



**ABNT-Associação
Brasileira de
Normas Técnicas**

Sede:
Rio de Janeiro
Av. Treze de Maio, 13 - 28º andar
CEP 20003 - Caixa Postal 1680
Rio de Janeiro - RJ
Tel.: PABX (021) 210-3122
Telex: (021) 34333 ABNT - BR
Endereço telegráfico:
NORMATÉCNICA

Copyright © 1990,
ABNT - Associação Brasileira
de Normas Técnicas
Printed in Brazil/
Impresso no Brasil
Todos os direitos reservados

AGO./1992

NBR 12676

Métodos para análise de documentos - Determinação de seus assuntos e seleção de termos de indexação

Procedimento

Origem: Projeto 14:001.01-011/91
CB-14 - Comitê Brasileiro de Finanças, Bancos, Seguros, Comércio, Administração e Documentação
CE-14:001.01 - Comissão de Estudo de Documentação
NBR 12676 - Documentation - Method for examining documents, determining their subjects, and selecting indexing terms - Procedure
Descriptor: Documentation
Válida a partir de 29/10/92

Palavra-chave: Documentação

4 páginas

1 Objetivo

1.1 Esta Norma fixa as condições exigíveis para a prática normalizada do exame de documentos, da determinação de seus assuntos e da seleção de termos de indexação. Destina-se aos estágios preliminares da indexação, não tratando das práticas de qualquer tipo de sistema de indexação, pré ou pós-coordenado. É dirigida para sistemas de indexação nos quais os assuntos dos documentos são expressos de forma resumida, e os conceitos são registrados através dos termos de uma linguagem de indexação. Aplica-se especialmente a serviços de indexação independentes e a serviços de indexação em rede.

1.2 Esta Norma pode também orientar os resumidores durante a fase de preparação de resumos e auxiliar na tradução dos pedidos dos usuários para os termos de uma linguagem de indexação.

1.3 Esta Norma não se aplica a serviços que empregam técnicas de indexação automática; os termos que ocorrem nos textos são organizados em conjuntos ou classes segundo critérios estabelecidos por computador (p. ex.: frequência de ocorrência e/ou de co-ocorrência no texto).

2 Documentos complementares

Na aplicação desta Norma é necessário consultar:

ISO 2788 - Documentação - Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngües

ISO 5864 - Documentação - Diretrizes para o estabelecimento e desenvolvimento de tesouros multilíngües

3 Definições

Para os efeitos desta Norma são adotadas as definições de 3.1 a 3.6.

3.1 Documento

Qualquer unidade, impressa ou não, que seja passível de catalogação ou indexação.

Nota: Esta definição se refere não apenas a materiais escritos ou impressos em papel ou suas versões em microforma (p. ex.: livros, jornais, diagramas, mapas), mas também a suportes não-impressos (p. ex.: registros legíveis por máquina, filmes, gravações sonoras), objetos tridimensionais e *realia* usadas como espécimens.

3.2 Conceito

Qualquer unidade de pensamento. O conceito pode ter o seu conteúdo semântico reexpresso pela combinação de outros conceitos, que podem variar de uma língua ou de uma cultura para outra.

3.3 Assunto

Tema representado num documento por um conceito ou combinação de conceitos.

3.4 Indexação

Ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos e que constituem uma linguagem de indexação.

3.5 Termo de indexação

Representação de um conceito sob uma das seguintes formas:

- a) termo derivado da linguagem natural, de preferência um nome ou uma locução nominal;
- b) símbolo de classificação.

3.5.1 Termo preferido

Termo utilizado consistentemente na indexação para representar um conceito; algumas vezes conhecido como "descriptor".

3.5.2 Termo não-preferido

Sinônimo ou quase sinônimo do termo preferido (descriptor); também conhecido como "não-descriptor". Não é atribuído aos documentos mas utilizado como remissiva no índice, para instruir o usuário (Use ou Ver) na procura do termo preferido.

3.6 Índice de assuntos

Listagem alfabética ou sistemática de assuntos que indica a posição de cada assunto num documento ou numa coleção de documentos.

4 Condições específicas

4.1 Estágios da indexação

4.1.1 A indexação consiste basicamente nos três estágios seguintes que, na realidade, tendem a se sobrepor:

- a) exame do documento e estabelecimento do assunto de seu conteúdo;
- b) identificação dos conceitos presentes no assunto;
- c) tradução desses conceitos nos termos de uma linguagem de indexação.

4.1.2 Em cada um destes estágios deve-se contar com a ajuda de instrumentos de indexação, tais como: tesouros, códigos de classificação, cabeçalhos de assunto, etc.

4.2 Análise do documento

4.2.1 O modo pelo qual um documento pode ser examinado depende em grande parte de sua forma física. Podem-se distinguir dois casos: documentos impressos e documentos não-impressos.

4.2.1.1 Os documentos impressos representam o caso comum nas bibliotecas e centros de informação onde o acervo consiste principalmente em monografias, periódicos, relatórios, anais de congressos, etc.

4.2.1.1.1 A compreensão destes documentos depende de

uma leitura que garanta que nenhuma informação foi negligenciada. Para isso, deve-se considerar especialmente:

- a) título e subtítulo;
- b) resumo, se houver;
- c) sumário;
- d) introdução;
- e) ilustrações, diagramas, tabelas e seus títulos explicativos;
- f) palavras ou grupos de palavras em destaque (sublinhadas, impressas em tipo diferente, etc.);
- g) referências bibliográficas.

Nota: Não se recomenda indexar apenas por qualquer um destes elementos isoladamente, pois em muitos casos eles não representam fonte confiável do tipo de informação necessária ao indexador, exigindo uma leitura mais aprofundada do documento.

4.2.1.2 Os documentos não impressos, tais como os multimeios, realia, etc., pedem procedimentos diferentes. Nem sempre é possível, na prática, examiná-los integralmente (p.ex.: projetar um filme, etc.). A indexação, então, é geralmente feita a partir do título e/ou da sinopse. Entretanto, o indexador deve ter acesso direto ao documento se o título e/ou a sinopse lhe parecem inadequados ou imprecisos.

4.3 Identificação dos conceitos

4.3.1 Após examinar o documento, o indexador deve adotar uma abordagem sistemática para identificar aqueles conceitos que são os elementos essenciais na descrição do assunto. Os serviços de indexação devem elaborar listas dos aspectos que forem identificados como importantes na área coberta pelo índice. As perguntas seguintes ilustram os aspectos gerais que devem ser incluídos numa listagem desse tipo:

- a) qual o assunto de que trata o documento?
- b) como se define o assunto em termos de teorias, hipóteses, etc.?
- c) o assunto contém uma ação, uma operação, um processo?
- d) o documento trata do agente dessa ação, operação, processo, etc.?
- e) o documento se refere a métodos, técnicas e instrumentos especiais?
- f) esses aspectos foram considerados no contexto de um local ou ambiente especial?
- g) foram identificadas variáveis dependentes ou independentes?
- h) o assunto foi considerado sob um ponto de vista interdisciplinar? (p.ex.: um estudo sociológico da religião)

4.3.1.1 Na lista anterior estão indicados alguns exemplos dos aspectos sob os quais qualquer assunto pode ser analisado. Outras perguntas podem ser formuladas para disciplinas específicas.

4.3.2 Na seleção de conceitos, o principal critério deve ser sempre o valor de um conceito para a expressão e recuperação do assunto do documento. Ao fazer a escolha dos conceitos, o indexador deve ter em mente as consultas que podem ser feitas ao sistema de informação. Neste contexto, o indexador deve:

- a) escolher os conceitos que forem considerados os mais apropriados para uma determinada comunidade de usuários;
- b) adaptar tanto os instrumentos de indexação como os próprios procedimentos em função da retroalimentação obtida através dos pedidos de informação. Esta adaptação, entretanto, não deve alterar a estrutura ou a lógica da linguagem de indexação.

4.3.3 O indexador não precisa necessariamente representar com termos de indexação todos os conceitos identificados durante o exame do documento. A escolha dos conceitos que devem ser selecionados depende da finalidade para a qual são usados os termos de indexação. Neste caso, as características da indexação mais afetadas são: seu grau de exaustividade e a especificidade dos termos selecionados.

4.3.3.1 A exaustividade se refere ao número de conceitos (como os associados às perguntas de 4.3.1) representados pelos termos atribuídos a um documento pelo indexador.

4.3.3.1.1 Os limites da abrangência do assunto não devem ser interpretados de modo muito restrito. Com o crescimento das redes de informação, é preciso ter em mente que os termos de indexação, inicialmente criados para um grupo de usuários (p.ex.: cientistas ou técnicos), podem ser utilizados por outros grupos de usuários (p.ex.: economistas).

4.3.3.1.2 Não se deve estabelecer nenhum limite arbitrário ao número de termos ou descritores que possam ser atribuídos a um documento. Isso deveria ser determinado inteiramente pelo volume de informações contido no documento e relacionado com as necessidades esperadas dos usuários. A imposição deste tipo de limite pode levar a alguma perda de objetividade na indexação e à distorção da informação que poderia vir a ser de valor na recuperação. Se um serviço de indexação necessita de limitar o número de termos, a seleção dos conceitos deve ser orientada pelo julgamento do indexador relativamente ao papel que cada conceito desempenha na expressão do assunto global do documento.

4.3.3.2 A especificidade se refere ao grau de precisão com que um termo define determinado conceito no documento. Ocorre perda de especificidade quando um conceito é representado por um termo com significado mais genérico. Os conceitos devem ser identificados o mais especificamente possível. Em certas ocasiões, porém, preferem-se termos que definem conceitos mais genéricos, dependendo dos seguintes fatores:

- a) o grau com que o indexador considera que a maior especificidade pode afetar negativamente o desempenho do sistema de indexação. Um indexador pode decidir, por exemplo, que modelos muito específicos de equipamento podem ser representados por termos mais genéricos, como o nome do fabricante e talvez da família de modelos, especialmente nos casos em que esses conceitos ocorrem apenas em áreas marginais do assunto coberto pelo índice;
- b) o peso atribuído ao conceito: se o indexador considera que uma idéia não está completamente desenvolvida, é justificável fazer-se a indexação num nível mais genérico.

4.3.4 A identificação dos conceitos pode também ser afetada pelo tipo de documento que está sendo indexado. Por exemplo, a indexação resultante de textos de livros, artigos de periódicos, etc. provavelmente diferirá da de resumos ou sinopses.

4.4 Seleção de termos de indexação

4.4.1 Ao expressar os conceitos por termos de indexação, o indexador deve observar as seguintes práticas:

- a) usar os descritores cabíveis já existentes na linguagem de indexação utilizada;
- b) para os termos que representam novos conceitos, deve-se verificar sua precisão e aceitabilidade em instrumentos de referência, tais como:
 - dicionários e enciclopédias de autoridade reconhecida nas suas especialidades;
 - tesouros, especialmente os elaborados de acordo com as ISO 2788 ou ISO 5964;
 - tabelas de classificação.

4.4.1.1 Pode-se também consultar especialistas no assunto, principalmente os que tenham conhecimento de indexação ou documentação.

4.4.2 O indexador deve estar familiarizado com estes tipos de instrumentos bem como com suas regras, procedimentos de trabalho e limitações.

4.4.3 O indexador freqüentemente encontra conceitos que não estão presentes nos tesouros ou nas tabelas de classificação existentes. Dependendo do sistema em uso, estes conceitos podem ser trabalhados de vários modos, por exemplo:

- a) expressos por termos ou descritores que são imediatamente admitidos na linguagem de indexação;
- b) representados provisoriamente por termos mais genéricos, deixando os novos conceitos para inclusão ulterior na linguagem de indexação.

5 Controle de qualidade

5.1 A qualidade de indexação depende de fatores relativos a:

a) consistência na especificidade dos termos atribuídos a um documento e no nível de exaustividade atingido na indexação (ver 4.3.3);

b) qualificações do indexador (imparcialidade, conhecimento, etc.);

c) qualidade dos instrumentos de indexação.

5.1.1 A imparcialidade por parte do indexador é um fator necessário para se obter consistência na indexação. Julgamentos subjetivos na identificação de conceitos e a conseqüente escolha de termos de indexação afetam inevitavelmente o desempenho do sistema de indexação. A consistência é mais dificilmente obtida quando a equipe de indexadores é grande, ou quando a indexação é feita por equipes de indexadores que trabalham em locais diferentes, como ocorre em sistemas descentralizados. Nestes casos, recomenda-se uma etapa de verificação cujos resultados devem ser repassados aos indexadores.

5.1.2 O indexador deve ter um conhecimento adequado do assunto coberto pelos documentos que está indexando. Deve compreender os termos encontrados nos docu-

mentos, bem como as regras e procedimentos da linguagem de indexação específica que está utilizando.

5.1.3 A qualidade da indexação depende da hospitalidade da linguagem de indexação utilizada. Esta deve admitir livremente novos termos ou mudanças na terminologia, bem como atender a novas necessidades dos usuários. Uma política de atualização freqüente é considerada essencial.

5.2 Quando possível, a qualidade da indexação deve ser testada:

a) pela análise dos resultados da recuperação, por exemplo, através do cálculo da relação entre a revocação e a precisão, ou do cálculo da relação entre o número de documentos relevantes recuperados e o número de documentos pertinentes recuperados;

b) pelo contato direto com os usuários. Neste caso, eles podem determinar, por exemplo, se certos termos ou descritores são passíveis de produzir falsas combinações e conseqüentemente gerar recuperações irrelevantes.