gabriela de Carvalho Cafruni^{1,2}, Vera Regina Casari Boccato³

¹Estudante de Biblioteconomia e Ciências da Informação. Universidade Federal de São Carlos, SP, Brasil

²Bolsista do Programa de Educação Tutorial em Biblioteconomia e Ciências da Informação. PET BCI.

Universidade Federal de São Carlos, SP, Brasil. E-mail: gaby_cafruninha@hotmail.com

³ Professora Associada do Departamento de Ciências da Informação. Universidade Federal de São Carlos, SP, Brasil. E-mail: vboccato@ufscar.br

Resumen. Na contemporaneidade, o bibliotecário tem inúmeras oportunidades de atuação profissional em diferentes áreas de informação, especialmente na Coordenação de Comunicação Social (CCSS), de universidades públicas brasileiras. Estes objetivos de gestão CCSS dos processos de comunicação e o fluxo de informações das instituições. Os recortes, considerados como fontes de notícias, proporcionam o acesso para a recuperação e uso das informações contidas em publicações como jornais diários, revistas científicas e da mídia, a fim de preservar e divulgar a memória histórica através de instituições de notícias. Uma vez que leva a indexação de conteúdo recortes para recuperação rápida e eficaz de informações do usuário, o objetivo deste estudo é uma análise comparativa entre a indexação manual e automática para identificar o tipo mais apropriado para a representação dos recortes em CCSS automatizado universidades brasileiras. A base de investigação é o CCS, da Universidade Federal de São Carlos (CCS-UFSCar), na realização da indexação, de 30 de recortes, cobrindo as áreas de Ciências Exatas, Biológica e Humanas. A indexação manual será realizada teorias de Lancaster (2003) e a automática de Gil Leiva (2008). Dado que essa investigação está em andamento, os resultados esperados referem-se à recomendação do tipo mais apropriado para indexação e recuperação de desempenho em sistemas automatizados da CCS-UFSCar, aplicável a outros contextos de informação nacionais e internacionais.

Introdução

A temática desta pesquisa é a realização de um estudo comparativo entre a indexação manual e a indexação automática de *clippings* de coordenadorias de comunicação social em ambiente universitário.

Segundo Le Coadic (1996, p. 26), a Ciência da informação

Tem por objeto o estudo das propriedades gerais da informação (natureza, gênese e efeitos), ou seja, mais precisamente: a análise dos processos de construção, comunicação e uso da informação; e a concepção dos produtos e sistemas que permitem sua construção, comunicação, armazenamento e uso.

Nesse sentido, a Ciência da Informação estuda a informação desde a sua formação até o processo de modificação de dados em conhecimento e a aplicação da informação em organizações. Deu-se origem para solucionar problemas como o de reunir, organizar e tornar acessível o conhecimento cultural, científico e tecnológico produzindo em todo mundo.

A Ciência da Comunicação passa a existir no contexto das sociedades modernas, com o objetivo de analisar os elementos que envolvem a produção, o tratamento e as conseqüências pela comunicação, e avaliar as relações entre as pessoas e os meios de comunicação, oferece um desenvolvimento científico que compõe uma visão característica para entender as transformações da sociedade.

A partir das considerações de Januário (2010, p. 8)

A efetiva relação entre a CI e Comunicação mostra-se a partir de várias dimensões, citando-se acima apenas duas delas. De um modo geral, para a CI a informação é um fenômeno, enquanto que para Comunicação é um processo. Porém, um fenômeno precisa de um processo para se efetivar, enquanto que um processo precisa de uma origem – nesse caso o fenômeno informação – para ter o quê processar. Ambas estão interessadas na comunicação humana e este é um dos motivos pelos quais elas devem ser estudadas em conjunto, pois visam o mesmo fim – em aspectos diversos e/ou correlatos – mas com objetivos comuns.

Desse modo, na interlocução entre a Ciência da Informação e as Ciências da Comunicação, respectivamente, a hemeroteca e o *clipping*, possibilitam o acesso para a recuperação e uso de informação contida em periódicos, tais como jornais diários, revistas científicas e de divulgação com o intuito de disseminar e preservar a memória histórica institucional por meio de notícias.

Para Medeiros, Melo e Nascimento (2008, p. 8) a “[...] hemeroteca refere-se a um acervo de jornais e revistas, de modo que apresente uma determinada organização técnica que facilite o processo de busca e recuperação da informação”.

A hemeroteca, vista como uma sessão de biblioteca disponibiliza recortes de matérias publicadas em jornais e revistas armazenados com a finalidade de proporcionar ao usuário novas informações e fontes de pesquisa. Muitas vezes, ela tem por objetivo suprir a carência de atualização da biblioteca, em que os assuntos guardados são selecionados, visando complementar as informações contidas nos livros, entre outras fontes, por meio da elaboração que facilite a consulta de seus usuários para a leitura e a pesquisa.

Sobre o *clipping*, Rabaça e Barbosa (2001, p. 138) definem como um

[...] Serviço de apuração, coleção e fornecimento de recortes de jornais e revistas sobre determinado assunto, sobre as atividades de uma empresa ou instituição, sobre determinada pessoa etc. É realizado geralmente pela área de comunicação (relações públicas, imprensa ou marketing institucional) da organização, pela agência de RP ou de publicidade que atende à empresa ou por uma agência especializada nesse tipo de serviço, conhecida como **agência clipper**. Diz-se tb. *clippagem*. **2.** Recorte de jornal. **3.** O conjunto de recortes fornecidos ao interessado e/ou arquivados.

Os clippings são importantes “[...] para a compreensão do fluxo informacional, já que serve de *feedback* aos releases enviados à imprensa; e é ferramenta de gestão estratégica da informação, pois também é um subsidiador para o gerenciamento da imagem da instituição”. (Bocato, 2010).

Em virtude dos avanços tecnológicos ocorridos no tratamento, na disseminação e recuperação da informação, tem-se, na atualidade, a hemeroteca e o clipping em formatos digitais, constituindo-se, assim, a hemeroteca digital e o *clipping* eletrônico (*eletronic clippings/e-clipping*).

Sobre a indexação, vista como um processo de tratamento temático da informação, ela possui pressupostos teóricos distintos dependendo da corrente teórica em que esta inserida. Na corrente espanhola a indexação possui dois níveis de tratamento, sendo o da forma, com a realização da descrição bibliográfica do documento e a de conteúdo com a sua representação temática. Para Gil Leiva (2008) a indexação é um processo técnico documental, isto é, um conjunto de operações dirigidas para a seleção, aquisição, registro e tratamento dos documentos com finalidade de possibilitar seu armazenamento e recuperação, e posteriormente sua propagação.

A corrente inglesa não faz distinção entre os processos de análise documentária e de indexação, em que segundo Lancaster (2004), a indexação de assuntos de um documento é eficiente quando ela vai além dos assuntos tratados no documento. O indexador deve tomar sua decisão de indexar baseada na causa do documento, além de observar prováveis interesses para um determinado grupo de usuários.

Pela perspectiva de Chaumier (1988, p. 63), teórico da corrente francesa, a indexação é considerada como a atividade de representar o conteúdo do documento com a utilização de linguagens documentárias para a geração de produtos documentários como os índices e as notações de classificação definem indexação como a “operação que consiste em escrever e caracterizar um documento, com o auxílio da representação dos conceitos nela contidos”.

Em complementação aos postulados teóricos apresentados, a norma técnica 12676 (1992) da Associação Brasileira de Normas Técnicas (ABNT)¹ expõe os estágios da indexação, que é composta por três etapas, sendo elas:

1. exame do documento e estabelecimento do assunto de seu conteúdo;
2. identificação dos conceitos presentes no assunto;
3. tradução desses conceitos nos termos de uma linguagem de indexação.

Acerca da qualidade da indexação, a respectiva norma técnica 12676 (1992, p. 4) atribui três fatores para o alcance da qualidade na indexação. São eles:

1. consistência na especificidade dos termos atribuídos a um documento e no nível de exaustividade atingido na indexação;
2. qualificações do indexador (imparcialidade, conhecimento, etc.);
3. qualidade dos instrumentos de indexação.

A indexação pode ser classificada em três tipos: 1) manual; 2) automática; e 3) semi-automática.

A indexação manual é feita pelo indexador no desenvolvimento da análise de assunto, na identificação e seleção de conceitos, para a representação por meio de termos de uma linguagem de indexação adotada pelo sistema de recuperação da informação (bases de dados, catálogos on-line e demais sistemas automatizados). Esse processo visa garantir a qualidade na indexação a partir do grau de especificidade do assunto tratado no documento que a indexação manual possibilita, tendo em vista os objetivos organizacionais, os interesses de pesquisa dos usuários e a política de indexação estabelecida pela unidade de

¹

informação. Além disso, ela também tem a possibilidade de selecionar assuntos do texto que mostram ser bons indicadores do que se trata o documento, porém caracterizando-se, muitas vezes, como lenta, subjetiva e de alto custo.

Segundo Rubi (2008, p. 19) a política de indexação

[...] deve ser representada por meio de uma filosofia que reflita os objetivos do sistema. Um dos aspectos concernentes à política de indexação diz respeito à conversão retrospectiva de dados e à compatibilidade de linguagem de indexação entre integrantes de um mesmo sistema cooperativo.

Com a ampliação e uso de novas tecnologias, passa a ser discutido pela literatura científica (Anderson, 1994; Lancaster, 2004; Gil Leiva, 1999, 2008, entre outros) o desenvolvimento de métodos de tratamento temático automatizados, visando, principalmente, a agilização na realização de atividades, tais como a indexação.

Por volta do final dos anos 1950, Hans Peter Luhn deu os passos iniciais em direção à indexação automática, durante suas atividades notou que tinha um “princípio do mínimo esforço” na comunicação escrita ou falada, referente à tendência de repetição de certas palavras em vez da utilização de palavras diferentes na comunicação oral ou escrita e propôs o primeiro método de indexação automática, que considerava a frequência das palavras dos títulos dos documentos, compondo um índice permutado, chamado KWIC (*Keyword in Context*).

No processo por extração automática, as palavras ou expressões que aparecem no texto são extraídas, por um programa de computador, e utilizadas para representar o conteúdo do texto como um todo, adotando-se critérios de frequência, posição e contexto (Lancaster, 2004, p. 286).

No processo por atribuição automática é necessário desenvolver, para cada termo a ser atribuído, um “perfil” de palavras ou expressões que costumam ocorrer frequentemente nos documentos às quais um indexador humano atribuiria esse termo. Se a cada termo de um vocabulário controlado correspondesse um perfil desses, seria possível utilizar programas de computador para cortejar as expressões importantes num documento (essencialmente aquelas que fossem extraídas segundo critérios de frequência) com essa coleção de perfis, atribuindo um termo ao documento sempre que esse perfil do documento coincidissem com o perfil dos termos. (Lancaster, 2004, p. 289).

Na indexação semi-automática, temos os procedimentos da automática com possibilidade de interferência do indexador. O programa realiza a indexação automática, em seguida o profissional indexador ajusta os termos de acordo com as necessidades.

Nessa perspectiva, a indexação tem por finalidade descrever os assuntos dos documentos, visando à recuperação da informação em sistemas automatizados. O indexador analisa os conteúdos a partir da realização da leitura documentária para a identificação e seleção de conceitos, representados em termos escolhidos dentre a linguagem de indexação utilizada pelo sistema.

Para tanto, definimos como o nosso problema de pesquisa a necessidade de verificação da qualidade da indexação de *clippings* a partir da prática manual e da agilização desse processo com o uso da indexação automática, frente a um grande volume de *clippings* a serem indexados e pela perspectiva dos princípios teóricos e metodológicos da área de Organização e Representação do Conhecimento.

A prática da indexação requer profissional especializado com conhecimento sobre o processo e procedimentos que envolvem a realização dessa atividade. O bibliotecário indexador possui as habilidades e competências necessárias adquiridas, a partir, não só de sua formação inicial, como também, muitas vezes, com o exercício da educação continuada. No contexto de Coordenadorias de Comunicação Social em ambientes universitários torna-se de suma importância a requisição pela qualidade da indexação para a recuperação relevante e ágil da informação.

Diante do exposto, propomos realizar um estudo comparativo entre as práticas da indexação manual e automática, visando identificar o tipo e a metodologia mais apropriados para a representação de *clippings* em sistemas automatizados de coordenadorias de comunicação social em ambientes universitários.

Essa pesquisa justifica-se pela importância do avanço tecnológico ocorrido no tratamento temático da informação e suas contribuições relevantes (Anderson, 1994; Lancaster, 2004; Gil Leiva, 1999, 2008) que ocasionaram na agilização do processo de indexação, com vistas ao acesso, recuperação e uso da informação jornalística disponibilizada por *clippings* eletrônicos em diversos ambientes e sistemas automatizados.

Uma vez que esta pesquisa encontra-se articulada a atividade de extensão “Organização e gestão da inovação em processos e produtos informacionais para a comunicação na

UFSCar”, coordenada pela Profa. Dra. Vera Regina Casari Boccato, ela revela uma motivação e um interesse pessoal desta pesquisadora pelo fato de ser bolsista do Programa de Educação Tutorial em Biblioteconomia e Ciência da Informação – Grupo PET BCI – UFSCar e estar atuando como estagiária voluntária na CCS-UFSCar mediante a realização da indexação de clippings de informações jornalísticas disponibilizadas para recuperação por assuntos pelos usuários internos e externos da UFSCar, locais e remotos.

Objetivo

Tem-se como objetivo contribuir com o processo de disseminação da informação jornalística no contexto universitário, a partir da realização de uma indexação consistente e ágil de clippings em sistemas automatizados, possibilitando a credibilidade e a visibilidade da imagem institucional.

Metodologia

Esta pesquisa qualitativa tem caráter descritivo-aplicado a partir de pesquisa bibliográfica na construção dos postulados teórico-metodológicos para a análise in loco do objeto em estudo.

A pesquisa iniciar-se-á pela elaboração de levantamentos bibliográficos em fontes de informação impressas e eletrônicas, exemplificadas pelas bases de dados, catálogos online, portais de revistas eletrônicas, entre outras na busca por assunto representativos dos temas Interdisciplinaridade entre Ciência da Informação e Ciências da Comunicação, hemeroteca, *clipping*, tratamento temático da informação, indexação, indexação semi-automática, indexação automática, linguagem natural, vocabulários controlados recuperação da informação, análise documentária, privilegiando a conceituação e as características de cada temática arrolada.

Os resultados alcançados pela realização dos levantamentos bibliográficos serão analisados e, posteriormente, sistematizados, configurando-se em parâmetros de análise para a realização de um estudo comparativo entre as práticas da indexação manual e automática, visando identificar o tipo e a metodologia mais apropriados para a indexação de *clippings* em sistemas automatizados de coordenadorias de comunicação social em ambientes universitários.

Além disso, a aquisição de referenciais teóricos e metodológicos contribuirão, também, na definição do software mais apropriado para a realização da indexação automática dos clippings previamente selecionados.

Resultados esperados

Espera-se que por meio deste estudo que se possa verificar se qual o método de indexação mais adequado para a representação e recuperação de *clippings* em sistemas automatizados de coordenadorias de comunicação social. Tendo em vista o contexto sócio-histórico-organizacional que permeia as universidades públicas na promoção do ensino, da pesquisa e da extensão com qualidade na formação de discentes com competência e habilidades necessárias para o exercício profissional, acredita-se que os resultados obtidos nos processos de indexação manual, automática e semi-automática possam ser diretivos nessa escolha para o tratamento temático da informação jornalística de diversos contextos de produção e uso.

A pesquisa iniciar-se-á pela elaboração de levantamentos bibliográficos em fontes de informação impressas e eletrônicas, exemplificadas pelas bases de dados, catálogos online, portais de revistas eletrônicas, entre outras na busca por assunto representativos dos temas Interdisciplinaridade entre Ciência da Informação e Ciências da Comunicação, hemeroteca, *clipping*, tratamento temático da informação, indexação, indexação semi-automática, indexação automática, linguagem natural, vocabulários controlados recuperação da informação, análise de conteúdo, privilegiando a conceituação e as características de cada temática arrolada.

Os resultados alcançados pela realização dos levantamentos bibliográficos serão analisados e, posteriormente, sistematizados, configurando-se em parâmetros de análise para a realização de um estudo comparativo entre as práticas da indexação manual e automática, visando identificar o tipo e a metodologia mais apropriados para a representação de *clippings* em sistemas automatizados de coordenadorias de comunicação social em ambientes universitários.

A avaliação desses diferentes tipos de indexação realizados *in loco*, isto é, a partir da indexação manual, automática e semi-automática possibilitará a identificação do processo

mais adequado para a representação temática de *clippings*, bem como a realização, caso necessário, de recomendações sobre o aprimoramento e a modelagem do tratamento temático de *clippings* para a recuperação da informação jornalística no Sistema de Apoio à Comunicação Integrada por usuários internos e externos da UFSCar.

Notas

A norma brasileira NBR 12676 (1992) da Associação Brasileira de Normas Técnicas teve sua elaboração baseada na norma ISO 5693 (1985) da International Organization Standardization, que por sua vez, esta última teve sua construção subsidiada pelos Princípios do UNISIST (1981).

Referências Bibliográficas

- Anderson, J. D. 1994. Standards for indexing: revising the American National Standard Guidelines Z39.4. *Journal of the American Society for Information Science*, v. 45, n.º 8, p. 628-636.
- Associação Brasileira de Normas Técnicas. 1992. NBR 12676: métodos para
- análise de documentos: determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro: ABNT, p. 4
- Boccato, V. R. C. (Coord.). 2011. Organização e gestão da inovação em processos e produtos informacionais para a comunicação na UFSCar. São Carlos, 2011. Relatório de atividade de extensão apresentado e aprovado à Pró-reitoria de extensão da Universidade Federal de São Carlos (PROEX-UFSCar).
- Chaumier, J. 1988. Indexação: conceitos, etapas e instrumentos. *Revista Brasileira de Biblioteconomia e Documentação*, v. 21, nº 1/2 (jan./jun.), p. 63-75.
- Gil Leiva, I. 1999. La automatización de la indización de documentos. Gijón: Trea.
- Gil Leiva, I. 2008. Manual de indización: teoría y práctica. Gijón: Trea, 2008.
- Lancaster, F. W. 1993. Indexação e resumos: teoria e prática. Brasília: Briquet de
- Lemos/Livros.
- Lancaster, F. W. 2004. Indexação e resumos: teoria e prática. 2ª ed. Brasília: Briquet de

- Lemos/Livros.
- Le Coadic, Yves-François. 1996. A ciência da informação. Brasília: Briquet de Lemos/Livros.
- Medeiros, R.; Melo, E. S. F.; Nascimento, M. S. 2008. Hemeroteca digital temática: socialização da informação em cinema. In: XV Seminário Nacional de Bibliotecas Universitárias (SNBU). Empreendedorismo e Inovação: desafios da biblioteca universitária. São Paulo, UNICAMP, UNESP, USP. São Paulo: CRUESP Bibliotecas
- Rabaça, C. A.; Barbosa, Gustavo Guimarães. 1998. Dicionário de comunicação. 3a ed. São Paulo: Ática.
- Rubi, M. P. 2008. Política de indexação para construção de catálogos coletivos em bibliotecas universitárias. Marília: Faculdade de Filosofia e Ciências. Universidade Estadual Paulista, 169 f. Tese de doutorado em ciência da informação.
- UNISIST (Sistema de Información de Ciencia y Tecnología de las Naciones Unidas). 1981. Princípios de indexação. Separata de Revista da Escola de Biblioteconomia da UFMG, v. 10, n. 1 (mar.), p. 83-94.