



XIX encontro nacional
de pesquisa em
ENANCIB ciência da informação

// SUJEITO INFORMACIONAL E AS
PERSPECTIVAS ATUAIS EM CIÊNCIA
DA INFORMAÇÃO. //

22-26
OUTUBRO
2018
LONDRINA/PR



XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2018

GT-2 – Organização e Representação do Conhecimento

APROXIMAÇÕES ENTRE A TEORIA DA CLASSIFICAÇÃO E O PROCESSO DE INDEXAÇÃO: AS CONTRIBUIÇÕES DO CLASSIFICATION RESEARCH GROUP

Raquel Luise Pret (Universidade Federal Fluminense)

Rosa Inês de Novais Cordeiro (Universidade Federal Fluminense)

APPROACHES BETWEEN THEORY OF CLASSIFICATION AND INDEXING PROCESS: CLASSIFICATION RESEARCH GROUP CONTRIBUTIONS

Modalidade da Apresentação: Comunicação Oral

Resumo: Estudo sobre as contribuições do *Classification Research Group (CRG)* no tocante às abordagens e propostas de seus membros que aproximaram à teoria da classificação do processo de indexação. O objetivo deste trabalho é apresentar, de forma panorâmica, as pesquisas desenvolvidas pela corrente inglesa de tratamento temático da informação que influenciaram as aplicações da indexação no campo da Ciência da Informação. Esta análise está baseada, além de uma ampla revisão de literatura acerca da temática proposta, no exame dos boletins do CRG entre os anos de 1950 e 1960, editados pelo *Journal of Documentation* e das publicações de seus membros como Vickery, Farradane, Foskett, entre outros. Trata-se de um recorte na pesquisa desenvolvida em nível de doutorado sobre o processo de indexação nos arquivos de universidades. Os resultados evidenciam a influência dos princípios classificatórios de Ranganathan nas abordagens, realizadas pelo CRG, da classificação visando à organização das áreas do conhecimento especializadas e a indexação de grandes volumes bibliográficos, nas décadas de 1950 e 1960, nos sistemas de informação. Estas procuravam atender as demandas de instituições governamentais e grandes empresas voltadas para a disseminação da produção científica. Conclui-se que as aproximações entre a teoria da classificação e o processo de indexação realizados pelo CRG procuraram estabelecer padrões para a representação da informação que facilitaram o desenvolvimento de sistemas automatizados de recuperação da informação.

Palavras-chave: Classification Research Group; Teoria da Classificação; Indexação; Tratamento Temático da Informação.

Abstract: Study on the contributions of the Classification Research Group (CRG) regarding the approaches and proposals of its members that approached the classification theory of the indexing process. This work aims to present, in a panoramic way, the researches developed by the English current of the information subject treatment that influenced the indexing applications in the Information Science field. Besides a broad literature review on the proposed theme, this analysis is also based on the evaluation of CRG bulletins between the 1950s and 1960s, which were edited by the *Journal of Documentation*, and on publications from members such as Vickery, Farradane, Foskett, among others. It is about a clipping in the research developed at the doctorate level on the indexing process in university archives. Results show the influence of Ranganathan's classification principles in CRG approaches towards the classification aiming at organizing specialized areas of knowledge and indexing large bibliographic volumes, in the 1950s and 1960s, in the information systems. These approaches aimed to meet the demands of government institutions and large companies focused on the dissemination of the scientific production. It was concluded that the approaches between the classification theory and the indexing process performed by the CRG aimed to stablish standards for the representation of information that facilitated the development of automated information retrieval systems.

Keywords: Classification Research Group; Classification Theory; Indexing; Information Subject Treatment.

1 INTRODUÇÃO

A multiplicidade das complexas abordagens acerca da teoria da classificação e do processo de indexação leva a inúmeros caminhos que por vezes ocasionam omissões ou confusões pela imensidão dos territórios que podem ser explorados sobre tais práticas discursivas, enunciados e diferentes escalas de análise possíveis. Assim, sem a pretensão de apresentar uma totalidade sobre tal aproximação, pois a tentativa de uma exegese apenas se converteria em um desejo megalomaniaco não possível de ser concretizado, nossa proposta é realizar um olhar sobre um contexto particular, considerando as repercussões das propostas surgidas e/ou consolidadas no âmbito do *Classification Research Group* (CRG) nas próprias práticas da indexação.

Diante dessa perspectiva, apresentamos algumas considerações sobre a indexação pela perspectiva da corrente inglesa de tratamento temático da informação¹. Os teóricos que elaboravam caminhos e possibilidades para o tratamento temático da informação nesta abordagem procuravam solucionar demandas dos centros de documentação e bibliotecas

¹ Guimarães (2009), ao abordar o tratamento temático da informação no âmbito da CI, apresenta três vertentes: a catalogação de assunto (*subject cataloguing*), de matriz norte-americana; a indexação (*indexing*), de ascendência inglesa; e a análise documentária (*analyse documentaire*), de influência francesa.

especializadas em desenvolver áreas do conhecimento estratégicas para grandes potências após a Segunda Guerra Mundial.²

No contexto da Guerra Fria, mais precisamente entre os anos 1950 e 1960, os esforços com a gestão e a disseminação da produção científica, elaboração de esquemas classificatórios e tornar mais eficaz o processo de indexação estavam associados ao tratamento de grandes volumes documentais e à necessidade das editoras e sistemas de informação recuperarem rapidamente registros, informações, conteúdos, assuntos que facilitassem o desenvolvimento de pesquisas científicas e tecnológicas em áreas como a Química, a Engenharia Civil e Aeroespacial, a Biomedicina e a Indústria Armamentista.

Desse modo, a classificação, mais próxima à Bibliologia de Paul Otlet, deveria ser capaz de organizar, gerenciar e recuperar a informação existente na documentação em seu sistema de informação (NUNES; TÁLAMO, 2009, p. 42). De forma semelhante, a indexação teria como função nessas instituições “facilitar a pesquisa de documentos ou de informações” (NEET, 1989, p. 7 *apud* GUIMARÃES, 2009, p. 107).

Considerando esse contexto sócio-histórico é que nos propomos a apresentar algumas contribuições do *Classification Research Group*, destacando a sua atuação na elaboração de abordagens acerca do processo de indexação que dialogassem com a teoria da classificação e de suas aplicações nos sistemas de informação nas décadas de 1950 e 1960.

2 PROCEDIMENTOS METODOLÓGICOS

Este trabalho apresenta os resultados parciais acerca da pesquisa bibliográfica, de caráter qualitativo, realizada em nível de doutoramento, que procurou compreender as principais abordagens para o tratamento temático da informação e suas relações com a organização do conhecimento no campo da Ciência da Informação.

Além de uma ampla revisão de literatura, esta análise está baseada no estudo dos boletins do *Classification Research Group*, registrados durante os anos de sua existência (1952-1969) e publicados pelo *Journal of Documentation*.

Outro procedimento metodológico utilizado foi o exame das publicações dos membros do CRG como Brian Vickery, Douglas Foskett, Jason Farradane e Derek Austin. Utilizamos como

2 Este trabalho faz parte dos resultados obtidos na pesquisa realizada para a produção da tese, **Do quê se trata? Documento, indexação e produção de arquivo**, a ser defendida no Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense, cuja qualificação foi aprovada.

recorte cronológico os anos de 1950 e 1960, período de atuação do CRG. Ademais, selecionamos os trabalhos que abordavam a indexação associada à teoria da classificação e os resultados das aplicações dos modelos propostos por esses autores.

3 CLASSIFICATION RESEARCH GROUP – APROXIMAÇÕES ENTRE A CLASSIFICAÇÃO E A INDEXAÇÃO

O *Classification Research Group* (CRG), formado na Inglaterra em 1952 por A. J. Wells e B. C. Vickery, surgiu a partir das discussões realizadas na *Royal Society Scientific Information Conference*, em 1948, acerca do grande aumento da produção científica e da preocupação com as classificações bibliográficas. Essa produção expansiva impactava nos procedimentos de gerenciamento e disseminação da informação, como também nos assuntos das classificações bibliográficas. Após a Conferência foi criado um grupo de trabalho sobre classificação de assunto.

Vanda Broughton (2012) esclarece:

O *Classification Research Group* do Reino Unido teve suas origens na espítome científica do período pós-guerra, especificamente em 1948 na *Royal Society Conference* sobre informação científica, que foi convocada como resultado das preocupações concernentes ao grande aumento da informação científica e aos problemas relativos à sua gestão e disseminação. Após a Conferência, foi criado um grupo de trabalho sobre classificação de assuntos, mas seu secretário, o físico J. D. Bernal, concluiu que o problema da classificação era ‘muito mais complexo do que se imaginava no início’ (BROUGHTON, 2012, p. 318, tradução nossa).

Sua proposta inicial era reunir bibliotecários especializados e *information officers* interessados em discutir os sistemas de classificação de suas instituições, sobretudo a aplicabilidade de princípios classificatórios e a recuperação da informação. O grupo sentia a necessidade de compartilhar suas experiências na organização sistemática de documentos, divulgando abordagens teóricas e aplicações práticas bem sucedidas. ‘

Eles não estavam muito interessados em organizar prateleiras de livros como em fornecer acesso rápido e eficiente às informações dos assuntos por meio de índices. Recuperação da informação era o seu objetivo, na definição eminentemente satisfatória de Calvin Mooers: ‘busca e recuperação de informações em seu armazenamento de acordo com a especificação por assunto’ (FOSKETT, 1970a, p. 8, tradução nossa)

Os debates registrados em seus boletins foram importantes referenciais para os estudos do tratamento temático da informação, em especial para a indexação. Segundo Foskett, os membros do CRG "[...] estavam ansiosos para explorar a base da indexação de assuntos a partir das discussões e do compartilhar das experiências" (FOSKETT, 1962, p. 127, tradução nossa). Nomes como Vickery, Farrandane, Faithorne, Coates, Foskett, Langridge, Austin e Kyle procuraram analisar e debater a teoria da classificação, os esquemas classificatórios e suas operações no âmbito das práticas realizadas nas instituições que trabalhavam.

Dessa forma, os princípios da classificação facetada, o uso de atributos, os termos (conceitos) coordenados, a classificação cruzada, a ordem combinada de facetas foram alguns temas abordados com o intuito de organizar a documentação, objetivando a representação e recuperação da informação. O grupo ficou ativo até 1969 com diversos artigos de seus membros publicados neste período, além do registro de suas discussões em seus boletins anuais.

De acordo com Guimarães (2003), a proposta do CRG era problematizar e elaborar metodologias de tratamento temático da informação. O primeiro esforço estava centrado em questões de representação temática e a avaliação da performance dos instrumentos tais como as linguagens documentárias.

Em seu quarto boletim, o CRG, ao debater os problemas relativos à recuperação da informação, aproximou a classificação da indexação. Era consenso no grupo que a recuperação da informação deveria ser consequência do tratamento temático dado aos documentos no momento da entrada nos sistemas. Nesse sentido, a base de um sistema de recuperação de informação seria uma rede de termos interligados. As interconexões de cada termo seriam potencialmente ilimitadas, e cada sistema de recuperação deveria selecionar certas interconexões para exibição. Esta seleção baseava-se em postulados quanto aos tipos de termo e de relação que deviam ser exibidos - postulados relativos ao nível semântico dos termos, categorias de termos a serem utilizados, relações genéricas, coordenadas e conjuntivas a serem exibidas, os tipos de operação de busca a serem conduzidos.

A especificação por assunto pode existir, no entanto, em mais de um modo: especificação por um autor, do que ele considera ser a realidade do que está investigando; especificação por um indexador, do que ele por sua vez considera ser a intenção do autor, a qual pode ser dada por uma 'inclinação' particular aos propósitos da própria organização do indexador; especificação por um usuário, que se tornou consciente de uma lacuna no conhecimento

que ele requer para lidar com uma situação atual. O problema em centro de informação é efetuar essas especificações em segundos. Aqui, todas as três especificações são idênticas e não há imprecisão ou ambiguidade (FOSKETT, 1970a, tradução nossa).

A primeira meta estabelecida para representar os documentos com o propósito da sua recuperação foi explorar a variedade de postulados que poderiam ser consolidados e avaliar a sua adequação em vários sistemas de recuperação da informação (CLASSIFICATION..., 1958, p. 137).

Destarte, nesse primeiro momento, o CRG dedicou-se em analisar os problemas semânticos na entrada dos documentos em um sistema, associando o processo de indexação à classificação. O CRG, ao discutir a técnica de análise de assuntos, deliberou que "a maneira correta de indexar um assunto específico é analisá-lo em termos classificatórios - conceitos que têm uma interpretação contemporânea e reconhecidamente sem ambiguidade - e mostrar as relações entre esses termos" (CLASSIFICATION..., 1958, p. 137). Para o grupo, a forma mais útil de esquema de classificação para a indexação seria aquela capaz de agrupar os termos em categorias bem definidas, que poderiam ser usadas de maneira independente para formar compostos.

Chegou-se à conclusão que haveria pelo menos duas maneiras de expressar relações entre termos classificatórios: (i) inserindo operadores entre os termos que atribuiriam significados diretos a essas relações, e (ii), agrupando os termos em categorias, citando estes em uma ordem particular que permitisse inferir a relação entre eles a partir do contexto. Estas duas técnicas de análise não são mutuamente exclusivas, e poderia ser desejável utilizar ambas.

Faithorne, uma das lideranças do grupo, ampliou a discussão acerca das técnicas que aproximavam a indexação da classificação, contrastando a identificação de um documento por descrição e por relevâncias. A descrição foi relativamente estável, no entanto, de acordo com o pesquisador, as relevâncias variaram de acordo com tempo e lugar (CLASSIFICATION..., 1958, p. 141).

A aproximação do processo de classificação da indexação feita pelo CRG baseou-se nas discussões sobre a aplicação do esquema de facetas, elaborado na **Colon Classification**, elaborada em 1933 por Ranganathan. Desde de suas primeiras reuniões, os membros do CRG relatavam a suas experiências e propunham métodos quem pudessem ser mais efetivos no tratamento temático dos documentos, visando a sua entrada nos sistemas.

Kyle, por exemplo, procurou trabalhar com a questão geral das divisões de forma. Ela levantou práticas possíveis para a entrada dos documentos considerando a forma: a) realizar as subdivisões dominantes, de maneira a recolher todos os materiais sob formas especiais, quer no início (materiais de abordagem), quer no final (materiais suplementares) de todo o índice, (b) torná-los secundários ao assunto, mas assim mesmo anteriorizando ou posteriorizando sua inclusão nos SRIs (Classification Research Group Bulletin, 1958, p. 142).

A preocupação do CRG nas décadas de 1950 e 1960 estava centrada em elaborar princípios para o estabelecimento das hierarquias nos sistemas classificatórios que fosse facilmente compreendidas por seus usuários. No sexto boletim, o grupo alertava para os problemas na elaboração de classificações especializadas, particularmente questões de ordem das facetas, que eram resolvidas por referência à natureza do assunto central ou ao ponto de vista predominante de uma “clientela” relativamente homogênea. No entanto, nenhum destes fatores estava disponível para uma classificação geral e era evidente que seria necessário um novo princípio para ordená-la.

As facetas eram exaustivamente discutidas considerando sempre sua aplicação em sistemas de informação. Ao procurar qualificar as categorias “coisa” e “atividade”, o CRG propôs o uso de níveis integrativos. O grupo percebeu a existência de “formas sucessivas de ordem numa escala de complexidade e organização” (CLASSIFICATION..., 1961, p. 158) e de uma sucessão de níveis em que as agregações se tornavam inteiras, agindo como unidades envolvidas em processos característicos de níveis particulares.

Foi no âmbito do CRG que surgiu a teoria dos níveis integrativos, oferecendo uma repartição preliminar viável da lista de coisas físicas e materiais, a partir da ordem das formas sucessivas, embora ainda não estivessem elaboradas, naquele momento do início dos anos 1960, as formas em que as facetas de energia seriam montadas nas séries dos níveis integrativos.

A teoria de níveis integrativos propunha soluções para a questão dos termos compostos baseados em termos isolados extraídos de classificações existentes em várias classes principais. Este era um dilema frequente na classificação de uma bibliografia geral ou na construção de índices. Segundo os relatos do sexto boletim do CRG, tal teoria surgiu pela necessidade de encontrar uma solução acerca de tais termos compostos que não era oferecida por nenhum tipo de faceta dentro de um sistema de classe que usasse a ***Colon Classification***.

O fator semântico foi um tópico que surgiu pela primeira vez nas discussões em grupo sobre classificações especializadas e apareceu novamente em conexão com a proposta de classificação geral. A teoria dos níveis integrativos realiza essa aproximação com a semântica ao analisar o significado de um termo complexo em um número de termos mais generalizados que denotam o próprio conceito. O valor da análise semântica para fins de recuperação significava a capacidade de um termo de exibir as relações existentes entre os conceitos que possuem propriedades em comum. Segundo o CRG, uma cadeia de classes genéricas que são reveladas a partir de um conceito específico representa uma série de propriedades do conceito em si e, conseqüentemente, representa uma análise semântica (CLASSIFICATION..., 1961, p. 160).

O CRG considerava as técnicas de faceta uma ajuda real na análise sistemática de documentos de áreas técnicas. O indexador logo formaria o hábito de classificar rapidamente o assunto em categorias, selecionando os números de classe de unidade e combinando os símbolos de acordo com a ordem preferida da comunidade de usuários de seu sistema de informação. Este procedimento, juntamente com a construção semiautomática do índice de assunto, aceleraria o trabalho da catalogação de assunto, uma vez que o conteúdo técnico do documento seria mais facilmente compreendido (CLASSIFICATION..., 1962, p. 86).

Foskett (1962) relata que os membros do CRG em suas discussões chegaram à conclusão de que a indexação por assunto, fosse alfabética, facetada, seleção mecânica ou não, dependeria de sua eficiência na coordenação de conceitos ou ideias expressos nos documentos, de modo que um conjunto de conceitos declarados como o requerimento de um pesquisador/usuário (*searcher*) possa ser combinado com o mesmo ou similar conjunto já existente na biblioteca ou sistema de informação. Logo, "o bibliotecário [indexador] tem que fazer uma tradução ou transformação do conjunto de termos em que a solicitação é declarada em um conjunto principal de estereótipos de significados em que os documentos estão organizados" (FOSKETT, 1962, p. 128)

4 RESULTADOS – AS ABORDAGENS E USOS DOS PRINCÍPIOS CLASSIFICATÓRIOS DE RANGANATHAN PELO CRG

De acordo com o CRG, a organização da documentação para a recuperação da informação passa por importantes etapas como: primeiro, a definição do assunto do

documento e a terminologia a ser adotada para classificar todo universo documental, depois o levantamento das facetas a serem adotadas. Ranganathan ao elaborar os princípios da **Colon Classification** estabeleceu cinco facetas, chamadas de *PMEST* - personalidade, matéria, energia, espaço, tempo, respectivamente.

Embora o CRG tenha rejeitado fidelidade a qualquer esquema de classificação existente, reconheceu o enorme poder e versatilidade do método de análise de facetas, e sempre o utilizou desde a primeira publicação do CRG no *Library Association Record*, de 1955. A necessidade de uma classificação facetada era compreendida como base para todos os métodos de recuperação de informação (FOSKETT, 1970a, p. 12)

O próprio Ranganathan teorizou acerca do uso das facetas no processo de indexação, com menção aos cabeçalhos de assunto. Para ele, a escolha do nome do assunto de um documento e a representação do nome no título da entrada de um assunto específico pode ser obtido por análise de faceta com base em postulados e princípios da *Colon Classification*. Conforme Ranganathan (1964), a análise das facetas é um trabalho feito no plano das ideias, no plano verbal e no plano notacional. Ele critica ainda a concepção que trata a análise das facetas como um método exclusivamente concebido para a classificação.

Depois de demonstrar que os cabeçalhos de assunto constituem uma linguagem artificial, aponta que o uso da análise de faceta para o cabeçalho de assunto não equivale a usar o número de notação, mas aplicar a própria teoria facetada. Neste trabalho, ele estabelece três usos possíveis para as facetas: a) determinação do número da classe; b) Serviço de Referência; e c) determinação do título do assunto. De acordo com Ranganathan, o êxito da **Colon Classification** no processo de classificação dos documentos permitiu a sua adaptação em outras operações como o serviço de referência - as exigências dos leitores, o prognóstico de suas necessidades de informação, sobretudo atendendo às demandas de leitores especialistas.

No momento em que é formulada a busca por assunto, de forma pontual, ela pode ser atendida pela análise de faceta - isto é, averiguando, através da interação entre o bibliotecário e o usuário, as facetas de exigência do requerente, o foco exato no assunto básico e cada uma de suas facetas isoladas, respectivamente. A determinação do conteúdo de um documento e seu título de assunto também tem sido uma operação em que a aplicação das facetas tem se revelado como método bem sucedido. Ela pode ser feita de uma forma sistemática e

padronizada a partir da análise das facetas e atribuição dos títulos a partir de suas combinações (RANGANATHAN, 1964, p. 111).

No entanto, Ranganathan alerta que, mesmo seguindo a tais critérios para as cinco facetas selecionadas, há 120 arranjos possíveis para representar um assunto específico. Diante dessa perspectiva, ele apresenta princípios que podem nortear, estabelecer diretrizes para a escolha dos assuntos de um documento. Apresenta o princípio da predominância, que estabelece ser a predominância do termo que decide a posição de cada entrada no índice; o princípio do sentido procurado, estabelece que o primeiro termo a constar no vocabulário de assunto deve ser o que dá razão, sentido, à existência do documento; o princípio da estatística de pesquisas ou opiniões, esse princípio aponta para a conjectura e a experiência dos sistemas de informação, a definição da posição dos termos pode ser baseada em um levantamento objetivo estatístico dos termos utilizados para a busca do documento ou nas opiniões dos usuários. Segundo Ranganathan (1964), só havia duas escolas de pensamento dominantes - uma escolhendo os termos do cabeçalho de assunto de forma perspectiva (*forward rendering*) e a outra escolhendo os termos de forma retrospectiva (*reverse rendering*).

Uma vez estabelecida a escolha e a representação do assunto específico de um documento para um cabeçalho com base no conjunto de postulados da análise das facetas e dos princípios para a sequência facetada, o número mínimo de referências cruzadas, a exemplo do “ver também”, nos cabeçalhos de assunto seria igualmente e mecanicamente derivado do cabeçalho de assunto com o auxílio do procedimento em cadeia. Nessa teoria, as decisões acerca da escolha das facetas que irão representar o assunto deverão ser feitas primeiramente no plano de ideias. Em seguida, as regras do procedimento em cadeia podem ser enquadradas para implementar no plano verbal, implementar a política adotada no plano das ideias (RANGANATHAN, 1964, p. 118).

As entradas de cabeçalho de assunto, em conjunto, formariam uma linguagem artificial. Sem dúvida, as palavras usadas para a determinação dos assuntos seriam retiradas da linguagem natural, mas a sintaxe de posição seria diferente da sintaxe da linguagem natural a partir da qual as palavras são tomadas. É sabido que a sintaxe varia de uma linguagem natural para outra. Para Ranganathan, os princípios para a análise das facetas e de sua ordenação ajudariam em grande medida a atribuição de assuntos específicos do documento por fornecer uma espécie de sintaxe universal, transformando-os em cabeçalhos de assunto, e, inclusive, estabelecendo seus números de classe.

As tarefas de catalogação de assunto e de classificação são igualmente beneficiárias desses postulados e princípios. A utilização de um único e mesmo procedimento na catalogação e na classificação não garante que a presunção do cabeçalho de assunto seja derivada do número de classe ou do número de classe sendo derivado do título de assunto, mas facilita uma padronização na representação do documento (RANGANATHAN, 1964, p. 119, tradução nossa).

Para Spiteri (1998, p. 19), o CRG ao longo de seus debates e troca de experiências percebeu que era necessário ampliar a quantidade de facetas para dez e com a seguinte ordem: tipo de produto final, partes, materiais, propriedade, processo, operação, agente, espaço, tempo e forma de apresentação. É importante destacar que o CRG, ao contrário dos postulados de Ranganathan, chegou a conclusão de que a escolha das categorias fundamentais e a ordem de citação deveriam derivar da natureza dos assuntos encontrados no acervo a ser representado.

A natureza do assunto seria revelada pelo próprio contexto em que o assunto estava inserido. Assim sendo, não deveria haver qualquer imposição de listas de categorias fundamentais apriorísticas, ainda que fossem exaustivas. Essa imposição invalidaria o benefício das possibilidades de múltiplas combinações na representação dos assuntos que o uso das facetas permite.

Segundo Foscett (1962, p. 137), o CRG, a partir dessa perspectiva, procurou contemplar uma espécie de *UR-classification*, uma espécie de sistema classificatório generalista que poderia ser adaptado por diversos sistemas classificatórios por meio dos níveis integrativos, ou seja, um conjunto primordial de termos em facetas, sub-arranjados de acordo com a teoria dos níveis integrativos, mas não indo mais fundo na síntese de complexos termos do que a própria teoria. Poderia traçar-se listas de termos, análises de termos complexos que precisariam ser usados como conjuntos em certos campos, relações básicas e atividades.

Vickery (1955), um dos autores de destaque no CRG, preocupou-se com as mudanças da linguagem e o surgimento de novos conhecimentos, alertando que um esquema classificatório jamais estaria inteiramente pronto. Dessa forma, procurou elaborar possibilidades de futuras expansões nos esquemas classificatórios, considerando sub-facetas na aplicação de novos princípios de divisão, com base na teoria dos níveis integrativos.

Vickery ao analisar a trajetória da prática de indexação entre 1945 e 1955, em seu artigo *Developments in subject indexing* (1955), afirma que houve uma enxurrada de artigos,

livros, palestras, simpósios e discussões sobre indexação, classificação, recuperação mecânica, análise de assuntos, recuperação da informação nesses dez anos. De acordo com o autor, essas publicações demonstram que uma geração anterior usou palavras e cabeçalhos de assunto, já os estudos da década de 1950 passou a usar termos isolados, facetas, analetos, descritores e unitermos.

A indexação alfabética, a classificação e a recuperação mecânica estavam sendo abordadas como processos completamente diferentes e descritos em termos tão idiossincráticos que os profissionais de um campo eram às vezes ininteligíveis para aqueles em outros. Vickery procurou aproximar tais tendências, revelando que todas essas formas de classificação, seleção mecânica e indexação de assunto estavam enfrentando problemas semelhantes e estavam começando a resolvê-los de maneiras semelhantes.

O primeiro problema citado por Vickery foi a eliminação de títulos sinônimos ou sobrepostos. O fato dos cabeçalhos estarem interligados levava o indexador a fazer um número considerável de entradas “ver também” (VICKERY, 1955, p. 1). Conforme a lista crescia, a necessidade de alguma ação para assegurar consistência e uniformidade tornava-se óbvia.

Para Vickery (1955, p.2), existia apenas uma maneira certa de indexar o novo termo: relacionando-o com os seus genéricos e coordenados: elaborando uma tabela sinóptica, como Cutter a chamava, mostrando os termos e suas interconexões em um mapa classificatório. Desse modo, o novo termo poderia ser ajustado e as referências cruzadas apropriadas feitas. Vickery aludiu Cutter ainda para afirmar que esse imenso esforço não era recomendado pelo bibliotecário norte-americano para um único catálogo. No entanto, "o que parece ser demais para um catálogo pode ser lucrativo para todos os catálogos" (CUTTER *apud* VICKERY, 1955, p.3) na forma de um sistema universal de classificação, complementa Vickery.

Vickery aprofundou as relações entre a teoria da classificação e a indexação, sobretudo na década de 1960, como demonstram Cordeiro e La Barre (2011, p. 182). Com o intuito de possibilitar as possibilidades de organização e representação do conhecimento, aprimorando a recuperação da informação e do documento, consequentemente, Vickery passou a teorizar sobre a normalização da linguagem dos documentos (conteúdos registrados), considerando também a normalização das perguntas (*queries*) dos usuários nos sistemas de informação. Vickery procurava acertar a combinação entre as representações dos documentos, realizadas

na entrada dos sistemas pela técnica da indexação, e as questões formuladas pelo usuário na saída dos sistemas de informação.

A linguagem da indexação teria a função de representar o conteúdo do documento e, nessa perspectiva, os esquemas classificatórios forneceriam parâmetros para a definição dos conceitos e suas interrelações, facilitando a escolha dos termos (CORDEIRO; LA BARRE, 2011, p. 182).

A indexação alfabética e a seleção mecânica são técnicas de indexação largamente difundidas na década de 1950 nos sistemas de informação cuja seleção dos assuntos começava de forma descendente, a partir dos termos individuais. Sua abordagem era chamada de indutiva. Ambas as técnicas haviam sido conduzidas, no decurso do seu desenvolvimento, a introduzir métodos de agrupamento de termos individuais.

O efeito destas operações figurava, em princípio, na construção de uma série de mapas classificatórios, nos quais os termos eram agrupados primeiramente em várias categorias, e dentro de cada categoria eram arranjados em árvores de família conceituais. A partir desses mapas, os termos de indexação e suas interconexões eram derivadas. Segundo Vickery (1955, p. 7), esse marco na transformação da indexação que passou a ser baseada na adoção de mapas conceituais já era nítido nos sistemas europeus e norte-americanos, onde essa técnica já era aplicada na indexação alfabética e na seleção mecânica dos termos.

Farradane, outro importante nome no CRG, propôs que as relações entre os diferentes termos dos cabeçalhos de assunto pudessem ser expressas por meio de um pequeno número de operadores-padrão, definidos uma única vez. Essas relações operacionais poderiam ser definidas de modo que, na maioria dos casos, só existisse uma possível ordem de citação que correspondesse ao que Farradane chamou de ordem totalmente dedutiva (FARRADANE, 1952, p.74).

Ao aproximar a classificação da indexação, Farradane justifica tal feito argumentando que a classificação representa uma teoria da estrutura do conhecimento, isto é, das relações entre as diferentes áreas do conhecimento. O problema básico é determinar quais são essas relações e como elas ligam os diferentes conceitos do conhecimento em uma estrutura coerente (FARRADANE, 1952, p.74).

A subdivisão arbitrária ou dedutiva de um conjunto de conhecimentos, segundo o autor, não poderia dar uma representação verdadeira dessas relações, pois estes não

consistiriam apenas em agrupamentos de uma classe e seus membros, ou na divisão de um todo em suas partes.

Farradane (1952) argumenta que uma classificação deve ser construída de forma indutiva, ou de baixo para cima (*bottom-up*), reunindo fragmentos conhecidos de relações. Partindo de elementos de conhecimento singularmente definíveis, denominados “isolados”, estabelecendo as relações entre eles, chamadas de “operadores”, sendo representados linearmente por símbolos denominados analetos (*analets*) (FARRADANE, 1952, p.74). Estes analetos poderiam ser rearranjados de acordo com uma técnica simples para proporcionar várias permutações desejáveis e permissíveis que, dispostas em ordem alfabética dos primeiros isolados, produziriam um índice de assunto.

Segundo Farradane (1952), quando podemos alocar um novo objeto ou uma nova experiência em uma classe de conceitos previamente conhecidos, ou quando podemos encontrar um fator comum em um conjunto de fenômenos elencados previamente, removemos um elemento de estranheza, uma possível fonte de medo do incompreensível, do nosso ambiente. "É uma maneira de alcançar a harmonia em nosso processo contínuo de adaptação a um mundo de outra forma hostil e incalculável" (FARRADANE, 1952, p. 74).

Para Farradane, nossa linguagem comum é um meio muito pobre para a expressão dos novos conceitos apresentados pela ciência. Por esta razão os cientistas necessitam em grande parte criar novas palavras para novos conceitos. No campo da classificação, no entanto, poderia ter sido demasiado confuso elaborar muitas palavras novas para descrever os operadores, e os nomes usados deveriam ser considerados como sendo apenas típicos dos significados possíveis de cada operador, tendo em conta a lógica do significado de cada operador.

No âmbito do CRG, os debates acerca do tratamento da informação e dos documentos procurando a melhor eficiência dos sistemas de informação passaram a incorporar as discussões acerca das possibilidades de interconexões entre diferentes centros de documentação e informação. O quarto boletim (1958) já apresentava a preocupação de seus membros com a construção de sistemas interconectados.

Para os membros do CRG, a base de um sistema de recuperação de informação é uma rede de termos interligados. As interconexões de cada termo são potencialmente ilimitadas, e cada sistema de recuperação deve selecionar certas interconexões para acesso aos documentos. A padronização e seleção dessas interconexões deveriam basear-se em

princípios quanto aos tipos de termo e às suas relações - princípios relativos ao nível semântico dos termos, das categorias a serem utilizadas, das relações genéricas, coordenadas e conjuntivas a serem representadas e dos tipos de operação de busca a serem conduzidos. O CRG preocupava-se em explorar a variedade de princípios que poderiam ser estabelecidos para a construção de sistemas classificatórios e a ordenação na construção de vocabulários controlados, além de avaliar a sua adequação em várias situações ocorridas nos sistemas de recuperação da informação.

No sétimo boletim (1962), os relatos sobre a utilização da indexação por cadeia de facetas tal como foi sugerido por Ranganathan mostrou-se bem sucedida para os centros de documentação que contavam com publicações altamente especializadas como nas áreas da saúde e da segurança. A prática era preparar as entradas de cadeia necessárias e referências cruzadas para novos códigos ao mesmo tempo em que o sistema de classificação estava sendo elaborado (CLASSIFICATION..., p. 67).

O CRG preocupou-se com a dispersão das facetas numa indexação em cadeia, pois a ordem de combinação das categorias geralmente dava prioridade aos produtos e não a seus materiais. Para superar esta dificuldade com a ordem preferida, na tentativa de reduzir a dispersão, as regras para a construção das cadeias na indexação deveriam ser estabelecidas em certas partes das tabelas que limitavam estritamente o uso das subdivisões previstas. Deveria haver uma preocupação em limitar o estabelecimento das entradas de assunto por cadeia, uma vez que um amplo índice em cadeia poderia levar a dispersão dos assuntos e, por conseguinte, a impossibilidade da recuperação do documento.

O grupo percebeu que diversos programas haviam abandonado a indexação em cadeia, pois as dificuldades encontradas quando as buscas eram feitas por determinadas combinações de termos causavam uma fadiga excessiva pela quantidade de documentos recuperados. Embora esta crítica da indexação em cadeia tenha sido aceita como comprovada em diversos relatos de seus membros como Cranfield e Cleverdon, isto não fez com que esta fosse inviabilizada completamente (CLASSIFICATION..., 1964, p. 147).

Dois pontos emergiram desta discussão: o primeiro foi que a linguagem descritiva mais detalhada trazia maiores chances de erro tanto para o indexador, quanto para o usuário; e o segundo foi a conclusão de que o conceito relevância, da forma em que estava sendo aplicado nos sistemas de indexação, era bastante amplo na atribuição dos assuntos dos documentos. Um cabeçalho de assunto poderia corresponder a um termo de pesquisa, mas a referência

recuperada ainda poderia ser não pertinente se, por exemplo, o documento estivesse em um nível inadequado para o leitor ou fosse, por algum motivo, uma apresentação incompreensível para ele.

Assim, recomendava-se que a indexação deveria primar pela especificidade e dessa forma ajudar o pesquisador a tornar seus requisitos de busca mais claros para si mesmo. Langridge levantou a questão do uso dos termos “genérico” e “hierárquico”. Verificou que estes termos criavam dificuldades e sugeriu a utilização do termo “relação parte-todo” com o argumento de que parte-todo e espécie-gênero eram ambas relações de inclusão e podiam, por conseguinte, ser descritas pelo mesmo termo (CLASSIFICATION..., 1964, p. 148). No entanto, as relações não poderiam ser resumidas a apenas este aspecto.

O Grupo procurou realizar uma análise para descobrir até que ponto o uso desses termos poderia alcançar certa consistência entre eles para: (1) para atribuir títulos de assuntos acordados a partir da literatura da área e (2) para analisar as relações entre os conceitos. Os relatos mostraram que desacordos surgiam na fase de análise conceitual, mas eram resolvidos no processo de tradução para a linguagem do sistema, uma vez que os conceitos haviam sido acordados (CLASSIFICATION..., 1964, p. 149).

Ao final da década de 1960, essas lideranças do CRG como Vickery, Foskett, Farradane, Langridge, Lynch e Cleverdon avaliaram esses métodos que articulavam os processos de classificação com a indexação e suas aplicações nos sistemas de informação como no Centro de Documentação do *Engineers Joint Council* (BLAGDEN, 1966), na *Chemical Abstracts* (LYNCH, 1966), no Centro de Informação e Documentação da *Euratom* (MAUPERON, 1967) e no *American Institute of Physics* (FOSKETT, 1970), sobretudo com a adoção cada vez maior e rotineira dos computadores para a organização e tratamento da documentação.

Os resultados dessas avaliações fogem ao escopo deste artigo, no entanto podemos apontar que havia consenso entre os autores que os processos de classificação da documentação e de indexação de seus conteúdos eram mais eficazes quando articulados nos sistemas de informação. Outras questões surgiram como o papel do classificador e do indexador no tratamento temático da informação, além da preocupação com o usuário e suas demandas informacionais. Esses aspectos aproximaram as abordagens da classificação e da indexação à teoria cognitiva, especialmente dos trabalhos de Van Dijk e Kintsch (1983).

Entretanto, entendemos que os estudos de classificação e indexação que se aproximam da compreensão do processo cognitivo no tratamento temático da informação

fazem parte de outro contexto sócio-histórico com práticas e metodologias que se distanciam da análise proposta por esse trabalho. Portanto, somente apontamos tais caminhos à guisa de delimitar o nosso objeto e fornecer indícios sobre esta temática complexa, que são as relações entre a classificação e a indexação, longe da pretensão de esgotá-la.

5 CONSIDERAÇÕES FINAIS

Destarte, procuramos apresentar algumas contribuições do *Classification Research Group* na abordagem do tratamento temático da informação que articularam os processos da classificação e da indexação.

Cabe destacar que essas propostas realizadas no âmbito do *CRG* estavam em consonância com as demandas dos centros de documentação e das bibliotecas especializadas para o desenvolvimento de áreas consideradas estratégicas por governos e grandes empresas no contexto do capitalismo desenvolvimentista, período da Guerra Fria, entre as décadas de 1950 e 1960. Os teóricos do *CRG* estavam elaborando pesquisas e soluções para atender as demandas de suas próprias realidades nas instituições que faziam parte.

A ***Colon Classification*** e a teoria facetada de Ranganathan foram importantes constructos nas aproximações entre a classificação e a indexação, realizadas pelo *CRG*, apresentando inclusive modelos para os sistemas de informação. Este artigo procurou destacar, nos resultados apresentados, os usos e aplicações propostas pelo *CRG* para o tratamento temático da informação, baseados, sobretudo, nos princípios classificatórios de Ranganathan que tiveram grande influência no grupo.

A incorporação dos computadores na Segunda Guerra Mundial e o aprofundamento de sua utilização nos meios de comunicação impactaram também as formas de organização da informação e no tratamento documental. Os primeiros centros de documentação e sistemas de informação automatizados trabalhavam com a lógica de estabelecer parâmetros para a representação de sua documentação, isto é, pensar a padronização de códigos, de dados, de informações para o funcionamento adequado, eficaz e eficiente dos sistemas, no sentido, em que as máquinas pudessem compreender e executar os comandos dados e que o fluxo de informações ocorresse sem interrupções.

As padronizações dos esquemas classificatórios e das técnicas de indexação facilitavam o desenvolvimento de sistemas automáticos de recuperação da informação. Esses complexos

processos intelectuais poderiam ser transformados em operações sistemáticas e rotineiras de representação executadas por máquinas a partir das matrizes dadas pelos profissionais da informação .

Essa aproximação das abordagens da classificação daquelas da indexação, procurando o estabelecimento de padrões para a representação dos documentos e de suas informações, convergiu para a elaboração de linguagens documentárias que, além de facilitar a interface entre os procedimentos para o tratamento temático da informação realizados pelo trabalho humano e pelos computadores, permitiam maior interação entre os centros de documentação e as bibliotecas especializadas.

Assim, concluímos esta análise acerca das contribuições do *Classification Research Group* no tocante às aproximações realizadas entre a classificação e a indexação para o tratamento temático da informação e a organização do conhecimento. Entendemos que a representação documentária e a recuperação da informação devem ser analisadas como processos comunicacionais interativos e inseridos em um contexto situacional de múltiplas relações sociais, culturais, políticas, fazendo parte de acordos e tensionamentos da sociedade à qual pertencem.

Ressaltamos que este resultado cartográfico é um sobrevoo em complexas abordagens sobre fundamentos do tratamento temático da informação e da organização do conhecimento sem qualquer pretensão de apresentar totalidades. Trata-se de um esforço para evidenciar as contribuições do CRG, sobretudo para melhor compreender e realizar processos complexos como a classificação e a indexação, buscando suas possíveis aproximações. Tais abordagens seguem influenciando estudos, práticas e aplicações na área.

REFERÊNCIAS

BLAGDEN, J.F. How much noise in a role-free and link-free co-ordinate indexing system? **Journal of Documentation**, v. 22, n. 3, p. 203-209, 1966.

BROUGHTON, V. Brian Vickery and the Classification Research Group: the legacy of faceted classification. In: GILCHRIST, A; VERNAU, J. (Ed.). **Facets of knowledge organization**. Bingley: Emerald, 2012. p 315-326.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 4. **Journal of Documentation**, v. 14, n. 3, p. 136-143, 1958.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n.6. **Journal of Documentation**, v. 17, n. 2, p. 156-172, 1961.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 7. **Journal of Documentation**, v. 18, n. 2, p. 65-88, 1962.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 8. **Journal of Documentation**, v. 20, n. 3, p. 146-169, 1964.

CLEVERDON, C. W. Automation in indexing, **Aslib Proceedings**, v. 13, n. 4, p. 107-109, 1961.

CORDEIRO, R. I. N.; LA BARRE, K. Análise de facetas e obra fílmica. **Informação & Informação**, Londrina, v. 16, n. 3, p. 180-201, jan./jun. 2011.

FARRADANE, L.E.L. A scientific theory of classification and indexing: further considerations. **Journal of Documentation**, v. 8, n. 2, p. 73 - 92, 1952.

FOSKETT, D. J. Classification and indexing in the Social Sciences, **Aslib Proceedings**, v. 22, n. 3, p.90-10, 1970.

_____. **Classification for a general index language**. London: The Library Association, 1970a.

_____. Two notes on indexing techniques. **Journal of Documentation**, v. 18, n. 2, p. 188-192, 1962.

GUIMARÃES, J. A. Abordagens teóricas no tratamento temático da informação (TTI): catalogação de assunto, indexação e análise documental. **Ibersid**, p. 105-117, 2009.

_____. A 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, jan./jun. 2008.

_____. A análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais". In: RODRIGUES, G. M.; LOPES, I. L. (Org.). **Organização e representação do conhecimento na perspectiva da ciência da informação**. Brasília: Thesaurus, p. 100-117, 2003.

LA BARRE, Kathryn. Facet analysis. **ARIST**, v. 44, p. 243-284, 2010.

LYNCH, M. F. Subject indexes and automatic document retrieval: the structure of entries in Chemical Abstracts subject indexes. **Journal of Documentation**, v. 22, n.3, p.167-185, 1966.

MAUPERON, A. Adaptation to users' needs. **Aslib Proceedings**, v. 19, n. 7, p.232-240, 1967.

NUNES, L.; TÁLAMO, M. F. G. M. Da filosofia da classificação à classificação bibliográfica. **Revista Digital de Biblioteconomia e Ciência da Informação**, v. 7, n. 1, p. 30-48, jul./dez. 2009.

SPITERI, L. A simplified model for facet analysis: Ranganathan 101. **Canadian Journal of Information and Library Science**. v. 23, p. 1-30, 1998.

VAN DIJK, T.A; KINTSCH, W. **Strategies of discourse comprehension**. New York: Academic Press, 1983.

VICKERY, B. C. Developments in subject indexing, **Journal of Documentation**, v. 11, n. 1, p. 1 - 11, 1955.