

Indexação e resumos: Teoria e prática

I. Introdução

- O propósito principal da elaboração de índices e resumos é contruir representações de documentos publicados numa forma que se preste a sua inclusão em alguma base de dados.
- Indexação: Em primeiro lugar, o produtor da base de dados seleciona da população de documentos recém-publicados aqueles que têm certos critérios para a sua inclusão na base de dados. Entre esses critérios estão o assunto de que trata o documento, o tipo de documento, a língua em que está escrito, ou a sua origem. No caso de artigos periódicos, os critérios de seleção estarão centrados no periódico e não no artigo; ou seja, alguns periódicos serão incluídos e outros não.
- Os itens selecionados para inclusão na base de dados serão “descritos” de várias formas.
- Os processos de catalogação descritiva identificam autores, títulos, fontes e outros elementos bibliográficos;
- Os processos de indexação identificam o assunto de que trata o documento; e o resumo serve para sintetizar o conteúdo do item.
- Tesouro → Tipo de vocabulário controlado.

II. Princípios da Indexação

A indexação de assuntos e a redação de resumos são atividades intimamente relacionadas, dado que ambas implicam a preparação de uma representação do conteúdo temático dos documentos. O resumidor redige uma descrição narrativa ou síntese do documento, e o indexador descreve o seu conteúdo ao empregar um ou vários termos de indexação, comumente selecionados de algum tipo de vocabulário controlado (tesauros). Os termos atribuídos pelo indexador servem como pontos de acesso mediante os quais um item é localizado e recuperado.

Etapas da indexação de assuntos:

1. Análise Concetual

2. Tradução

Análise Concetual → implica decidir do que se trata um documento, ou seja, o seu assunto.

Uma indexação de assuntos eficiente implica que se tome uma decisão não apenas quanto ao que é tratado num documento, mas também porque que ele se reveste de provável interesse para determinado número de utilizadores. Assim, não há um conjunto “correto” de termos de indexação para documentos.

O indexador, deste modo, deve formular várias perguntas sobre um documento:

1. De que se trata?
2. Por que foi incorporado no acervo?
3. Qual dos seus aspetos serão de interesse para os utilizadores?

Quanto mais especializada a clientela de um centro de informação, maior a probabilidade de a indexação ser feita sob medida, ajustando-se com precisão aos interesses do grupo. Assim, a indexação deve ser moldada para se ajustar às necessidades de determinada clientela. Para certos tipos de materiais, a indexação orientada para o utilizador pode ser até mais importante do que o é no caso de artigos de periódicos, livros ou relatórios médicos. Há uma importante lição a tirar dos princípios de indexação orientada para o utilizador: é preciso que os indexadores saibam muito mais do que os princípios da indexação. Devem, em especial, estar inteiramente a par dos interesses da comunidade atendida e das necessidades de informação dos seus membros.

Tradução → segunda etapa da indexação de assuntos, envolve a conversão da análise concetual de um documento num determinado conjunto de termos de indexação. Assim, faz-se uma distinção entre indexação por extração (indexação derivada) e indexação por atribuição. Na indexação por extração, palavras ou expressões que realmente ocorrem no documento são

selecionadas para representar seu conteúdo temático. A indexação por atribuição de termos ao documento a partir de uma fonte que não é o próprio documento. Envolve a representação da análise conceitual mediante um termo ou termos extraídos de um vocabulário.

Vocabulários controlados

Um vocabulário controlado é essencialmente uma lista de termos autorizados. O indexador apenas pode atribuir a um documento termos que contem da lista adotada pela instituição para a qual trabalha. No entanto, o vocabulário controlado é mais do que uma mera lista. Inclui, em geral, uma forma de estrutura semântica. Essa estrutura destina-se, especialmente, a:

- Controlar sinónimos, optando por uma única forma padronizada, com remissivas de todas as outras;
- Diferenciar homógrafos. Por exemplo, Perú (país) é um termo bastante diferente de Perú (ave);
- Reunir ou ligar termos cujos significados apresentem uma relação mais estreita entre si. Dois tipos de relações são identificados explicitamente: as hierárquicas e as não hierárquicas (ou associativas)

São três os tipos principais de vocabulários controlados:

1. Esquemas de classificação bibliográfica (como a Classificação Decimal de Dewey)
2. Listas de cabeçalhos de assuntos
3. Tesauros

A tradicional lista de cabeçalhos de assuntos é similar ao tesouro por ser de base alfabética, mas difere dele porque incorpora uma estrutura hierárquica imperfeita e por não distinguir claramente as relações hierárquicas das associativas. Os três tipos de vocabulário controlam sinónimos, distinguem homógrafos e agrupam termos afins, mas empregam métodos um tanto diferentes para alcançar estes objetivos.

Indexação como classificação

Há uma distinção entre três expressões: *Indexação de assuntos*, *catalogação de assuntos* e *classificação*.

Catalogação de assuntos → atribuição de cabeçalhos de assuntos para representar o conteúdo total de itens bibliográficos inteiros. (livros, relatórios, periódicos, etc) no catálogo das bibliotecas.

Indexação de assuntos → representação do conteúdo temático de partes de itens bibliográficos, como é o caso do índice de final de livro.

Classificação → atribuição de números de classificação (extraídos de um esquema de classificação – por exemplo, o decimal de Dewey (CDD), o da Library of Congress (LC)) – a itens bibliográficos, especialmente com a finalidade de arrumá-los nas estantes das bibliotecas. Deste modo, permeia todas as atividades pertinentes ao armazenamento e recuperação da informação. No campo de armazenamento e recuperação da informação, a classificação de documentos refere-se à formação de classes de itens com base no conteúdo temático. Os vocabulários controlados são essencialmente listas dos rótulos com os quais se identificam e, por ventura, se organizam essas classes.

III. A prática da indexação

Para indexação, as partes a serem lidas atentamente são as que apresentam maior probabilidade de dizer o máximo sobre o conteúdo no menor tempo: título, resumo, sinopse e conclusões. Os títulos das secções e as legendas das ilustrações ou tabelas também merecem maior atenção. Certas partes de um documento são especialmente gratificantes para um indexador como por exemplo parágrafos de abertura (de capítulos ou secções) ou frases de abertura e encerramento de parágrafos, dado que são ricos em palavras indexáveis.

Uma função importante do indexador é contribuir para o aperfeiçoamento do vocabulário controlado, comunicando as suas deficiências aos responsáveis pela sua manutenção.