

A INDEXAÇÃO

Astério Tavares Campos

Departamento de Biblioteconomia

Universidade de Brasília

70910 Brasília, DF

Discute o objeto da indexação e a sua finalidade. Mostra a dificuldade de se identificarem elementos verdadeiramente representativos no texto por falta de suporte teórico, e aponta elementos polêmicos para discussão.

1. INTRODUÇÃO

A indexação consiste, fundamentalmente, na captação do conteúdo informativo do documento e na tradução desse conteúdo numa linguagem que sirva de intermédio entre o usuário e o documento. Trata-se de operação delicada e, por vezes, complicada que não teve, até agora, grandes suportes teóricos. Fairthorne chega mesmo a dizer que o que se fez até agora, neste campo, é simplesmente vergonhoso. Não dispomos ainda de meios que possam ser considerados verdadeiramente científicos para efetuar o processo básico da indexação. Hutchins chega quase às mesmas conclusões e faz algumas indicações para uma possível solução do problema. Entretanto, dizer que o esforço por sumarizar um texto correspondente ao desenrolar em sentido inverso do enunciado básico de que partiu o autor corresponde à verdade, mas não parece dar qualquer ajuda de ordem prática.

2. A INDEXAÇÃO

A indexação, entendida como processo básico na recuperação da informação, consiste, fundamentalmente, na captação do conteúdo informativo do documento e na tradução do mesmo numa linguagem que deve servir de intermediário entre o usuário — com as respectivas exigências — e o documento*. Trata-se de

* Obviamente, ao falarmos aqui em recuperação da informação, entendemos a recuperação das referências bibliográficas que nos conduzem à informação contida no documento.

uma operação que, no dizer de Lancaster,¹ se realiza em duas etapas ou momentos:

1. determinação, através de uma análise conceitual, do conteúdo informativo do documento;
2. conversão dos termos dessa análise numa linguagem intermediária entre o indexador e o leitor.

Embora se trate, efetivamente, de duas operações distintas — que necessariamente se sucedem numa ordem cronológica — vale a pena acentuar o caráter integrativo das mesmas, de modo que se pode, ou talvez se deva admitir que se trata de uma única operação em dois momentos. Não há exagero em acentuar a unicidade destas duas operações, sobretudo tendo em vista que a teoria da indexação deveria não apenas indicar o modelo de elaboração de uma linguagem de informação, mas deveria também apresentar instrumentos para interpretação do documento.

Este segundo aspecto do problema — ou seja, a ajuda que a teoria da indexação deveria prestar na interpretação do conteúdo do documento — não tem sido suficientemente posto em relevo, mesmo porque as dificuldades que normalmente se encontram são maiores do que à primeira vista se poderia imaginar. A situação é, sob certos aspectos, desconcertante. Fairthorne, no seu estimulante livro *Towards Information Retrieval*,² expressa-se de maneira bastante incisiva. Diz ele que a classificação bibliográfica apresenta uma lista de rótulos a serem atribuídos aos textos, mas raramente diz qual o rótulo correto a ser atribuído a determinado texto. A teoria da classificação esforça-se por orientar nessa escolha, mas ainda se mostra muito modesta nesse trabalho. Conclui o autor por dizer que é algo demasiado vergonhoso para que possa ser mencionado. Esta observação poderia ter sido expressa em linguagem menos dura. Infelizmente reflete uma realidade bastante desagradável. Não dispomos ainda de meios que possam ser considerados verdadeiramente científicos para efetuar o processo básico da indexação. Pode-se dizer que esta vem sendo processada mais como uma arte, ou, mais corretamente, um artesanato, do que propriamente uma técnica que se deixe iluminar pelos postulados ou princípios de uma autêntica teoria.

Há que reconhecer que as dificuldades neste setor não são pequenas. Temos que investigar, em primeiro lugar, até que ponto o êxito desse trabalho depende do grau e extensão da cultura do indexador. Serão necessários estudos especializados por parte do indexador a fim de obter pleno resultado ao indexar áreas especiais do conhecimento humano? Um mínimo de conhecimentos da área em que se está indexando é obviamente indispensável. Bastará esse mínimo fornecido pelos estudos de preparação na graduação? Seja, porém, qual for o grau de conhecimento especializado a ser exigido do indexador, é certo que uma boa teoria da indexação poderá fornecer-lhe inestimável auxílio. Fairthorne, no trecho acima citado, reconhece esta possibilidade. O que se pode acrescentar é que, em alguns casos, essa teoria é quase inexistente, o que equivale dizer que algumas linguagens

de indexação foram construídas em bases quase intuitivas, sem grande preocupação teórica.

Mas o problema apresenta ainda outras dificuldades. Hutchins, no seu livro **Languages of indexing and classification**, analisa o problema com certo cuidado.³ Examina inicialmente algumas objeções ou dificuldades relacionadas com o mesmo processo de indexação. A primeira, diz ele, consiste na pretensão de que o mesmo documento deva, sempre e em qualquer lugar, receber a mesma indexação. Parte-se, obviamente, do presuposto falso de que o documento possua qualquer classificação antes de ser classificado. Há o esquecimento grave de que qualquer documento deve ser sempre classificado tendo em vista o possível leitor ou usuário e de que as necessidades destes são sempre variáveis. Numa palavra, fica esquecido o fato de que a indexação, como processo lingüístico, submete-se às exigências não só de uma sintaxe e de uma semântica, mas também de uma pragmática, para utilizar a conhecida nomenclatura de Charles Morris. Além do mais há que lembrar a variedade de processos que o indexador pode utilizar, além da diversidade de sistemas ou de linguagens empregadas, sem contar com a possível defasagem entre a capacidade efetiva do indexador e o seu real desempenho.

A conclusão a que o mencionado autor chega é que não devemos esperar que a lingüística sozinha venha a apresentar solução satisfatória para o problema. Acreditamos também que a indexação não seja um problema exclusivamente lingüístico, a não ser que se atribua à lingüística uma extensão tão globalizante que venha compreender a lógica, a teoria da ciência e até a sociologia da ciência. O que parece mais correto será dizer que, no caso, a lingüística é apenas um ponto de partida para alcançar outros momentos básicos do pensar humano.

A seguir, Hutchins procura indicar algumas possíveis soluções para a técnica de análise do texto. Lembra, em primeiro lugar, que um texto nada mais é do que uma organização sintagmática de enunciados. E como essa organização pode ser vista em vários níveis (fonético, sintático, semântico, etc.), assim também a análise do texto pode receber vários enfoques. O nível que no nosso caso mais interessa é precisamente o semântico. A hipótese que, nesse nível, se levanta é a de que o autor, ao desenvolver o texto, partiu de um enunciado básico (*Topic statement*) Trata-se agora de descobrir qual terá sido esse enunciado. Mesmo admitindo como válida esta hipótese, teremos dado algum passo adiante na formulação do conteúdo do documento? Acredito que não. Mesmo o recurso às chamadas gramáticas transformacionais não parece resolver o problema. Continuamos na mesma situação. Não sabemos como chegar até lá. Dizer que o esforço por sumarizar um texto corresponde a desenrolar, em sentido inverso, do enunciado básico pode corresponder à verdade, mas não parecer dar qualquer ajuda de ordem prática. Em síntese, o problema é de suma importância mas não temos, por enquanto, soluções definitivas e infalíveis.

Comunicação recebida em 27.02.87

A Indexação

Abstract:

The indexing

Discusses about the object of the indexing process and its objective. Points out the difficulties to identify truly representative elements within the text, due to the lack of theoretical foundations. Points out controversial aspects for discussion.

REFERÊNCIAS

01. LANCASTER, F. S. *Information Retrieval Systems*. New York, John Wiley, 1968. p. 3.
02. FAIRTHONE, R. A. *Towards Information Retrieval* s.l.p. Archon Books, 1968. p. 124.
03. HUTCHINS, W. J. *Languages of Indexing and Classification*. London, Peter Peregrinus, 1975.c.7.