

As bibliotecas e a organização do conhecimento: evolução e perspectivas

Maria Inês Lopes

Fundação Calouste Gulbenkian

I. Introdução

Pela sua natureza e missão, a actividade biblioteconómica sempre ocupou um lugar visível ao lado das diversas ciências que têm a ver com a organização do conhecimento e da sua expressão. É, porventura, nessa relação que se encontra o lado intelectualmente mais interessante, e também o mais difícil, da actividade própria das bibliotecas. É principalmente da relação que detêm com o domínio do conhecimento e das suas formas de organização e transmissão, que as bibliotecas obtêm a importância, que sempre tiveram, ao longo da sua história.

Primeiro, antes da industrialização da imprensa, por constituírem em si repositórios do conhecimento humano existente, que reflectiam na organização física dos seus acervos, e cuja importância era directamente proporcional à raridade e pouca acessibilidade desses repositórios. Depois, já de certa forma como agentes da democratização no acesso ao conhecimento, e mais tarde como agentes indispensáveis da sua intermediação, as bibliotecas adoptaram, em complemento ou substituição da organização física das suas colecções, formas de organização lógica dessas mesmas colecções, através de instrumentos de pesquisa separados.

A elaboração desses instrumentos, cuja aplicação ao universo de recursos de informação disponíveis produz os chamados catálogos de assuntos,

constitui o resultado daquilo que designamos por organização do conhecimento em bibliotecas.

2. Conhecimento e organização: duas noções fundamentais

Conhecimento e organização são duas noções fundamentais e indissociáveis. Não simplesmente porque o tema que focamos é organização do conhecimento, mas porque a própria existência de conhecimento humano pressupõe sempre fenómenos de organização.

Genericamente toda a actividade biblioteconómica se orienta, de alguma forma, para a resolução das questões que se colocam à utilização dos registos do conhecimento humano. Mas a relação específica da biblioteca com o conhecimento reside na produção de acessos organizados aos conteúdos de informação desses registos. Essa relação não é linear nem deve ser vista como exterior ou alheia aos próprios processos de formação, registo e utilização do conhecimento. Importa conhecer a natureza e modelos desses processos para relativizar o alcance do que tem sido, e do que pode ser, aquilo que designamos por actividade de organização do conhecimento em bibliotecas.

Por definição, a primeira função da biblioteca é angariar recursos de informação que são produzidos individualmente, e organizá-los para uso colectivo. Deste ponto de vista podemos considerar duas acepções: o conhecimento individual, possuído por um indivíduo ou acessível através dele, e o conhecimento social, possuído colectivamente por uma sociedade ou sistema de grupo. Não existe propriamente dicotomia ou fronteira entre estas duas acepções de conhecimento: grande parte do conhecimento humano pertence simultaneamente a estes dois níveis, e o conhecimento social – aquele que, por definição, está registado e disponível publicamente – é sempre de alguma forma originado e originador de conhecimento individual.

O conhecimento em si tem uma dimensão essencialmente humana. Não há informação sem estruturas cognitivas, sem memória. O processamento humano da informação é um sistema contínuo e sempre inacabado de organização de informação, a partir do qual se produz o conhecimento abstracto – a formação dos conceitos e a sua inserção numa estrutura relacional dinâmica. A memória alimenta-se dessa estrutura relacional que também é fluida, e não estática: a informação armazenada na memória humana evolui no tempo, não é constante, e ao ser utilizada é continuamente reinterpretada e modificada.

Um aspecto fundamental a destacar é o facto de o processamento humano da informação, e portanto a própria memória, se desenvolver no contexto de actos de comunicação com a realidade exterior ao indivíduo. O plano da

comunicação é, assim, indissociável da formação e utilização de conhecimento, ou seja, o conhecimento não se pode desligar das formas de representação (exteriorização) usadas na comunicação.

A eficácia dessa comunicação depende não só da paridade das estruturas cognitivas dos parceiros dessa comunicação mas também do domínio das formas e mecanismos de transmissão de informação. Por sua vez, essas formas de transmissão reflectem, na sua diversidade, a própria diversidade das estruturas cognitivas dos indivíduos, produzindo interpretações, reproduções e recriações multifacetadas da informação. A dinâmica e plasticidade da linguagem humana – principal sistema de comunicação – são ao mesmo tempo reflexo e origem dos processos de formação e utilização do conhecimento.

Quando falamos da organização do conhecimento em bibliotecas falamos sobretudo da organização de representações do conhecimento – a sua exteriorização em informação contida nos documentos. Isto é, falamos de mecanismos para gerir, ou melhor controlar, a dispersão, a complexidade e a diversidade que caracterizam o universo dos objectos transmissores (documentos) de representações de conhecimento (informação), tal como existem na sua ordem natural.

Para viabilizar uma utilização controlada desse universo de representações a biblioteca constrói uma memória artificial que visa suplementar as memórias individuais. Essa memória artificial é organizada segundo uma estrutura conceptual – mapas de conhecimentos, ou seja, conceitos e suas relações – que se sobrepõe e deve ser independente relativamente às estruturas cognitivas/quadros de referência e formas de comunicação dos indivíduos que produziram (autores) essas representações.

Para isso são elaborados instrumentos de organização da informação – as linguagens documentais – que produzem novas representações que se pretendem com níveis de estabilidade e validade adequados para uso comum de comunidades, mais ou menos definidas e mais ou menos alargadas. Enquanto instrumentos de comunicação – isto é, formas e mecanismos de representação dos conceitos e suas relações – as linguagens documentais visam ultrapassar os fenómenos que na comunicação natural são geradores de dispersão, incoerência e ambiguidade, mas de preferência dentro dos limites de uma razoável capacidade de utilização dessas formas, ou seja, desde que os mecanismos de controlo não originem novas ambiguidades ou novas dificuldades de comunicação.

Basicamente, existem dois tipos principais de sistemas de organização e representação controlada de conhecimento em bibliotecas: sistemas de classificação, de base conceptual hierárquica e com formas de representação codificadas (numéricas e/ou alfanuméricas), e sistemas terminológicos,

assim designados por utilizarem formas de representação baseadas em componentes da linguagem natural.

Em qualquer dos casos, esses sistemas de representação são linguagens de comunicação artificiais: têm um léxico próprio, normalmente composto por unidades de significação ao nível dos conceitos, com uma morfologia e semântica definidas; têm uma estrutura paradigmática própria, que estabelece ligações semânticas entre as unidades desse léxico, hierárquica e associativamente; e, nalguns casos, têm também uma estrutura sintagmática própria, que domina as formas de representação complexas, de associação de conceitos, através de expedientes sintácticos convencionados que combinam, na aplicação, duas ou mais unidades do léxico.

De acordo com o nível de controlo que permitem e com as capacidades de representação que encerram, esses sistemas são mais ou menos eficazes a resolver os problemas da linguagem natural, e cumprem melhor ou pior as suas funções de comunicação. Na conciliação destes dois aspectos residem sempre soluções de compromisso: normalmente, na medida em que as linguagens documentais são mais sofisticadas em termos de capacidades de representação controlada, acarretando por isso maior número de convenções, também diminui a sua facilidade de utilização: a aplicação e manutenção do sistema é mais cara e mais exigente, e a sua exploração pelo utilizador, mais afastada dos modelos da linguagem natural enquanto meio de comunicação, requer maior aprendizagem.

Dentro de cada um dos grandes tipos de linguagens documentais referidos, existiu sempre uma diversidade de soluções, com variadas metodologias e características de construção e representação, sejam elas consubstanciadas em catálogos sistemáticos, usando classificações, ou em catálogos alfabéticos. E assim continuará a ser pois essas memórias artificiais são sempre actividades, produtos, e resultados de determinados contextos, com opções e contingências diferentes, no espaço e no tempo.

Diferentes contextos produzem sistemas de organização com qualidades e efeitos diferentes, em função de factores nem sempre coincidentes e com pesos variáveis, quer de ordem conceptual, quer de gestão, quer humana. A nível conceptual, destaca-se o grau de adequação e solidez das bases teóricas em que um sistema assenta, e que afectam o grau de coerência e durabilidade (validade no tempo) das suas soluções. Em termos de gestão, é crucial o nível de investimento, isto é, a quantidade/qualidade de recursos afectos ao desenho, implementação e manutenção de um sistema, que influencia os níveis de sofisticação e qualidade que são possíveis de alcançar na sua aplicação.

Finalmente, há a considerar os factores humanos, que são mais diversos, menos controláveis e que têm consequências mais profundas, e efeitos mais alargados; por um lado, a incontornável subjectividade inerente à condição

humana das actividades de análise, interpretação, ponderação, selecção e mesmo representação de conteúdos de informação. A subjectividade é originadora do maior ou menor grau de correcção e consistência dos conteúdos de informação de um sistema, e só parcialmente é explicável pelas diferenças entre os indivíduos em matéria de educação e treino, literacia e capacidades cognitivas, uma vez que a inconsistência também existe como fenómeno individual. Por outro lado, existem as inevitáveis diferenças de adaptação e comunicação dos utilizadores com os sistemas, por razões idênticas às já expostas. Essas diferenças existem sempre, por melhor definida que tenha sido à partida a audiência de um sistema, sobretudo actualmente, em que a acessibilidade de informação em rede conduz à globalização de recursos.

3. Evolução dos sistemas de recuperação por assuntos

Os mais antigos sistemas de organização e recuperação por assuntos que ainda hoje se usam nas bibliotecas nasceram há cerca de 100 anos, derivados das classificações filosóficas: são as classificações bibliográficas enciclopédicas, de que se destacam a *Classificação Dewey* e *CDU*¹, universos conceptuais concebidos para encerrar, numa estrutura abrangente, a organização do conhecimento humano existente. São sistemas de base hierárquica e essencialmente enumerativa, que permitem uma ordenação sistemática de conjuntos de objectos de informação, de acordo com categorias lógicas e através de sistemas notacionais que são independentes das línguas.

Cientificamente validados para a época em que surgiram, esses sistemas têm como primeiro ponto fraco as dificuldades de manutenção: rapidamente se tornam mapas de conhecimento desactualizados ou controversos, sendo também onerosos e complicados de gerir em termos da sua capacidade para acomodar novas áreas de conhecimento, novos conceitos e pontos de vista, sem implicar alteração retrospectiva da sua estrutura base.

Por outro lado, a unidimensionalidade é outra característica negativa desses sistemas que só à custa de complicados léxicos e expedientes sintácticos conseguem representar significações muito específicas e complexas, incluindo relações não hierárquicas. A verdade é que quando a especificidade é um requisito, as classificações aparecem ao utilizador como sistemas cabalísticos.

Essa terá sido uma das razões que explicam que sistemas de classificação bem mais recentes, de natureza mais dinâmica, e com uma base intelectualmente mais sólida e mais sofisticada, como a *Colon Classification* e a *Bliss Classification* – que assumiram como essencial a noção de multidimensionalidade do conhecimento integrando na sua construção a noção de facetas e relações – não lograram uma implantação tão alargada como as classificações *Dewey* e

CDU. Também por isso, a CDU, que se distingue da Dewey precisamente por incluir expedientes relacionais e sintácticos, observou nas últimas décadas uma tendência de utilização prática mais «económica», isto é, sem levar ao extremo todas as suas possibilidades de especificação.

Simultaneamente com as classificações surgem sistemas alfabéticos constituídos por índices fraseológicos, muitas das vezes também de carácter classificatório, que em muitas das bibliotecas pretendiam ser uma alternativa ou complemento às classificações decimais. São índices que fizeram uma época e que evoluíram para os chamados sistemas de cabeçalhos pré-coordenados², assim designados posteriormente, para os distinguir dos sistemas de indexação por palavras ou conceitos, que surgiram depois. Eram ainda índices complexos, difíceis de gerir e, de uma maneira geral, sem uma estrutura de controlo de vocabulário.

Mais tarde, já na segunda metade deste século, à medida que as bibliotecas se confrontam com um universo de informação em crescimento exponencial, tanto em volume como na diversificação e complexidade das áreas de conhecimento e actividade, cresce a importância do acesso à informação, sobrepondo-se à da sua organização sistemática.

Nesta fase, a capacidade de intermediação da biblioteca passa a depender cada vez mais da acuidade, rapidez de produção e facilidade de utilização dos respectivos instrumentos de pesquisa. Os sistemas de classificação têm dificuldade em responder a estes requisitos, e os índices alfabéticos de cabeçalhos pré-coordenados de então não possuem uma estrutura de controlo e de orientação semântica, nem permitem uma recuperação não linear, isto é, por qualquer elemento componente que não o elemento ordenador.

É por isso que nesta fase, sobretudo a partir dos anos 70, as bibliotecas adoptam uma nova abordagem, a da organização de conteúdos através do controlo das suas formas de expressão linguística. A terminologia passou a fazer parte do quadro interdisciplinar em que as bibliotecas operam para produzir os seus instrumentos de pesquisa, exigindo novas competências, reconhecidas mas dificilmente alcançáveis, sobretudo em bibliotecas gerais de investigação, não circunscritas a domínios especializados do conhecimento.

A corresponder a estas necessidades surgem directrizes profissionais internacionais³ orientadoras dos processos de desenvolvimento e manutenção de tesouros, que passam a dominar a preocupação dos profissionais: a construção de linguagens documentais no mínimo constituídas por vocabulários controlados, e em sentido mais completo incluindo estruturas de relações hierárquicas e associativas.

Na prática, estes sistemas de indexação terminológica assumiram como fundamental a noção de relação, abrindo a possibilidade de coordenação de conceitos, não na aplicação mas na fase de recuperação, pelo que receberam

a designação de sistemas pós-coordenados. A popularidade deste tipo de sistema advém de uma maior facilidade de construção, aplicação e, aparentemente, de utilização. Mas a preocupação fundamental com a construção de léxicos controlados e estruturados, se por um lado introduziu mais rigor terminológico, também acabou por instalar uma visão reducionista do processo de indexação, relegando na prática para segundo plano as questões de análise e de representação, no sentido em que assumem a análise da informação e a sua representação apenas ao nível dos conceitos isolados.

Mais recentemente, a introdução da informática nas bibliotecas revalorizou a pré-coordenação, isto é, a coordenação de conceitos na fase de análise e a representação da síntese correspondente, através de expedientes de sintaxe controlada, assumida, que passou também a ser para estes sistemas a importância de se basearem em vocabulários-base controlados e estruturados.

De facto, os sistemas informatizados não só proporcionam uma gestão mais eficaz de cabeçalhos complexos, como permitem a sua pesquisa multidireccional, isto é, por cada elemento componente independentemente da sua posição. O resultado é uma maior especificidade na representação, maior coextensividade com os conteúdos dos documentos indexados, e, ao mesmo tempo, maior integridade na informação colhida na fase de análise, expressa através de uma síntese. Essa síntese é o valor acrescentado que estes sistemas possuem face à indexação pós-coordenada.

A opção pela pré ou pós-coordenação não é ainda hoje uma questão resolvida, nem parece que possa sê-lo senão em termos meramente intelectuais. De facto, essa opção tem muito a ver com factores de gestão, uma vez que para processar com a igual rapidez o mesmo volume de informação, a pré-coordenação requer muito maior investimento, qualitativo e quantitativo, de recursos humanos.

De uma forma geral, no entanto, existe a noção e a experiência de que os sistemas pós-coordenados se aplicam com melhores resultados (ou, se quisermos, com menos inconvenientes) em bibliotecas especializadas, onde a indexação por conceitos é menos susceptível de produzir recuperações não pertinentes; e que os sistemas pré-coordenados são essenciais em bibliotecas gerais, em que, do ponto de vista da pesquisa, há maior necessidade de servir a informação específica em contexto, uma vez que não lhes está subjacente nenhum domínio de específico conhecimento.

Com uma opção ou outra, os catálogos alfabéticos de assuntos passaram a suprir melhor as necessidades de acesso rápido e directo à informação. Mas, apesar de em muitos casos possuírem uma estrutura de relações semânticas, não lograram substituir a função de ordenação global, lógica e hierarquizada, nem a eliminação das barreiras linguísticas, baseada numa expressão codificada universal, pontos fortes das classificações.

Por isso, ainda hoje as duas abordagens são complementares, na elaboração dos instrumentos de acesso por assunto à informação produzidos pelas bibliotecas. Mais uma vez, é sobretudo nas bibliotecas gerais que essa complementaridade é necessária, não só pela diversidade das áreas de conhecimento que encerram, mas também pelo volume das respectivas colecções.

4. Tendências e questões actuais

Do muito pouco que fica dito sobre uma matéria tão complexa como a que abordamos, ressalta antes de mais a noção de que a validade universal, e temporal, de qualquer sistema de organização do conhecimento – enquanto estrutura conceptual – é uma utopia. Do mesmo modo que não há sistemas de comunicação de informação com validade e eficácia universais, isto é, para todos os contextos, indivíduos, espaços e tempos.

Longe de pretender ser uma negação da validade do trabalho de organização do conhecimento que se tem praticado nas bibliotecas, esta noção permite-nos considerar esse trabalho numa dimensão de processo aberto, essencial para, sem se perder o pragmatismo necessário à velocidade do presente, se poderem questionar as suas bases e reorientar as suas estratégias. Permite, enfim, encarar e compreender, sem ansiedades, a insatisfação que neste domínio existe com o que o presente hoje nos oferece.

Essa noção, relativizadora, perpassa por algumas tendências e questões que hoje se colocam em vários aspectos: normalização, compatibilidade e metodologias de trabalho, fundamentos teóricos, estudo e investigação e, finalmente, estratégias orientadas para as novas tecnologias e para o domínio da globalização da informação na era electrónica.

a) Normalização e compatibilidade

Nas últimas décadas, o crescimento massivo da informação produzida e distribuída sob as mais variadas formas, tornou cada vez mais indispensável a função de intermediação da biblioteca, reforçando a importância dos instrumentos de pesquisa. Essa evolução produziu uma complexificação da organização interna da biblioteca, e conduziu a uma normalização progressiva de quase todas as suas actividades, que hoje se verifica tanto a nível local como internacional, desenvolvida sobretudo nos últimos trinta anos.

De todas essas actividades, a que se relaciona com a produção de instrumentos de recuperação por assuntos é ainda aquela que reúne menos consenso profissional, tanto na teoria como na prática. A análise e recuperação por assuntos continua a ser o menos compreendido dos aspectos do controlo bibliográfico, sendo também a área mais difícil para a cooperação e transfe-

rência de informação, dada a falta – ou, até agora, impossibilidade – de consistência e compatibilidade intersistemas que a esse nível se verifica.

Não quer isto dizer que não existam nesse campo algumas normas nacionais ou internacionais, mas as que existem não têm a função de assegurar resultados normalizados no processamento da informação, como na catalogação descritiva, sendo antes orientadoras de aspectos gerais ou de procedimentos parciais. É o caso das já referidas directrizes sobre a construção de *tesauros*, ou para a análise e indexação de documentos⁴, neste caso um instrumento que pretende ser de cariz normativo que mais não é que uma incipiente orientação sem outra utilidade prática que a de ser uma primeira abordagem útil para o ensino. Mais recentemente, no âmbito da IFLA, foram desenvolvidas e aprovadas directrizes para a constituição de ficheiros de autoridade de assuntos (*GSARE – Guidelines for Subject Authority and Reference Entries*, 1994) que, como o nome indica, são importantes para uniformizar o registo dos conteúdos e suas formas de apresentação, mas não têm a ver com a concepção e critérios desses conteúdos.

Apenas alguns instrumentos de representação organizada e controlada do conhecimento podem ser considerados *normas de facto*, até um certo ponto, dado serem utilizados por conjuntos significativos de instituições. É sobretudo o caso de algumas classificações bibliográficas universais, maioritariamente usadas em bibliotecas gerais, ou de certos *tesauros* ou listas equivalentes, neste caso mais frequentes em áreas de conhecimento e actividade específicas, mas normalmente com limitações de utilização por diferentes comunidades linguísticas.

No entanto, na maior parte das vezes essa adopção não se faz pelas mesmas razões e com os mesmos mecanismos de consenso e correspondência a necessidades comuns com que se produz e adopta uma norma, no sentido formal do termo. Adoptados e interpretados fora do contexto da sua produção, isto é, sem o domínio das suas bases conceptuais de origem, esses instrumentos são, na maioria dos casos, modificados e readaptados às necessidades locais, sem uma gestão conjunta com a fonte de origem, com resultados adversos para a compatibilidade.

Daí uma tendência dos últimos anos para o trabalho de cooperação entre instituições, visando não só sustentar a compatibilidade mas também angariar competências e partilhar custos de desenvolvimento: são exemplos disso, o sistema *RAMEAU – Répertoire d'autorités matière encyclopédique et alphabétique unifié*, no âmbito das bibliotecas francesas, e também do *Projecto CLIP – Compatibilização de Linguagens de Indexação em Português*, em Portugal⁵. Mais do que a procura conjunta de uma única solução, estas iniciativas são antes de mais uma estratégia de trabalho, que evidenciam uma atitude aberta, tendente a admitir que todas as soluções são de compromisso e não definitivas.

Também no âmbito das classificações, contrariamente ao que há alguns anos se previa, existe hoje um redobrado interesse sobre as suas potencialidades a nível da compatibilidade e transferência internacional de informação, estando neste momento iniciada a elaboração de um formato MARC internacional de troca para registo electrónico de sistemas de classificação⁶, de forma a facilitar o seu uso na gestão automatizada dos acessos por assunto em sistemas de gestão de informação bibliográfica.

Importantes, sem dúvida, estas tendências são uma resposta apenas ao lado pragmático dos problemas. Não são o tipo de resposta para aprofundar, rever ou dar novas bases teóricas à actividade, porque se ocupam de aspectos parciais e intermédios.

b) Fundamentos teóricos e interdisciplinaridade

Nas últimas décadas a gestão das bibliotecas tem sido sujeita a pressões que produziram uma maior tendência para os aspectos operacionais e procedimentais, em detrimento da fundamentação dos aspectos estruturais. Talvez por essa razão ainda hoje a maioria dos sistemas de recuperação por assuntos, nomeadamente os que são terminológicos, não têm uma base estabelecida de forma explícita, organizada e orientadora, de princípios e regras de construção, manutenção e aplicação, assim como também raramente as instituições têm, também definida de forma explícita, uma política de indexação. Na maior parte dos casos, de facto, os sistemas consubstanciam-se apenas no seu instrumento de representação (a lista de cabeçalhos, *tesauro* ou esquema de classificação).

Esta situação não favorece o mútuo entendimento entre os sistemas e, entre os profissionais, assim como não ajuda a desenvolver bases conceptuais comuns. Não é por acaso que a terminologia técnica nesta área está pouco consolidada, diferindo entre comunidades profissionais com experiências e tradições técnicas diferentes.

Para corresponder à necessidade de uma base comum de entendimento internacional sobre os princípios subjacentes à construção de catálogos alfabéticos de assuntos, assim como para identificar os aspectos comuns e as diferenças na forma como diferentes sistemas resolvem ou não esses princípios, foi criado em 1990 um grupo de trabalho no âmbito da IFLA, cujo resultado é um estudo⁷, actualmente em fase de discussão internacional, que será publicado em breve. Elaborado em colaboração internacional, esse estudo revê, à luz de um conjunto de princípios identificados, a literatura de onze sistemas de indexação, originários de dez países e nove línguas diferentes. Não pretende ser mais do que é – um ponto de situação – mas é um passo prévio indispensável a futuros desenvolvimentos, conceptuais e normativos.

Um outro aspecto que importa realçar, e que constitui a maior falha, em termos da literatura dos instrumentos profissionais, situa-se ao nível dos processos individuais de análise e de selecção dos conteúdos a representar: é uma verdadeira caixa negra, que urge abrir, analisar, descodificar. Não há dúvida que no âmbito das bibliotecas a experiência existe e é longa, e que a competência no sector é incontornável mas, de uma maneira geral, carece de mais partilha, discussão e aprofundamento, porque são os processos de interpretação e análise de conteúdos que iniciam e determinam o exercício de organização do conhecimento nas bibliotecas.

É por essa razão que achamos especialmente importante e útil o paralelismo de modelos e a inter-relação de processos que existe entre o sistema de conhecimento individual e os sistemas de organização de conhecimento em bibliotecas, como realçámos no início. É com base nesse conhecimento do conhecimento que se podem desenvolver concepções e percepções mais estruturadas sobre o que são e devem ser os sistemas de recuperação de informação por assuntos em bibliotecas.

Para isso é fundamental que se assumam competências interdisciplinares, que facilitem a fundamentação explícita, e não vagamente implícita ou decorativa, em áreas como as das ciências cognitivas, da epistemologia, da linguística e das ciências da comunicação, quer nos aspectos teóricos quer aplicados, nomeadamente à computação. Esta é uma tendência que há algum tempo se constata ao nível da literatura de opinião profissional, e que já se verifica no plano da investigação transdisciplinar com interesse para a biblioteconomia. Mas que está ainda longe de se reflectir na prática corrente do ensino e do exercício profissionais, isto é, dos sistemas e das competências reais das pessoas e das organizações.

c) Tecnologias de informação e novas áreas de intervenção

Por fim, um último aspecto, que actualmente levanta questões fundamentais, mas que tem sido também gerador de alguns equívocos: a evolução das tecnologias de informação.

Nos últimos 15 anos, as tecnologias de informação vieram valorizar drasticamente os instrumentos de pesquisa de informação produzidos pelas bibliotecas, melhorando os seus métodos e processos de exploração, introduzindo facilidades de pesquisa que em sistemas manuais eram praticamente impossíveis.

A rapidez e baixo custo de manipulação automática de grandes volumes de dados, através de facilidades de pesquisa baseadas em índices de palavras-chave geradas por computador a partir de títulos, resumos, etc., em confronto com os elevados custos em recursos humanos e tempo para produzir indexação, alimentou durante anos a falsa ideia de esta função poder ser substituída por processos puramente automáticos.

A confusão gerada numa primeira fase pelas novas tecnologias teve que ver, essencialmente, com factores económicos de gestão: a pressão na redução dos custos e a sedução das novas facilidades tecnológicas, criaram a ilusão de que oferecer serviços de pesquisa em texto livre nos títulos e resumos, com facilidades de utilização de operadores booleanos, de proximidade, etc., era uma alternativa à clássica indexação intelectualmente produzida. Hoje já é claro que essas facilidades de pesquisa não constituem só por si organização de conhecimento, e que elas não são uma alternativa, mas antes um excelente complemento.

Novas facilidades tecnológicas estão sempre a surgir, que optimizam recursos, aliviando aspectos acessórios, que não necessitam de constituir um trabalho de construção intelectual local, como aliar aos catálogos dicionários de variantes ortográficas ou de sinónimos. Outro exemplo é a análise estatística das chaves de pesquisa usadas pelos utilizadores, ou dos tipos de transacções de pesquisa mais efectuadas, hoje ajudas preciosas à manutenção e melhoria de um catálogo de assuntos, que alguns sistemas comerciais de gestão de informação bibliográfica já disponibilizam. Todas estas facilidades ajudam à exploração e manutenção de catálogos, mas não substituem a sua elaboração intelectual.

Por outro lado, essa confusão inicial foi também gerada pelas enormes expectativas criadas em torno da construção de sistemas inteligentes na área das ciências humanas: aqueles que integram vocabulários controlados, gramáticas formais, modelos de ligações semânticas, etc., permitindo ao computador identificar situações-padrão e aplicar diferentes algoritmos de processamento. Mas, nos domínios de interesse para as actividades correntes das bibliotecas, esses sistemas são ainda experimentais. Até hoje, por exemplo, ainda não se tornou realidade uma das mais antigas promessas da computação aplicada ao processamento da linguagem natural: a criação de programas perfeitos e completos de tradução automática de textos, isto é, que não necessitem de intervenção humana posterior sobre os resultados.

Em termos do que interessa para a produção de informação de descrição de conteúdos, a evolução dos chamados sistemas inteligentes só tem vindo confirmar as características de imprevisibilidade e de indeterminação inerentes aos processos naturais de comunicação humana, quando se trata da transmissão de conhecimento.

Mesmo que fosse possível programar um computador para efectuar toda a série de interpretações e juízos discriminatórios que o indexador humano desenvolve ao fazer indexação atribuída – e que supõem um substrato complexo de conhecimentos, de domínio da língua, de recurso a terceiras fontes de informação, etc. – é óbvio que a sua utilidade só faria sentido se as bibliotecas dispusessem de versões electrónicas de todos os documentos

que necessitam de processar, o que não tem sido aplicável, e dificilmente virá a ser na totalidade, à realidade das bibliotecas.

Não quer isto dizer que a evolução dos sistemas inteligentes, nomeadamente dos que processam informação textual, não seja do interesse das bibliotecas. Pelo contrário, essa evolução é altamente interessante para as bibliotecas, não no âmbito da produção de catálogos mas no que respeita ao novíssimo domínio da gestão e exploração de documentos em formato digital.

É precisamente a evolução das tecnologias de comunicação, nomeadamente a *Internet*, que permitindo a disponibilização directa e universal da informação em rede, vem alterar os formatos e os circuitos tradicionais entre os produtores, distribuidores e utilizadores da informação.

Num primeiro tempo, tudo pareceu o melhor dos mundos – onde aparentemente as bibliotecas pouco já teriam a acrescentar. A própria tecnologia da *Internet*, intuitiva e acessível, as facilidades do hipertexto, assim como o êxito, na maioria dos casos efêmero, dos motores de pesquisa, contribuíram para uma primeira ilusão de que tudo estava resolvido em termos de liberdade e capacidade de acesso à informação. Por pouco tempo, como sabemos.

Rapidamente se reconheceu que essa liberdade só existe na medida da real capacidade de acesso individual à informação, e que essa capacidade não se pode apenas confinar aos motores de pesquisa, muito úteis mas que se limitam a indexar automaticamente todos os textos – e, convém sublinhar, apenas os textos – das páginas *Web* existentes na *Internet*, mesmo quando esses gigantescos índices permitem facilidades de pesquisa sofisticadas, como operadores Booleanos, de proximidade e de relevância. A utilidade dos motores de pesquisa é inegável, mas sobretudo para pesquisar a localização de uma informação precisa que se conhece. Para explorar com determinados critérios, sobretudo temáticos, o que existe disponível e se desconhece, os resultados são, na maior parte dos casos, decepcionantes.

Explorar a *Internet* significa, na sua própria terminologia, navegar. Para navegar, a não ser que se queira andar à deriva, são necessários instrumentos de navegação, orientadores. É aí que os sites de apontadores – organizados e estruturados, após pesquisa, identificação, selecção, categorização e descrição – rapidamente passaram a ser orientadores indispensáveis dos utilizadores da *Internet*. A indexação humana da *Internet* começou por iniciativas individuais – quem não conhece a lista de Yanoff – para em seguida ser assumida por instituições, normalmente de investigação e ensino, e hoje já é uma actividade comercial.

No quadro da *Internet*, que alimenta a verdadeira internacionalização da actividade científica e cria expectativas mais exigentes que nunca da parte dos utilizadores, a intermediação da biblioteca já é reconhecida como funda-

mental. Por um lado, por imperativos de integração: integração tecnológica, integração dos diversos tipos de recursos de informação, integração de serviços. Por outro lado, por necessidades efectivas do acesso organizado aos recursos de informação externos às Bibliotecas e disponíveis na Internet. Nos últimos dois anos assistiu-se a uma integração dos catálogos das bibliotecas na Internet. Actualmente, o que se espera das bibliotecas é que esses catálogos passem também a incluir informação secundária sobre recursos existentes na rede e apontadores activos que permitam, para além da identificação, o acesso directo a esses mesmos recursos.

De forma integrada nos catálogos convencionais, ou constituindo listas estruturadas de apontadores ou bases de dados separadas, a emergência é de que as bibliotecas transponham para a rede a sua experiência e as suas técnicas de organização de conhecimento.

Esta súbita emergência traz uma série de novas questões, que se desdobram a partir de duas vertentes principais: como vão as bibliotecas corresponder neste novo contexto electrónico à missão de angariar (significando aqui seleccionar) os recursos electrónicos existentes em rede que sejam potencialmente úteis para as suas comunidades de utilizadores? Como vão, numa segunda etapa, gerir a diversidade e dispersão desses recursos electrónicos, oferecendo acessos por assunto de forma estruturada, estável, abrangente e fiável, como já o fazem para os recursos de informação que possuem localmente nas suas colecções?

A mera transposição de critérios e metodologias de trabalho até agora usados nas bibliotecas – e moldados pela tecnologia e modos de distribuição do documento impresso e principalmente textual – não é fácil nem linear, se considerarmos que os documentos electrónicos têm uma natureza completamente diferente, são compostos e dinâmicos, de conteúdos não estáveis, e que actualmente a sua identificação reside apenas na localização (URL), que também é variável no tempo.

Desenvolvimentos actualmente em curso em matéria de normalização dos elementos de identificação (URIS), localização e descrição de recursos electrónicos, nomeadamente incorporados nos próprios documentos, resolverão em breve alguns destes problemas de ordem técnica, ao mesmo tempo que se desenvolvem instrumentos tecnológicos para facilitar a gestão e controlo dos registos e dos acessos por assunto aos recursos em rede, de que é exemplo o Projecto ROADS⁸, desenvolvido no âmbito do *Electronic Libraries Programme* (eLib, UK).

No entanto, as questões de organização intelectual, só o trabalho prático e a experiência acumulada resolverão. Nesta vertente, a abordagem possível das bibliotecas neste momento é de ordem pragmática, e em si própria já é um enorme valor acrescentado: é a de começar por aplicar aos recursos

electrónicos em rede a mesma experiência de análise e os mesmos instrumentos de representação organizada de conhecimento que já atravessaram gerações, nas bibliotecas.

Existem disso muito boas experiências na *Internet*, que fazem com que os nossos utilizadores de longos anos nas bibliotecas se sintam, finalmente, em casa. Por muitas críticas que os catálogos de assuntos tradicionais suscitem, eles são oásis na *Internet*. As classificações universais, por exemplo, dada a sua natureza de sistemas gerais de ordenação hierárquica, a sua alargada difusão internacional e a sua independência das línguas, são um instrumento que tem sido considerado adequado e utilizado com sucesso. Basta visitar, por exemplo, o site *Beyond Bookmarks: Schemes for Organizing the Web*⁹, para se obter uma extensa lista de sites onde os recursos da *Internet* são indexados com os instrumentos de organização de conhecimento que tradicionalmente são usados em bibliotecas.

Ou seja – parafraseando uma veterana especialista de sistemas de classificação bibliográfica – podemos, apesar da incerteza e assumindo a transitoriedade, começar por meter «vinho novo nas nossas velhas garrafas».

Notas

1 Surgidos respectivamente em 1876 e 1895, os sistemas de classificação Dewey e CDU são ainda os mais utilizados a nível mundial. A *Dewey Decimal Classification (DDC)*, utilizada com maior incidência nas comunidades e regiões anglo-saxónicas, tem uma implantação mais alargada e, conseqüentemente, mais traduções do que a *Classificação Decimal Universal (CDU)*, cuja utilização é mais frequente nos países e línguas europeus. A DDC, actualmente na 21.ª edição, é da responsabilidade da Forest Press, subsidiária da OCLC (<http://www.oclc.org>), sendo mantida editorialmente com o apoio da Library of Congress (<http://www.loc.gov>). Ao longo de décadas da responsabilidade da FID, a CDU é actualmente gerida pelo UDC Consortium, fundado em 1991, pela FID e editores da CDU de diversos países. Em Portugal, o sistema de classificação quase exclusivamente utilizado é a CDU, remontando a 1898 a sua primeira aplicação bibliográfica publicada, e a 1950 a criação de uma comissão nacional para a preparação de uma versão portuguesa abreviada do sistema, que veio a ser publicada em 1961. Actualmente, a entidade oficialmente licenciada para edição da CDU em Portugal é a Biblioteca Nacional, que tem em preparação uma versão electrónica deste sistema de classificação.

2 O mais universal – e mais conhecido – destes sistemas é o *LCSH – Library of Congress Subject Headings*, com uma história e evolução de cerca de 100 anos, cuja filosofia e/ou conteúdo ainda hoje, directa ou indirectamente, é a base de muitos sistemas de indexação utilizados em bibliotecas gerais de investigação de todo o mundo, com especial relevância para o âmbito das bibliotecas nacionais. Em Portugal, o sistema de indexação SIPORbase, desenvolvido pela Biblioteca Nacional no final dos anos 80, foi concebido tendo em conta a experiência do sistema LCSH, entre outras fontes, e contou com a colaboração de técnicos da própria Library of Congress.

3 Essas directrizes, intituladas *Guidelines for the establishment and development of monolingual thesauri* surgiram primeiro no âmbito da UNESCO (UNISIST), em 1970. Mais tarde foram publicadas no âmbito da ISO, com o mesmo título, em 1974, com o número ISO 2788. A versão mais recente da ISO 2788 é de 1986, e a correspondente norma portuguesa é de 1992, com a identificação NP 4036.

4 ISO 5963. 1985 – *Documentation. Méthodes pour l'analyse des documents, la détermination de leur contenu et la sélection des termes d'indexation*. Genève: ISO, 1985. A origem deste documento normativo remonta a 1975, quando a UNESCO publica *Principes d'Indexation*, mais tarde assumido pela ISO, sob a identificação ISO 5963 1981. A correspondente norma portuguesa é de 1989, com a identificação NP 3715.

5 Maria Inês LOPES – *Subject Authority Control in a Union Database: Perspectives for a national co-operative Project in Portugal*. Knowledge Organization. Frankfurt: ISKO. ISSN 3-0943-7444. Vol. 20, n.º 3 (1993), p. 133-138.

6 Foi no ano de 1997 decidido no âmbito da IFLA, pelo Permanent UNIMARC Committee, o desenvolvimento desse formato, após um estudo prévio, realizado por Elaine WOODS – *Requirements for a format for classification data. Final report*. [S. l.: s. n.]: IFLA Sections on Classification and Indexing and Information Technology, Joint Working Group on a Classification Format, July 1996.

7 IFLA Working Group on Principles Underlying Subject Heading Languages – *Principles Underlying Subject Heading Languages (SHLs)*. Final Draft. Edited by Maria Inês Lopes and Julianne Beall. [S. l.: s. n.], March 1997.

8 Roads – *Resource Organization And Discovery in Subject-based services*. Endereço na www: <<http://www.ukoln.ac.uk/roads/>>

9 Endereço na www: <<http://www.public.iastate.edu/~CYBERSTACKS/CT>>

Bibliografia

ALLEN, Bryce – *Cognitive Abilities and Information System Usability*. Information Processing & Management. New York: Pergamon Press Ltd. ISSN 0306-4573. Vol. 30, N.º 2 (1994), p. 177-191.

AMERICAN SOCIETY OF INDEXERS – *Indexing the Web* [online]. [Acedido em 1997.11.14]. Disponível na www: <<http://www.well.com/user/asi/webindex.htm>>

BIES, Werner – *Thinking with the help of Images: On the Metaphors of Knowledge Organization*. Knowledge Organization. Frankfurt: ISKO. ISSN 3-0943-7444. Vol. 23, n.º 1 (1996), p. 3-8.

CAYWOOD, Carolin – *Library Selection Criteria for WWW Resources* [online]. [Acedido 1997.11.14]. Disponível na www: <<http://www6.pilot.infi.net/~carolyn/criteria.html>>

Cognitive Paradigms in Knowledge Organization. Proc. of the Second International ISKO Meeting, Madras, 26-28 August 1992. Madras: Sarada Ranganathan Endowment for Library Science, 1992.

COLLINS, Harry M. – *Humans, machines and the structure of knowledge*. Stanford Electronic Humanities Review. ISSN 1048-3721. Vol. 4, Issue 2: *Constructions of the mind* [online]. Updated 20 July 1995. [Acedido em 1997.11.14]. Disponível na www: <<http://shr.stanford.edu/shreview/4-2/text/collins.html>>

DAHLBERG, Ingetraut – *The Network of knowledge fields: Conceptual Systematisation in Action*. In *Cognitive Paradigms in Knowledge Organization*. Proc. of the Second International ISKO Meeting, Madras, 26-28 August 1992. Madras: Sarada Ranganathan Endowment for Library Science, 1992. p. 323-333.

EKBOTE, E. R.; SEKHAR, M. – *Cognitive Skills of Conceptualisation Process and Types of Concepts*. In *Cognitive Paradigms in Knowledge Organization*. Proc. of the Second International ISKO Meeting, Madras, 26-28 August 1992. Madras: Sarada Ranganathan Endowment for Library Science, 1992. p. 23-34.

FROHMAN, Bernd – *Cognitive paradigms and users needs*. In *Cognitive Paradigms in Knowledge Organization*. Proc. of the Second International ISKO Meeting, Madras, 26-28 August 1992. Madras: Sarada Ranganathan Endowment for Library Science, 1992. p. 36-50.

FUGMAN, Robert – *Subject Analysis and Indexing: Theoretical Foundation and Practical Advice*. Frankfurt/Main: INDEKS Verlag, 1993. ISBN 3-88672-500-6.

FUGMAN, Robert, ed. lit. – *Tools for Knowledge Organization and the Human Interface*. Proc. 1st International ISKO Conference, Darmstad, 14-17 August 1990. Frankfurt/Main: INDEKS Verlag, 1990. 2 vol. ISBN 3-88672-020-9; ISBN 3-88672-021-7.

- GARCIA MARCO, Francisco Javier; ESTEBAN NAVARRO, Miguel – *On Some Contributions of the Cognitive Sciences and Epistemology to a Theory of Classification*. Knowledge Organization Frankfurt: ISKO. ISSN 3-0943-7444. Vol. 20, n.º 3 (1993), p. 126-132.
- GILCHRIST, Alan, ed.lit. – *From Classification to «Knowledge Organization»: Dorking Revisited or «Past in Prelude»*. A collection of reprints to commemorate the forty year span between the Dorking Conference (First International Study Conference on Classification Research 1957) and the Sixth International Study Conference on Classification Research (London, UK) 1997. The Hague: FID, 1997. ISBN 92 66 00 714 5.
- GODERT, Winfried – *Information as a Cognitive Construction: A Communication-Theoretic Model and Consequences for Information Systems*. Knowledge Organization. Frankfurt: ISKO. ISSN 3-0943-7444. Vol. 23, n.º 4 (1996), p. 206-212.
- GREEN, Rebecca, ed. lit. – *Knowledge Organization and Change*. Proc. 4th International ISKO Conference, 15-18 July 1996, Washington, DC. Frankfurt/Main: INDEKS Verlag, 1996. ISBN 3-88672-024-1.
- HUTCHINS, W. J. – *Languages of Indexing and Classification: a Linguistic Study of Structures and Functions*. Stevenage: Peter Peregrinus, 1975. ISBN 090 1223 68 9
- KHO, Cristhopher Soo-Guan – *The Use of Relation Matching in Information Retrieval*. LIBRES: Library and Information Science Research [online]. ISSN 1058-6768. Vol. 7, Issue 2, Sept. 30 (1997). Disponível na www: <<http://aztec.lib.utk.edu/libres/>>
- LOPES, Maria Inês – *Principles Underlying Subject Heading Languages: an International Approach*. (Paper presented to the 61st IFLA Council and General Conference, Istanbul, 1995). International Cataloguing and Bibliographic Control. Frankfurt: IFLA UBCIM. ISSN 1011-8829. Vol. 25, n.º 1 (Jan.-Mar.1996), p.10-12. Disponível na www: <<http://www.nlc-bnc.ca/ifla/IV/ifla61/61-lop.htm>>
- LOPES, Maria Inês – *Subject Indexing in Portuguese Libraries: a new approach with SIPORbase*. In *Subject Indexing: Principles and Practices in the 90's*. Proc. of the IFLA Satellite Meeting held in Lisbon, Portugal, 17-18 August 1993. Munchen: K. G. Saur, 1995. ISBN 3-598-11251-3. p. 121-143.
- LOPES, Maria Inês – *The UDC in Portugal: Background and Prospects*. Extensions and Corrections to the UDC. The Hague: UDC Consortium. Vol. 17 (1995), p. 81-82.
- LYNCH, Clifford – *Searching the Internet: Combining the skills of the librarian and the computer scientist may help organize the anarchy of the Internet*. Scientific American [online]. March 1997. [acedido em 1997.11.14]. Disponível na www: <<http://www.sciam.com/0397issue/0397intro.html>>
- PALOWITZ, Casey; HOROWITZ, Lisa – *Meta Information Structures for Networked Information Resources*. Cataloging and Classification Quarterly. Binghamton: The Haworth Press, Inc. ISSN 0163-9374. Vol. 21, N.º 3/4 (1996), p. 1091-30.
- RIGGS, Fred W. – *Onomastics and Terminology – Pt.º I: Their Contributions to Knowledge Organization*. Knowledge Organization Frankfurt: ISKO. ISSN 3-0943-7444. Vol. 23, n.º 1 (1996), p. 25-30.
- Subject Indexing: Principles and Practices in the 90's*. Proc. of the IFLA Satellite Meeting held in Lisbon, Portugal, 17-18 August 1993. Munchen: K. G. Saur, 1995. ISBN 3-598-11251-3.
- SVENONIUS, Elaine – *The Impact of Computer Technology in Knowledge Representations*. In *Cognitive Paradigms in Knowledge Organization*. Proc. of the Second International ISKO Meetings, Madras, 26-28 August 1992. Madras: Sarada Ranganathan Endowment for Library Science, 1992. p. 1-14.
- TAUBES, Gary – *Indexing the Internet*. Science Online [Online]. Washington: American Association for the Advancement of Science. ISSN 1095-9203. Special Report Computers'95: Fluid Dynamics. [Acedido 1997.11.14]. Disponível na www: <<http://www.edoc.com/aaas/computers/webindex.html>>
- UMARANI, Asha – *Knowledge Classification – A Permanent Structure for Dynamic Knowledge*. In *Cognitive Paradigms in Knowledge Organization*. Proc. of the Second International ISKO Meetings, Madras, 26--28 August 1992. Madras: Sarada Ranganathan Endowment for Library Science, 1992. p. 212-222.
- WEINBERG, Bella Hass – *Complexity In Indexing Systems – Abandonment and Failure: Implications for Organizing the Internet*. In *ASIS 1996 Annual Conference Proceedings* [online]. [Acedido 1997.11.14]. Disponível na www: <<http://www.asis.org/annual-96/ElectronicProceedings/weinberg.html>>
- WILLIAMSON, Nancy J. – *Standards and Rules for Subject Access*. Cataloging and Classification Quarterly. Binghamton: The Haworth Press, Inc. ISSN 0163-9374. Vol. 21, n.º 3/4 (1996), p.155-176.