

# PROPOSTA DE UM MANUAL DE INDEXAÇÃO PARA BIBLIOTECAS UNIVERSITÁRIAS

## PROPOSAL OF A MANUAL INDEX FOR UNIVERSITY LIBRARIES

Rafaela Carolina da Silva<sup>1</sup>  
Jean Fernandes Brito<sup>2</sup>

### RESUMO

**Introdução:** A eficácia da recuperação depende uma boa indexação e a precisão das linguagens documentárias, que são determinadas neste processo. A indexação é uma atividade que visa não só determinar o assunto de um documento, mas representá-lo por meio de uma linguagem documentária. **Objetivos:** Assim objetiva-se, neste artigo, propor um guia de indexação para bibliotecas universitárias, a fim de contribuir para com o desempenho das atividades de organização e disseminação da informação já implantados nessas localidades. **Metodologia:** O estudo caracteriza-se por uma revisão de literatura e pelo estudo do Vocabulário Controlado utilizado pela Universidade de São Paulo (USP), dado à grande replicação desse vocabulário por outras instituições universitárias. **Resultados:** Temos como resultado dessa pesquisa um manual com recomendações de política de indexação para bibliotecas universitárias. **Conclusões:** Para um efetivo desempenho no processo de indexação as medidas pré-estabelecidas deverão ser tomadas para a realização de um entrosamento entre o sistema de informação e a relação com os usuários. A eficácia da recuperação depende de uma boa indexação e da precisão das linguagens documentárias, que são determinantes nesse processo.

**Descritores:** Indexação. Manual de Indexação. Bibliotecas. Bibliotecas Universitárias.

---

<sup>1</sup> Doutoranda em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP). E-mail: rafaelacarolinasilva@gmail.com

<sup>2</sup> Mestrando em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC). E-mail: jjeanfernandes@gmail.com

## 1 INTRODUÇÃO

A atividade de indexar é vista como o ato de descrever e identificar um documento de acordo com o seu assunto, pois através de um processo de análise, conceitos são extraídos do documento e então traduzidos para os termos de instrumentos de indexação, tais como: tesouros, listas de cabeçalhos de assunto, esquemas de classificação etc.

A partir da escolha do instrumento de indexação, os conceitos antes extraídos do documento, são registrados como “dados de informação” e, conseqüentemente, organizados de forma a facilitar o acesso na recuperação da informação. As técnicas de indexação podem tanto ser usadas visando organizar os conceitos em instrumentos de recuperação da informação quanto por analogia, cujo cerne é analisar e organizar as perguntas em conceitos representados como descritores ou combinação de descritores, símbolos de classificação, etc.

Substancialmente, a indexação compreende dois estágios: estabelecimento dos conceitos tratados num documento, ou seja, o assunto; e tradução dos conceitos nos termos da linguagem de indexação. O primeiro estágio, condizente com a determinação do assunto, evidencia que o estabelecimento do assunto de um documento pode ser subdividido em três etapas: compreensão do conteúdo do documento como um todo, os objetivos dos autores, dentre outros; identificação dos conceitos que representam este conteúdo, objetivos; seleção dos conceitos válidos para a recuperação.

Porém, essas três etapas se sobrepõem em: compreensão do documento (sua compreensão total dependerá até certo ponto da forma do documento, podendo ser ele gráfico ou não-gráfico), identificação de conceitos (após o exame do documento feito pelo indexador, cabe a ele adotar uma abordagem lógica, selecionando os conceitos que melhor representarão o assunto do documento) e seleção de conceitos (os conceitos identificados não necessitam ser todos registrados como termos indexadores, apenas os que forem de encontro com o objetivo para o qual as informações são indexadas (a exaustividade e especificidade estão contidas neste processo). O segundo estágio, referente à representação de conceitos por termos de uma linguagem

de indexação, deixa claro que, para assegurar a organização de conceitos de uma forma útil e acessível, é necessário o conhecimento profundo dos instrumentos de indexação, que normalmente fazem parte de duas categorias: verbais (representados por tesouros, listas de cabeçalhos de assunto) e simbólicos (representados por símbolos de classificação).

Assim, para garantir o controle de qualidade da indexação, dois são os fatores preponderantes: as qualificações do indexador (devendo ele entender os termos encontrados bem como as regras e procedimentos do sistema de indexação em questão; ser imparcial, de modo a evitar elementos de julgamento subjetivos, que afetam o processo de indexação) e a qualidade dos sistemas de indexação. Logo, o produto do processo de indexação são descritores já traduzidos para a linguagem documental.

As universidades contribuem para a geração de novos conhecimentos por meio do ensino e da pesquisa científica e tecnológica. A comunicação desses novos conhecimentos é essencial para a manutenção de um ciclo de produção e de conhecimento. Na medida de em que novos conhecimentos viram registros, torna-se essencial garantir o acesso à informação contida neles.

Sendo assim, objetiva-se, neste artigo, propor um guia de indexação para bibliotecas universitárias, a fim de contribuir para com o desempenho das atividades de organização e disseminação da informação já implantados nessas localidades. O estudo caracteriza-se por uma revisão de literatura e pelo estudo do Vocabulário Controlado utilizado pela Universidade de São Paulo (USP), dado à grande replicação desse vocabulário por outras instituições universitárias.

Nesse contexto, para um efetivo desempenho, medidas pré-estabelecidas, deverão ser tomadas para a realização de um entrosamento entre o sistema de informação e relação com os usuários. A eficácia da recuperação depende uma boa indexação e a precisão das linguagens documentárias, que são determinadas neste processo. A indexação é uma atividade que visa não só determinar o assunto de um documento, mas representá-lo por meio de uma linguagem documentária. Assim, o processo envolve duas etapas: análise conceitual e uma análise de tradução.

Em um primeiro momento a indexação é realizada a análise do documento, visando a identificação do assunto e dos conceitos que elas apresentam. Em um segundo momento, os conceitos identificados são traduzidos para uma linguagem documentária, que poderá ser verbal ou notacional. Em uma biblioteca, a atividade de indexação pode ser resumida como uma representação do documento, cujo objetivo final é representar o assunto do documento, com vistas à recuperação da informação nela contida. Nesse contexto, todos os esforços devem ser direcionados para uma boa indexação e para que isto ocorra é necessário que muitas singularidades sejam observadas, como o indexador e as técnicas de indexação, os usuários do catálogo, o acesso e a própria instituição, como no caso a biblioteca universitária. Todos estes aspectos são importantes para a utilização nos elementos do processo de indexação.

Fujita e Rubi (2006) afirmam que para que haja uma boa política de indexação, o processo depende um todo dentro da unidade de informação, obtendo uma visibilidade e encontrando a necessidade que existe dentro do processo de indexação. Logo, as bibliotecas devem ter iniciativas e verificar a importância de se aplicar uma política de indexação.

## **2 LINGUAGEM DOCUMENTÁRIA EM NÍVEIS DE ESPECIFICIDADE E EXAUSTIVIDADE: ESCOLHA DA POLÍTICA DA INDEXAÇÃO**

Lancaster (1968) destaca que a linguagem de indexação afeta o desempenho de um sistema de recuperação de informações em dois pontos: na estratégia de busca, estabelecendo a precisão com que o indexador descreve os interesses do usuário e na indexação, estabelecendo a cobertura de assuntos com que o indexador pode descrever o assunto dos documentos. Segundo Carneiro (1985), o indexador é responsável pelas decisões com relação à linguagem de indexação, que terão importantes efeitos sobre o desempenho do sistema.

A Linguagem Natural, nesse sentido, é a linguagem do autor, uma linguagem não controlada, que os seres humanos se utilizam no seu dia a dia. Assim, por não necessitar de um material padronizado em sua indexação, o

processo acaba sendo mais rápido do que aquele feito sob linguagens controladas; porém, a precisão da indexação fica comprometida, na medida em que a variedade de vocabulário entre contextos diferentes pode ser dificultar a recuperação da informação pelos usuários de uma determinada unidade de informação. A Linguagem Documentária, ao contrário, constitui-se de um conjunto de termos padronizados que representam um termo da linguagem natural, com vistas a uma maior padronização do sistema de indexação. São as linguagens de indexação, caracterizadas por fornecer uma indexação mais precisa do item analisado.

Tais linguagens podem ser pré-estabelecidas pelo profissional da informação, ou seja, as combinações de busca são feitas pelo indexador, ou pós-coordenadas, onde o usuário tem a possibilidade de, por meio de combinações entre termos, realizar sua busca. Um sistema de indexação ideal compartilharia do uso das duas estratégias, de modo a facilitar a busca e recuperação informacional de um usuário, portanto, neste manual, serão adotadas as duas medidas de forma combinada.

A informação pode ser entendida de maneira social quando transformada em conhecimento (BARITÉ, 2001). Uma vez registrado, esse conhecimento necessita ser organizado de modo a agrupar coisas iguais separar as diferentes. Sendo assim, para que ocorra a recuperação dessa informação, objetivo fim do processamento de dados de uma unidade de informação, é necessário que o tratamento dessas temáticas se ocupem de algumas técnicas e ferramentas. Nesse sentido, a representação do conhecimento deve ser exaustiva para abrigar todas as áreas do conhecimento que determinado documento abriga, mas também específica, para representar exatamente a temática desse documento. Dessa maneira, “A exaustividade relaciona-se à capacidade do sistema de se indexar o documento em profundidade, ou seja, além do assunto principal, são indexados também os assuntos secundários.” (NAVES, 2004, p. 9). Já “A especificidade é a capacidade do sistema de ser preciso ao especificar um assunto de um documento, a fim de abranger somente o conteúdo temático principal do documento.” (FUJITA, 2012, não paginado).

Logo, a Indexação exaustiva emprega termos para abranger o conteúdo temático do documento de modo bastante completo, sendo maior quando se aumenta o número de palavras (profundidade) presentes na representação. Visa, portanto, abranger o conteúdo temático do documento de modo bastante completo, focando o usuário e a unidade de informação. Destarte, a exaustividade abarca a indexação de todos os assuntos contemplados no documento, o que demanda alta revocação (quantidade de documentos relevantes recuperados) do sistema. Exemplo: se o usuário fizer uma busca no catálogo com o assunto “História do Brasil”, o mesmo recuperaria todos os assuntos sobre esse tema.

A especificidade refere-se à indexação somente do assunto específico tratado no documento, delimitando a busca, ou recuperando termos mais específicos sobre o assunto. Exemplo: se o usuário fizer uma busca no catálogo com o assunto “História do Brasil”, o mesmo poderia abranger somente o tema “Brasil Colônia”. Sendo assim, é necessário que a indexação seja o mais consistente e criteriosa possível, para que não se perca algum tipo de informação no momento da busca. Considera-se, portanto, a cobertura de assuntos e o uso das linguagens controladas como a identificação das áreas de assunto a serem indexadas e, posteriormente, representadas, demandando um tratamento mais ou menos aprofundado dos assuntos a serem tratados.

Quanto à sua função, as linguagens documentárias são vistas de diversas maneiras. São construídas para indexação, armazenamento e recuperação da informação e correspondem a sistemas de símbolos, destinadas a traduzir os conteúdos dos documentos (CINTRA, 1994). A linguagem documental ainda pode ser pré ou pós-coordenada, e o que irá determinar a adoção de uma ou outra é a política de indexação do sistema em questão.

Toda a linguagem documentária tem um controle, ou seja, normas e diretrizes para serem adotadas. Os vocabulários controlados apresentam, geralmente, três elementos básicos: uma lista de termos (descritos e não-descritores) uma rede de relações entre os descritores (hierárquicos e não-hierárquicos e regras de uso.

Os descritores são os termos autorizados para o uso de indexação. Ao atribuir um descritor a um documento, assume-se que este descritor irá representar de forma relevante o conteúdo do documento. Exemplo: **ARQUITETURA DA INFORMAÇÃO** - Descritor.

Os qualificadores são descritores usados sempre em associação com o descritor principal, de modo a especificar um determinado assunto. Exemplo: **ARQUITETURA DA INFORMAÇÃO (USABILIDADE) - Qualificador - Descritor.**

A remissivas não são descritores, mas remetem ao descritor. As remissivas são termos importantes equivalentes aos descritores. Exemplo: **ARQUITETURA DA INFORMAÇÃO; VER TAMBÉM ARQUITETURA DIGITAL.**

Os elos falsos são não-descritores que agrupam termos mais específicos. Exemplos: **ARQUITETURA DA INFORMAÇÃO; ARQUITETURA DA INFORMAÇÃO DIGITAL; ARQUITETURA DA WEB; ARQUITETURA INFORMACIONAL; ARQUITETURA DIGITAL.**

Um dos vocabulários controlados mais utilizados e conhecidos é o da USP. Integram as Bibliotecas do Sistema Integrado de Bibliotecas da Universidade de São Paulo (SIBIUSP), no Catálogo Coletivo da USP (DEDALUS). Ele abrange várias áreas do conhecimento inerentes às atividades de ensino, pesquisa e extensão. As especificidades do sistema são feitas de acordo com a combinação de descritores e qualificadores. O vocabulário, ao mesmo tempo em que possui alta exaustividade, também é bem específico, o que delimita alto nível de exaustividade e precisão do sistema.

O conteúdo dos trabalhos é expresso pela combinação ou coordenação de descritores, qualificadores, remissivas e elos que podem ser ou não falsos e, para facilitar a busca, podem ser usados os operadores booleanos *AND*, *OR*, *NOT* e o processo de truncagem. O manual deixa bem explícito a questão da possibilidade de combinação de termos pelo usuário, e é, por tais razões até agora mencionadas, que esta política tomou como base o Vocabulário Controlado da USP para o gerenciamento do vocabulário a ser utilizado.

A escolha da linguagem apresenta-se de acordo com o sistema de informação e com os seus usuários, podendo ser livre (natural), controlada, pré ou pós-coordenada. Nesta política, optou-se pela linguagem controlada, combinando-as pré e pós coordenadamente. Quanto ao significado de revocação, a mesma destina-se a recuperar documentos úteis, enquanto a precisão evita a recuperação de documentos considerados desnecessários. Portanto, a revocação, como exemplifica a imagem abaixo, refere-se à exaustividade, ao número total de documentos recuperados na busca; já a precisão está ligada à especificidade dos sistemas, ou seja, ao número total de documentos recuperados em relação à sua relevância.

Assim, “Quanto mais exaustivamente um sistema indexa seus documentos, maior será a revocação (número de documentos recuperados) na busca e, inversamente proporcional, a precisão será menor.” (FUJITA, 2012, grifo nosso). Os fatores de especificidade e exaustividade têm influência em todo o processo de recuperação da informação, desse modo, a decisão da política de exaustividade não deve assumir a forma de limites absolutos à quantidade de termos a serem atribuídos, mas “[...] ao invés disso, essa política poderia sugerir uma faixa de termos.” (SOARES *et al.*, 2014, p. [7]). Para tanto, seria preciso classificar os itens abrangidos e os não abrangidos, segundo certas características, como tipo de documento, língua, lugar de publicação e título. (LANCASTER, 2004, p. 139). Logo, as estratégias de busca podem ser determinadas diretamente pelo usuário, sugeridas por um profissional da informação, ou acompanhadas pela sugestão do usuário e o apoio do bibliotecário.

Nessa perspectiva, Fujita (2012) disserta que o indexador possui três formas de julgamento quanto aos níveis de especificidade e exaustividade do sistema. São eles: 1) Opção de julgamento 1 com termos específicos e genéricos (ESPECIFICIDADE E EXAUSTIVIDADE) = recuperação com alta precisão e também alta revocação; 2) Opção de julgamento 2 somente para termos específicos (ESPECIFICIDADE) = recuperação com alta precisão e baixa revocação; 3) Opção de julgamento 3 somente para termos genéricos (EXAUSTIVIDADE) = recuperação com baixa precisão e alta revocação. (FUJITA, 2012).

### **3 RESULTADOS: DESENVOLVIMENTO DO MANUAL**

Como este manual trata de uma biblioteca universitária, optou-se pela utilização da opção 1 de indexação, já que o foco está na abrangência dos cursos oferecidos pela unidade em questão, que podem demandar alta ou baixa revocação. “O tempo de resposta do sistema demarca o retorno de um sistema no momento do pedido das informações e recuperação, de acordo com as exigências de revocação e precisão do um sistema” (RODRIGUES, 2014, p. 19) e sua forma de saída é determinada pela satisfação das necessidades dos usuários. Portanto, a combinação da exaustividade com a especificidade é, como dito anteriormente, importante para a recuperação da informação, na medida em que cada unidade de informação insere-se em um contexto diferenciado, demandando diferentes níveis de cobertura do conteúdo de um documento. A revocação e a precisão, nesse contexto, relacionam-se aos assuntos de um documento contemplados no processo de indexação e que, posteriormente, serão recuperados por meio de estratégias de busca.

#### **3.1 Tempo de resposta e formas de saída do sistema**

O tempo de resposta de um sistema informacional, de forma sucinta e de fácil compreensão, compreende o tempo gasto entre a requisição feita pelo usuário ao sistema e o momento em que a resposta é exibida ao mesmo, ou seja, o tempo em que a informação leva para ser processada. Através dos mecanismos de busca, tal resposta pode ser de maior especificidade (precisão do sistema ao especificar o assunto de um documento, abordando apenas o conteúdo temático principal do mesmo, prezando o aspecto qualitativo) ou de maior exaustividade (capacidade do sistema de indexar o documento em profundidade, além do assunto principal, dando ênfase no aspecto quantitativo). Porém, o usuário pode optar por uma busca onde ambas as opções são utilizadas (especificidade e exaustividade), dependendo de sua necessidade informacional.

Já as formas de saída do sistema compreendem o formato de apresentação dos resultados de busca aos usuários, e deve se levar em conta qual a preferência do usuário quanto à apresentação dos mesmos, uma vez que podem ser dispostos referentes à separação dos itens bibliográficos de

acordo com sua tipologia, de ordenação por data (do mais recente para o mais antigo e vice-versa), por ordem alfabética de autor ou de título. São exemplos de formatos de apresentação: longo (referência e resumo), detalhado (referência, resumo e assunto, título (somente o título do documento) e referência (somente a referência do documento). Nesse contexto, a recuperação de documentos é realizada com a apresentação dos registros recuperados (na tela do computador ou impresso), com formatos de saída e com níveis de exaustividade e especificidades previamente selecionados pelo usuário. Esses, combinados com o tempo de saída do sistema, envolvem o processo de recuperação da informação.

### **3.2 Estratégias de busca, recuperabilidade e previsibilidade do sistema**

No planejamento da estratégia de busca, algumas ações ou táticas precisam ser criteriosamente relacionadas, principalmente com a seleção de termos ou conceitos e com as restrições e ou limites determinados pelo solicitante entre outras questões. Oldroyd e Citroen (1977) identificaram três grandes etapas para a decisão no processo de planejamento da estratégia de busca: decisão no processo de planejamento da estratégia de busca: decisão sobre qual a 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.

Ajudar o pesquisador a definir clara e precisamente o problema é muito importante para formulação de uma efetiva estratégia de busca, sendo raras as ocasiões em que as mesmas demandem total conhecimento do usuário. Logo, a maneira como o usuário busca suas informações afeta as atividades de busca em linha ou nos índices impressos.

Para a preparação da estratégia de busca, o usuário deve fornecer preferencialmente os seguintes dados: escrever um título sucinto; escrever uma pequena definição do problema, listar os termos que não são desejados. A partir desses dados, o intermediário pode executar a busca de informação de uma forma mais simples.

A seguir, são apresentadas orientações organizadas em etapas, que podem ajudar o intermediário da busca e o pesquisador a definir o problema,

contribuindo, portanto, para acentuar a qualidade da busca. 1ª Etapa- Discussão do tópico geral da pesquisa: Essa etapa tem como objetivo definir o que o usuário procura através de uma entrevista; 2ª Etapa- Conhecimentos básicos sobre os instrumentos de busca: Se o usuário tiver conhecimento sobre bases de dados, esse deve sozinho desenvolver a pesquisa para definir as palavras chaves e termos para a busca; o bibliotecário deve auxiliar, mas não exercer sua opinião própria; 3ª Etapa- Formulação “provisória” da estratégia de busca: Consiste em uma busca inicial provisória, onde os termos similares são agrupados, com o objetivo de uma tentativa da recuperação da informação; 4ª Etapa - Compreensão da lógica dos conjuntos de termos: O profissional da informação deve servir de auxílio para que o usuário entenda a lógica da seleção dos termos mais qualificados para recuperação da informação; 5ª Etapa - Interdisciplinaridade: O bibliotecário deve expandir a visão do usuário, para que esse possa aumentar a possibilidade de recuperar a informação; 6ª Etapa- Eliminação de termos indesejados: O usuário examinará a necessidade de exclusão de termos afins com o tema pesquisado; 7ª Etapa - Especificação dos parâmetros relevantes para a execução da busca: Determinação dos limites da busca, como gastos, bases de dados possíveis para a pesquisa etc.

Todas essas etapas podem ser categorizadas como pré-requisitos indispensáveis para o planejamento das estratégias de busca, visando à recuperação das informações, em suas variáveis formas de saída, dentro do sistema, pelo usuário (previsibilidade). As estratégias de busca pelo usuário dependem de processos realizados nas atividades de indexação, podendo utilizar operadores booleanos para a sua busca (*AND*, *OR*, *NOT*), e as proximidades e truncamentos, (+, -, “”, (), []). Portanto, a atividade de indexação é de suma importância em todas as etapas e processos, pois o usuário selecionará aquilo que é pertinente em sua pesquisa. Sendo assim, torna-se possível reconhecer se um item é ou não relevante a partir das informações sobre esse documento contidas no sistema.

### 3.3 Critérios de avaliação

Nesta seção é apresentada a importância da avaliação do sistema, uma vez que é a partir dela que se obtém conhecimento da satisfação quanto às necessidades informacionais do usuário. Nesse sentido, em caso de repercussão negativa, a instituição pode identificar e corrigir as falhas, bem como ver meios de melhorar o sistema por meio de suas avaliações satisfatórias. Os critérios de avaliação também visam verificar se a política menciona a data da última atualização/revisão, qual é o período entre uma avaliação e outra e se existe uma comissão de avaliação para o desenvolvimento de tais atividades. Várias são as formas de se obter tais informações, dentre elas: questionários, caixas de sugestão, estudos de usuários, realizações de observações.

Segundo Lancaster (2004, p. 135) , há quatro medidas para avaliar a recuperação da informação: cobertura ( se a quantidade de documentos sobre um assunto publicado durante um período se encontram na base de dados), recuperabilidade (quantos documentos sobre um determinado assunto incluídos na base de dados são encontrados com o emprego de estratégias de busca “razoáveis”), previsibilidade (com que eficiência o usuário pode aferir quais itens serão ou não úteis ao utilizar informações da base de dados), e atualidade (se itens recém publicados são recuperáveis ou atrasos na indexação de resumos provocam uma situação em que os itens recuperados mostram resultados de pesquisa antigos ao invés de novos). Logo, “A indexação e a redação de resumos não são atividades que devam ser consideradas fins em si mesmas” (LANCASTER, 2004, p. 135).

Sendo assim, esta política preza pelos critérios estabelecidos acima, a fim de melhor disponibilizar as informações indexadas pela unidade, com períodos de avaliação de seis em seis meses, critérios de cobertura bem definidos e atualização semanal. Por conseguinte, a recuperação da informação pelo usuário se mostrará mais adequada às suas necessidades informacionais.

#### 4 CONSIDERAÇÕES FINAIS

A comunicação, para o ser humano, caracteriza-se um meio de sobrevivência, pois, ao se comunicar uns com os outros, os homens expõem suas necessidades informacionais e seus conhecimentos para fins determinados. A padronização da comunicação, de acordo com o contexto em questão, é essencial para que o emissor e o receptor que conheçam tal padronização consigam decodificar as informações contidas em tal processo de comunicação.

Contudo, para que o conhecimento seja durável, é preciso registrá-lo em suportes documentários. A partir do momento em que essa informação é registrada, tornam-se necessárias operações de coleta (seleção e aquisição de documentos); tratamento (processamento dos documentos coletados com relação ao suporte material ao conteúdo: representação descritiva dos aspectos físicos do documento (processos como a catalogação); representação temática de conteúdo, ou seja, uma condensação do assunto do documento (como a elaboração de resumos, indexação - criação de índices e a Classificação); e difusão da informação (para que esses registros sejam disseminados de forma precisa, as atividades de tratamento da informação devem ser realizadas por profissionais que tenham conhecimento teórico e metodológico do material e do conteúdo a ser trabalhado - os profissionais da Ciência da Informação - gerando produtos e serviços do sistema documentário, de acordo com a comunidade usuária: consultas, empréstimo e comutação de documentos).

Para que o processo de análise seja validado de maneira a contribuir com a representação da informação, é preciso que haja uma leitura sistemática, diferente da leitura do dia a dia: a leitura documentária. Esses procedimentos de tratamento devem ser precedidos de políticas advindas das unidades informacionais, para que haja a devida padronização das informações, de acordo com os padrões institucionais.

Cada biblioteca possui sua peculiaridade em relação ao acervo, a equipe e a própria organização o que interfere no processo da indexação. Portanto é importante salientar que uma política de indexação se faz necessária para

auxiliar os profissionais e a própria instituição evitando assim perda de tempo no fazer. (RODRIGUES, 2014, p. 17).

Percebe-se, assim, que os processos de tratamento da informação baseiam-se em linguagens documentárias advindas de padrões internacionais, nacionais ou locais, possuindo interfaces com a Linguística (Análise do Discurso, Linguística Textual e Terminologia), com a Semiótica (importância da significação e da tipologia textual no processo de análise, síntese e representação da informação), com a Diplomática e com a Lógica (rede lógico-semântica na organização da informação). Nesse contexto, as atividades realizadas na Análise Documentária compreendem a Indexação (criação de índices), a Classificação (análise, síntese e representação de conteúdo com base em linguagens documentárias advindas de padrões internacionais (como a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU), por exemplo e a elaboração de resumos.

Sendo assim, a área de tratamento da informação circunscrita à área de Tratamento da informação, a Análise Documentária possui um quadro conceitual com ligações teóricas em Linguística, Lógica e Semiótica, no que se refere à semântica e análise do texto, sendo representado pelas suas linguagens, partindo do pressuposto de significante e significado. No que se refere à abordagem do texto, alguns estudos confirmam conceitos da Análise do Discurso da Aplicação das Linguagem textual e da Terminologia (Estudo dos Termos).

A indexação, por ser um processo cujo objetivo final é a criação de índices, serve como meio de acesso entre os recursos informacionais de uma unidade de informação e seu respectivo público. Nesse sentido, não só a forma, mas o conteúdo do documento devem ser analisados. Em relação aos estudos sobre a indexação, a Classificação Facetada de Ranganathan possibilitou a maior especificidade do assunto, por meio das categorias: personalidade, matéria, energia, espaço e tempo. Baseando-se nessas facetas, o *Classification Research Group* ampliou essa quantidade de facetas para: tipo, estrutura, constituintes, propriedades, processos, operações, técnicas e generalidades, relacionando conceitos (que, no contexto da indexação, possuem um certo grau de relação) a partir de operadores relacionais.

Dessa forma, os indexadores devem examinar atentamente o texto em que estão trabalhando, para que seja possível selecionar os conceitos que melhor representem o documento de análise. Para uma maior padronização do trabalho, os conceitos selecionados serão traduzidos em linguagens controladas (interface dos bibliotecários durante o processo de indexação), servindo para o trabalho do indexador no processo de indexação, o que torna desnecessário o conhecimento dessas estratégias pelo usuário leigo.

O processo de indexação passa por: análise (compreensão, identificação e conceitos importantes do texto), seleção dos conceitos válidos para a representação no devido contexto e tradução desse conteúdo (representação de conceitos por meio de linguagens de indexação). Nesse contexto, a *World Information System for Science and Technology*, geradora da primeira norma para o processo de indexação, afirma que o indexador pode se deparar com dois tipos de documentos: os gráficos (monografias, periódicos) e os não gráficos. No Brasil, a Associação Brasileira de Normas Técnicas (ABNT) publicou a tradução dessa mesma norma, onde não há distinção dos tipos de documentos.

Sendo o assunto a informação relevante abordada no texto, é preciso ressaltar que a seleção dessa informação sofre a influência da política de indexação do sistema de informação ao qual se insere o indexador. A instituição decidirá se o tema extraído do documento será mais específico ou se considerará níveis mais genéricos.

Para a análise dinâmica do conteúdo de um documento, deve-se, devido à impraticabilidade de se ler um documento no todo, observar o título, a introdução, as primeiras frases de capítulos e parágrafos, as ilustrações, as tabelas, os diagramas e suas explicações, bem como a conclusão e as palavras sublinhadas. A atenção maior deve estar nas partes finais do documento, onde são comunicados os alcances obtidos.

Cabe destacar que, dependendo das condições de análise, se necessário, como na produção de índices internos ou alfabéticos-remissivos, a leitura do documento, em sua íntegra, é aconselhada. Recomenda-se fazer um questionamento do texto, a partir da relação entre as facetas conceituais válidas em um sistema e questões anteriormente preparadas ao processo de

busca, que sirvam como base para a seleção dos conceitos essenciais do documento. Logo, a identificação de conceitos vem do conhecimento da área de cobertura do documento, seguida da seleção dos mesmos e determinada pelo objetivo de representação do documento - seu contexto, excluindo processos isolados. Os conceitos selecionados passam por um processo de tradução da linguagem natural para a linguagem controlada - linguagens de indexação (descritores, termos gerais, termos específicos, entre outros).

Nesse contexto, pode-se considerar que o indexador faz dois processos de análise: a análise do documento e a tradução da linguagem do autor; e a análise dos questionamentos dos usuários, passando as linguagens deles para a linguagem do sistema. Tais processos presumem o controle de qualidade, predestinados pelo máximo possível de imparcialidade do indexador (o homem é um ser subjetivo, tornando-se quase que impossível uma análise totalmente objetiva da situação), pelo conhecimento desse profissional (com formação superior na área) em relação à análise de assunto e pelo seu relacionamento com o usuário.

A seleção do assunto e a prática profissional do indexador é condicionada pelas políticas de indexação da instituição, ou seja, pelo contexto em que o serviço está inserido. A qualidade do trabalho pode ser observada nos resultados obtidos pelos usuários no momento da recuperação da informação. A abordagem do conteúdo pode ser feita de forma simplista (assuntos como unidades absolutamente objetivas; léxico e números; aspectos textuais explícitos; indexação automática), de forma orientada para o usuário (aderência de processos interpretativos do ser humano à análise textual; destaque de aspectos textuais implícitos) ou de forma orientada pela demanda (onde o indexador pensa em modos de melhor mediar a informação registrada com o usuário de uma unidade de informação). Verifica-se que as três concepções possuem vantagens e desvantagens: a principal vantagem de adotar uma concepção simplista se refere ao barateamento de computadores e de softwares, contudo, essas abordagens permeiam o aspecto do menor esforço do profissional, que demanda menor valor comunicativo.

O leitor - indexador, por ter que seguir parâmetros e tempo delimitados pelas organizações, procura selecionar assuntos que estejam no âmbito de seu

conhecimento profissional, mesmo que tenha que abranger novas informações à sua lista classificatória. Já o leitor comum busca recuperar dados que acrescentem informações aos conhecimentos que já possuem, ou seja, informações complementares ao seu saber. Um assunto possui conteúdos intrínsecos, inerentes ao texto (*aboutness*), e os seus variáveis significados, que mudam de acordo com a situação de análise e leitura (*meanings*). Portanto, as classificações de assunto podem variar de acordo com a época, já que esse assunto pode estar sendo relevante em um dado momento, mas não em outro. Tais abordagens se sobrepõem, na medida em que dependem umas das outras para que a qualidade do processo seja alcançada.

As interpretações do processo de leitura são vistas, para alguns, como um jogo onde o leitor prediz o conteúdo do texto apenas por trechos. Para outros, a compreensão e a interpretação das palavras acontece pela junção das letras e do som. Dessa maneira, são diversos os autores que produziram modelos explicativos do ato de ler. Nesse sentido, as concepções de leitura e do modo como o homem lê decorrem da concepção linear, em que o processo de leitura pressupõe que o homem lê da esquerda para a direita, de cima para baixo, em um contexto onde só se observa a forma estrutural da linguagem (lê-se letra por letra, palavra por palavra, para se chegar a um resultado semântico: princípio do alfabeto). No entanto, percorrem a concepção de que a leitura é um processo comunicativo, no qual aspectos extrínsecos ao texto influenciam no ato de ler - é a relação autor-texto, ou seja, o contexto do leitor e o contexto da realização da tarefa de leitura relacionam-se com as restrições textuais. O sentido semântico do texto não se encontra somente na estrutura textual, mas também no contexto psicológico e contextual do indivíduo que lê.

O leitor racional busca um texto conciso (quantidade de caracteres adequada a um texto informativo), de qualidade (veracidade da informação contida), relacional (relevância do assunto) e escrito com clareza (modo de escrita). Para tal, utiliza tanto esquemas de leitura inconscientes, como estratégias de leitura conscientes. As estratégias conscientes, ou metacognitivas, possuem um objetivo, são passíveis de observação, ensináveis e flexíveis, adquiridas por um processo de aprendizagem e variam de pessoa para pessoa. Já os esquemas inconscientes, ou cognitivos, partem do princípio

da canocidade (ordem natural sintática e semântica) e do princípio da coerência (global: busca da relação entre a intenção do autor e a produção do texto; local: relação de sentido entre trechos específicos; e temática: uso da informação tópica para fazer sentido no texto), ocorrendo no processo fluído de leitura.

Na área da Documentação, (onde há profissionais que lidam com o tratamento documentário de conteúdo, com a finalidade de facilitar a recuperação desses documentos por meio de representações concisas, como índices e resumos - processo da Análise Documentária) a leitura deve destacar somente as partes relevantes do texto. O caminho é ter conhecimento de uma superestrutura textual, que envolve não só a leitura linear, mas também a capacidade de fazer inferências a partir de conhecimentos prévios. Portanto, tal profissional deve se ocupar de não perder tempo com assuntos ou parágrafos irrelevantes aos tópicos principais do texto.

Consideramos importante abordar as noções de leitura em Documentação apresentadas por Cintra (1987), nas quais são apontados vários fatores que concorrem para a qualidade de um texto: a manutenção do tema, a correção gramatical, a adequação lexical, a estrutura do texto. Desse modo, o leitor indexador deve buscar entender a estrutura textual, sendo que o apoio estrutural do texto permite ao leitor ser seletivo e alocar atenção a trechos importantes.

O autor da mensagem (emissor) se comunica com seu leitor (receptor) por meio do texto (canal de comunicação, determinado pela alfabetização ou não - princípio para se entender a língua em que se lê), que, por sua vez, é capaz de produzir novos conhecimentos a partir da negociação da significação da leitura (*feedback*) e vice - versa. Importante frisar que esse *feedback*, por ser uma interpretação, pode conter ruídos, por excesso de informação incluída ou silêncio, pela falta de informação disponibilizada (o que atrapalha a compreensão do texto). Assim, o indexador utiliza-se de esquemas (processo natural do ser humana quando lê) e de estratégias (habilidades as quais o leitor recorre para melhor compreender o que lê). O foco está na leitura como ato do trabalho do profissional da informação, que lê de forma a entender processos terminológicos, históricos ou literários da informação registrada em suportes

documentários - possuem objetivos pré-definidos, estratégias de análise, políticas de indexação e de demanda do usuário. É o processo de Análise Documentária, ou seja, a elaboração de classificação, resumos e índices, onde as três variáveis são indissociáveis no processo de compreensão.

Em relação ao texto, o conhecimento prévio sobre a estrutura da disponibilização das informações textuais ajuda na compreensão das partes principais do texto. O conhecimento do código linguístico determina se haverá ou não decodificação da mensagem. A nível de conteúdo, ou tema a que se trabalha, tem-se as ideias consideradas importantes pelo leitor e as ideias que foram concebidas como importantes desde o início - as ideias do autor. Cabe dizer que tais processos variam de acordo com tipo de estrutura textual - o artigo destaca o texto científico.

Um texto argumentativo se divide em categorias de argumentação (justificativa e conclusão), justificativa (marco e circunstância), circunstância (pontos de partida e fatos), pontos de partida (legitimidade e reforço). O texto científico gira em torno da problematização do tema, da hipótese levantada, da metodologia de pesquisa e dos resultados obtidos, que serão questionados na conclusão. Já o texto expositivo expõe os problemas que levaram ao desenvolvimento do texto, as coisas e as soluções disponíveis.

O serviço de indexação deve ser sempre realizado por profissionais formados na área da Ciência da Informação, que saibam se utilizar de linguagens pré e pós-coordenadas. A capacitação desses profissionais influencia na disponibilização de documentos em uma unidade de informação. Assim, a instituição tem um lugar importante na capacitação de profissionais, que, muitas vezes, desconhecem esses métodos.

No Brasil, o ensino da indexação está presente nos cursos de nível superior e em disciplinas de tratamento da informação, documentação, representação temática e controle bibliográfico. A leitura documentária, antes mais dispersa na área de Documentação geral, agora está ligada ao controle e representação da informação. No entanto, a presença das empresas na capacitação de profissionais ainda não é o bastante. Assim, o profissional que trabalha com o tratamento da informação para fins de representação necessita ter pensamento crítico, conhecimento técnico, estrutural e contextual sobre

como se comportar em relação aos processos de indexação. Esses processos nada mais são do que estratégias de leitura, nas quais o tratamento de assuntos é feito com ferramentas e técnicas apropriadas para representar tais informações: e é exatamente nesse momento que se ressalta a importância de uma política de indexação bem estruturada, ou seja, um auxílio aos indexadores em suas tomadas de decisões.

## REFERÊNCIAS

- BARITÉ, M. Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In: CARRARA, K. (Org.). **Educação, universidade e pesquisa**. Marília: Unesp-Marília-Publicações; São Paulo: FAPESP, 2001.
- CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 221-41, set. 1985.
- CINTRA, A. M. M. **Para entender as linguagens documentárias**. São Paulo: Polis/APB, 1994.
- CINTRA, A. M. M. Estratégias de leitura em Documentação. In: SMIT, J. (Coord.). **Análise documentária: a análise da síntese**. Brasília: IBICT, 1987.
- FUJITA, M. S. L. **Política de indexação de assuntos para bibliotecas universitárias**. 2012. Disponível em: <[http://www.snbu2012.com.br/minicursos-e-oficinas/pdf/Mariangela\\_Spotti\\_Lopes\\_Fujita.pdf](http://www.snbu2012.com.br/minicursos-e-oficinas/pdf/Mariangela_Spotti_Lopes_Fujita.pdf)>. Acesso em: 09 abr. 2017.
- FUJITA, M. S. L.; RUBI, M. P. O ensino de procedimentos de política de indexação na perspectiva do conhecimento organizacional: uma proposta de programa para a educação à distância do bibliotecário. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 1, p. 48-66, jan./abr. 2006.
- LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.
- LANCASTER, F. W. **Information retrieval systems: characteristics, testing, and evaluation**. New Yourk, J. Willey, 1968.
- NAVES, M. M. L. Curso de Indexação: Princípios e técnicas de indexação, com vistas à recuperação da informação. Belo Horizonte: Biblioteca da Universidade Federal de Minas Gerais, 2004. 22 p.
- OLDROYD, B. K; CITROEN, C. L. Study of strategies used in online searching. **Online Review**, v. 1, n. 4, p. 295-310, 1977.

RODRIGUES, T. A. **Política de indexação em bibliotecas destinadas às necessidades dos usuários deficientes visuais**. 2014. 98 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

SOARES, K. de M. *et al.* **A prática da indexação**: Capítulo 3. Universidade Estadual da Paraíba. 2014. 14 slides. Disponível em:  
<<http://www.ebah.com.br/content/ABAAAgSPwAJ/pratica-indexacao-slide>>. Acesso em: 09 abr. 2014.

## PROPOSAL OF A MANUAL INDEX FOR UNIVERSITY LIBRARIES

### ABSTRACT

**Introduction:** The effectiveness of Information retrieval depends on coherent indexing and the accuracy of documentary languages, which are determined in this process. Indexing is an activity that aims not only to determine the subject of a document, but also to represent it through a documentary language. **Objective:** From this perspective, the article aims to propose an indexing guide for university libraries, in order to contribute to the performance of the information organization and dissemination activities already in place. **Methodology:** The study is characterized by a review of the literature and by the study of Controlled Vocabulary used by the University of Sao Paulo, given the replication of this vocabulary by other university institutions. **Results:** We have as a result of this research a manual with index policy recommendations for university libraries. **Conclusions:** For an effective performance in the indexing process, the pre-established measures must be taken to establish a relationship between the information system and the relationship with users. The effectiveness of retrieval depends on good indexing and the accuracy of documentary languages, which are decisive in this process.

**Descriptors:** Indexing. Indexing Manual. Libraries. University Libraries.

## PROPUESTA DE UN MANUAL DE INDEXACIÓN PARA BIBLIOTECAS UNIVERSITARIAS

### RESUMEN

**Introducción:** La efectividad de la recuperación de información depende de la indexación coherente y la precisión de los lenguajes documentales, que se determinan en este proceso. La indexación es una actividad que tiene como objetivo no solo determinar el tema de un documento, sino también representarlo a través de un lenguaje documental. **Objetivo:** Desde esta perspectiva, el artículo tiene como objetivo proponer una guía de indexación para las bibliotecas universitarias, con el fin de contribuir al desempeño de la organización de la información y las actividades de difusión ya en marcha. **Metodología:** El estudio se caracteriza por una revisión de la literatura y por el estudio de Vocabulario Controlado utilizado por la Universidad de São Paulo, dada la replicación de este vocabulario por parte de otras instituciones universitarias. **Resultados:** Como resultado de esta investigación, tenemos un manual

con recomendaciones de política de índice para bibliotecas universitarias.  
**Conclusiones:** Para un desempeño efectivo en el proceso de indexación, se deben tomar las medidas preestablecidas para establecer una relación entre el sistema de información y la relación con los usuarios. La efectividad de la recuperación depende de una buena indexación y la precisión de los lenguajes documentales, que son decisivos en este proceso.

**Descriptores:** Indexación. Manual de indexación. Bibliotecas. Bibliotecas universitarias.

**Recibido:** 04.03.2018

**Aceitos:** 31.05.2018