



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
Campus de Marília



CULTURA  
ACADÊMICA  
*Editora*

# Do Conceito e dos Princípios Metodológicos de Indexação. Subsídios para uma Análise Crítica

Maria Da Graça De Melo Simões

**Como citar:** SIMÕES, M. da G. de M. Do Conceito e dos Princípios Metodológicos de Indexação. Subsídios para uma Análise Crítica. *In:* JORENTE, M. J. V.; PADRÓN, D. I. (org.). **Una Mirada a la ciência de la información desde los nuevos contextos paradigmáticos de la posmodernidad.** Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2017. p. 151-178.  
DOI: <https://doi.org/10.36311/2017.78-85-7983-904-7.p151-178>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported.

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição - Uso Não Comercial - Partilha nos Mesmos Termos 3.0 Não adaptada.

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 3.0 Unported.

## CAPÍTULO 6

### DO CONCEITO E DOS PRINCÍPIOS METODOLÓGICOS DE INDEXAÇÃO. SUBSÍDIOS PARA UMA ANÁLISE CRÍTICA

*Maria da Graça de Melo Simões*

#### 1 CONSIDERAÇÕES GERAIS SOBRE O PRELÚDIO DE UMA PRÁTICA

De um modo linear e redutor pode afirmar-se que a prática da indexação por assunto surge associada à necessidade que é sentida pelo ser humano de organizar a informação com vista à sua localização e recuperação, de um modo célere e preciso, com o objetivo de transferir e construir um novo conhecimento. Esta operação, na maioria dos casos é realizada em serviços de informação, em particular em bibliotecas e arquivos, por profissionais especializados aos quais são exigidas competências técnicas, cognitivas e outras habilitações de índole intelectual e pessoal, como possuem um espírito analítico-sintético que lhes permita realizar o processo dinâmico e concomitante da seleção/omissão concetual, meio através do qual se efetiva a indexação. Neste sentido, pode afirmar-se que a indexação é um processo de escolha concetual e terminológica.

A indexação por assunto, tarefa capital do tratamento técnico de um documento, numa perspetiva concetual ampla contextualiza-se na

<https://doi.org/10.36311/2017.78-85-7983-904-7.p151-178>

organização do conhecimento; quando considerada num domínio conceitual restrito insere-se na representação da informação por assunto, que se formaliza na atribuição de termos aos conceitos de um documento, que foram selecionados recorrendo-se para o efeito a linguagens de indexação, como é o caso dos sistemas de classificações (ORTEGA; LARA, 2010).

As provas documentais legadas pela história são um testemunho credível para se afirmar que foi no âmbito da representação da informação por assunto que esta prática se desenvolveu. Fê-lo, umas vezes deixando cair determinados procedimentos, outras introduzindo novos, procurando nesta dialética de rejeição/ inovação, atualizar-se e recriar-se, de modo a responder às novas exigências informacionais da sociedade. Na conjuntura histórica, entre outros acontecimentos que concorreram para estes momentos de metamorfoses e inovação, salientam-se: a descoberta do papel (105 a.C.), o aparecimento da imprensa (séc. XV), a emergência das publicações periódicas (séc. XVII-XVIII), a mecanização da imprensa (séc. XIX) e a informática (séc. XX), em particular a Internet. Por tudo isto pode afirmar-se que a indexação por assunto se caracteriza por ser uma prática dinâmica que se ajusta às conjunturas sociais e se pode considerar uma prática ainda milenar. No que se refere à primeira particularidade, é-o na medida em que se vai adaptando às alterações e inovações que, naturalmente, concorrem para novas necessidades informacionais. Fá-lo, por um lado através da adoção de novos processos e da elaboração de novos produtos moldados às necessidades informacionais emergentes de quem os procura, por outro através do desenvolvimento e ajustamento dos seus princípios metodológicos. Apesar de durante séculos ser considerada uma atividade exercida com fins exclusivamente práticos, a indexação por assunto nunca deixou de ter bases metodológicas, assentes em fundamentos empíricos e arbitrários. Como irá observar-se, os princípios sistematizados surgirão apenas em finais do século XIX. Tal como se depreende da designação, a segunda característica mencionada, tem a ver com o facto de ser uma prática que se exerce ao longo de milénios.

A elaboração de produtos documentais ao longo deste período, embora num estado muito incipiente, que formalmente pouco ou nada

tem a ver com os atuais<sup>84</sup>, é uma prova documental irrefutável e indelével da existência desta prática. Neste âmbito contam-se os resumos, os índices e os catálogos de assuntos, elaborados pelos produtores dos textos originais ou por quem detinha a sua guarda.

Segundo Witty (1973), é nas civilizações da Mesopotâmia (segundo milénio a.C.) que se encontram as primeiras anotações aos textos escritos em argila. Este procedimento é também observado na civilização egípcia, nesta altura em textos de pergaminho e papiro<sup>85</sup>. De acordo com o mesmo autor, o seu objetivo era a imediata recuperação da informação, quando ela se encontrava em textos longos.

Alavancada nas mesmas razões e prosseguindo os mesmos objetivos, esta prática veio também a observar-se nas civilizações romana e grega. Foi através dos “resumos”<sup>86</sup>, muitos dos quais depositados nas bibliotecas de Pérgamo e de Alexandria, que foi possível construir a história destas civilizações (SHOLNIK, 1979). A prática de representar a informação condensada estendeu-se à Idade Média (desenvolvida sobretudo em mosteiros), afirmando-se na Idade Moderna com o florescimento das Academias científicas e das revistas científicas<sup>87</sup>. Continuou a crescer e a especializar-se até à atualidade<sup>88</sup>. Outros testemunhos dela são os índices que no Médio Oriente, remontam à era pré-cristã e os catálogos de assuntos, que tiveram ao longo dos tempos um desenvolvimento muito similar aos dos resumos e aos dos índices, na medida em que se baseiam, de uma forma geral, na mesma prática e prosseguem o mesmo objetivo: a localização e recuperação da informação.

A primeira notícia conhecida sobre “catálogos de assuntos” data de 1300 a. C., na Assíria, onde foram encontrados, em tábuas de argila,

---

<sup>84</sup> Durante séculos os catálogos e os índices não passaram de listas ou inventários de assuntos e os resumos eram breves anotações às margens do texto ou descrições, mais ou menos extensas, do conteúdo de um documento.

<sup>85</sup> Tal prática prendia-se com questões de preservação e segurança dos documentos, e com o seu tamanho e formato (SIMÕES, 2015a).

<sup>86</sup> Descrições abreviadas do conteúdo dos documentos com uma estrutura indefinida.

<sup>87</sup> Segundo Kronick (1962) *Le Journal des sçavans* é considerado o primeiro periódico (1665) onde aparece este tipo de documento num formato que pode considerar-se um prelúdio do atual.

<sup>88</sup> Para um maior desenvolvimento sobre o assunto ver: SIMÕES, M. G. M. Resumo Documental: um incursão à (des)construção concetual na literatura científica, 2015b). Disponível em: <<http://dx.doi.org/10.14195/978-989-26-1035-1>>. Consultado em: 14 dez. 2015.

registos de títulos de livros e informações relativas ao seu conteúdo, considerando-se estas anotações rudimentares como o seu prelúdio. Outro registo meritório é o Catálogo de Calímaco, na Biblioteca de Alexandria, que se encontrava organizado alfabeticamente por autores e assuntos gerais. A Idade Média e o Renascimento foram pródigos na elaboração de catálogos. Na primeira, contribuiu para isso a atividade intelectual desenvolvida nos mosteiros<sup>89</sup>. Na segunda, com o advento do Renascimento a invenção da imprensa de Gutenberg, que trouxe o livro impresso e o consequente aumento do número de exemplares, contribuindo para uma maior disseminação da informação e o desenvolvimento das práticas do tratamento técnico dos documentos, imprescindíveis para a organização do conhecimento. Neste contexto proliferaram os catálogos impressos de livreiros e outros tipos de catálogos, nomeadamente nas instituições de ensino, como é o caso das universidades.

## 2 UMA ABORDAGEM DA INDEXAÇÃO POR ASSUNTO NO CONTEXTO CUSTODIAL

A necessidade de recuperar a informação de um modo condensado e rápido foi uma constante ao longo do tempo, tal como dá conta Collison (1969), numa breve história sobre índices e indexação. A prática da construção de produtos através dos quais se acesse à informação (índices, catálogos, resumos, etc.) veio a consolidar-se essencialmente entre finais do século XVIII e meados do século XX, embora apresentassem uma estrutura indefinida e carecessem de fundamentos teóricos. Para a sua elaboração contribuíram a produção vertiginosa da documentação, o acesso e o consumo da literatura científica e outras razões de natureza técnica.

Ao longo deste período, esta prática caracterizou-se por ser essencialmente operacional, movida por razões pragmáticas, baseada em técnicas pontuais e empíricas e destituída de fundamentos epistemológicos, como atesta a parca bibliografia de base teórica produzida no período considerado. Caracterizava esta bibliografia o facto de ser, geralmente, associada a necessidades empíricas da organização da informação, normalmente em

---

<sup>89</sup> Neste contexto apontam-se os catálogos do Convento de St. Martin em Dover (1389) e o do Mosteiro de Syon (Inglaterra, séc. XVI).

arquivos e bibliotecas, e não a objetivos epistemológicos, como constata a publicação das Regras do catálogo-dicionário de Cutter, as classificações bibliográficas enciclopédicas e as listas de encabeçamentos de matérias<sup>90</sup>.

Num contexto em que a materialização do documento (físico e manuseável) assume uma importância capital, dado o facto de este assumir um valor probatório irrefutável e patrimonial, aspetos muito considerados no paradigma positivista<sup>91</sup>, concorre para a parca valorização da informação nele contido. Deste modo, compreende-se que a prática da indexação tenha sido pautada por razões de natureza administrativa e cultural, bem como por motivos de preservação e conservação do próprio documento. Contudo, a partir dos inícios do século XX, o desenvolvimento do conhecimento científico, a sua produção e consumo, obrigavam a construção de documentos secundários, construídos com bases técnicas e teóricas mais sólidas, que permitissem uma maior divulgação, localização e acesso ao documento primário<sup>92</sup>. Deste modo impõe-se uma abordagem integralista do documento que considerasse o documento na sua dimensão física e na informacional. Esta nova visão partilhada e enfatizada à época por alguns estudiosos, como Henri de la Fontaine e Paul Otlet, este último na obra *Traité de Documentation* (1934), veio contextualizar uma teoria e uma prática documental orientada para a difusão, partilha e cooperação da informação.

Neste âmbito, afirmam-se os produtos secundários resultantes da indexação, como os índices, agora realizados sob consistentes orientações metodológicas, técnicas sistemáticas alavancadas em motivos de natureza pedagógica e científica<sup>93</sup>.

---

<sup>90</sup> Classificação Decimal Dewey; Classificação da Biblioteca do Congresso, Classificação Decimal Universal e Lista de encabeçamentos da Biblioteca do Congresso.

<sup>91</sup> Tal posição advém do facto de ele ser o elemento por excelência que legitima a emergência de novos estados e que é a base para a construção da memória e identidade cultural e nacional. Por esta razão, a partir dos finais do séc. XVIII, em particular após a Revolução Francesa, começam-se a valorizar e a investir em bibliotecas e arquivos, não só como depósitos de documentos, mas também como centros de tratamento da informação, cuja finalidade era “[...] servir interesses historiográficos e culturais” (RIBEIRO, 2004).

<sup>92</sup> Neste contexto, bibliotecas e arquivos passam de meros depósitos de conservação de documentos a entidades produtoras embora, como refere (RIBEIRO, 2004), tal situação sempre “[...] ocorreu por razões instrumentais e operativas, servindo os interesses dos seus utilizadores ... em regra, as respetivas entidades produtoras”.

<sup>93</sup> No século XVIII, numa tentativa de se ajustarem aos novos quadros mentais emergentes, a apresentação por

Para permitir uma seleção célere e um breve acesso à informação, os periódicos começaram a incluir índices. O que foi uma prática incipiente e casual, de início, passou a ser um procedimento frequente entre os séculos XVIII e meados do século XX, sendo usual haver periódicos que indexavam cada artigo pelo nome do autor e também pelo assunto específico (COLLISON, 1969)<sup>94</sup>. Um índice que atesta a evolução deste recurso documental é o *Poole's Index to periodical literature*<sup>95</sup>.

Nos finais do século XIX, na tentativa de domar a explosão bibliográfica, em grande parte causada pela imprensa mecanizada, que a par do desenvolvimento científico veio contribuir para um aumento das publicações periódicas, os índices expandiram-se a uma velocidade alucinante<sup>96</sup>. Eram considerados um dos instrumentos mais eficazes e credíveis para localizar a informação, mediadores privilegiados entre o documento e o utilizador. Entre os vários tipos de índices destacam-se os de assunto, pela sua função na recuperação da informação e complexidade inerente à sua elaboração. Segundo Silva e Fujita (2004), é importante considerar-se o índice de assuntos como um dos instrumentos de pesquisa mais relevantes. São-no, na medida em que um utilizador através de um termo ou conjunto de termos consegue localizar e aceder cirurgicamente à informação. Tal facto prende-se com dois postulados inerentes à sua elaboração: a

---

ordem sistemática do conhecimento, que prevaleceu ao longo da Idade Clássica e Medieval, é preterida em prol da ordem alfabética (SIMÕES, 2011), ordenação que vai influenciar a construção de índices de assuntos e catálogos alfabéticos de assuntos. Este século é considerado a primeira idade de ouro dos índices. Para uma localização mais fácil da informação, os livros escolares em Inglaterra (séc. XIX), passaram a incluir índices, o que concorreu para que este se autonomizasse relativamente a outros elementos estruturais de um livro, com funções idênticas, como é o caso do Sumário (*table*).

<sup>94</sup> É nos Estados Unidos e em Inglaterra que se vem a consolidar e a formalizar a indexação nas publicações periódicas. Além dos periódicos e dos livros escolares, destacam-se os índices dos livros de contas de grandes empresas comerciais e ainda as enciclopédias, entre as quais salienta-se a *Encyclopaediae Britannica*, pela qualidade que apresentavam.

<sup>95</sup> Este índice, criado em 1882 por William Frederick Poole, bibliotecário da Yale University, apresentava duas características inovadoras: era cumulativo e as entradas dos índices dos artigos eram palavras-chave extraídas dos títulos dos mesmos, prática retomada pelo *Reader's guide to periodical literature* (1901), onde cada artigo aparecia indexado pelo autor e pelo assunto. Ainda de acordo com Collison (1969), este primava por ser uniforme e preciso, na medida em que apresentava uma teia de remissivas que relacionava assuntos semanticamente associados entre si. Nesta prática, embora de modo incipiente, antevemos a dos índices estruturados.

<sup>96</sup> Assim, continuou a assistir-se à elaboração contínua de índices de todos os tipos: individuais, coletivos, nacionais e/ou em colaboração internacional, cumulativos, setoriais etc. Para a evolução dos índices salientam-se dois aspetos: o progresso das revistas científicas, que continuaram a ser a base privilegiada da sua construção e a emergência das novas tecnologias, que vieram ancorar este tipo de documentos: na elaboração, no armazenamento, na difusão e no acesso.

exaustividade e a especificidade concetual. Neste contexto, e porque a sua elaboração pressupõe uma indexação muito precisa e refinada, ela situa-se no estágio da indexação exaustiva (LANCASTER, 1993)<sup>97</sup>.

Dadas as suas características, como a ordem alfabética, a especificidade e a exaustividade, os índices são um excelente recurso na localização e acesso à informação, seja em ambientes tradicionais seja em digitais, nestes últimos, em especial no que se refere à localização de documentos em bases de dados. Apesar de existirem ferramentas de pesquisas em texto integral, contudo no que concerne às áreas das ciências sociais os índices têm um papel importante, dado que por questões semânticas há conceitos implícitos que escapam a estas ferramentas de busca.

Outro recurso de mediação entre o documento e o utilizador que se desenvolveu graças à emergência de novos instrumentos de representação da informação (listas de encabeçamento de matérias e classificações bibliográficas) foram os catálogos de assuntos<sup>98</sup>. Tal como aconteceu nos índices, o seu desenvolvimento deve-se ao aperfeiçoamento de técnicas metodológicas e à criação de um corpus de princípios sistemáticos, que subjazem à sua realização. A partir de meados do século XIX, os catálogos começaram a assumir um formato e uma função mais próxima da atual, relevando-se na sua estrutura o uso de remissivas e entradas analíticas. Estas listas, que normalmente não eram mais do que um inventário, passam agora a cumprir a função de localizar essas obras para um melhor acesso e divulgação da informação. Este foi um período no qual se consolidaram os catálogos de assuntos, quer os alfabéticos, quer os sistemáticos. Para tal contribuíram, por um lado as *Cutter's rules for a dictionary catalogue*

---

<sup>97</sup> Este tipo de indexação proporciona ao utilizador um número significativo de pontos de acesso, o que é um dos propósitos mais imediatos de um índice.

<sup>98</sup> *Grosso modo* pode afirmar-se que o acesso à informação através dos catálogos de assuntos pode efetuar-se por meio de duas vias: os catálogos alfabéticos de assuntos e os catálogos sistemáticos. Os primeiros caracterizam-se segundo Foskett (1977, p. 247) por “[...] contains subject entries and cross-references arranged alphabetically in on sequence”. O mesmo autor (*Ibidem*, 250) define o segundo tipo de catálogo como “Systematic arrangement brings related subjects together by using notation at its code vocabular [...] Entries consist of a heading, which in this case is a notational symbol”. O primeiro garante a recuperação da informação de modo específico, o segundo, organiza o conhecimento em grandes sistemas epistemológicos, permite a recuperação da informação de um modo muito geral. Neste sentido, Buchanan (1979, p. 10), refere o seguinte: “[...] Les systèmes de classification sont donc des langages documentaires qui permettent la représentation la plus générale. Par nature, ils facilitent le bouquinage et l'accès au non-spécialiste. Ao proporcionarem estes dois tipos de acesso irão completar as respostas às questões colocadas, facto que será um valor acrescentado para o utilizador (SIMÕES, 2011).

(1876), que vieram regular as entradas dos catálogos de assuntos, por outro a publicação da 1ª edição da Classificação Decimal de Dewey, no mesmo ano (FOSKETT, 1977) que veio a incentivar e a vulgarizar a construção do catálogo sistemático. Como expoente máximo dos catálogos de assuntos alfabéticos refere-se o Catálogo da Biblioteca do Congresso, elaborado com base na *List of Subject Headings for Use in Dictionary Catalogs* (1<sup>st</sup> ed., 1895; 2<sup>nd</sup> ed., 1898), à qual os catalogadores iam acrescentando novas entradas de acordo com as suas necessidades (STONE, 2000)<sup>99</sup>.

A título de síntese, relativamente às práticas de indexação no período considerado, infere-se que, embora assentes e focadas em fundamentos e motivos de natureza funcional, apresentam um desenvolvimento técnico e metodológico cujo propósito era ajustarem-se e darem resposta a uma sociedade motivada pelas ideias de progresso e valorização individual características da mentalidade positivista. Outro ponto a salientar é o facto de, pela primeira vez, se assistir à preocupação de compilar de forma sistemática um corpus de princípios e normas técnicas conducentes à representação da informação e à organização do conhecimento, com vista à uniformidade e consistência.

### 3 UMA ABORDAGEM DA INDEXAÇÃO POR ASSUNTO NO CONTEXTO PÓS-CUSTODIAL

A partir da segunda metade do século XX, as alterações ocorridas a nível sociocultural, com o crescimento da produção científica e com a revolução operada nas tecnologias, concorreram para uma nova abordagem da indexação por assunto, quer no que respeita ao aspeto prático, quer no que se refere ao aspeto teórico.

O contexto pós-custodial caracteriza-se, sobretudo pela desmaterialização do documento, a emergência da Ciência da informação (*Information science*), a interdisciplinaridade entre a Ciência da informação e outras áreas do saber (Ciências da computação), a aplicação das novas tecnologias da informação e da comunicação (TIC) em particular a *world*

---

<sup>99</sup> Ao longo do século XX, esta lista, agora com a nomenclatura *Library of Congress Subject Headings* (1909), foi uma referência na prática da indexação por assunto para uma tipologia alargada de bibliotecas, como pode observar-se pela citação que se segue: *The list was also considered appropriate for the very largest public libraries, some colleges and many university libraries* (STONE, 2000).

*wide web* (*www*) à Ciência a informação; o investimento na formação de profissionais da área da Ciência da informação (ensino superior) , a produção científica e académica sobre a identidade epistemológica da Ciência da informação<sup>100</sup> e a gestão dos fluxos de informação na perspetiva do uso generalizado.

*Grosso modo* este modelo desenvolve-se num contexto socio-tecnológico que vulgarmente se generalizou como sociedade da informação. Relativamente à indexação, este modelo pauta-se por dois momentos: um, que se localiza entre os anos sessenta e os anos noventa do século XX, e outro, que se situa dos alvares do século XXI até à atualidade. Caraterizam o primeiro momento as ideias de indexação automática (algoritmos, formulas estatísticas, linguística computacional), produção de literatura de natureza teórico-metodológica sobre indexação (textos normativos internacionais e textos de cariz epistemológico sobre os fundamentos e propósitos da indexação), aplicação das tecnologias emergentes a esta área específica (construção de linguagens, divulgação e acesso dos seus produtos). O segundo carateriza-se pela simbiose da indexação com os ambientes de informação digital (novos sistemas de organização de conhecimento (mapas concetuais, taxonomias, ontologias), indexação pós-coordenada (*uniterm*, descritores, tesauros), indexação colaborativa (folksonomias)<sup>101</sup> e pela semântica.

### 3.1 CONCEITO E PRINCÍPIOS METODOLÓGICOS

Para um melhor entendimento da evolução desta prática no decorrer destes dois paradigmas optou-se por apresentar uma breve reflexão crítica sobre a evolução do conceito e dos seus princípios metodológicos. Esta escolha deve-se ao facto de estes conceitos traduzirem, de uma forma direta e/ou indireta, os aspetos epistemológicos que a fundamentam e apresentarem a evolução dos principais pontos metodológicos que a conformam.

<sup>100</sup> RIBEIRO, F. (2004). Gestão da Informação/Preservação da Memória na era pós-custodial: um equilíbrio precário?

<sup>101</sup> Sistema que permite ao utilizador propor termos para representar determinado conceito. Ver: Wal, T. V. (2007). *Folksonomy*.

### 3.1.1 CONCEITO

Toda a dinâmica observada nas estruturas mentais e tecnológicas nos finais do século XIX e inícios do século XX, concorreu, não só, para o aperfeiçoamento das técnicas de indexação mas também para um investimento teórico, nunca antes verificado. Neste sentido, interessa relevar que a abordagem teórica da indexação é muito recente quando comparada com os quase quatro milénios de prática. Os primeiros estudos de sistematização teórica datam de finais do século XIX. Até então, na maioria dos casos esta prática era exercida de forma intuitiva, procurando dar respostas pontuais a questões levantadas mais pela documentação que pelas necessidades reais de informação dos utilizadores. Para o surgimento das preocupações teóricas e metodológicas da indexação, apontam-se entre outras razões as questões colocadas pelos utilizadores que reclamavam pesquisas mais céleres e específicas, em parte devido à especialização do conhecimento científico e da multidisciplinaridade que começavam a emergir. Neste contexto exigiam produtos que respondessem de forma cabal às questões que levantavam e instrumentos que lhes facultassem a realização de pesquisas com um nível elevado de pertinência.

Guimarães (2008), refere que a abordagem teórica à indexação é feita a partir de três aspetos: os processos, os instrumentos e os produtos, dado o facto de serem estes elementos que sustentam esta operação. De acordo com este autor (2009), é no final do século XIX e início do século XX, nos Estados Unidos, e na primeira metade do século XX, em Inglaterra, que se desenvolvem os estudos metodológicos relativos aos produtos e aos instrumentos da indexação. No final da década de 60 e início da de 70 do século referido desenvolveram-se em França os referenciais teórico-metodológicos relativos ao processo. Sobre as abordagens teóricas, o mesmo autor (2008), identifica três orientações que influenciaram a indexação: a *Subject cataloguing*, de influência norte-americana, a *indexing*, de influência inglesa, e a *Analyse documentaire*, de influência francesa. A primeira aceção é centrada na atividade profissional realizada em bibliotecas destacando-se o uso dos cabeçalhos de assunto. Na segunda tendência, orientada para os centros de documentação e para o mundo editorial, assume particular relevância o uso dos tesouros. A terceira, como se infere pela

própria designação, centra-se nos procedimentos para a identificação e a seleção dos conceitos, com o objetivo de criar produtos documentais. É no âmbito destas abordagens que se constrói e evolui o conceito de indexação.

Ao longo do século XX, os estudiosos que contribuíram com subsídios para o processo de construção epistemológica nesta área, encontram-se distribuídos por estas três correntes teóricas. Mais do que ligá-los individualmente a uma corrente, interessa saber o entendimento que têm desta noção e como a formalizam metodologicamente.

Gil Leiva (1999), apresenta uma definição de indexação que demonstra a sua complexidade. Fá-lo recorrendo a dois elementos: à indexação da informação contida no documento e à indexação das perguntas que são formuladas pelo utilizador, ideia também implícita em Chaumier (1986), quando refere que a indexação não só intervém na entrada de informação, mas também na sua saída (recuperação da informação). Segundo o primeiro autor, a indexação é um processo que consiste na análise e na seleção dos conceitos essenciais e no armazenamento dos mesmos em linguagem controlada ou natural, que permite recuperar a informação. De acordo com este autor concluiu-se que o principal objetivo da indexação é o armazenamento e a recuperação da informação. Próxima desta definição coloca-se Pinto Molina (1993), ao mencionar que a indexação é a técnica de caracterizar o conteúdo de um documento e/ou as perguntas documentais, extraíndo as ideias mais significativas para as representar por termos de indexação adequados, provenientes de uma linguagem natural empregue pelo autor, ou por uma linguagem controlada. Guinchat e Menou (1984), descrevem este processo colocando a tónica na representação, quando referem que ela é uma operação que permite escolher os termos mais adequados para representar o conteúdo de um documento. Por outro lado, Chaumier (1986), afirma que a análise da informação é a principal operação sem a qual é completamente impossível utilizar a informação.

Ainda no contexto concetual, pela importância que tiveram na prática da indexação, é de salientar a definição apresentada pelos *Princípios de indexação da UNISIST* (1975) e pela *Norma ISO 5963-1985* (1988, 4.1 e 4.2). O primeiro texto diz que a indexação é a operação que descreve e caracteriza o conteúdo de um documento através da representação dos seus

conceitos com a ajuda de uma linguagem documental; menciona ainda, que esses conceitos são extraídos por meio de um processo de análise. O segundo refere que a indexação não consiste na descrição de um documento enquanto entidade física e que através dela se extraem os conceitos dos documentos por um processo de análise intelectual, sendo convertidos em termos, recorrendo para tal a instrumentos de indexação.

A diversidade concetual observada nas definições expostas encontra-se igualmente presente quando se abordam as questões de natureza operacional, em particular no que se refere às etapas de realização. Tal como referem Brown, Langridge e Mills (1976), a indexação e a classificação baseiam-se no conteúdo dos documentos e o seu cumprimento passa pela aplicação de uma metodologia faseada. Partindo de uma revisão de literatura, Gil Leiva (1999), conclui que este processo pode desenvolver-se entre duas a oito etapas<sup>102</sup>.

Pelo exposto conclui-se que a indexação é um processo que consiste na análise conceitual, identificação de assuntos de um documento e tradução desses assuntos em termos representativos, que podem ou não ser controlados, dependendo da linguagem de indexação que se usa; além disso, tem como objetivo disponibilizar, com objetividade, o conteúdo dos documentos o pensamento dos autores, isto é: abrir caminho para que se venha a obter uma informação de qualidade (MENDES; SIMÕES, 2002).

Naturalmente que as ideias expressas se associam à noção de indexação contextualizada na indexação humana, que apresenta duas fases distintas e complementares: análise e representação. No que se refere à indexação automática, o processo é diferente. Segundo Lancaster (1993),

---

<sup>102</sup> Assim, para alguns autores, entre os quais Lancaster (1993, p. 8) e Langridge (1989, p. 98), as etapas da indexação efetuam-se a dois tempos, o primeiro divide o processo em: a) análise conceptual e b) tradução; o segundo em: a) análise do assunto e, b) tradução para termos de uma linguagem. No entanto, outros autores como J. Rowley (1982, 45-46), Cleveland e Cleveland (1990, 104-112) e a própria Norma ISO 5963 (4.3) dividem-no em três fases. J. Rowley refere as seguintes etapas: a) Familiarização; b) Análise; c) Conversão dos conceitos em termos de indexação. Os segundos autores mencionam: a) Análise de conteúdo; b) Determinação do assunto; c) Conversão dos assuntos para uma linguagem documental. A Norma ISO determina: a) Exame do documento e definições do seu conteúdo; b) identificação e seleção das noções principais do conteúdo; c) Escolha dos termos de indexação. Refere, contudo, que a primeira e a segunda, na prática se podem sobrepor. A metodologia apresentada pelos autores referidos é seguidora do que é postulado nos Princípios de indexação da *UNISIST* (1975, p. 2), quando registam que o processo de indexação se desenvolve em dois pontos, a saber: a) Reconhecimento dos conceitos contidos na informação dos documentos a indexar; b) Representação dos conceitos numa linguagem documental.

podem identificar-se duas modalidades: a) indexação por extração automática; b) indexação por atribuição automática<sup>103</sup>. Em qualquer dos casos descritos os *softwares* realizam a “análise” com base nos termos do texto e não nos conceitos. Nesta lógica, pode designar-se à indexação automática como uma indexação terminológica.

A partir de meados do século XX os produtos da indexação, nomeadamente os índices, podiam ser elaborados sem a intervenção humana, recorrendo-se a regras estatísticas, linguísticas e informáticas. Assim, o esforço intelectual despendido, sobretudo nas áreas das ciências puras e aplicadas, caracterizadas pelo uso de terminologia denotativa, que não concorre para entropias semânticas, dá lugar ao algoritmo. As técnicas de indexação automática que tiveram início na década de 50 do último século, com os trabalhos baseados na frequência de Luuhn e Baxendale, em 1957 e em 1958 respetivamente, foram-se reinventando para se ajustarem aos ambientes digitais do novo século.

### 3.1.2 PRINCÍPIOS METODOLÓGICOS

Tal como se observou relativamente à noção de indexação, também o conceito de metodologia, enquanto um conjunto de fases e normas para regular uma prática, é uma ideia contemporânea. Até finais do século XIX esta prática baseava-se num conjunto de princípios avulsos e pontuais elaborados com base no bom senso de quem indexava. Comprova esta afirmação o facto de os assuntos entrarem, na maioria dos casos, pelo genérico, sem se observarem regras de uniformidade e consistência formal<sup>104</sup>. Princípios como a exaustividade e a especificidade dos assuntos não eram tidas em conta, pelo menos de um modo sistemático. Para que tal circunstância se verificasse era necessário que a indexação privilegiasse a análise documental, situação que apenas vem a ocorrer no século XX, em função da especialização do conhecimento científico e de um utilizador que pro-

<sup>103</sup> Para um maior entendimento deste assunto ver os conceitos e a contextualização dos modelos: *SLIC* (*Selective listing in combination*) e o *PRECIS* e ainda os métodos *KWIC* (*Keyword in context*) e *KWOC* (*Keyword out of context*). Ver: LANCASTER, F. W. (1993). *Indexação e resumos: teoria e prática*.

<sup>104</sup> Em parte, esta situação denuncia a herança legada pelo catálogo dicionário, cuja entrada dos assuntos era feita pelos termos genéricos e/ou pelas primeiras palavras significativas do título do documento a indexar a *catchword titol* (HANSON; DAILY, 1958) no caso de não existir autor.

cura o específico em detrimento do geral. O compromisso entre utilizador e a área científico-técnica trouxe consigo um elevado nível de exigência no que se refere à pertinência na recuperação da informação por assunto, posição que vai concorrer para repensar, entre outros pontos, o *modus operandi* e os princípios metodológicos. Tal posição irá repercutir-se nos produtos daí resultantes, em particular nos índices e nos catálogos.

Neste contexto, pode afirmar-se que a exaustividade e a especificidade, princípios metodológicos que determinam e garantem, em grande medida, a qualidade dos catálogos de assuntos, essencialmente a partir da década de 60 do séc. XX se encontram alavancadas sobretudo na evolução tecnológica e científica e num utilizador altamente especializado que prefere os subtemas em detrimento dos temas, a qualidade da informação em prol da quantidade que é simultaneamente consumidor e produtor de informação, e que procura bases de dados e catálogos de serviços de informação especializados. Esta nova atitude concorre para que se proceda à dissecação dos conteúdos dos documentos tornando-se o profissional da informação num anatomista do saber (SIMÕES, 2008). Esta disposição indicia uma mudança de paradigma, focada no novo perfil de utilizador e no uso da informação, posição diversa da vigente, cunhada no positivismo e centrada na custódia do documento<sup>105</sup>.

Estes dois princípios não emergiram em meados do século XX de forma natural. Cutter, na obra *Rules for a dictionary catalogue* (1876), no capítulo III *Subjects*, ao apresentar um conjunto de orientações metodológicas para a entrada dos assuntos num catálogo, já se refere à especificidade<sup>106</sup>. No que respeita a este princípio, Cutter preconiza que os assuntos deverão dar entrada no catálogo pelo específico. Contudo, esta recomendação baseia-se mais em razões de ordem prática e económica do que em fundamentos teóricos. De acordo com Foskett (1973), este autor nunca assumiu a especificidade de modo efetivo, na medida em que recomendava que se fizessem ocasionalmente remissivas do específico para o geral, para superar a carência de especificidade (CUTTER, 1876); todavia, não pre-

---

<sup>105</sup> Esta nova atitude concorreu para a criação de novas linguagens de indexação com um elevado nível de estruturação dos termos, como foi o ca-\*so dos tesouros. 49

<sup>106</sup> Cutter refere ainda outros dois princípios: o do uso e o sindético.

conizava o procedimento inverso, isto é: estabelecer uma remissiva do geral para o específico<sup>107</sup>. De entre os princípios enunciados por Cutter, este foi o que mais se desenvolveu e consolidou, tendo-se fortalecido e afirmado depois da década de 70, sendo uma referência permanente nas obras de alguns dos mais consagrados teóricos da indexação como Foskett (1973), Haykin (1985), Rowley (1982), Chaumier (1986), Langridge (1989) e Lancaster (1993).

Outro princípio metodológico preconizado por Cutter, que continua a ser considerado um pilar na prática da indexação no paradigma pós-custodial, mesmo em ambientes de informação digital, é o do uso, também designado por princípio da conveniência<sup>108</sup>, apesar da sua natureza dedutiva e subjetiva e de todas as dificuldades sentidas na determinação do perfil do utilizador, agravadas com a emergência da Internet. Atesta a sua importância a produção de literatura científica e académica sobre o comportamento dos utilizadores face ao uso da informação e à determinação das suas áreas de interesse. Acresce ainda a alusão que dele fazem alguns dos mais conceituados teóricos,

83que veem nele um alicerce desta prática. A título de exemplo referem-se, entre outros, os já mencionados no que respeita ao princípio da especificidade. A mesma importância lhe é dada pela *Norma ISO 5963* nos pontos 6.3 e 6.4, quando postula que a análise deverá ser a mais exaustiva e específica possível, para responder não apenas ao utilizador que se conhece, mas ao potencial utilizador; de tal modo, os assuntos periféricos devem também ser considerados.

O princípio sindético, tal como Cutter o preconizou, apresentava mais razões de natureza económica, funcional e empírica do que teóricas. As remissivas eram usadas essencialmente em duas situações: para colmatar a distância entre os assuntos que tinham entre si relações semânticas que, por motivos de alfabetação, se encontravam dispersas no catálogo, e nas entradas dos assuntos específicos. As remissivas foram-se alterando ao lon-

---

<sup>107</sup> Apesar de não se observar esta situação, e o princípio subjacente à elaboração deste tipo de remissivas não ser o mesmo que hoje as norteia, antevê-se neste procedimento o prelúdio das relações hierárquicas.

<sup>108</sup> Entre dois ou mais termos que podem representar um conceito deve escolher-se aquele que é mais familiar ao utilizador.

go do tempo, quer no que respeita à nomenclatura, ao princípio e à função<sup>109</sup>. Para tal contribuíram as tecnologias de informação que trouxeram consigo a solução para a questão das distâncias semânticas e dos problemas de alfabetação, assim como os estudos teóricos desenvolvidos sobre este assunto, que vieram permitir o seu desenvolvimento de um modo sustentável, na ótica de um novo perfil de utilizador. O princípio sindético veio potenciar o estabelecimento das relações semânticas entre os assuntos, em particular as hierárquicas e as associativas, tipo de relações adotadas quer pelas estruturas de organização do conhecimento tradicional (tesauros e listas de encabeçamentos de matérias), quer nas surgidas no âmbito da informação digital (ontologias e taxonomias).

Embora os princípios enunciados por Cutter não tenham constituído uma base teórica para a construção dos catálogos de assuntos e índices, elas foram durante muitos anos, a base da prática norte-americana na construção de catálogos de assuntos, encontrando-se formalizados na lista *Sears* e na *Library of Congress Subjects Headings*, (Foskett, 1973) e também o prelúdio de determinados recursos metodológicos aplicados à construção de vocabulários controlados e estruturados, quer em ambientes tradicionais, quer em ambientes de informação digital.

Apesar das diferenças entre os princípios metodológicos apresentados têm em comum o facto de nortearem uma prática de indexação uniforme e precisa. Em vez de se observar uma rutura epistemológica, observa-se uma linha de continuidade relativamente aos pressupostos e às práticas de indexação consideradas nos dois paradigmas, na medida em que são resultado de um processo evolutivo, quer no que respeita às questões empíricas, quer às epistemológicas. Assim, quando se fala de princípios metodológicos, pela repercussão que assumiram nesta prática, não podem deixar de referir-se os trabalhos de J. Kaiser, S. R. Ranganathan e E. J. Coates. O primeiro, na sua obra *Systematic Indexing* (1911) apresenta uma metodologia para a análise e a disposição dos assuntos compostos. Esta ordem baseava-se em três categorias: *concreto*, *processo* e *lugar*, na qual a primeira assume a primazia. O segundo, na obra *Classified Catalogue*

---

<sup>109</sup> A nomenclatura alterou-se para relações semânticas, com a emergência dos tesauros e dos *softwares* aplicados à construção e gestão dos vocabulários de indexação, passando a função empírica a ser de natureza teórico/prática.

*Code with Additional Rules for Dictionary Catalogue Code* (1958) concebe um sistema baseado em cinco categorias: personalidade, matéria, energia, espaço e tempo (PMEST), base para a análise dos assuntos e determinação da ordem dos cabeçalhos compostos, recurso metodológico muito eficaz na análise de assunto por facetas, na construção de tesouros e classificações facetadas. O último, E. J. Coates, tomando como ponto de partida as contribuições de Cutter, Kaiser e Ranganathan, definiu um conjunto de princípios normativos para os encabeçamentos de matérias, com o propósito de uniformizar os encabeçamentos da *British National Bibliography*. Estes princípios metodológicos formalizam-se na obra *Subject Catalogues: headings and structure* (1960). A sua teoria metodológica, que foi aplicada no *British Technology Index*, incide essencialmente sobre a formulação da ordem de cabeçalhos compostos de assuntos, e baseia-se em quatro categorias: *coisa, parte, material, ação*.

A contribuição das teorias destes três autores recai sobretudo na construção das linguagens de indexação, em particular no que se refere ao uso das remissivas<sup>110</sup> e das facetas, sendo a sua referência apresentada umas vezes de forma implícita outras de forma explícita. A noção de faceta, também designada por categoria, assumiu um papel importante na indexação, tanto na análise, como na representação concetual<sup>111</sup>. Os seus fundamentos, concebidos em ambientes tradicionais e associados às classificações bibliográficas, na atualidade são revisitados e ajustados aos novos sistemas de organização de conhecimento como as ontologias e as taxonomias, sendo nas últimas o seu pilar estrutural. O mesmo se observa no que diz respeito às remissivas (relações semânticas) que estão mais desenvolvidas nos tesouros e nas ontologias, nestas últimas com o propósito de se atingir a formalização e especificação caraterís-

---

<sup>110</sup> Embora tenham sido utilizadas pelos três autores, Ranganathan foi o mais parco no seu uso.

<sup>111</sup> A faceta no processo de análise tem como função interrogar um documento sob um determinado número de perspetivas, concorrendo deste modo para a sua desconstrução concetual. Devido a tal circunstância, o assunto pode ser representado de forma específica e precisa. No âmbito da indexação pré-coordenada, a noção de faceta postulada por estes teóricos tinha como propósito contribuir para a construção de encabeçamentos de assuntos compostos, em particular no que se refere à disposição dos assuntos específicos que os constituíam. Este princípio veio a desenvolver uma técnica nova na construção das linguagens documentais, as classificações facetadas, das quais se salienta a Classificação Colon e, mais tarde, os tesouros facetados, cuja construção se baseou no princípio das classificações facetadas e no modelo dos tesouros tradicionais. Deste tipo de tesouros salienta-se o *Thesaurofacet : a thesaurus & faceted classification for engineering & related subjects*, publicado em 1969. De acordo com Foskett (1973, 352) este tesouro não é mais do que dois instrumentos juntos num único: uma classificação e um tesouro que funcionam numa dinâmica de complementaridade.

ticas deste tipo de estruturas de organização do conhecimento dos ambientes de informação digital<sup>112</sup>.

Após a emergência das novas tecnologias, outros teóricos refletiram nas suas obras a importância dos princípios na indexação; entre eles destacam-se Fugmann (1993), que apresenta cinco axiomas.<sup>113</sup> Lancaster (1993), vê na maioria destes axiomas um conjunto de elementos que poderão condicionar o desempenho dos sistemas de recuperação e propõe dois princípios para o processo de indexação: a) que inclua todos os tópicos reconhecidamente de interesse para os usuários do serviço de informação que sejam tratados substancialmente no documento; b) que indexe cada um desses tópicos tão especificamente quanto o permita o vocabulário do sistema e o justifiquem as necessidades ou interesses dos usuários. Relativamente ao enunciado da alínea a) interessa referir que, as alterações trazidas pelas novas tecnologias no que se refere aos quadros cognitivos, à partilha de conhecimento nas redes de informação, à crescente permuta e cooperação dos vários serviços e à interdisciplinaridade e transdisciplinaridade nas investigações, concorrem para que esta recomendação não seja a mais adequada, na medida em que não se pode apenas pensar nos interesses do utilizador de um serviço particular.

A completar o conjunto de princípios metodológicos referidos, Mendes e Simões (2002), com base num documento proposto pela IFLA (1999) - *Principles Underlying Subject Languages*, recomendam que se devem ter em conta a qualidade da análise, a simplicidade formal, a coerência e a uniformidade<sup>114</sup>.

A automatização e informatização da indexação trouxeram novos princípios metodológicos. De eles destacam-se os preconizados do Sistema de indexação PRECIS<sup>115</sup>, que preceituava a análise dos assuntos de acordo

<sup>112</sup> Ver: Gruber, T. R. (1993). *A Translation Approach to Portable Ontology Specifications*.

<sup>113</sup> São os seguintes: definição, ordenação, grau de ordenação, previsibilidade e fidelidade.

<sup>114</sup> A *qualidade da análise do conteúdo* - extrair de forma fidedigna a ideia do autor; a *garantia de uso* - os termos devem ser definidos a partir da forma como são utilizados na comunidade dos utilizadores; a *simplicidade formal* - representar o termo de indexação de forma simples; a *coerência e uniformidade* - usar sempre o mesmo critério para a mesma situação ou situações análogas e fazer corresponder ao mesmo conceito o mesmo termo, por forma a evitar a ocorrência da ambiguidade no momento da pesquisa; as autoras reforçam que a escolha dos termos é uma decisão terminológica como tal, a precisão dever-se-á ter sempre presente.

<sup>115</sup> Sistema concebido por D. W. Austin e usado pela primeira vez em 1971 pela *British National Bibliography*

com duas categorias: *entidades* e *atributos*; baseia-se ainda num conjunto de relações sintáticas e semânticas, as últimas em número de quatro. Pela importância que vieram a ter na construção dos vocabulários de indexação, destacam-se as relações semânticas, em particular as de tipo hierárquico - relações género-espécie, a relação atributiva assente no binómio entidade e atributo e a relação foco-diferença, que irá corresponder ao termo e ao modificador, sendo o *termo* o foco e o *modificador* a diferença. Compunha ainda este sistema um conjunto de operadores.

Entre os vários desafios que as novas tecnologias trouxeram salienta-se a cooperação bibliográfica, centrada na migração de dados bibliográficos e que ocorreu primeiro a nível interno depois a nível internacional. Neste contexto, muitos foram os projetos e programas informáticos que se desenvolveram a partir da década de 70 do século XX<sup>116</sup>, entre os quais o projeto *MARC*, que entre outras vantagens veio potenciar a normalização nacional e internacional, em particular no que se refere às orientações metodológicas para a estrutura dos conteúdos, meio através do qual apenas é possível serem legíveis por máquina e, por isso, poderem ser localizados e recuperados.

No âmbito deste espírito normativo surgiram também os *Princípios de indexação*, que foram publicados pela UNISIST em 1975. A designação não deixa margem para dúvidas relativamente ao objetivo; tratava-se de um texto normativo, que apresentava um conjunto de princípios metodológicos que deveriam nortear a prática da indexação. Dada a sua flexibilidade, este texto era adaptável ao maior número possível de casos, concorrendo, deste modo, para a consistência e uniformidade da indexação, facto que iria, naturalmente, resultar numa mais-valia para os resultados da pesquisa, que se desejam pertinentes e precisos<sup>117</sup>.

---

(BNB). Ver: Austin, D. W. (1974). *PRECIS: a manual of concept analysis and subject indexing*.

<sup>116</sup> Um dos projetos pioneiros foi o projeto *MARC* (1968) (*MACHine Readable Cataloguing* - catalogação legível por máquina), ao qual se seguiram, na década de 70, vários formatos *MARC* nacionais, criados com base no padrão internacional *ISO 2709*. Em 1977 a IFLA cria o formato *UNIMARC* (*Universal Machine Readable Cataloguing*). Este formato teve como objetivo criar uma linguagem comum, que permitisse a troca internacional de registos bibliográficos, colmatando a incompatibilidade entre os vários formatos *MARC* nacionais. Este formato permitia a migração de dados, o desenvolvimento dos sistemas integrados de informação e potenciava uma realidade que veio a observar-se na década de 90 - o tratamento da informação em linha.

<sup>117</sup> A grande novidade que trouxe esta Norma prende-se com o facto de pela primeira vez num texto normativo de âmbito internacional sobre indexação aparecer, de forma explícita, a alusão direta às duas fases do processo

Apesar de, *Os Princípios de Indexação* apresentarem e desenvolverem uma metodologia de base funcional, na medida em que apresentam orientações técnicas a seguir na prática da indexação<sup>118</sup>, são o primeiro texto normativo e sistematizado de um conjunto de recomendações metodológicas a seguir à escala internacional, que pela sua pertinência técnica foi desenvolvido nas décadas seguintes, quer pelos teóricos desta especialidade, quer pelas normas internacionais de indexação<sup>119</sup>.

Ainda no âmbito do paradigma pós-custodial interessa relevar o princípio da garantia literária, em inglês *literary warrant*, que apesar de ter sido enunciado pela primeira vez por Hulm, na obra *Principles of book classification* (1911-1912), ele veio a desenvolver-se sobretudo a partir da década de 60 do século XX. Para tal renascimento contribuíram, por um lado o desenvolvimento da literatura científica, quer no que respeita aos produtores, quer aos consumidores deste tipo de literatura que exigem um vocabulário igual àquele que é usado nos textos, por outro a indexação automática que, na maioria dos casos usa palavras extraídas do próprio texto. Enquanto o princípio do uso se refere ao modo como os utilizadores formulam a sua pesquisa, o princípio da garantia literária tem a ver com a terminologia usada na construção das linguagens de indexação: tesouros, ontologias e taxonomias, que se constroem, geralmente com base no vocabulário de um domínio.

De acordo com Barité *et al.* (2010), entende-se por garantia literária o facto de se privilegiar, para extração e validação da terminologia a ser incorporada num sistema de classificação ou em qualquer outro sistema

---

de indexação pela sua relevância cita-se o ponto 3 (análise do documento). Embora alguns autores já tivessem mencionado este ponto, o seu objetivo não era diretamente a análise do assunto do documento, mas sim a sua formalização nos encabeçamentos das listas de matérias, portanto o seu propósito prendia-se com uma questão formal.

<sup>118</sup> Este texto apresenta duas fases para a elaboração da indexação: Ponto 3: Reconhecimento dos conceitos contidos na informação e ponto 4: Representação dos conceitos através de uma linguagem documental. O ponto 3, relativo à análise, é subdividido em três; 3.1-Apreensão do documento, 3.2- Identificação dos conceitos, 3.3-Seleção dos conceitos. Neste último são referidos de forma explícita dois princípios da indexação: a) a exaustividade dos termos; b) a especificidade (*Principes d'indexation*, 1975, p. 2-8).

<sup>119</sup> Ver entre outros autores: Lancaster (1993), Langridge (1989), J. Rowley (1982) e Cleveland e Cleveland (1990). As orientações metodológicas relativas às fases e os princípios da exaustividade e especificidade foram também objeto de desenvolvimento em outras normas: a Norma ISO 5963: *Documentation. Méthodes pour l'analyse des documents, la détermination de leur contenu et la sélection des termes d'indexation* (1985) e a ISO 214 *Documentation. Analyse pour les publications et la documentation* (1976).

de organização do conhecimento, a literatura de um domínio do saber; entendimento já observado na *NISO* (1994, 3.6.2.1) e em Foskett (1973), embora para este autor o sistema deva basear-se no material que nele introduzimos e não em condições puramente teóricas.

Se numa primeira leitura este princípio metodológico parece pacífico, no seu entendimento e aceitação coloca algumas questões.

Partindo-se do fundamento de que os documentos são a fonte privilegiada da extração e validação dos termos usados na construção dos instrumentos de representação da informação, infere-se que este princípio é o garante da atualização e legitimação dos vocabulários de indexação, sobretudo no que respeita à terminologia científico-técnica, na medida em que este tipo de terminologia é consensual, especializada e evolutiva. A conceção tradicional deste princípio, de incidir o seu foco na terminologia do documento, tem sofrido alterações, tendo-se, por isso, observado uma extensão semântica do mesmo. Nesta perspetiva, Begthol (1986), destaca três tipos de garantia literária: a científica, a educacional e a cultural, sendo a última considerada como um garante do multiculturalismo e dos valores éticos, extravasando deste modo, a função de natureza meramente técnica e científica que lhe era atribuída na conceção tradicional. Por isso, Hudon (1997), levanta a questão do multiculturalismo relativamente aos tesauros, ao referir que um tesouro multilingue deve ser aquele em que todas as culturas a que se destina se sintam representadas. Os instrumentos de representação da informação e do conhecimento são estruturas dinâmicas, que se ajustam e reinventam de acordo com as necessidades sociais e reais de quem deles necessita e também de quem os constrói, como tal devem ter sempre em conta o modo como se produz e como se consome a informação. Por isso, é legítimo lembrar que os termos representam conhecimento, na medida em que formalizam conceitos e, como tal, a sua representação nunca poderá ser efetuada fora de um contexto, independentemente da sua natureza. O reinventar deste princípio, que aparentemente foi esquecido durante décadas, prende-se neste novo século com a discussão em torno das questões relacionadas com o acesso à informação, ancorado nos ambientes digitais veiculado pela INTERNET e, em especial, pelas redes sociais, que permitem um acesso local e global, indepen-

dente do indivíduo, da cultura e da área do saber. Esta questão é tão mais complexa quando se considera que um indivíduo no mesmo país pode ser recetor e assimilador de vários valores culturais e, como tal, ter direito a aceder à mesma informação utilizando diferentes termos, isto é, de acordo com as suas orientações culturais.

Por todas estas questões, este princípio, hoje mais do nunca, é um elemento metodológico imprescindível na prática da indexação.

A título de síntese dir-se-á que, apesar das alterações ocorridas nas estruturas mentais, sociais e tecnológicas, não se observa uma rutura na prática da indexação, quando observada nos dois paradigmas considerados; verifica-se uma evolução contínua nos princípios metodológicos e nos fundamentos teóricos que sustentam esta prática. Outro ponto a considerar é o facto de se presenciar, essencialmente no modelo pós-custodial, a preocupação de se identificar e contextualizar epistemologicamente esta área. Atestam esta ideia a literatura de cariz científico e académico, que foi publicada sobretudo a partir dos anos 60. Considera-se também como ponto a reter o facto de ser uma prática instituída e normalizada por organismos internacionais de relevo, como a *ISO* e a *NISO*.

Na atualidade encaram-se como grandes desafios da indexação rever e consolidar a sua identidade epistemológica face à interdisciplinaridade envolvente, nomeadamente no que se refere às ciências da computação. Rever a sua prática de acordo com o atual contexto cultural e tecnológico, sem deixar cair os princípios metodológicos que a norteiam. Neste ponto recomendam-se situações de compromisso entre os princípios de uso, de garantia literária e o princípio cultural, recorrendo-se, para a sua efetivação, à articulação dos seus referenciais teóricos com as técnicas informáticas.

Por último, interessa promover ainda o compromisso entre a indexação dita tradicional e a colaborativa (*folksonomia*), esta realizada pelo cidadão anónimo, com o objetivo de contribuir para uma indexação mais flexível e multidimensional, pois expressa os modelos culturais do utilizador. É na contextualização e no ajustamento da indexação à sociedade contemporânea, que busca respostas para o pluralismo cultural e a afirmação

dos valores éticos e deontológicos, que esta prática se assume como essencial, pois só nesta perspectiva poderá responder com o rigor às necessidades informacionais de uma comunidade culturalmente heterogénea. Apenas com base neste fundamento a indexação poderá contribuir para a construção de novo conhecimento, com vista a uma sociedade mais equitativa.

## REFERÊNCIAS

AUSTIN, D. W. **PRECIS: a manual of concept analysis and subject indexing.** London: British Library Bibliographic Services Division, 1974.

BARITÉ, M. Et al. Garantia literária: elementos para uma revisão crítica após um século. **TransInformação**, Campinas, v. 22, n. 2, p. 123-138, maio/ago. 2010. Disponível em: <<http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/490/470>>. Acesso em: 8 nov. 2015.

BEGHTOL, C. Semantic validity: Concepts of warrant in bibliographic classification systems. **Library Resources & Technical Services**, v. 30, n. 2, p. 109-125, 1986.

BROWN, A. G.; LANGRIDGE, D. W.; MILLS, J. **An introduction to subject indexing.** London: Clive Bingley, 1976.

BUCHANAN, B. *Theory of library classification.* London: Clive Bingley, 1979.

CHAUMIER, J. **Analisis y lenguajes documentales: el tratamiento lingüístico de la información documental.** Tradução de L. Porta. Madrid: Editorial Mitre, 1986.

CLEVELAND, B. D.; CLEVELAND, A. D. **Introduction to indexing and abstracting.** Englewood: Libraries Unlimited, 1990.

COATES, E. J. **Subject catalogues: headings and structures.** London: The Library Association, 1960.

COLLISON, R. L. **Indexes and indexing: guide to the indexing of books.** London: ABC – Clio Press, 1969.

CUTTER, C. A. **Rules for a printed dictionary catalog.** Washington: Government Printing Office, 1876. Disponível em <<https://archive.org/stream/cu31924029518978#page/n9/mode/2up>>. Acesso em: 8 nov. 2015.

FOSKETT, A. C. **A abordagem temática da informação.** São Paulo: Polígno, 1973.

\_\_\_\_\_. **The subject approach to information.** Londres: Clive Bingley, 1977.

FUGMANN, R. **Subject analysis and indexing: theoretical foundation and practical advice.** Frankfurt - Main: Indeks Verlag, 1993.

GIL LEIVA, I. **La automatización de la indización de documentos.** Girona: Trea, 1999.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge Systems Laboratory**, California, v. 5, n. 2, p. 1-23, abr. 1993. Disponível em: <<http://eecs.ceas.uc.edu/~mazlack/ECE.716.Sp2011/Semantic.Web.Ontology.Papers/Gruber.93a.pdf>>. Acesso em: 15 dez. 2015.

GUIMARÃES, A. Abordagens teóricas de tratamento temático da informação (TTI): catalogação de assunto, indexação e análise documental. **IBERSID: revista de sistemas de información y documentación**, v. 3, p. 105-117, 2009. Disponível em: <<http://www.ibersid.eu/ojs/index.php/ibersid/article/viewArticle/3730>>. Acesso em: 29 jan. 2015.

\_\_\_\_\_. dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-americana de Ciência da Informação (RICI)**, v. 1, n. 1, p. 77-99, 2008. Disponível em: <<http://periodicos.unb.br/index.php/RICI/article/view/2761/2331>>. Acesso em: 8 nov. 2015.

GUINCHAT, C.; MENOUE, M. **Introduction générale aux sciences et techniques de l'information et de la documentation**. Paris: Les Presses de l'Unesco, 1984.

HANSON, E. R.; DAILY, J. E. Catalogues and cataloguing. **Encyclopedia of library and information science**, v. 4, p. 262-263, 1958.

HAYKIN, D. J. Subject headings: fundamental concepts. In: CHAN, L. M.; RICHMOND, P. A.; SVENONIUS, E. (Ed.). **Theory of subject analysis: a sourcebook**. Colorado: Libraries Unlimited, 1985.

HUDON ; M. Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts. **Information services & use**, v. 17, n. 2-3, p. 111-123, 1997.

IFLA . **Principles Underlying Subject Languages (SHLs)**. Munchen: K. G. Sauer, 1999.

ISO 5963. 1985. Documentation – Méthodes pour l'analyse des documents, la détermination de leur contenu et la sélection des termes d'indexation. In: **Documentation et information: recueil de normes ISO I**. Genève: ISO, p. 575-579, 1988.

KAISER, J. **Systematic Indexing**. London: Isaac Pitman & Sons, 1911. Disponível em: <<https://archive.org/details/systematicindexi00kaisuoft>>. Acesso em: 12 nov. 2015.

KRONICK, D. A. **A history of scientific and technical periodicals**. Nova Iorque: Scarecrow Press, 1962.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos: Livros, 1993.

1. LANGRIDGE, D. W. **Subject analysis: principles and procedures**. London: Bowker-Saur, 1989.

MENDES, M. T. P.; SIMÕES, M. G. **Indexação por assuntos: princípios gerais e normas**. Lisboa: Gabinete de Estudos a&b, 2002.

NISO, National Information Standards Organisation. **ANSI/NISO Z39.19-1993 Guidelines for the construction, format and management of monolingual thesauri**. Bethesda, MD: NISO Press, 1994.

ORTEGA, C. D.; LARA, M. L. G. A noção de documento: de Otlet aos dias de hoje. **DataGramZero**, v. 11, n. 2, 2010. Disponível em: <[http://www.dgz.org.br/abr10/Art\\_03.htm](http://www.dgz.org.br/abr10/Art_03.htm)>. Acesso em 8 nov. 2015.

OTLET, P. **Traité de documentation: le livre sur le livre, théorie et pratique**. Bruxelles: Editions Mundaneum, 1934.

PINTO MOLINA, M. **Fundamentos y procedimientos**. Madrid: EUDEMA, 1993.

RANGANATHAN, S. R. **Classified Catalogue Code with Additional Rules for Dictionary Catalogue Code**. [S.l.]: Madras Library Association, 1958.

RIBEIRO, F. Gestão da Informação/Preservação da Memória na era pós-custodial: um equilíbrio precário?. **Livro de Atas de Conferência Nacional**, Porto: Universidade do Porto - Faculdade de Letras, 2004. Disponível em: <<http://ler.letras.up.pt/uploads/ficheiros/artigo8861.PDF>>. Acesso em: 12 dez. 2015.

ROWLEY, J. E. **Abstracting and indexing**. London: Clive Bingley, 1982.

SHOLNIK, H. Historical development of abstracting. **Journal of Chemical Information and computer sciences**, v. 19, n. 4, p. 215-228, 1979.

SILVA, M. R.; FUJITA, M. S. L. A prática de indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, v. 16, n. 2, 2004.

SIMÕES, M. G. **Classificações bibliográficas: percurso de uma teoria**. Coimbra: Almedina, 2011.

\_\_\_\_\_. **Da abstração à complexidade formal**. Coimbra: Almedina, 2008.

\_\_\_\_\_. **Resumo documental e literatura científica: origem, desenvolvimento e consolidação**. Lisboa: Páginas a&b, v. 3, n. 3, p.15-36, 2015.

\_\_\_\_\_. **Resumo Documental: um incursão à (des)construção concetual na literatura científica**. Coimbra: Imprensa da Universidade de Coimbra, 2015. Disponível em: <<http://dx.doi.org/10.14195/978-989-26-1035-1>>. Acesso em: 14 dez. 2015.

STONE, A. The LCSH century: a brief history of congress subject headings, and introduction to the centennial essays. **Cataloging & Classification Quarterly**, Reino Unido, v. 29, n. 1-2, 2000. Disponível em: <<http://www.catalogingandclassificationquarterly.com/ccq29nr1-2ed.htm>>. Acesso em: 8 nov. 2015.

UNISIST. **Principes de indexation**. Paris: UNESCO, 1975.

WAL, T. V. Folksonomy. **Vanderwal.net**, [S.l.], 2007. Disponível em: <<http://vanderwal.net/folksonomy.html>>. Acesso em 15 dez. 2015.

WITTY, F. J. The beginnings of indexing and abstracting: some notes towards a history of indexing and abstracting in antiquity and the middle ages. **The Indexer**, Sheffield, v. 8, n. 4, 1973.

# PARTE III

COMPLEJIDADES TECNOLÓGICAS EN  
EL ENTORNO DE LA CIENCIAS DE LA INFORMACIÓN