

A CIBERNÉTICA APLICADA AO CONTROLE BIBLIOGRÁFICO

*Tarcísio Zandonade**

1. INTRODUÇÃO

Uma obra clássica, Louis Couffignal define a cibernética como aquele modo de ação que faz com que ela seja eficaz.¹ A cada domínio da atividade humana, por conseguinte, poder-se-á aplicar uma cibernética especial, com vistas a que esse campo específico de ação adquira o valor fundamental da eficácia.

No presente encontro, vamos considerar os fenômenos básicos do controle bibliográfico sob o aspecto cibernético, na expectativa de que essa análise possa revelar relações analógicas entre a cibernética e sua aplicação às ciências contábeis e atuárias.

A aplicabilidade da cibernética a cada um dos campos da ciência decor-

re de sua própria natureza interdisciplinar. Assim como a estatística e a teoria geral de sistemas, a disciplina da cibernética, embora oriunda de pesquisas matemáticas e biológicas, não tem suas conclusões restritas pelos limites de uma determinada ciência.

Nos últimos anos do século XIX, os belgas Paul Otlet e Henri La Fontana desenvolveram a atividade científica da "documentação", com vistas à organização do conhecimento registrado. Elaboraram o *Répertoire Bibliographique Universel*, para o qual compilaram cerca de 11 milhões de fichas catalográficas, indexando os artigos de grande parte dos 10 mil periódicos científicos correntes no início deste século².

Quase à mesma época, uma atividade semelhante foi desenvolvida em

* Professor assistente CID/FA/UnB.

1 Louis Couffignal, *Les notions de base*, Paris, Gauthier-Villars, 1958. (Information et cybernetique).

2 Maria Cristina Bello Ferreira Pinto, "Catálogo & bibliografias: evolução histórica do trabalho de controle bibliográfico", R. Esc. Bibliotecon, B. Horizonte, UFMG, Vol. 16, n. 2, set. 1987, pp. 143-158.

Munique, na Alemanha, pelos intelectuais suíços Karl Wilhelm Bühler e Adolf Saager. Esse movimento, conhecido como Movimento da Ponte – *Bridge Movement* –, ganhou o apoio e a liderança do renomado químico alemão Wilhelm Ostwald,³ e dessa organização participou também Otlet. O Movimento da Ponte tinha por objetivo interligar o conhecimento das várias disciplinas científicas, de forma a consorciar aque-

las teorias que se ativessem a mais de um domínio científico. O próprio Ostwald elaborou um Tratado de Organização do Conhecimento. Esse movimento cooperativo e integrador das ciências foi precocemente abortado pela Primeira Guerra Mundial mas, a partir deste e de outros movimentos de cooperação entre as ciências, resultaram áreas interdisciplinares de conhecimento científico, como a cibernética.



2. CONTROLE BIBLIOGRÁFICO E CIBERNÉTICA

O controle bibliográfico já se encontrava presente desde o aparecimento dos mais remotos elementos da organização dos registros gráficos do conhecimento humano, ou seja, desde os primeiros sinais concretos de organização da escrita. Esses produtos gráficos eram constituídos – tanto quanto se conhece das civilizações mesopotâmicas – de

registros contábeis e inventariais da administração pública dos povos antigos.

O termo controle bibliográfico, no entanto, veio a ser cunhado por Egan e Shera, em um trabalho pioneiro publicado em 1949,⁴ poucos meses depois da edição da obra-mãe da cibernética por Norbert Wiener.⁵ Não se sabe ao certo se Jesse Hauk Shera, então professor da primeira escola de pós-graduação em Biblioteconomia, da Universidade de Chicago, e sua colaboradora Margareth Egan, tiveram prévio

conhecimento do trabalho desenvolvido pela equipe de Norbert Wiener. A relação íntima, entretanto, entre os conceitos de controle bibliográfico e os de cibernética é muito evidente na definição de controle bibliográfico, oferecida pelos pesquisadores de Chicago:

... os controles bibliográfico são aqueles mecanismos utilizados para orientar a energia intelectual na extração, a partir da totalidade da informação registrada, daquelas porções relevantes para uma determinada tarefa, com a maior rapidez e economia⁶.

Além de haverem registrado com esse trabalho o nascimento do termo “controle bibliográfico”, Egan e Shera parecem haver inaugurado uma nova posição paradigmática dentro da pesquisa biblioteconômica, uma vez que passaram a investigar “a recuperação da informação⁷ como um campo da atividade de engenharia” [da informação].⁸ Essa nova linha de pesquisa pas-

sou a focalizar, preferencialmente, os mecanismos – *engine*, daí *engineering* – de controle bibliográfico. Esse paradigma alternativo adquiriu relevância, nas décadas recentes, com a revolução tecnológica da automação e culmina hoje com os “mecanismos de busca” – *search engines* –, que elaboram a indexação de recursos da Internet.

Os mecanismos de controle bibliográfico foram concebidos por Egan e Shera como sendo aqueles mecanismos capazes de oferecer uma acessibilidade física e uma acessibilidade temática à informação. Essas duas vias de recuperação da informação são complementares e, embora excluam a possibilidade de uma terceira via, não são mutuamente excludentes. A natureza e a capacidade desses dois modos de recuperação da informação foram posteriormente estudadas, entre outros, por Patrick Wilson⁹ – “controle des-

3 Takashi Satoh, *The bridge movement in Munich and Ostwald's treatise on the organization of knowledge*, Libri, Vol. 37, n. 1, 1987, pp. 1-24. Wilhelm Ostwald (1853-1932) foi o fundador da disciplina científica da físico-química e veio a ser o ganhador do Prêmio Nobel de Química, em 1909.

4 M. E. Egan e J. H. Shera, “Prolegomena to bibliographic control”, *Journal of Cataloging and Classification*, Vol. 5, n. 2, Winter 1949, pp.17-19. Tradução, para fins acadêmicos, por Tarcísio Zandonade, com o título “Introdução ao controle bibliográfico”.

5 Norbert Wiener, *Cybernetics: or control an communication in the animal and the machine*, Nova York, John Wiley, 1951, 194p.

6 M. E. Egan e J. H. Shera, *idem*.

7 O termo “recuperação da informação” – em inglês *information retrieval* – foi cunhado por Calvin N. Mooers, em 1950. O novo termo, de acordo com Herbert Coblans – *Use of mechanised methods in documentation work*, Londres, Aslib, 1966, 89 p. – foi registrado no artigo de C. N. Mooers, “Comments on the paper by Bar-Hillel”, publicado no *American Documentation*, Vol. 8, 1957, pp. 114-116. O novo termo de Mooers corresponde ao conceito de “extração”, utilizado por Egan e Shera no citado trabalho pioneiro. É oportuno observar que a recuperação da informação na Internet vem sendo denominada de *information discovery*.

8 Gernot Wersig, “Information science: the study of postmodern knowledge usage”, *Information Processing & Management*, Vol. 26, n. 2, 1993, pp. 229-239.

9 Patrick Wilson, *Two kinds of power: an essay on bibliographic control*, Berkeley, University of California Press, 1968, 155p.

critivo" e "controle exploratório" –, e por Alfredo Serrai¹⁰ – "controle decisivo" e "algorítmico" e "controle semântico" ou "heurístico".



3. CONTROLE BIBLIOGRÁFICO E CIBERNÉTICA

Hans Wellisch estabeleceu a profunda relação entre o controle bibliográfico e a cibernética.¹¹ Para Wellisch, um sistema de controle bibliográfico se constitui de um mecanismo que exerce basicamente a função de controle descritivo e a função de controle exploratório, bem como a operação de transcrição dos dados do documento-fonte para o documento-substituto e a operação de arranjo desses documentos-substitutos. Wellisch comprova que a função descritiva, bem como as operações de transcrição e de arranjo podem ser submetidas integralmente ao controle da cibernética, enquanto que a função exploratória não é passível de subordinação total a essas leis.

Para uma compreensão do estudo de Wellisch, apresentamos, a seguir, uma síntese do seu trabalho. Parte da premissa de que um sistema de recuperação da informação, como ademais todo e qualquer sistema, deve ser analisado à luz da teoria geral de sistemas e das leis fundamentais da cibernética.

Poderemos entender, assim, um sistema de recuperação da informação se tomarmos como exemplo um modelo físico, com estrutura e funcionamento simples, do qual possamos visualizar os elementos constitutivos, como, por exemplo, um sistema hidráulico. Uma caixa d'água (S = sistema), no alto de um edifício residencial (E = ambiente), abastecida por uma entrada de água proveniente da rede pública (D = variedade da entrada), fornecendo água para as várias unidades residenciais do edifício (O = saída), é um sis-

tema regulado ciberneticamente. Para que esse sistema hidráulico permaneça em estado viável (m), deve ser dotado de um subsistema de controle (C = controle), constituído por um ou mais reguladores (R = regulador), como bóias, registros, ladrões, torneiras, sen-

sores etc. Para funcionar satisfatoriamente – isto é, para que seja viável –, deve obedecer à Lei da Variedade Requerida, formulada por Ashby.¹² Qualquer sistema de recuperação da informação também está sujeito a essa Lei da Variedade Requerida.¹³



4. A RECUPERAÇÃO DA INFORMAÇÃO NÃO INTERMEDIADA POR UM MECANISMO

Busca de um determinado documento numa coleção desorganizada – por exemplo, num depósito caótico de documentos – permite a eventual localização desse documento. O procedimento primário para se localizar esse documento, nessas condições, consiste em "inspecionar todos os documentos da coleção, um a um, até que se localize o documento desejado", operação a que podemos chamar "algoritmo de exploração". O resultado será quase sempre

eficaz – ou seja, se o documento desejado se encontra na coleção, será eventualmente recuperado –, mas extremamente demorado – por exemplo, a inspeção de um por um, à razão de 400 títulos por hora, de todos os livros de uma pequena biblioteca particular tomaria em torno de quatro minutos, enquanto que a inspeção dos livros da biblioteca de uma pequena universidade poderia tomar quatro semanas e os de uma biblioteca nacional poderia tomar de quatro a 400 anos.¹⁴

Um nível mais elevado de controle, mas ainda não intermediado por um mecanismo de recuperação da informação, poderá ser obtido mediante a

10 Alfredo Serrai, *La biblioteconomia come scienza. introduzione ai problemi e alla metodologia*, Florença, Leo O. Olshki, 1973, 123p. (Biblioteconomia e bibliografia, 9).

11 Hans H. Wellisch, "The cybernetics of bibliographic control: Toward a theory of document retrieval systems", *Journal of the American Society for Information Science*, Vol. 31, n. 1, Jan. 1980, pp. 41-50. Esse trabalho foi publicado, em forma monográfica, pelo Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT: Hans H. Wellisch, *A cibernética do controle bibliográfico: Para uma teoria dos sistemas de recuperação da informação*, trad. Tarcísio Zandonade, Brasília, IBICT, 1987, 58p.

12 W. Ashby, *Uma introdução à cibernética*, Trad. de Geraldo Gerson de Souza, "Coleção Estudos, n. 1", São Paulo, Perspetiva, 1970, 345p., cf. p. 244: "Apenas a variedade em R pode forçar a baixa da variedade devida a D; somente a variedade pode destruir a variedade".

13 É até irônico verificar que a inviabilidade de um sistema é mais observável do que sua viabilidade. A existência da maioria dos sistemas com os quais convivemos – até mesmo e especialmente os que mantemos dentro de nós mesmos, como o sistema respiratório, o sistema circulatório sanguíneo, o sistema digestivo – é empiricamente observada somente quando esses se tornam inviáveis!

14 A. J. Evans, R. G. Rhodes, S. Keenan, *Formation des utilisateurs de l'information scientifique e technique: Guide de l'Unisist pour les enseignants*, Paris, Unesco, 1977, 201p., cf. p. 66.

organização de todos os itens de uma coleção, por exemplo, pela ordem alfabética dos títulos de seus documentos constitutivos. Esse procedimento será útil somente para aqueles usuários que venham em busca de documentos, conhecendo previamente seus

títulos. Mas, para aqueles usuários que tenham, *a priori*, somente o nome do autor ou o tema tratado pelo documento desejado, a reunião organizada do acervo será igualmente inútil do ponto de vista da eficácia da recuperação da informação.



5. A FUNÇÃO DESCRITIVA E AS OPERAÇÕES DE TRANSCRIÇÃO E DE ORDENAÇÃO

Esta forma, o acesso a um documento somente será eficaz mediante o estabelecimento, entre a coleção de documentos e os seus potenciais usuários, de um sistema de representação analógica do acervo. Elabora-se um mecanismo formado por

documentos-substitutos – fichas catalográficas¹⁵ – de cada um dos documentos-fonte – livros – existentes nessa coleção. A esse mecanismo chamamos “catálogo bibliográfico”. Embora mais complexo do que o modelo hidráulico analisado acima, o catálogo possui uma estrutura semelhante ao modelo e obedece às mesmas leis cibernéticas. Como todo sistema de recuperação da informação, o catálogo tem por objetivo localizar um

15 Obviamente, o “registro bibliográfico”, de há algum tempo, pode ser elaborado de forma mais eficiente mediante as técnicas da informática e das bases de dados. Utilizamos, entretanto, o exemplo da ficha catalográfica, tanto por sua visibilidade, como para prestar uma homenagem aos mais de duzentos anos de bons serviços prestados à recuperação da informação por essa singela, mas singular invenção. A ficha catalográfica foi institucionalizada como suporte para o registro bibliográfico mediante a utilização de cartas de baralho, na Revolução Francesa, com a promulgação do primeiro código de catalogação francês, em 1791. Para mais detalhes sobre o assunto, cf. Joseph Smalley, “The French cataloging code of 1791: A translation”, *The Library Quarterly*, Vol. 61, n. 1, Jan. 1991, pp. 1-14. A maior originalidade introduzida pela ficha catalográfica consistiu no fato de que, ao contrário do que ocorria até fins do século XVIII, com os catálogos em livros de folhas encadernadas, a nova modalidade de suporte físico propiciou a criação de uma forma de catálogo em que o registro de cada livro passou a ser elaborado sobre uma carta individual de baralho, o que veio a lhe conferir uma capacidade de ser inserida em qualquer ordenação no catálogo. O que é o princípio fundamental das cartas de baralho: a sua “baralhabilidade”, ou seja, a capacidade das cartas serem ordenadas em qualquer arranjo que se queira!

documento dentro de uma determinada coleção, quer através das características de identificação e de reconhecimento do documento, quer mediante o conteúdo temático do documento.

Para que o catálogo – esse mecanismo formado pelos substitutos dos documentos da coleção – funcione eficazmente, três requisitos básicos devem ser preenchidos.

IDENTIDADE DO DOCUMENTO

Os substitutos dos documentos devem identificar cada documento de maneira única, o que se obtém seguindo-se às *regras de um código de catalogação*. Analogamente à identificação de um cidadão, obtida através de seu registro de nascimento, a identificação bibliográfica é obtida por intermédio da elaboração de seu registro bibliográfico, de que resulta, no nosso exemplo, a ficha catalográfica.

TRANSCRIÇÃO DOS DADOS

A transcrição consiste na operação de registrar ou de copiar os dados a partir do documento-fonte para o documento-substituto. Os substitutos devem ser registrados por grafemas conhecidos tanto pelo catalogador como pelo usuário. Esse requisito será atendido mediante as *regras ortográficas da língua do documento* e, para escritas desconhecidas dos usuários, por meio das *regras de transliteração*.

ARRANJO DOS SUBSTITUTOS DOS DOCUMENTOS

Os substitutos devem ser ordenados numa seqüência preestabelecida, de modo que cada substituto ocupe exclusivamente uma posição previsível, que o torne recuperável de forma única, o que se obtém mediante as *regras de arquivamento e ordenação dos substitutos* no sistema.



6. A FUNÇÃO EXPLORATÓRIA

A mesma eficácia cibernética não ocorre com a segunda via de acesso aos documentos-fonte, ou seja, a busca de um documento de uma coleção por intermédio de um *tópico*, de um *tema*, de um *assunto*. Para que essa forma de recupe-

ração seja possível, o conteúdo temático do documento-fonte deve ser traduzido, na entrada do sistema, numa *linguagem documentária* – ou seja, por palavras-chave retiradas da linguagem natural do próprio documento, por descritores extraídos de tesouros, por códigos obtidos de esquemas de classificação bibliográfica etc. – , e es-

ses representantes semânticos são atribuídos ao substituto do documento. Assim, os documentos sobre um determinado assunto serão recuperáveis, como itens físicos, por meio de seus substitutos, na medida em que esses sejam portadores da mesma representação semântica com que o usuário do sistema venha a procurá-lo.

Obviamente, seguindo-se essa via, a capacidade da recuperação estará limitada pela extensão da indexação fornecida pelo classificador na entrada do sistema. Como a variedade de assuntos tratados por um documento é sempre maior do que qualquer categorização que se tenha atribuído a cada documento, nem sempre se pode obter uma união exata entre os temas da pergunta e os documentos recuperados. Na prática, nem todos os documentos recuperados serão relevantes, bem como nem todos os documentos relevantes serão recuperados, se não tiverem sido "rotulados" previamente como atinentes ao assunto pesquisado.

A dificuldade reside primariamente no caráter subjetivo e instável da *relevância* de cada documento para res-

ponder a uma questão temática. No momento de atribuir a cada documento uma categoria, o classificador não pode, *a priori*, prever todas as alternativas para as quais o referido documento será relevante para os diferentes usuários finais. Nem a indexação de todos os termos significativos do texto completo, ou mesmo de uma parte privilegiada do documento, poderá ser de grande valia, uma vez que, dessa forma, obter-se-á, em razão inversamente proporcional, um elevado "índice de revocação", à custa de um baixo "índice de precisão".

Dessa forma, como na busca de um documento por um assunto a variedade da entrada é sempre ilimitada, o processo de recuperação de documentos por assunto será necessariamente um processo de tentativa e erro — ou seja, um controle exploratório —, em que todos os resultados devem ser analisados um a um para confirmação de sua relevância para o tema em questão. Essa talvez seja hoje a maior limitação a que está restrita a busca de informação na Internet, pela intratabilidade da busca semântica na *selva selvaggia* dos recursos registrados na Internet.



7. CONCLUSÃO

Os sistemas de recuperação da informação, também conhecidos como instrumentos, ferramentas ou mecanismos de controle bibliográficos, bem como os "mecanismos de busca" para a "descoberta de recursos" na Internet, constituem-se em engenhos subordinados às leis gerais da cibernética. Nesses sistemas, por força da Lei da Variedade Requerida, sabe-se que o controle descritivo pode ser realizado integralmente, enquanto que o controle exploratório é inevitavelmente limitado pela infinita variedade semântica da entrada. Com

o desenvolvimento da informática, muitos pesquisadores ganharam uma grande esperança de que também o controle exploratório poderá ser integralmente submetido a controle, enquanto outros têm assumido uma postura cética, em nome de uma inviabilidade intrínseca entre revocação e precisão integrais. O problema, hoje, está entregue aos estudos da "inteligência artificial" e à construção de "sistemas especialistas", domínios do saber nos quais as ciências contábeis e atuariais — em razão da natureza da informação processada nessa área da atividade científica e tecnológica — estão obtendo resultados concretos e conclusivos.

