
O POTENCIAL DA CLASSIFICAÇÃO DECIMAL UNIVERSAL (CDU) NA RECUPERAÇÃO DA INFORMAÇÃO NA WEB

THE POTENTIAL OF UNIVERSAL DECIMAL CLASSIFICATION (UDC) IN THE INFORMATION RETRIEVAL ON THE WEB

Ítalo Rodrigues Castro

*Doutorando em Gestão e Organização do Conhecimento - UFMG
italocastro@gmail.com*

Gercina Ângela de Lima

*Doutora em Ciência da Informação
Professora na Escola de Ciência da Informação e no PPGGOC/ECI/UFMG
glima@eci.ufmg.br*

Resumo

Os Sistemas de Organização do Conhecimento, especificamente a Classificação Decimal Universal, objetiva aumentar a qualidade dos relacionamentos dos conceitos de um domínio, tendo potencial de uso nos sistemas de recuperação on-line. O propósito deste artigo é apresentar metodologias utilizadas com a Classificação Universal atrelada a um tesouro, haja vista melhorar a recuperação da informação. Os procedimentos metodológicos utilizados classificam a pesquisa como descritiva, qualitativa e exploratória, recorrendo à aplicação de uma revisão da literatura. Para seleção das fontes de consulta, utilizaram-se as principais bases de dados da área da Ciência da Informação, sem recorte temporal, sendo escolhidos, apenas, os trabalhos que abordavam o uso da Classificação Decimal Universal para melhorar a recuperação da informação no ambiente web. Os resultados revelaram vários aspectos relativos ao contexto desse tema, as bases teóricas e fundamentos conceituais, identificando estudos antigos e recentes, o que possibilitou apontar práticas e perspectivas futuras relativas ao uso dos sistemas de classificação bibliográfica. Com o mapeamento da literatura, concluiu-se que o potencial da Classificação Universal está em correspondência com os termos de um tesouro relacionado. Percebeu-se que a temática ainda foi pouco explorada, com predominância de artigos publicados em língua estrangeira, motivo pelo qual é necessária maior atenção no que se diz respeito ao potencial desse sistema de classificação para a recuperação da informação na web.

Palavras-chave: Sistemas de Organização do Conhecimento. Sistemas de Classificação. Classificação Decimal Universal. Recuperação da Informação na web.

Abstract

The Knowledge Organization Systems, specifically the Universal Decimal Classification, aims to increase the quality of the relationships of the concepts of a domain, having potential use in on-line recovery systems. The purpose of this article is to present methodologies used with the Universal Classification linked to a thesaurus, in order to improve information retrieval. The methodological procedures used classify the research as descriptive, qualitative and exploratory, using a literature review. To select the sources of consultation, the main databases in the area of Information Science were used, without time frame, being chosen only the works that addressed the use of the Universal Decimal Classification to improve the retrieval of information in the web environment. The results revealed several aspects related to the context of this theme, the theoretical bases and conceptual foundations, identifying old and recent studies, which made it possible to point

out future practices and perspectives regarding the use of bibliographic classification systems. With the mapping of the literature, it was concluded that the potential of the Universal Classification is in correspondence with the terms of a related thesaurus. It was noticed that the theme has still been little explored, with a predominance of articles published in a foreign language, which is why it is necessary to pay more attention to the potential of this classification system for the information retrieval on the web.

Keywords: Knowledge Organization Systems. Classification Systems. Universal Decimal Classification. Information Retrieval on the web.

1 INTRODUÇÃO

Com o avanço das tecnologias, é possível solucionar problemas causados pela falta de organização da informação em diversos ambientes informacionais. Para organizar a informação, o uso dos Sistemas de Organização do Conhecimento (SOCs) é fundamental, mesmo os tradicionais, cujo objetivo desse processo é facilitar a recuperação da informação on-line. Todavia, o potencial do uso de estruturas de conhecimento na interface do usuário ainda não tem sido muito disseminado, sobretudo no ambiente web.

Os SOCs, quando incorporados aos serviços da web, podem funcionar como mapas semânticos que orientam os usuários nas buscas. Embora, essa orientação também é facilitada com o tratamento temático dos documentos, com uso de instrumentos de controle terminológico, como os tesouros. Esses instrumentos, segundo Gomes (1996), são criados para organizar os assuntos tratados nos documentos. A autora salienta que os tesouros orientam os procedimentos de busca em uma base de dados, haja vista padronizar o uso da terminologia. Portanto, os SOCs e os tesouros não auxiliam, apenas, a indexação das informações, mas também a recuperação.

Nesse contexto, este artigo objetiva realizar uma revisão da literatura referente à produção científica e tecnológica sobre o potencial do uso da Classificação Decimal Universal (CDU) como instrumento para facilitar a recuperação da informação na web. Foi formulada a seguinte questão de pesquisa para esta revisão: “Quais pesquisas foram realizadas sobre o potencial da CDU com um tesouro relacionado para melhorar a recuperação da informação na web?”.

A partir desta delimitação, o presente artigo foi estruturado em seis seções, já incluída esta parte introdutória, a saber: referencial teórico, procedimentos metodológicos utilizados, os resultados e discussão, as considerações finais e a lista de referências. No referencial teórico, foram abordados os fundamentos teóricos e conceituais referentes aos seguintes aspectos: SOCs, CDU, tesouro e recuperação da informação.

2 REFERENCIAL TEÓRICO

As subseções seguintes abordam assuntos associados ao problema de investigação deste artigo. Especificamente, são descritos conceitos e fundamentos inerentes à CDU e os tesouros como tipos de SOCs, com foco no processo de recuperação da informação.

2.1 SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

A expressão SOC designa todos os tipos de esquemas para organizar informações e promover a gestão do conhecimento, como os sistemas de classificação, *gazetteers*, bancos de dados lexicais, taxonomias, tesouros e ontologias (HODGE, 2000). Esses sistemas modelam a estrutura semântica subjacente de um domínio, capaz de auxiliar a navegação e tradução por meio de rótulos, definições, digitação, relacionamentos e propriedades para os conceitos (HILL *et al.*, 2004; KOCH; TUDHOPE, 2004; ZENG, 2008). Utilizados como serviços web, tais sistemas facilitam a busca e recuperação de recursos, agindo como mapas semânticos, o que possibilita uma orientação comum para indexadores e futuros usuários, humanos ou máquinas (KOCH; TUDHOPE, 2003, 2004).

A representação da informação é constituída por um conjunto de atributos capaz de representar um determinado recurso, a partir da descrição do conteúdo informacional, traduzido em termos que representam esse conteúdo. Os SOCs são utilizados para a representação de um domínio, cujos recursos informacionais estão inseridos, para que o usuário consiga recuperar documentos relevantes. Sendo assim, torna-se necessário verificar se o sistema utilizado está compatível com a linguagem do usuário,

considerando a necessidade desse sujeito. Para isso, os SOCs, como os sistemas de classificação bibliográfica e os tesouros, podem sofrer alterações estruturais para representar um determinado domínio do conhecimento, de modo a atender a demanda dos usuários (PRASAD; MADALLI, 2009).

Brascher (2009) corrobora o pensamento de Prasad e Madalli (2009), reforçando a importância do uso dos SOCs para a recuperação de documentos confiáveis, nas buscas em bases de dados, seja em ambiente digital ou na web. No contexto digital, Souza e Alvarenga (2004) descrevem os padrões da web semântica, como por exemplo, o padrão *Simple Knowledge Organization System* (SKOS) que utilizam as linguagens *Xtensible Markup Language* (XML) e *Hyper Text Markup Language* (HTML) para representar o documento em termos semânticos e os significados contextuais, a partir de *tags*, as quais descrevem o conteúdo de um documento.

O SKOS apoia o uso dos SOCs para a organização e representação da estrutura básica e conteúdo de sistemas conceituais com diretrizes para a construção de tesouros, sistemas de classificação, entre outros tipos de vocabulário controlado de forma estruturada no âmbito da web semântica. A ideia geral inerente a esses padrões é a disponibilidade e a troca de vocabulários controlados em ambiente web, podendo contribuir para a descoberta de novos recursos por meio de referências, resolução de ambiguidades da linguagem natural e o fornecimento de contexto semântico para o processamento de texto (BAKER *et al.*, 2013).

Os sistemas de classificação bibliográfica manifestam-se como o agrupamento de entidades ou conceitos com as mesmas características ou que possuem alguma relação, por pertencerem aos mesmos níveis na hierarquia ou nas adjacências. Eles permitem selecionar os termos que representam o conhecimento humano, a partir da combinação dos conceitos que melhor definam um assunto, por meio da combinação de termos identificados (PIEIDADE, 1977).

Nesse contexto, existem diversos SOCs que permitem classificar o conhecimento nas unidades de informação, tal como a Classificação da Biblioteca do Congresso, que é um sistema flexível e que se pode expandir. Outros sistemas de classificação bibliográfica que são bem conhecidos são a Classificação Decimal de Dewey (CDD), um sistema de classificação rígido, enumerativo, de dez classes numéricas com extensões decimais; e a CDU, um sistema também de natureza enumerativa, porém com maiores possibilidades de associações e relacionamentos (SOUZA, 2012). A CDU, que é o objeto de análise desta pesquisa, será detalhada na próxima subseção.

2.2 CLASSIFICAÇÃO DECIMAL UNIVERSAL

A CDU é reconhecida por possuir uma estrutura enumerativa e hierárquica, que pode ser incrementada e, frequentemente, precisa ser adaptada e modificada para acompanhar as mudanças e a evolução do conhecimento. A expansão dos assuntos é advinda das diversas possibilidades de relacionamento entre eles, como acontece com a classificação facetada (HODGE, 2000).

O sistema de classificação facetada foi proposto por Ranganathan, conceituado como um método de classificar objetos, coisas e/ou processos do mundo (CAMPOS, 2001). É na obra *Prolegomena*, de Ranganathan (1967), que estão elencados os princípios e cânones, considerando três distintos planos: plano das ideias, plano verbal e o plano notacional. Portanto, entende-se que essa classificação é frequentemente utilizada na construção de diferentes tipos de SOCs (GOMES; CAMPOS, 2019).

Por sua vez, a CDU também se classifica como um sistema enumerativo e hierárquico, muito semelhante à CDD, sendo que essa possui uma estrutura mais rígida e com assuntos mais amplos. A Classificação Decimal Universal (CDU) surgiu no final do século XIX, durante a Conferência Internacional de Bibliografia, quando os belgas Paul Otlet e Henry La Fountain receberam a tarefa de organizar o Repertório Bibliográfico Universal. Com essa missão, os autores utilizaram como base a CDD, sobretudo por ela facilitar a organização dos documentos de forma impressa ou digital. Salienta-se que, em ambos os sistemas, os documentos podem ser organizados de forma lógica, sendo classificados por assunto e agrupados por semelhança para posterior recuperação da informação pelos usuários em um sistema de recuperação da informação (PIEIDADE, 1977).

Slavic (2005) afirma que a CDU é um sistema internacional para classificação de documentos, haja vista representar todo o conhecimento humano dentro de uma hierarquia decimal, e também possui um índice alfabético. A autora revela que a CDU é utilizada para a organização de documentos nas estantes e bibliografias universais e especializadas de acordo com assuntos e notações, auxiliando na classificação dos documentos em diversas áreas do conhecimento.

Na CDU, a representação das categorias (entidades) e dos termos utiliza um código único, permitindo a interoperabilidade e reuso dos dados. Dessa forma, as práticas de organização do conhecimento, mediante a utilização desse sistema e de um tesauro, possibilitaram um aumento significativo na quantidade de termos e relações, com a finalidade de permitir a representação de um domínio com maior eficiência na recuperação da informação (FRÂNCU, 2003b).

2.3 TESAUROS

Os tesauros são instrumentos de controle terminológico utilizados em sistemas de informação por indexadores e pesquisadores, tendo em vista traduzir a linguagem dos documentos em uma linguagem controlada. Esses instrumentos podem ser utilizados tanto na indexação quanto na recuperação de informações (BRASCHER, 2009). Para Hjørland (2016), o tesauro é uma ferramenta semântica flexível (vocabulário controlado) que, além de possuir uma estrutura com relações de hierarquia existente nos sistemas de classificação, inclui diferentes conjuntos de relações semânticas significativas entre os termos, tais como: associativas, hierárquicas, de equivalência, entre outras.

Os termos de um tesauro são utilizados para representar um determinado domínio, eliminando a ambiguidade da linguagem natural, por conseguinte, facilitando a recuperação. Desse modo, o tesauro pode ser definido como uma lista de termos autorizados, com a função de conduzir o usuário de um conceito para outro de forma intuitiva (FOSKETT, 1973). O tesauro é utilizado em um domínio específico, onde os termos representam uma linguagem padronizada e pós-coordenada. Essa linguagem contribui para a recuperação da informação, sendo composta por termos simples ou compostos que estão relacionados de modo sintático e semântico (CURRÁS, 1995).

As normas *internacionais International Organization for Standardization (ISO) 25964-1 (2011) e ISO 25964 (2013)* são dedicadas à construção de tesauros com outros tipos de SOCs, levando em conta a interoperabilidade na web. Essas normas visam à conexão do tesauro durante os processos de construção, manutenção, disseminação e troca de dados na web. Além disso, elas são importantes para a normalização dos tipos de mapeamento, haja vista a integração de tesauros que são capazes de associar um conceito a uma ou mais linguagens, mono ou multilíngue. Para isso, são estabelecidas relações semânticas adaptadas com a terminologia de cada linguagem com outros SOCs, para fins de recuperação da informação na web.

2.4 RECUPERAÇÃO DA INFORMAÇÃO

Com o aumento exponencial do volume de informação disponível, é necessária, cada vez mais, a utilização de métodos avançados de recuperação, de modo a atender, com mais eficiência, às necessidades dos usuários. Saracevic (1999) menciona a recuperação como um dos campos mais importantes da Ciência da Informação, devido às relações multi e interdisciplinares dessa área com outras áreas do conhecimento. O autor destaca a recuperação como um processo de comunicação entre o emissor (usuário) e o receptor (sistema de recuperação da informação). Assim, o usuário emite uma pergunta para o sistema de recuperação da informação, o qual disponibiliza as informações pertinentes ou relevantes para satisfazer uma necessidade de informação (SARACEVIC, 1999).

Em sistemas de informação utilizados em ambientes digitais ou na web, torna-se necessária a organização da informação, por meio da indexação (análise de assunto e tradução) e da recuperação da informação (necessidade de informação e estratégias de busca). Em linhas gerais, essas atividades contemplam os processos de produção, armazenamento e disseminação da informação. É importante destacar que a organização da informação, no ambiente digital, usa metadados para representar e recuperar documentos armazenados, normalmente, em bases de dados. Esses documentos são representados por meio do tratamento descritivo (características extrínsecas) e temático (características intrínsecas ou de conteúdo), de modo a tornar esses recursos passíveis de recuperação, em face de um processo de busca pelo usuário (SARACEVIC, 1996).

Durante o processo de indexação, são produzidos índices, que serão utilizados para a representação do conteúdo informacional dos documentos, com o objetivo de facilitar a recuperação da informação. O índice permite aos usuários selecionarem documentos que tenham sido armazenados em bases de dados, utilizando na busca, as ferramentas e estratégias do sistema de recuperação da informação (SRI). Nas atividades de busca, podem ser utilizadas estruturas de dados conhecidas como arquivos invertidos, cujos termos apontam para os documentos, como também o número de ocorrências

e a posição desses termos em cada documento, auxiliando na precisão durante a recuperação da informação (BAEZA-YATES; RIBEIRO-NETO, 1999).

Na recuperação da informação, o usuário de um sistema pode dar um *feedback* acerca dos documentos recuperados, informando quais foram mais úteis. Durante a busca, os usuários podem formular uma expressão de busca em linguagem natural, controlada ou por palavras-chave. Outra estratégia de busca é por navegação (*browsing*), cujo usuário navega por páginas selecionando links que sejam pertinentes à necessidade. Nesse sentido, quanto maior a especificidade durante a busca, maior será a precisão, diminuindo a revocação. Os modelos clássicos de recuperação da informação (booleano, vetorial e probabilístico) utilizam a estatística, lógica booleana e teoria de conjuntos. Esses modelos são muito utilizados pelo SRI e para mecanismos de busca na web, em que os documentos são representados por um conjunto de termos que indicam o assunto dos documentos (SOUZA, 2006).

3 PROCEDIMENTOS METODOLÓGICOS

Este trabalho caracteriza-se como uma pesquisa descritiva, qualitativa e exploratória. Na coleta de dados, recorreu-se a uma extensa pesquisa bibliográfica, do tipo revisão da literatura, considerando como tema investigado o potencial da CDU com um tesouro relacionado para melhorar a recuperação da informação na web. Foi realizada uma pesquisa exaustiva nas principais bases de dados do campo da Ciência da Informação com conteúdo multidisciplinar: *Library and Information Science Abstracts (LISA)*, *Information Science & Technology Abstracts (ISTA)*, *Library, Information Science & Technology Abstracts with Full Text (LISTA)*, *Scopus* e *Web of Science*.

A partir da questão da pesquisa, foram elaboradas estratégias de busca com as seguintes palavras-chave: “*udc*”, “*information retrieval*”, “*thesaurus*”, “*thesauri*”, “*cdu*”, “*recuperación de información*”, “*tesouro*”, “*tesouros*”, “*recuperação da informação*”, “*universal decimal classification*”, “*clasificación decimal universal*” e “*classificação decimal universal*”. Nas expressões de busca, combinaram-se operadores booleanos (*AND* e *OR*), nos idiomas inglês, espanhol e português.

Quanto aos critérios de inclusão, citam-se: trabalhos do tipo artigos de periódicos, anais de congresso, teses e dissertações; com abordagem teórico-metodológica e prática; e sem recorte temporal, pois se trata de uma temática ainda pouco explorada. Por outro lado, os critérios de exclusão foram: monografias, editoriais, prefácios, sumários, entrevistas, notícias, revisões, tutoriais, *workshops*, painéis, pôsteres e publicações pagas, como tipologia documental; estudos que estejam em desacordo com a questão de pesquisa, referindo-se, mais especificamente: à utilização de outros sistemas de classificação bibliográfica, à indexação automática utilizando a CDU ou outros sistemas de classificação, à integração do vocabulário entre diferentes sistemas de classificação bibliográfica, a testes de usabilidade em interfaces com usuários, e a estudos comparativos entre sistemas de classificação bibliográfica; trabalhos que não estejam escritos nos idiomas português, espanhol e inglês; e publicações duplicadas.

4 RESULTADOS E DISCUSSÃO

A revisão da literatura teve como objetivo identificar os estudos relevantes a respeito do potencial da CDU com um tesouro relacionado para melhorar a recuperação da informação na web. No Quadro 1 são apresentados alguns dados coletados durante a pesquisa, indicando a presença de 44 publicações selecionadas, sendo a mais antiga publicada em 1964, até as mais recentes, no ano de 2017, com sua autoria (autores e coautores) e ano; título e os tipos de documentos.

Quadro 1 - Publicações selecionadas na revisão de literatura (1964-2017)

Nº	Autoria e ano	Título	Tipos de documento
1	Freeman (1964)	Computers and Classification Systems: Journal of Documentation	Journal of Documentation
2	Mills (1970)	Progress in Documentation: Library Classification	Journal of Documentation
3	Dahlberg (1971)	Possibilities for a new universal decimal Classification	Journal of Documentation
4	Wall (1973)	Indexing language structure for automated retrieval	Information Storage and Retrieval
5	Hindsoti (1979)	Reflections on the utilization of the Universal Decimal Classification	Information Storage and Retrieval

6	Wall (1980)	Intelligent indexing and retrieval: A man-machine partnership	Information Processing and Management
7	Fiałkowski, Muraszkiwicz e Tolwińska (1981)	Methodology of world clearinghouse activity	Information Storage and Retrieval
8	Buxton (1990)	Computer searching of UDC numbers	Journal of Documentation
9	Mcilwaine (1997)	The Universal Decimal Classification: Some Factors Concerning Its Origins, Development, and Influence	Journal of the American Society for Information Science
10	Chan e Zeng (2002)	Ensuring interoperability among subject vocabularies and knowledge organization schemes: a methodological analysis	IFLA Journal
11	Frâncu (2003a)	Multilingual access to information using an intermediate language	Tese
12	Frâncu (2003b)	The impact of specificity on the retrieval power of a UDC-based multilingual thesaurus	Cataloging & classification
13	Chudamani (2004)	Classification Model for Libraries in the Digital Environment	International Caliber
14	Schallier (2004)	What a subject search interface can do	Extensions & Corrections to the UDC
15	Slavic (2004)	UDC implementation: from library shelves to a structured indexing language	International Cataloging & Bibliographic Control
16	Slavic (2005)	Classification management and use in a networked environment: the case of the Universal Decimal Classification	Tese
17	Slavic (2006a)	Interface to classification: some objectives and options Item type Preprint Interface to Classification: Some Objectives and Options	Extensions and Corrections to the UDC
18	Slavic (2006b)	The level of exploitation of Universal Decimal Classification in library OPACs: A pilot Study	Vjesnik bibliotekara Hrvatske
19	Slavic (2006c)	UDC in subject gateways: experiment or opportunity?	Knowledge Organization
20	Frâncu (2007)	Does convenience trump accuracy? The avatars of the UDC in Romania	Extensions and Corrections to the UDC
21	Gnoli (2007)	Progress in synthetic classification: towards unique definition of concepts	Extensions and corrections to the UDC
22	Pika (2007)	Universal Decimal Classification at the ETH-Bibliothek Zürich-a Swiss perspective	UDC Consortium
23	Slavic, Cordeiro e Riesthus (2007)	Enhancement of UDC data for use and sharing in a networked environment Item type Conference Paper Enhancement of UDC data for use and sharing in a networked environment	The 31st Annual Conference of the German Classification Society on Data Analysis, Machine Learning, and Applications Artigo de conferência
24	La Barre (2007)	The heritage of early FC in document reference retrieval systems, 1920-1969	Library History
25	Colillas e Nicolau (2007)	Aplicación de un nuevo sistema de indización en una colección de recursos especializados en ciencias de la educación	El profesional de la información
26	Broughton e Slavic (2007)	Building a faceted classification for the humanities: principles and procedures	Journal of Documentation
27	Hjørland (2007)	Arguments for 'the bibliographical paradigm': Some thoughts inspired by the new English edition of the UDC	The Sixth International Conference on Conceptions of Library and Information Science-"Featuring the Future"

			Artigo de conferência
28	Colillas e Nicolau (2008)	Improving subject searching in databases through a combination of descriptors and UDC	Zadar: Bobcatsss
29	Kaosar (2008)	Merits and Demerits of using Universal Decimal Classification on Internet	Det Informationsvidenskabelige Akademi (IVA) Dissertação
30	Slavic (2008)	Faceted classification: management and use	Axiomathes
31	Slavic, Cordeiro e Riesthuis (2008)	Maintenance of the Universal Decimal Classification: overview of the past and preparations for the future	International Cataloguing and Bibliographic Control Journal
32	Dahlberg (2008)	The Information Coding Classification (ICC): A modern, theory-based fully-faceted, universal system of knowledge fields	Axiomathes
33	Balíková (2009)	The role of UDC classification in the Czech Subject Authority File	Extensions & Corrections to the UDC
34	Colillas (2009)	Tratamiento de los Metadatos de Contenido en la Web Semántica o cuando la Lógica se hace Evidente	ISKO Artigo de Conferência
35	Hajdu-Barát (2009)	Integration of a thesaurus and Universal Decimal Classification (UDC) to improve subject access: the Hungarian experience	Extensions & Corrections to the UDC
36	Segundo (2009)	Crossroads: Multiple Directions to Usability Using MARC classification format for UDC and mappings to other KO systems for an enriched authority file	Extensions & Corrections to the UDC
37	Broughton (2010)	Concepts and terms in the faceted classification: the case of UDC	Knowledge Organization
38	Frâncu e Sabo (2010)	Implementation of a UDC-Based Multilingual Thesaurus in a Library Catalogue: The Case of BiblioPhil	Knowledge Organization
39	Doorn e Polman (2010)	From classification to thesaurus ... and back? Subject indexing tools at the library of the Afrika-Studiecentrum Leiden	Knowledge Organization
40	McIlwaine (2010)	Classification at a Crossroads: Multiple Directions to Usability	International Cataloguing and Bibliographic Control: Quarterly Bulletin of the IFLA UBCIM Programme
41	Colillas (2011)	UDC on the Internet: Theory and project in evolution for use of indexing and retrieval systems	IFLA Journal
42	Vukadin e Slavic (2014)	Challenges of Facet analysis and Concept Placement in Universal Classifications: the Example of Architecture in UDC	Knowledge Organization
43	Lee (2017)	Numbers, instruments and hands: the impact of faceted analytical theory on classifying music ensembles	Knowledge Organization
44	Slavic e Davies (2017)	Facet Analysis in UDC: Questions of Structure, Functionality and Data Formality	Knowledge Organization

Fonte: Dados da pesquisa.

A seleção foi conduzida pelos critérios de inclusão e exclusão, anteriormente estabelecidos, evidenciando as principais metodologias, técnicas e resultados alcançados, em cada estudo. Esses trabalhos foram agrupados de acordo com as características semelhantes de cada um, seguido de uma análise de alguns desses estudos organizados em ordem cronológica.

O primeiro trabalho selecionado na literatura foi proposto por Freeman (1964), que apresentou a estrutura da CDU, porém com a ordem de armazenamento diferente da estrutura original. A metodologia mostrou que os caracteres dos dados poderiam ser substituídos por outros caracteres ou ordenados de outro modo, onde a busca poderia ser realizada apenas por assunto, utilizando uma terminologia exata em linguagem natural e por índices de assuntos, ordenados por ordem alfabética ou

catálogos de documentos em um SRI. Os resultados da pesquisa apontaram que a notação classificatória da CDU pode ser uma ferramenta bastante eficiente em sistemas de informação, principalmente para mostrar as relações hierárquicas utilizando um tesouro (FREEMAN, 1964). Esse aspecto também pode ser observado na pesquisa de Hindsoti (1979).

Mills (1970) propôs uma classificação moderna para bibliotecas, baseada em índices, com uso da CDU associada a um tesouro, sendo considerado como um sistema pós-coordenado que auxilia na recuperação da informação. Essa metodologia baseou-se na criação de um índice na CDU para os termos de um tesouro, levando em consideração os sinônimos, termos gerais e específicos e suas relações, cujo conceito deve estar em um único lugar e deve ser atribuída uma única notação. Os resultados mostraram a consistência interna e a eficácia dos termos da CDU relacionada com o tesouro, sendo conectados de forma explícita em uma base única com a definição dos conceitos relacionados (MILLS, 1970). Os mesmos procedimentos metodológicos foram aplicados, anos depois, por Wall (1973) e Wall (1980).

No artigo de Dahlberg (1971) realizou-se a realocação de assuntos das classes principais da CDU, de maneira que todos os conceitos e objetos existentes pudessem ser integrados em um sistema. Os procedimentos metodológicos utilizados foram por meio de um sistema estruturado, logicamente, para facilitar a utilização e leitura pelo computador. Além disso, o sistema possibilita subdividir os assuntos, prevenindo omissões de assuntos sem prejudicar a notação. A partir dos princípios da análise facetada existentes na CDU, foi possível incluir, na nova classificação, um maior número de categorias para contemplar os conceitos e vocabulários existentes em um, domínio, a partir de um tesouro. Os resultados demonstraram que a utilização da CDU e de um tesouro, podem também fornecer aos bibliotecários e documentalistas um número maior de categorias de assuntos (DAHLBERG, 1971). Buxton (1990), Dahlberg (2008) e Hjørland (2007) também concluíram que a metodologia utilizada por Dahlberg (1971) conecta todas as disciplinas do conhecimento humano com um vocabulário controlado.

No trabalho de Fiałkowski, Muraszkiwicz e Tołwińska (1981), estabeleceram-se os princípios e a metodologia do sistema *Clearinghouse* de Varsóvia, que fornece um serviço bibliográfico especializado para coletar, processar e disseminar informações sobre tesouros, sistemas de classificação, descritores, palavras-chave e listas de cabeçalho de assunto. Na recuperação, é fornecido o número da CDU e as palavras-chave, que estão armazenadas em uma base de dados. Como resultado, foi necessário automatizar a recuperação da informação (RI) da *Clearinghouse*, no sistema *Thesauri e CLAssification* (TEKLA). Tal sistema processa a bibliografias e descrições de tesouros, esquemas de classificação, palavras-chave, categorias de assuntos dos documentos, idiomas e anos de publicação, sendo esses elementos armazenados e processados pelo sistema (FIAŁKOWSKI; MURASZKIEWICZ; TOŁWIŃSKA, 1981).

Na literatura, um estudo semelhante aos resultados de Fiałkowski, Muraszkiwicz e Tołwińska (1981) é o desenvolvido por Mcilwaine (1997). Os procedimentos metodológicos utilizados na citada pesquisa foram a partir de números da CDU e por palavras-chave, em vez de utilizar um SRI. Os resultados revelaram que as pesquisas por palavras-chave mostram que a mesma palavra pode estar em mais de um número de classificação da CDU. Portanto, quanto menor o número no sistema de classificação, menor a quantidade de locais que estará à palavra-chave. Nesse sentido, com o advento da automação, é possível desenvolver um esquema eficiente na RI, que não seja, apenas, a simples ordenação de documentos nas estantes.

No trabalho de Chan e Zeng (2002), foi apresentada a interoperabilidade de diferentes sistemas de organização do conhecimento, como vocabulários controlados, sistemas de classificação e tesouros. Verificou-se que todos os instrumentos podem ser interoperáveis como se estivessem todos dentro de um único SRI. Desse modo, criou-se um esquema único de classificação baseado na CDU, mapeando os termos existentes desse sistema de classificação com os termos de um tesouro adicional para cada assunto seja um novo ou mais específico, criando-se mapas mentais do usuário que utilizam termos representativos para otimizar a RI. Os resultados alcançados foram a necessidade de interoperabilidade entre vocabulários multilíngues na organização do conhecimento em um SRI, sendo necessário construir uma concordância (*matching*) entre os vocabulários envolvidos.

O método anterior também foi proposto por Schallier (2004), com o uso da CDU em um *Online Public Access Catalog* (OPAC), para a recuperação dos assuntos dos títulos da Biblioteca do Congresso (LCSH) e da *Medicine Subject Headings* (MeSH). Ademais, o trabalho de Segundo (2009), também utilizou este método que foi chamado de metatesouro da Biblioteca Nacional de Medicina ou

Unified Medical Language System (UMLS), nos Estados Unidos, com a criação de vocabulários de assunto, termos de sistemas de classificação e do tesouro para torná-los interoperáveis.

No ano de 2003, foi apresentada a tese de Frâncu (2003a) e publicado o artigo Frâncu (2003b), onde os estudos objetivaram explorar os impactos da especificidade no potencial de recuperação de um tesouro multilíngue baseado na CDU com indicadores de desempenho, tais como: revocação, precisão e relevância. A metodologia adotada visou atender os requisitos de uma ferramenta de RI altamente eficiente e para satisfazer as necessidades do usuário e as expectativas do indexador, respeitando as regras de classificação. O tesouro adaptado com a CDU permitiu que o usuário tivesse a oportunidade de ampliar e restringir a busca de acordo com suas necessidades como dispositivo de desambiguação. O mesmo padrão de pesquisa também foi descrito no trabalho de Colillas e Nicolau (2008), Frâncu (2007), Frâncu e Sabo (2010), Kaosar (2008), Mcilwaine (2010), Pika (2007) e Slavic, Cordeiro e Riesthuis (2007).

Slavic (2004) mostrou que algumas bibliotecas estão construindo, gradualmente, sistemas com tesouros ou títulos de assuntos baseados nos dados existentes na CDU. Os métodos utilizados foram aplicados em usuários fora do domínio da biblioteca que esperam gerenciar os sistemas de classificação com dados disponíveis em arquivos de autoridade de assunto da CDU. Os resultados mostraram que as ferramentas de organização do conhecimento devem ser tratadas como dados independentes e separadas do próprio registro de metadados, podendo ser utilizados em ambiente multilíngue, visto que, geralmente, a classificação pode perder os relacionamentos de sintaxe se não forem codificados. Por isso, a CDU permite a codificação dos termos de forma pré-coordenada, ou pode ser utilizada como *Uniform Resource Identifiers* (URI), na web, expressando novos conceitos para garantir a integridade e consistência na recuperação (SLAVIC, 2004). A mesma metodologia pode ser observada em Balíková (2009).

A tese de Slavic (2005) e o artigo Slavic (2008) descreveram as preocupações existentes no *design*, construção e uso da classificação, de acordo com as características estruturais e funcionais para que a classificação facetada seja utilizada de modo eficiente na recuperação. Nos procedimentos metodológicos, definiram-se alguns requisitos funcionais comuns às classificações analítico-sintéticas, como por exemplo, a utilização de um vocabulário controlado. Os resultados demonstraram que nenhuns desses padrões atenderam aos requisitos típicos de linguagens de indexação pré-coordenada e não possui solução para expressar relacionamentos sintagmáticos entre conceitos em uma classificação analítico-sintética, o que exige um grande esforço intelectual dos classificionistas. O formato do banco de dados da CDU proposto foi uma sugestão para um modelo de dados mais robusto para um sistema baseado na CDU em ambiente web (SLAVIC, 2005, 2008). Esse mesmo método foi utilizado por Broughton (2010), Broughton e Slavic (2007), Chudamani (2004), Gnoli (2007), La Barre (2007), Slavic (2006a), Slavic (2006b) e Slavic (2006c).

Slavic (2006b) utilizou a CDU em uma interface de portais de assunto, verificando o impacto do uso da classificação na web. O método utilizado foi o mesmo do sistema *German Harvest Automated Retrieval and Directory* (GERHARD), por meio de um banco de dados com classificação automática e um serviço integrado de busca e navegação em um diretório de assunto. Os metadados gerados e o índice de documentos foram processados em um banco de dados relacional (ORACLE), e cada assunto da classe da CDU foi hierarquicamente exibido com uma legenda de classe curta, levando em conta a precisão na RI. Os resultados verificaram que as bibliotecas reconhecem e usam as classificações principalmente com o propósito de organização das coleções em ambiente web, possibilitando o controle do vocabulário e as relações semânticas entre os assuntos relevantes para a descoberta de recursos, utilizando ou não sua notação (SLAVIC, 2006b). Os mesmos aspectos foram considerados no estudo de Slavic (2006c).

Colillas e Nicolau (2008) apresentaram um sistema que permite a organização dos conceitos baseada nos relacionamentos das disciplinas de forma sistematizada por assunto e não por ordem alfabética, com maior flexibilidade para a combinação de assuntos. Os métodos escolhidos pelos autores foram baseados em um conjunto de descritores equivalentes, dispensando o índice temático da CDU, com uso de links para que as classes da CDU fossem equivalentes com as notações e descritores, representando o conteúdo dos documentos. Nos resultados utilizou-se o tesouro com diferentes termos de indexação que se relacionam sem gerar incompatibilidade de vocabulário com os termos da CDU (COLILLAS; NICOLAU, 2008). Hajdu-Barát (2009) utilizou a mesma técnica, onde criou-se um

sistema complexo para combinar termos da CDU e do tesouro em um projeto de construção de tesouros na Biblioteca Nacional Húngara (Széchényi).

No mesmo ano, Slavic, Cordeiro e Riesthuis (2008) descreveram uma visão geral da história e revisão da CDU, destacando a política de manutenção e desenvolvimento da CDU iniciado em 1993, de modo a atualizar o sistema. A metodologia possibilitou melhorar o controle sobre as mudanças, sua consistência em todo o sistema e a garantia de que os usuários pudessem utilizá-lo, por meio de rastreamento e automação dessas mudanças no ambiente de uso. Os resultados mostraram que as tabelas da CDU foram criadas com os conceitos simples, os códigos do sistema com números compostos e índices na forma de um tesouro. Assim sendo, foram despendidos esforços para fornecer o sistema com um índice alfabético e mapeamentos para outros sistemas de assuntos, a fim de melhorar a indexação e a recuperação (SLAVIC; CORDEIRO; RIESTHUIS, 2008).

Colillas (2009) propôs uma lógica nos metadados, baseada na CDU, como ferramenta contextualizada nos cabeçalhos de assuntos, concebidos na atribuição de descritores com um padrão lógico para representar os assuntos dos documentos. A proposta dessa metodologia foi permitir sucessivas seleções em três níveis de hierarquia com as classes e subclasses da CDU por meio de um menu, contendo as notações e a expressão alfabética, agrupando os assuntos por meio de operadores booleanos. A busca pós-coordenada está mais associada a RI, cujas linguagens controladas são mais associadas às linguagens pré-coordenadas, possibilitando maior flexibilidade na sintaxe em relação à busca pós-coordenada. Os resultados mostraram que para uma recuperação precisa foi necessário a existência de um tesouro multilíngue ou uma adaptação específica dos cabeçalhos de assunto da Biblioteca do Congresso. (COLILLAS, 2009).

Na pesquisa de Doorn e Polman (2010) também foi estabelecida a concordância entre os códigos da CDU com um tesouro, atribuindo os códigos desse sistema com os títulos do catálogo em descritores do tesouro. Os procedimentos metodológicos demonstraram como um tesouro de estudos africanos foi construído visando à indexação e recuperação de assunto na biblioteca, em conjunto com um sistema baseado em palavras, sendo uma alternativa mais amigável para os códigos da CDU. Os resultados demonstraram que, para cada descritor, foi atribuído um código da CDU, utilizando a edição mais recente do *Master Reference File* (MRF) para conectar a um tesouro em diferentes versões linguísticas do MRF da CDU, no futuro (DOORN; POLMAN, 2010).

Segundo Colillas (2011) existe uma ordem lógica para os termos de indexação, representando o conteúdo principal do documento, com uso dos descritores da CDU de forma universal, tendo em vista garantir maior precisão na recuperação da informação. A metodologia utilizou um sistema de indexação e recuperação baseado na CDU no projeto *Faceted Application of Subject Terminology* (FAST). Essa ferramenta possibilitou a criação de novas associações, baseadas na indexação e na classificação, que aparecem nos documentos pela adaptação do citado projeto em conjunto com a CDU. Os resultados geraram vários documentos reais na base de dados do catálogo da Biblioteca do Congresso, indexados de acordo com a versão on-line da CDU no sistema proposto, para serem recuperados de forma multidisciplinar na base de dados (COLILLAS, 2011).

Em 2014, Vukadin e Slavic (2014) mostraram a adequação da CDU para a indexação e recuperação de assuntos sobre a área da Arquitetura e sua estrutura hierárquica menos rígida, cujas relações são bem representadas e apoiadas por princípios analítico-sintéticos que permitem a descrição do assunto de modo detalhado e com menos dependência. O método utilizado para fins de melhorar a recuperação foi os princípios da análise facetada na CDU, com o tesouro *Art & Architecture Thesaurus On-line* (AAT), em 2013, com linguagem de indexação alfabética, preocupada com aspectos linguísticos para o controle de vocabulário da CDU. Os resultados apontaram que os conceitos e a representação notacional são mutualmente exclusivos, permitindo uma organização e recuperação mais precisa na CDU, com o recurso analítico-sintético dos conceitos listados neste domínio (VUKADIN; SLAVIC, 2014).

O ano mais recente abordado nesta revisão foi 2017, com o estudo de Lee (2017) que também utilizou a análise facetada para realizar uma análise aprofundada na área musical. A metodologia utilizou um número significativo de tipos de informações relacionadas à classificação na área da Música. Foram considerados a flexibilidade e o caráter universal da CDU, incluindo, apenas, um pequeno número de subfacetadas e seus relacionamentos, de acordo com a ordem de citação e modelos para a construção de um tesouro, haja vista facilitar a recuperação. Os resultados apresentaram os aspectos principais da

Música, por meio da análise das facetas e de suas relações, para analisar assuntos complexos (LEE, 2017).

Slavic e Davies (2017) utilizaram a mesma técnica anterior, por meio de requisitos de sistemas de informação para a modelagem de dados, com o objetivo de criar um projeto de bancos de dados e a interface de sistemas de classificação bibliográfica em ambientes on-line. A utilização da CDU possibilitou a garantia de um progresso apropriado, com as classes (entidades) exibidas de modo correto. As autoras evidenciaram que, dependendo dos requisitos, devem ser utilizadas ferramentas adicionais, como por exemplo, os tesouros para gerenciar as concordâncias entre as estruturas antigas e as novas (SLAVIC; DAVIES, 2017).

5 CONSIDERAÇÕES FINAIS

Esta revisão da literatura possibilitou conhecer o desenvolvimento dos estudos sobre o potencial da CDU na recuperação da informação na web. As publicações selecionadas na amostra permitiram conhecer os pesquisadores de diversos países, não sendo identificado nenhum estudo publicado no Brasil, cujo idioma predominante foi o inglês.

Dentre os 44 estudos analisados, percebeu-se que a autora mais produtiva é Aida Slavic. Os anos que obtiveram o maior número de publicações foi o ano de 2007 e 2008, com predominância em artigos de periódicos, no que tange aos tipos de documentos ou meios de publicação. A diminuição do número de publicações de 2008 a 2017 deve ser mais investigada, como também análises aos trabalhos publicados em outros idiomas, haja vista verificar novas possibilidades de uso da CDU relacionada a tesouro para recuperação na web.

No que tange ao conteúdo das 44 publicações, pelos resultados, foi possível constatar a potencialidade do uso da CDU que, geralmente, é utilizada para organização dos documentos em uma biblioteca ou centros de documentação. Os estudos reforçaram que, se a CDU for vinculada a um tesouro, essa é uma proposta viável, pois promoverá melhorias na recuperação da informação. Por outro lado, a recuperação pode ser mais eficiente quando combinada ao alto grau de especificidade e flexibilidade da CDU, cujos termos, notações e relações semânticas potencializam a recuperação, sobretudo em tesouros de domínios específicos.

Verificou-se, também, a contribuição do tesouro ao fornecer termos padronizados que contemplam os assuntos de um domínio, em conformidade com os termos utilizados pelos usuários nas atividades de busca. Nessas atividades, a recuperação será mais efetiva se a informação registrada nas bases de dados levarem em consideração as necessidades de informação dos usuários.

Para estudos futuros, é sugerida uma revisão sistemática, contemplando dados quantitativos e estudos em outras fontes de pesquisa, como livros e anais de evento. Além disso, sugere-se um estudo comparativo entre as diferentes metodologias utilizadas pelos autores desta temática e a aplicação de estudos de caso, de modo a reconhecer o uso da CDU, no cotidiano de profissionais que lidam com a informação no ambiente web.

REFERÊNCIAS

BAEZA-YATES, R.; RIBIERO-NETO, B. **Modern information retrieval**. New York: ACM press, 1999.

BAKER, Thomas *et al.* Key choices in the design of Simple Knowledge Organization System (SKOS). **Journal of Web Semantics**, v. 20, p. 35-49, 2013.

BALÍKOVÁ, M. The role of UDC classification in the Czech Subject Authority File Item type Meetings and Proceedings at a Crossroads: Multiple Directions to Usability The role of UDC classification in the Czech Subject Authority File. **Extensions & Corrections to the UDC**, v. 1, 2009.

BRASCHER, M. Tesouro, taxonomia e ontologia: uma evolução. Ciclo de conferências sobre Organização da Informação Legislativa e Jurídica. Brasília, 2009.

BROUGHTON, V. Concepts and terms in the faceted classification: the case of UDC. **Knowledge organization**, v. 37, n. 4, p. 270-279, 2010.

BROUGHTON, V.; SLAVIC, A. Building a faceted classification for the humanities: principles and procedures. **Journal of Documentation**, London, v. 63, n. 5, p. 727-754, 2007.

BUXTON, A. B. Computer searching of UDC numbers. **Journal of Documentation**, London, v. 46, n. 3, p. 193-217, 1990.

CAMPOS, Maria Luiza de Almeida. **Linguagem documentária: teorias que fundamentam sua elaboração**. Rio de Janeiro: EdUFF, 2001.

CHAN, L. M.; ZENG, M. L. Ensuring interoperability among subject vocabularies and knowledge organization schemes: a methodological analysis. **IFLA journal**, v. 28, n. 5-6, p. 323-327, 2002.

CHUDAMANI, K. S. Classification Model for Libraries in the Digital Environment. *In*: INTERNATIONAL CALIBER, 2., 2004, Índia. **Anais [...]**. India: New Delhi, 2004.

COLILLAS, M. G. Tratamiento de los Metadatos de Contenido en la Web Semántica o cuando la Lógica se hace Evidente. *In*: NUEVAS PERSPECTIVAS PARA LA DIFUSIÓN Y ORGANIZACIÓN DEL CONOCIMIENTO, 1., 2009, Barcelona. **Anais [...]**. Barcelona: ISKO, 2009.

COLILLAS, M. G. UDC on the internet: Theory and project in evolution for use of indexing and retrieval systems. **IFLA Journal**, v. 37, n. 4, p. 305-313, 2011.

COLILLAS, M. G.; NICOLAU, A. Improving subject searching in databases through a combination of descriptors and UDC. **Zadar**: Bobcatsss, Barcelona, v. 1, n. 8, p. 1-29, 2008.

CURRÁS, E. **Tesauros: linguagens terminológicas**. Brasília, DF: IBICT, 1995.

DAHLBERG, I. Possibilities for a new universal decimal Classification. **Journal of Documentation**, London, v. 27, n. 1, p. 18-36, 1971.

DAHLBERG, I. The Information Coding Classification (ICC): A modern, theory-based fully-faceted, universal system of knowledge fields. **Axiomathes**, v. 18, n. 2, p. 161-176, 2008.

DOORN, M. V.; POLMAN, K. At a Crossroads: Multiple Directions to Usability From classification to thesaurus ... and back? Subject indexing tools at the library of the Afrika-Studiecentrum Leiden [extended abstract]. **Knowledge Organization**, v. 37, n. 3, p. 203-208, 2010.

FIAŁKOWSKI, K. R.; MURASZKIEWICZ, M. R.; TOŁWIŃSKA, A. B. Methodology of world clearinghouse activity. **Information Processing & Management**, v. 17, n. 1, p. 21-25, 1981.

FOSKETT, D. J. **A Study of the Role of Categories in a Thesaurus for Educational Documentation**. Strasburg: Council of Europe, 1973.

FRÂNCU, V. Does convenience trump accuracy? The avatars of the UDC in Romania. **Extensions and Corrections to the UDC**, v. 29, n. 1, p. 263-272, 2007.

FRÂNCU, V. **Multilingual access to information using an intermediate language**. 2003a. 203f. Tese (Doutorado) - Universiteit Antwerpen, Antwerpen, 2003a.

FRÂNCU, V. The impact of specificity on the retrieval power of a UDC-based multilingual thesaurus. **Cataloging & classification**, v. 37, n. 1-2, p. 49-64, 2003b.

FRÂNCU, V.; SABO, C. Implementation of a UDC-Based Multilingual Thesaurus in a Library Catalogue: The Case of BiblioPhil. **Knowledge Organization**, v. 37, n. 3, p. 209-215, 2010.

FREEMAN, R. R. Computers and Classification Systems. **Journal of Documentation**, London, v. 20, n. 3, p. 137-145, 1964.

GNOLI, C. Progress in synthetic classification: towards unique definition of concepts Item type Conference Paper. **Extensions and corrections to the UDC**, Haia, v. 29, n. 1, p. 4-5, 2007.

GOMES, H. E. **Classificação, tesouro e terminologia**: fundamentos comuns. Palestra preparada para as Tertúlias do Departamento de Biblioteconomia da UNIRIO. Rio de Janeiro: UNIRIO, 1996.

GOMES, H. E.; CAMPOS, M. L. A. A organização do conhecimento na *Web*: contribuições de Shiyali Ramamrita Ranganathan e de Ingetraut Dahlberg . *In*: IACS/UFF, 2019, Niterói. **Anais [...]**. Niterói: UFF, 2019.

HAJDU-BARÁT, Á. Integration of a thesaurus and Universal Decimal Classification (UDC) to improve subject access: the Hungarian experience. **Extensions & Corrections to the UDC**, v. 31, p. 197-207, 2009.

HILL, L. *et al.* Integration of knowledge organization systems into digital library architectures. **Data Analysis and Knowledge Discovery**, v. 20, n. 1, p. 4-8, 2004.

HINDSOTI, R. Reflections on the utilization of the Universal Decimal Classification. **Aslib Proceedings**, v. 31, n. 6, p. 305-311, 1979.

HJØRLAND, B. Arguments for ‘The Bibliographical Paradigm’: Some Thoughts Inspired by the New English Edition of the UDC. *In*: INFORMATION RESEARCH, 12., 2007. **Anais [...]**. [S.l.]: [S.n.], 2007.

HJØRLAND, B. Does the traditional thesaurus have a place in modern information retrieval? **Knowledge Organization**, v. 43, n. 3, p. 145-159, 2016.

HODGE, G. **Systems of knowledge organization for digital libraries**: beyond traditional authorities files. Washington: Council on Library and Information Resources, 2000.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) 25964-1. **Information and documentation Thesauri and interoperability with other vocabularies**: Part 1: Thesauri for information retrieval. Geneva: ISO, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) 25964-2. **Thesauri and interoperability with other vocabularies**: Part 2: Interoperability with other vocabularies. Geneva: ISO, 2013.

KAOSAR, A. **Merit & Demerit of using Universal Decimal Classification on the Internet**. 2008. Dissertação (Mestrado em Ciência da Informação) - Royal School of Library and Information Science, Universidade de Copenhagen, Denmark, 2008.

KOCH, T.; TUDHOPE, D. New applications of knowledge organization systems: introduction to a special issue. **Journal of Digital Information**, v. 4, n. 4, 2003.

KOCH, T.; TUDHOPE, D. **User-centred approaches to Networked Knowledge Organization Systems/Services (NKOS)**. Background, 2004.

LA BARRE, K. The heritage of early FC in document reference retrieval systems, 1920-1969. **Library History**, Champaign, v. 23, n. 2, p. 129-149, 2007.

LEE, D. Numbers, instruments and hands: the impact of faceted analytical theory on classifying music ensembles. **Knowledge Organization**, v. 44, n. 6, p. 405-415, 2017.

MCILWAINE, I. C. Classification at a Crossroads: Multiple Directions to Usability. **International Cataloguing and Bibliographic Control: Quarterly Bulletin of the IFLA UBCIM Programme**, v. 39, n. 2, p. 41-42, 2010.

MCILWAINE, I. C. The Universal Decimal Classification: Some Factors Concerning Its Origins, Development, and Influence. **Journal of the American Society for Information Science**, London, v. 48, n. 4, p. 331-339, 1997.

MILLS, J. Progress in Documentation: Library Classification. **Journal of Documentation**, Califórnia, v. 26, n. 2, p. 120-160, 1970.

PIEDADE, M. A. R. **Introdução à teoria da classificação**. Rio de Janeiro: Interciência, 1977.

PIKA, J. Universal Decimal Classification at the ETH-Bibliothek Zürich-a Swiss perspective. **UDC Consortium**, London, v. 29, n. 1, p. 229-251, 2007.

PRASAD, A. R. D.; MADALLI, D. P. Classificatory ontologies. **Extensions & Corrections to the UDC**, v. 31, n. 1, p. 223-232, 2009.

RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SARACEVIC, T. Information Science. **Journal of the American Society for Information Science**, New Jersey, v. 50, n. 12, p. 1051-1063, 1999.

SCHALLIER, W. "What a subject search interface can do". **Extensions & Corrections to the UDC**, Bélgica, v. 26, n. 1, p. 18-28, 2004.

SEGUNDO, R. S. Crossroads: Multiple Directions to Usability Using MARC classification format for UDC and mappings to other KO systems for an enriched authority file. **Extensions & Corrections to the UDC**, v. 31, n. 1, p. 285-291, 2009.

SLAVIC, A. **Classification management and use in a networked environment**: the case of the Universal Decimal Classification. 2005. 463f. Tese (Doutorado em Filosofia) - University of London, London, 2005.

SLAVIC, A. Faceted classification: management and use. **Axiomathes**, v. 18, n. 2, p. 257, 2008.

SLAVIC, A. Interface to classification: some objectives and options Item type Preprint Interface to Classification: Some Objectives and Options. **Extensions and Corrections to the UDC**, Haia, v. 28, 2006a.

SLAVIC, A. The level of exploitation of Universal Decimal Classification in library OPACs: A pilot Study. **Vjesnik bibliotekara Hrvatske**, v. 49, n. 3-4, p. 155-182, 2006b.

SLAVIC, A. UDC implementation: from library shelves to a structured indexing language. **International cataloguing and bibliographic control**, London, v. 1, 2004.

SLAVIC, A. UDC in subject gateways: experiment or opportunity? **Knowledge Organization**, v. 2, n. 33, p. 67- 85, 2006c.

SLAVIC, A.; CORDEIRO, M. I.; RIESTHUIS, G. Enhancement of UDC data for use and sharing in a networked environment Paper presented at the Librarian Workshop in conjunction with. *In: ANNUAL CONFERENCE OF THE GERMAN CLASSIFICATION SOCIETY ON DATA ANALYSIS*, 31., 2007, Alemanha. **Anais [...].** Alemanha: Freiburg i. Br, 2007.

SLAVIC, A.; CORDEIRO, M. I.; RIESTHUIS, G. Maintenance of the Universal Decimal Classification: overview of the past and preparations for the future. **International Cataloguing and Bibliographic Control Journal International Cataloguing and Bibliographic Control**, London, v. 37, n. 2, p. 23-29, 2008.

SLAVIC, A.; DAVIES, S. Facet Analysis in UDC: Questions of Structure, Functionality and Data Formality. **Knowledge Organization**, v. 44, n. 6, p. 425-435, 2017.

SOUZA, R. R. Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. **Perspectivas em ciência da informação**, Belo Horizonte, v. 11, n. 2, p. 161-173, 2006.

SOUZA, R. R.; ALVARENGA, L. A web semântica e suas contribuições para ciência da informação. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 132-141, jan./abr. 2004.

SOUZA, S. **CDU: como entender e utilizar a segunda edição padrão internacional em Língua Portuguesa**. São Paulo: Zamboni livros, 2012.

VUKADIN, A.; SLAVIC, A. Challenges of Facet analysis and Concept Placement in Universal Classifications: the Example of Architecture in UDC. **Knowledge Organization**, v. 1, p. 236-243, 2014.

ZENG, M. L. Knowledge Organization Systems (KOS). **Knowledge Organization**, v. 35, n. 2-3, p. 160-182, 2008.

WALL, R. A. Indexing language structure for automated retrieval. **Information Storage and Retrieval**, v. 9, n. 11, p. 607-617, 1973.

WALL, R. A. Intelligent indexing and retrieval: A man-machine partnership. **Information Processing and Management**, Loughborough, v. 16, n. 1, p. 33-90, 1980.