

XX ENANCIB

21 a 25 Outubro/2019 – Florianópolis

A Ciência da Informação e a era da Ciência de Dados

ISSN 2177-3688

GT-2 – Organização e Representação do Conhecimento

**AVALIAÇÃO DA INDEXAÇÃO A PARTIR DOS ELEMENTOS DE EXAUSTIVIDADE E
ESPECIFICIDADE PELA REVOCAÇÃO E PRECISÃO NA RECUPERAÇÃO DA INFORMAÇÃO**

***EVALUATION OF INDEXING THROUGH THE ELEMENTS OF COMPLETENESS AND SPECIFICITY
BY RECALL AND PRECISION IN INFORMATION RETRIEVAL***

Valdenise César Garcia - Universidade Federal do Pará
Franciele Marques Redigolo - Instituição Univerisade Federal do Pará

Modalidade: Trabalho Completo

Resumo: Estuda a avaliação da indexação, visto que esse procedimento é imprescindível para proporcionar mais qualidade ao seu processo e permitir uma recuperação da informação eficiente para os usuários. Tem-se como objetivo geral realizar a avaliação da qualidade da indexação efetuada nos documentos das áreas de Ciências Biológicas, Humanas e Engenharias das bibliotecas da Universidade Federal do Pará nos quesitos de exaustividade e especificidade, mediante a recuperação da informação pelos elementos de revocação e precisão. Utiliza como metodologia a pesquisa exploratória, abordando o problema sob um enfoque qualitativo, em que aplicou entrevista semiestruturada a 55 usuários das bibliotecas distribuídos nos cursos de Biologia, Psicologia e Engenharia Civil. Os resultados mostraram que a indexação realizada nos documentos analisados dos três cursos foi considerada satisfatória, principalmente para Biologia, em que se verificou maior qualidade. Contudo, para os documentos dos cursos de Psicologia e Engenharia civil, torna-se necessário elaborar diretrizes especialmente para o quesito de especificidade. Conclui-se que avaliar a indexação por meio da recuperação da informação, incluindo os usuários, torna-se essencial quando se quer buscar melhorias para o processo, pois estes podem aperfeiçoar a representação da informação pelo seu *feedback*, mostrando o que é desejado para cada área específica dos documentos tratados.

Palavras-Chave: Indexação; Avaliação da Indexação; Recuperação da Informação.

Abstract: It studies the evaluation of indexing, as this procedure is essential to provide more quality to its execution and to provide efficient retrieval of information for users. The general objective is to evaluate the quality of the indexing performed in the documents of the fields of Biological Sciences, Humanities and Engineering of the libraries of the Federal University of Pará in terms of completeness and specificity, by retrieving the information by the recall and precision elements. It uses as exploratory research methodology, approaching the problem from a qualitative approach. It was applied a semi-structured interview to 55 users of libraries distributed in the Biology, Psychology and Civil Engineering courses. The results showed that the indexing performed in the analyzed documents of the three courses was considered satisfactory, especially for Biology, where it was found higher quality. However, for the Psychology and Civil Engineering courses documents, it is necessary to develop guidelines especially for the specificity question. It is concluded that evaluating indexing through information retrieval, including users, becomes essential when seeking improvements to the process, as it can improve the information representation through their feedback, showing what is desired for each specific field of the documents.

Keywords: Indexing; Evaluation of Indexing; Information Retrieval.

1 INTRODUÇÃO

A informação está presente em todos os ambientes e contextos. Ao longo dos tempos foi e é objeto de pesquisa de muitos estudiosos de diversas áreas do conhecimento com o intuito de investigar sua origem, conceito, função, formas de coleta, organização, armazenamento, disseminação, transmissão, interpretação e uso.

No contexto da organização e representação da informação isso não é diferente. Os serviços e processos realizados para organizar adequadamente a informação e disseminá-la para os usuários têm passado também por constantes estudos, mudanças e aperfeiçoamentos, principalmente no cenário das Tecnologias da Informação e Comunicação (TICs) e das grandes exigências quanto ao acesso imediato à informação.

Com tal importância, o desenvolvimento das atividades de organização da informação não pode passar despercebido aos olhos dos gestores das unidades de informações, mas esforços devem ser direcionados para que os processos sejam regularmente avaliados e aprimorados em prol de um desempenho mais eficaz.

Destaca-se, neste contexto, o processo de indexação como uma das atividades de organização do conhecimento e da informação, que trabalha exclusivamente com o tratamento do conteúdo dos documentos, e como tal carece também de melhorias e avaliações para que as unidades de informações possam disponibilizar com eficácia a informação aos seus usuários quando estes realizam nos catálogos a pesquisa por assunto.

Os elementos mais indicados para avaliação da indexação e de sua qualidade por meio da recuperação da informação são aqueles responsáveis diretamente pelos resultados na recuperação, como apontam Fujita (2012) e Fujita e Gil Leiva (2014): exaustividade, especificidade, linguagem, revocação e precisão. A exaustividade está ligada à etapa de análise de assunto (identificação e seleção dos conceitos), a especificidade à fase de tradução do conceito para a linguagem, e as medidas de revocação e precisão retratam, nos resultados da recuperação da informação, os efeitos da utilização desses dois elementos, exaustividade e especificidade, na indexação.

Diante disso, verifica-se o papel central dos usuários nesse processo de avaliação da indexação pela recuperação da informação, pois os níveis adequados dos elementos de exaustividade e especificidade para o processo de indexação são estabelecidos conforme as suas necessidades, o que só é possível verificar a partir da análise de suas preferências no

momento em que estes utilizam a base de dados para recuperação da informação.

Desse panorama é que se desencadeia o problema desta pesquisa, cujo ponto de partida é uma observação, captada pela autora, da não utilização de depoimentos dos usuários como forma de avaliar o processo de indexação e aperfeiçoar a recuperação da informação, no caso em especial das bibliotecas da Universidade Federal do Pará (UFPA).

No contexto específico dessas bibliotecas, a indexação realizada, apesar de ser uma grande ferramenta de suporte para a recuperação de documentos por assunto, consoante os interesses e necessidades informacionais dos usuários, apresenta seus entraves, principalmente pela inexistência de uma política de indexação.

Por isso, esforços devem ser direcionados a estudos relativos à indexação, principalmente a respeito da sua avaliação, de modo que o processo possa ser analisado para verificar eventuais falhas e lapidá-lo, sempre em busca da sua qualidade.

Assim, o objetivo desta pesquisa é realizar a avaliação da qualidade da indexação efetuada nos documentos das áreas de Ciências Biológicas, Humanas e Engenharias das bibliotecas da Universidade Federal do Pará nos quesitos de exaustividade e especificidade, mediante a recuperação da informação pelos elementos de revocação e precisão.

Desta maneira, este estudo vem trazer uma contribuição teórica e prática para a área de Tratamento Temático da Informação e para os estudos de organização do conhecimento e da informação a partir da investigação sobre a avaliação da qualidade da indexação, mediante os elementos de exaustividade, especificidade, revocação e precisão, com ênfase nos usuários.

Esta pesquisa formaliza-se como exploratória, pois sua finalidade é proporcionar mais informações sobre o assunto pesquisado, explicando os fenômenos observados, com base um estudo aplicado, abordando o problema sob um enfoque qualitativo.

2 O PROCESSO DE INDEXAÇÃO E A SUA AVALIAÇÃO

A indexação é compreendida como parte do ciclo documentário, pois se insere como um dos procedimentos do tratamento dos documentos para disponibilizá-los aos usuários, de modo que é uma atividade essencial no contexto das unidades de informação, principalmente em bibliotecas.

Como atividade de tratamento temático da informação da corrente teórica inglesa, a

indexação volta-se para a análise dos conceitos (unidades dos pensamentos) expressos pelos autores nos documentos, produzindo representações temáticas desses objetos, tendo com isso a finalidade de organizar os objetos informacionais através de seus conteúdos intrínsecos em sistemas de informações manuais ou automáticos, para então facilitar a busca e a recuperação da informação, tornando-a acessível para os usuários que dela necessitam (ARAÚJO JÚNIOR, 2007; RUBI, 2009).

Segundo os “Princípios de indexação”, a indexação é conceituada como “[...] a ação de descrever e identificar um documento de acordo com seu assunto. [...] Durante a indexação, os conceitos são extraídos do documento através de um processo de análise, e então traduzidos para os termos de instrumentos de indexação” (UNISIST, 1981, p. 84).

De modo mais completo, Rubi (2009, p. 81) aponta que a indexação se refere à “identificação do conteúdo do documento por meio do processo de análise de assunto e à representação desse conteúdo por meio de conceitos. Esses conceitos, por sua vez, serão representados ou traduzidos em termos advindos de uma linguagem documentária”. A referida autora acrescenta ainda que a finalidade desse processo é permitir a “intermediação entre o documento e o usuário no momento da recuperação da informação, seja em índices, catálogos ou bases de dados” (RUBI, 2009, p. 81).

Em vista disso, Gil Leiva (2012) aponta que a indexação possui dois objetivos; o primeiro é permitir o armazenamento quando a indexação é direcionada ao documento, enquanto o segundo é possibilitar a recuperação dos documentos quando a indexação é focada nas questões realizadas pelos usuários no momento da busca pela informação, visando a suas necessidades informacionais. Assim, esse autor afirma que o objetivo geral da indexação “é o armazenamento da informação para atender às necessidades de informação” (GIL LEIVA, 2012, p. 71).

Todavia, a indexação não é uma tarefa simples, mas complexa (FUJITA; RUBI, 2006), cujo processo envolve muitas variáveis (CARNEIRO, 1985), englobando desde o profissional indexador até o usuário que recebe as informações na etapa final do processo, de modo que é vista por Dias e Naves (2013, p. 20) como uma operação “delicada, por vezes, complicada”.

Diante disso, o ato de indexar é sistematizado em etapas para auxiliar os indexadores na execução da atividade. Na literatura, verifica-se uma divergência quanto ao número de etapas da indexação, no entanto, Rubi (2009) ressalta que os procedimentos da indexação envolvem basicamente as seguintes operações: análise, referente à leitura para identificação

e seleção de conceitos; síntese, relacionada à elaboração de resumos; e representação, momento da tradução para as linguagens documentárias.

A etapa inicial do processo de indexação é denominada de análise de assunto (RUBI, 2009). Para Fujita (2003) essa é a fase mais importante do trabalho do indexador, visto que é por meio dela que o assunto do documento é determinado através de três etapas: leitura documentária, identificação e seleção de conceitos válidos para recuperação.

Por intermédio da leitura documentária, o indexador tem o primeiro contato com o item bibliográfico a ser tratado. Conforme Silva e Fujita (2004, p. 148), ela possibilita, no primeiro momento, a “identificação de conceitos que caracterizam o assunto tratado no documento, e no segundo momento, a seleção dos conceitos”.

A identificação de conceitos é considerada por Silva e Fujita (2004, p. 147) como ponto principal da análise de assunto, uma vez que é nesse momento que o indexador “compreende os conceitos tratados em um documento, bem como verifica sua importância para seu sistema de informação”, de maneira que, a partir da identificação dos conceitos essenciais, o indexador seleciona aqueles que melhor representam sua temática.

O momento da seleção dos conceitos implica para o indexador uma questão importante: decidir quais conceitos serão escolhidos ou rejeitados. Essa decisão não deve ser tomada levando em consideração as suas preferências pessoais, mas está condicionada aos objetivos e finalidades do sistema de informação e às necessidades de informação da comunidade desse sistema (SILVA; FUJITA, 2004).

É nessa fase que os elementos de exaustividade e especificidade entram em ação, cujos princípios determinam o grau com que os documentos serão indexados, pois a exaustividade diz respeito à abrangência na indexação, ao qual corresponde, segundo Lancaster (2004, p. 27), “ao número de termos atribuídos em média” para representar o documento, podendo ser exaustiva, quando há o “emprego de termos em número suficiente para abranger o conteúdo temático do documento de modo bastante completo”, ou seletiva, quando é empregado um número menor de termos com a finalidade de abranger somente o conteúdo principal do documento (LANCASTER, 2004, p. 27). E a especificidade corresponde à profundidade na indexação, isto é, a “exatidão com que os descritores utilizados representam o conteúdo temático do documento” (PIECADE, 1983, p. 12).

Desse jeito, o indexador pode selecionar somente os conceitos principais (poucos termos) ou principais e secundários (mais termos) que englobam o conteúdo do documento

(exaustividade seletiva ou exaustiva) ou de maneira mais precisa escolher conceitos específicos ao seu tema (especificidade).

Nesse sentido, Fujita (2012) declara que o ideal na indexação é que um documento seja indexado com uma quantidade maior de termos que contemple a cobertura temática do documento, mesmo que para isso sejam utilizados vários termos específicos, em detrimento de apenas um único termo genérico, da mesma forma Lancaster (2004, p. 34) aponta que “seria melhor utilizar vários termos específicos, ao invés de um termo que seja mais genérico”, assim, o documento pode ser representado com uma quantidade de assuntos o suficiente para contemplar todos os seus aspectos temáticos específicos e proporcionar uma recuperação da informação mais precisa.

A etapa da tradução, procedimento final da indexação, implica que as palavras empregadas pelo autor do documento sejam convertidas em termos previamente definidos pertencentes a uma linguagem de indexação, isto é, os conceitos selecionados pelo indexador, em forma de frase de indexação, serão traduzidos em termos controlados.

Todas essas fases e procedimentos podem sofrer interferências e comprometer a qualidade da indexação, e conseqüentemente a recuperação da informação no momento da busca por assunto nas bases de dados, pois a qualidade da indexação, segundo Lancaster (2004, p. 92) é determinada pela “eficácia da recuperação”, isto é, quando são recuperados documentos desejados, e evitados os não desejados.

À vista disso, para que os processos que envolvem a atividade de indexação sejam executados com eficácia, é necessário que as bibliotecas disponham de manuais e políticas para orientar os profissionais quanto aos procedimentos a serem realizados, de maneira que haja uma padronização e por conseqüência uma qualidade nos processos executados.

A política de indexação, nessa conjuntura, aparece para sanar as necessidades de orientações para o ato de indexar, pois é composta por um conjunto de diretrizes formuladas para guiar o processo de indexação, visando à padronização dos atos e diminuindo a subjetividade de cada profissional no momento da tarefa.

Quando não há essa política nas bibliotecas, o processo acaba sendo prejudicado, porque os elementos da política direcionam o profissional em cada tarefa, como no caso dos elementos de exaustividade e especificidade que indicam respectivamente quantos conceitos serão identificados e selecionados no estágio da análise de assunto e quais serão

realmente traduzidos para a linguagem de indexação, conforme a realidade de cada biblioteca e as necessidades de seus usuários (RUBI; FUJITA, 2010).

Assim, a inexistência da política faz com que o processo seja realizado sem parâmetros, o que inevitavelmente minimiza a qualidade da atividade de indexação e por consequência a recuperação da informação na busca por assunto, gerando frustrações para os usuários quanto ao acesso preciso e imediato à informação.

Desta forma, para um eficaz processo de indexação, é necessária uma política de indexação nas bibliotecas, bem como estudos de avaliação para verificar a qualidade da atividade, de modo que o processo possa ser revisado, corrigido e aperfeiçoado, com intuito de oferecer uma recuperação da informação eficiente aos usuários.

No contexto da avaliação da indexação, a atividade pode ser avaliada de diferentes formas. Uma delas é a avaliação extrínseca, na qual os resultados da indexação são avaliados por meio da recuperação da informação (GIL LEIVA, 2008; FUJITA; GIL LEIVA, 2014).

Na avaliação extrínseca, a indexação pode ser avaliada por meio da interconsistência, quando se compara o resultado da indexação de uma base de dados com o de outra base utilizando como ponto de referência o mesmo documento que foi indexado nas duas bases, ou mediante a recuperação, quando se analisa a indexação pelos elementos de revocação e de precisão na recuperação da informação para verificar seus índices (GIL LEIVA, 2008).

Desse modo, a avaliação extrínseca por meio da recuperação nos permite inserir o usuário no processo de avaliação da indexação, já que se utilizam os elementos de revocação e de precisão, aspectos estes diretamente associados ao perfil dos usuários da biblioteca (GUIMARÃES, 2004). Esse tipo de avaliação permite também verificar as opiniões dos usuários quanto aos quesitos necessários para que a atividade de indexação seja realizada de forma a atender as suas exigências e necessidades de informação.

Nesta pesquisa, deu-se destaque para à avaliação da indexação mediante a recuperação da informação pelos elementos de revocação e precisão, pois a intenção é inserir os usuários no processo de avaliação da qualidade da indexação realizada no contexto específico de uma rede de bibliotecas de uma única instituição.

Segundo Lancaster (2004, p. 93), a qualidade da indexação é definida a partir da “eficácia de recuperação”, isto é, quando o sistema de recuperação da informação recupera documentos úteis conforme a necessidade de informação do usuário e evita documentos inúteis, ou seja, documentos não desejados, de modo que esta qualidade é definida pelo

usuário no momento em que este realiza a busca no catálogo e julga os documentos recuperados como relevantes ou irrelevantes.

Destacam-se, nesse contexto, os elementos de revocação e precisão, pois são as medidas que permitem, respectivamente, verificar a capacidade do sistema de recuperar documentos úteis e evitar documentos inúteis (LANCASTER, 2004), de forma que são consideradas por Lancaster (2004) as medidas mais adequadas para aferir os resultados de uma busca em documentos úteis ou inúteis.

Desta forma, este estudo utilizará como critérios de avaliação da qualidade da indexação os elementos descritos por Carneiro (1985) relacionados ao processo de indexação: exaustividade e especificidade; e referentes à recuperação da informação: revocação e precisão, pois, verificando os níveis de revocação e precisão sob a ótica do usuário, pode-se avaliar como estão os princípios de exaustividade e especificidade empregados na indexação e se estes estão adequados as suas necessidades de informação.

Esses quatro elementos são considerados os principais critérios para aferir a qualidade da indexação (UNISIST, 1981; VAN SLYPE, 1991; LANCASTER, 2004; NUNES, 2004), uma vez que a exaustividade mede a qualidade na escolha dos conceitos, isto é, quando todos os conceitos importantes do documento são bem identificados e selecionados, e a especificidade mede a qualidade com que os descritores selecionados da linguagem de indexação correspondem aos conceitos escolhidos (VAN SLYPE, 1991), e a revocação e precisão são os elementos para medir, através da recuperação da informação, a qualidade da exaustividade e especificidade realizada no processo de indexação (UNISIST, 1981; LOPES, 1985; LANCASTER, 2004; GIL LEIVA, 2012; FUJITA, 2016).

3 PROCEDIMENTOS METODOLÓGICOS

A partir de um estudo aplicado, abordando o problema sob um enfoque qualitativo, esta pesquisa formaliza-se como exploratória, pois sua finalidade é proporcionar mais informações sobre o assunto pesquisado, explicando os fenômenos observados.

Quanto aos procedimentos metodológicos, esta pesquisa se caracteriza como um estudo de casos múltiplos realizado com usuários das bibliotecas da UFPA de três cursos de graduação (Biologia, Psicologia e Engenharia civil), representando três distintas áreas, a saber: Ciências Biológicas, Ciências Humanas e Engenharias.

O universo desta pesquisa é constituído pelas bibliotecas universitárias da UFPA composto pelo conjunto de 36 bibliotecas que colaboram com os objetivos e missão da instituição ao promover e disseminar informações técnico-científicas à comunidade universitária e ao público externo.

Para coleta de dados, foi utilizada a técnica da entrevista semiestruturada com 55 usuários selecionados dos cursos de Engenharia Civil (18 usuários), Biologia (17 usuários) e Psicologia (20 usuários), os quais avaliaram a recuperação da informação no que diz respeito à busca por assunto, no catálogo *online* do sistema de bibliotecas da UFPA.

A entrevista foi acompanhada de um roteiro com 11 perguntas para guiar a pesquisadora na obtenção das informações desejadas. O roteiro de entrevista foi elaborado de acordo com o que propõe Carneiro (1985). Assim, as perguntas foram formuladas com base nos elementos de exaustividade e de especificidade, de acordo com os procedimentos realizados no processo de indexação, e nos elementos de revocação e precisão, conforme os documentos tenham sido analisados e julgados como relevantes ou irrelevantes na recuperação da informação.

Para a seleção dos alunos foram definidos os seguintes critérios: ser discente da graduação e estar cursando o último semestre, pois, acredita-se que estes discentes já estão familiarizados com os assuntos empregados nas suas áreas específicas e por isso podem analisar mais eficientemente os resultados das suas pesquisas.

Finalizada a entrevista, os dados foram organizados em quadros. A partir do conhecimento da quantidade total de documentos recuperados pelas pesquisas, do total de documentos considerados relevantes e do total de documentos existentes na coleção, foram aplicadas as fórmulas para verificar as taxas de revocação e de precisão do sistema, Figura 1 e Figura 2, para cada pesquisa de cada usuário.

Figura 1: Fórmula para medir a taxa de revocação utilizada na pesquisa.

$$\text{Taxa de Revocação} = \frac{\text{N}^\circ \text{ de referências relevantes encontradas na busca teste}}{\text{N}^\circ \text{ de referências conhecidas previamente}} \times 100$$

Fonte: Lancaster (1968 apud LOPES, 1985, p. 249).

Figura 2: Fórmula para medir a taxa de precisão utilizada na pesquisa

$$\text{Taxa de Precisão} = \frac{\text{N}^\circ \text{ de referências relevantes recuperadas}}{\text{N}^\circ \text{ total de referências recuperadas}} \times 100$$

Fonte: Lancaster (1968 apud LOPES, 1985, p. 249), e Gil Leiva (2008, p. 88).

As categorias para análise dos dados foram elaboradas conforme os objetivos desta pesquisa a fim de avaliar a qualidade da indexação, as quais são: exaustividade e especificidade, referentes aos aspectos do processo de indexação, e revocação e precisão, concernentes à recuperação da informação por assunto, todos baseados em Carneiro (1985).

4 RESULTADOS E DISCUSSÃO DA AVALIAÇÃO DA QUALIDADE DA INDEXAÇÃO

Para o **curso de Biologia**, com base na pesquisa realizada no catálogo pelos 17 usuários, foram recuperados e analisados no total 992 documentos, de modo que, conforme os seus interesses de informação, os usuários julgaram os documentos como relevantes ou irrelevantes. O Gráfico 1 apresenta as taxas de revocação e precisão da recuperação da informação de cada pesquisa realizada por cada usuário.



Fonte: Elaborado pelas autoras, 2019.

Conforme o Gráfico 1, foi verificada uma revocação e uma precisão com níveis altos e semelhantes, proporcionando uma resposta menos exaustiva e mais precisa na busca no catálogo, no qual foram recuperados muitos documentos relevantes correspondendo ao assunto pesquisado pelo usuário e com taxas baixas de documentos irrelevantes. Pode-se dizer, com isso, que a indexação realizada nesses documentos, de fato, está adequada, com níveis de exaustividade e de especificidade coerentes para proporcionar uma recuperação da informação eficiente, isto é, recuperação condizente com o que é esperado pelos usuários, como pode ser observado pelos resultados no Quadro 1.

Quadro 1: Resultados das quatro categorias de análise do curso de Biologia

Curso de Biologia – Área de Ciências Biológicas		
Categorias	Questões	Resultados
Revocação	-Itens relevantes na pesquisa -Quantidade de itens recuperados -Resultado da pesquisa	<ul style="list-style-type: none"> • Taxas altas de revocação. • Quantidade de itens recuperados adequada para 65% dos usuários encontrar os documentos relevantes. • Resultados condizentes com a temática, 88% de satisfação.
Exaustividade	-Quantidade de assuntos na indexação -Quantidade de assuntos desejada -Abrangência desejada	<ul style="list-style-type: none"> • Quantidade de assuntos satisfatória para 83%. • Indexação mais completa, desejam uma quantidade de 3 a 5 assuntos (59%) e 5 a 10 (41%). • Abrangência com termos principais e secundários (95%).
Precisão	-Precisão do assunto -Precisão desejada	<ul style="list-style-type: none"> • Precisão entre o assunto pesquisado e o assunto indexado. • Desejam maior precisão (59%) nos resultados.
Especificidade	-Assuntos na indexação -Utilização de assuntos -Preferências por assunto	<ul style="list-style-type: none"> • Alta especificidade, termos específicos na indexação (88%). • 70% utilizam termos específicos (buscas específicas). • 88% preferem assuntos mais específicos (Alta especificidade).

Fonte: Elaborado pelas autoras, 2019.

Diante desses resultados, para o curso de Biologia, foi verificada uma indexação realizada com mais qualidade, em que se visualizou uma recuperação da informação mais eficiente, visto que foram obtidas taxas equivalentes de revocação e de precisão, permitindo tanto a recuperação de documentos relevantes como impedindo os irrelevantes.

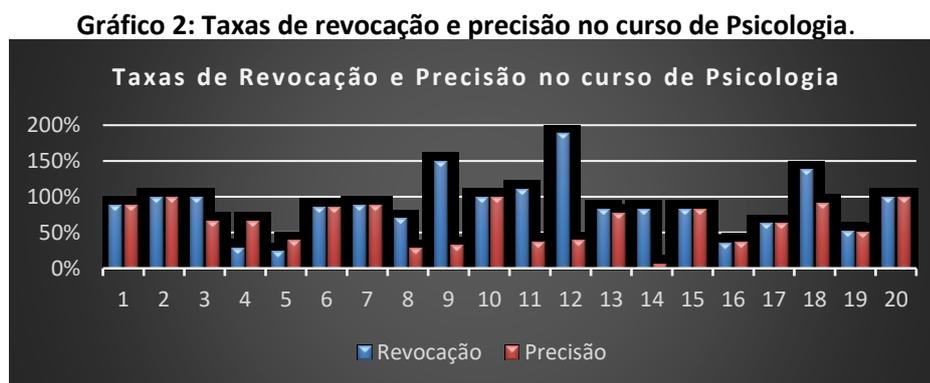
Isso demonstra uma exaustividade e especificidade empregadas de forma equilibrada em grande parte dos documentos, nos quais os conceitos para representá-los foram identificados e selecionados em quantidade e pertinência adequadas a sua temática e também corresponderam com precisão ao termo da linguagem de indexação no momento da tradução.

Deste modo, constatou-se que, apesar da inexistência de uma política de indexação para orientar os indexadores quanto à identificação e seleção de conceitos na exaustividade e tradução dos termos na especificidade para indexar os documentos, os resultados verificados até este momento mostram-se promissores e indicam uma qualidade na atividade de indexação realizada em grande parte dos documentos analisados do curso de Biologia, pois obtiveram-se: revocação e precisão com níveis altos e semelhantes, possibilitando uma resposta menos exaustiva e mais precisa, e níveis de exaustividade e de especificidade equilibrados para proporcionar uma recuperação da informação eficiente.

Assim, a indexação está cumprindo, como Lancaster (2004) e Fujita (2016) afirmam, sua finalidade de recuperação da informação de documentos desejados, recuperando os úteis e evitando o máximo de inúteis, de modo que, para esta análise, infere-se que há uma

qualidade na indexação, com exaustividade e especificidade apropriadas, em que os usuários se sentiram satisfeitos com os documentos recuperados.

Para o **curso de Psicologia**, foi analisado pelos 20 usuários o total de 668 documentos com base nas pesquisas realizadas no catálogo, de maneira que, conforme seus interesses, os usuários julgaram os documentos como relevantes ou irrelevantes. As taxas de revocação e precisão de cada pesquisa podem ser observadas no Gráfico 2.



Fonte: Elaborado pelas autoras, 2019.

Como observado no Gráfico 2, para os documentos recuperados do curso de Psicologia, verifica-se uma variação quanto aos níveis de revocação e precisão, nos quais algumas pesquisas apresentaram baixa precisão e alta revocação, demonstrando que o sistema não está impedindo a recuperação dos documentos irrelevantes.

Entretanto, algumas pesquisas apresentaram taxas de revocação e precisão semelhantes, mostrando uma coerência na identificação de termos pertinentes e em quantidade desejável na exaustividade e específicos na especificidade, ao ponto de trazer níveis satisfatórios de documentos recuperados para os usuários. Esse nível de equilíbrio entre exaustividade e especificidade proporciona resultados eficientes na recuperação da informação, segundo Nunes (2004) e Rubi (2009).

Diante da análise do elemento de revocação, observou-se que a exaustividade empregada está com níveis variados, com predominância para a alta exaustividade, em que se verificou que muitos documentos estão com quantidade de conceitos adequados, pois foram recuperados, em grande parte das pesquisas, documentos relevantes em comparação ao que existe no acervo da biblioteca, tendo taxas altas de revocação, mas também verificou-se a utilização de termos amplos provocando a baixa revocação.

As taxas de precisão, no entanto, não foram tão altas, uma vez que o sistema não

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

evitou a recuperação de documentos não relevantes, isto é, muitos itens recuperados não corresponderam ao assunto pesquisado, devido aos níveis variados de especificidade empregada na indexação, com predominância para o uso de termos não específicos, o que provocou taxas variadas de precisão, como pode ser observado no Quadro 2.

Quadro 2: Resultados das quatro categorias de análise do curso de Psicologia.

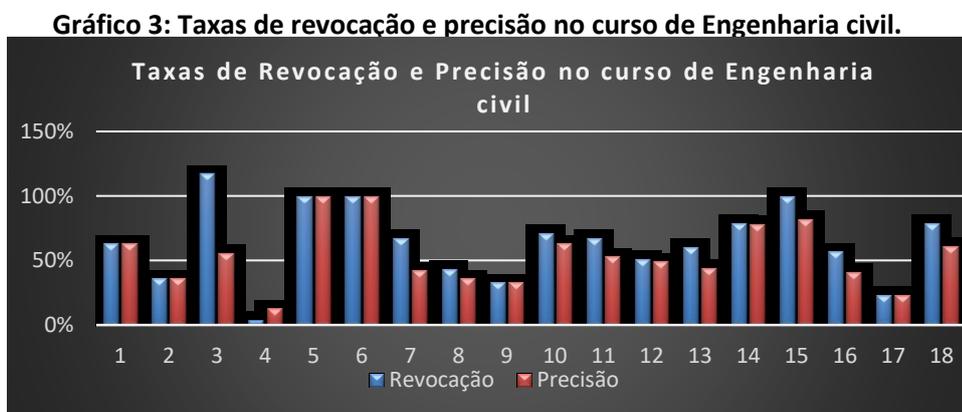
Curso de Psicologia – Área de Ciências Humanas		
Categories	Questões	Resultados
Revocação	-Itens relevantes na pesquisa -Quantidade de itens recuperados -Resultado da pesquisa	<ul style="list-style-type: none">• Taxas variadas de revocação, com predominância para alta.• Quantidade satisfatória (65%).• 75% de satisfação com a pesquisa, resultados relevantes.
Exaustividade	-Quantidade de assuntos na indexação -Quantidade de assuntos desejada -Abrangência desejada	<ul style="list-style-type: none">• Quantidade de assuntos satisfatória para 90%.• Indexação não exaustiva, desejam de 3 a 5 assuntos (70%).• Abrangência desejada com termos principais e secundários (75%), para identificar os implícitos.
Precisão	-Precisão do assunto -Precisão desejada	<ul style="list-style-type: none">• Níveis variados, resultados não condizentes ao pesquisado.• 85% desejam resultados mais precisos.
Especificidade	-Assuntos na indexação -Utilização de assuntos -Preferências por assunto	<ul style="list-style-type: none">• Uso de termos específicos (45%) e gerais (55%).• 65% utilizam termos específicos.• 70% preferem assuntos mais específicos.

Fonte: Elaborado pelas autoras, 2019.

Assim, com a análise dos resultados das quatro categorias, verificou uma indexação realizada com qualidade em poucos documentos e necessitando de ajustes em outros, pois obtiveram-se: taxas altas de revocação em oposição aos níveis baixos de precisão, isto é, uma recuperação da informação com documentos relevantes, mas também com irrelevantes; níveis de revocação e de precisão variados, apresentando resultados com respostas exaustivas e não precisas; exaustividade com termos pertinentes e em quantidade adequada, mas alguns documentos com termos amplos; especificidade com níveis variados e sem uniformidade.

Deste modo, acredita-se que a indexação pode ser moldada para usuários do curso de Psicologia, ao empregar uma quantidade suficiente de termos para não expandir demais o conteúdo do documento e ao mesmo tempo não limitar a sua abrangência, utilizando conceitos que representem tanto os assuntos principais como os secundários, obtendo, com isso, taxas ideais de revocação, e com a especificidade adequada para se obter mais precisão nos resultados, de forma que possa haver uma qualidade da indexação para uma recuperação da informação mais eficiente.

Para o **curso de Engenharia civil**, foi analisado o total de 924 documentos com base nas pesquisas realizadas no catálogo pelos 18 usuários. No Gráfico 3, pode-se observar as taxas de revocação e precisão de cada pesquisa.



Fonte: Elaborado pelas autoras, 2019.

De modo geral, para os documentos do curso de Engenharia civil, percebeu-se a necessidade de ajustes na representação dos documentos quanto ao elemento de exaustividade, pois foram observados níveis variados de revocação, altos e baixos, na recuperação da informação, com predominância de taxas altas, o que possibilitou recuperar documentos úteis. No entanto, com taxas baixas em algumas pesquisas, a quantidade de obras consideradas relevantes foi mais baixa. Isso demonstra uma variação dos níveis de exaustividade, principalmente na relação entre pertinência dos conceitos e conteúdo dos documentos, pois no quesito relativo à quantidade de termos utilizados, esta se mostrou satisfatória e suficiente para descrever o conteúdo do documento, segundo os usuários, como pode ser observado no Quadro 6 pelos resultados do curso.

Quanto à especificidade, verifica-se a necessidade de uma indexação em que os conceitos traduzidos em termos da linguagem de indexação sejam mais específicos, para que os assuntos pesquisados pelos usuários correspondam aos termos indexados, haja vista que poucos documentos corresponderam com precisão ao assunto pesquisado (Quadro 3).

Quadro 3: Resultados das quatro categorias de análise do curso de Engenharia civil

Curso de Engenharia civil – Área de Engenharias		
Categories	Questões	Resultados
Revocação	-Itens relevantes na pesquisa -Quantidade de itens recuperados -Resultado da pesquisa	<ul style="list-style-type: none"> • Taxas variadas de revocação, com predominância para alta. • Quantidade de itens recuperados satisfatória (72%) para análise do que seria útil. • 100% de satisfação com resultados da pesquisa.
Exaustividade	-Quantidade de assuntos na indexação -Quantidade de assuntos desejada -Abrangência desejada	<ul style="list-style-type: none"> • Quantidade de assuntos satisfatória para 90%. • Indexação não exaustiva, preferem de 1 a 5 assuntos (89%). • Abrangência com termos principais e secundários (89%), para não excluir temas importantes.
Precisão	-Precisão do assunto -Precisão desejada	<ul style="list-style-type: none"> • Predominância para a baixa (recuperando os irrelevantes). • 94% preferem resultados mais precisos (Mais precisão).
Especificidade	-Assuntos na indexação -Utilização de assuntos -Preferências por assunto	<ul style="list-style-type: none"> • Termos específicos (44%) e não específicos (56%). • 56% utilizam termos específicos. • 78% preferem assuntos mais específicos.

Fonte: Elaborado pelas autoras, 2019.

Diante desses resultados, foi observada uma indexação que necessita de ajustes em seus elementos de exaustividade e especificidade, pois poucas pesquisas tiveram níveis altos de revocação e precisão.

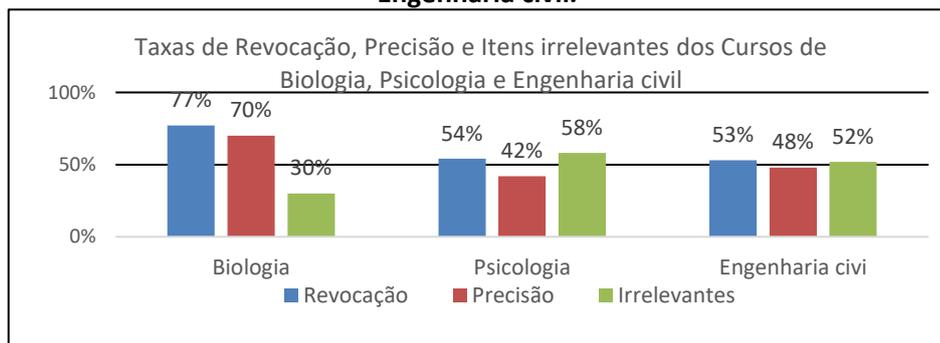
Verifica-se que, para este curso, a indexação, de certa forma, permite que os usuários acessem os documentos, mas não com a eficiência desejável para recuperar tantos documentos relevantes quanto possível e evitar o máximo de irrelevantes, de modo que se percebe uma menor qualidade na indexação desses documentos, visto que Lancaster (2004) afirma que a eficiência da recuperação da informação reflete a qualidade da indexação.

Assim, observou-se uma indexação com qualidade em poucos documentos, pois obtiveram-se níveis de revocação e de precisão variados e distintos, em que a alta revocação é acompanhada da baixa precisão; precisão com taxas variadas (predominância para a baixa), em que o sistema recuperou muitos documentos irrelevantes; exaustividade com níveis adequados, mas com termos amplos para a alta revocação; especificidade com níveis variados, predominando a baixa especificidade.

Diante dessa observação, constata-se que, para uma indexação adequada aos documentos do curso de Engenharia civil, com níveis de revocação e precisão que proporcionem uma recuperação da informação mais eficiente, a representação deve ser moldada nos seus elementos de exaustividade e especificidade, conforme os desejos dos usuários, seguindo diretrizes específicas para o tratamento de documentos de áreas específicas.

Com os resultados dos três cursos, verifica-se que a indexação realizada nos documentos analisados do curso de Biologia foi a mais adequada, com níveis de exaustividade e especificidade equilibrados e altos, ao ponto de permitir uma recuperação da informação mais eficiente, pois uma grande parte dos documentos recuperados foi considerada relevante e precisa aos assuntos pesquisados, com taxas de 77% e 70% respectivamente de revocação e precisão no total, como pode ser verificado no Gráfico 4. Isso demonstra uma qualidade na representação temática da informação em que os conceitos identificados e selecionados foram apropriados e descreveram o tema fundamental do conteúdo do documento em quantidade e pertinência adequada, e posteriormente traduzidos para a linguagem de indexação em termos suficientemente específicos para permitir a precisão e a exatidão dos resultados, sem restringir a sua temática, ao ponto de proporcionar uma alta revocação e alta precisão.

Gráfico 4: Taxas de revocação, precisão e itens irrelevantes dos cursos de Biologia, Psicologia e Engenharia civil.



Fonte: Elaborado pelas autoras, 2019.

Para os cursos de Psicologia e Engenharia civil, os documentos analisados apresentaram uma indexação não tão coerente em seus níveis de exaustividade e especificidade, pois algumas pesquisas apresentaram níveis bem elevados de revocação em oposição aos níveis baixos de precisão. Assim, para esses cursos, foi observada uma variação na exaustividade e especificidade, já que não se tem uma uniformidade na identificação, seleção e tradução dos conceitos, resultando em taxas altas e baixas de revocação e precisão na recuperação da informação, de modo que os usuários em algumas pesquisas conseguiram obter os documentos desejados e em outras não, sendo necessárias mudanças no processo de indexação para que os documentos sejam indexados adequadamente conforme as necessidades dos usuários.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

Deste modo, observou-se que, do total de documentos analisados pelos usuários dos três cursos (2.584 itens), mais da metade foram considerados relevantes na recuperação da informação (1.415), uma vez que os usuários conseguiram de certa forma encontrar o documento desejado, mas também muitos foram considerados irrelevantes (1.169), 45% no total em relação ao total de documentos recuperados e analisados, como observado no Quadro 4. Isso demonstra uma necessidade de melhorias no processo de indexação para que a quantidade de documentos considerados irrelevantes seja a menor possível.

Quadro 4: Total de documentos analisados dos cursos de Biologia, Psicologia e Engenharia civil

Cursos	Total de documentos recuperados e analisados	Doc. considerados relevantes	Doc. considerados irrelevantes	Porcentagem dos irrelevantes
Biologia	992	690	302	30%
Psicologia	668	282	386	58%
Engenharia civil	924	443	481	52%
Total	2584	1415	1169	45%

Fonte: Elaborado pelas autoras, 2019.

Diante desses resultados, constatou-se que, para mais qualidade na indexação e eficiência na recuperação da informação em catálogos de bibliotecas, torna-se necessário adotar medidas para que o processo de indexação seja realizado com mais eficácia e permita aos usuários o acesso eficiente aos documentos, como destacado no Quadro 5.

Quadro 5: Medidas para o processo de indexação ser realizado com mais qualidade

Medidas para o Processo de Indexação	Justificativa
Adoção de metodologias de leitura documentária	Para auxiliar na identificação e seleção dos conceitos mais relevantes na análise de assunto no momento da exaustividade.
Atualização da linguagem de indexação	Com termos atuais e pertinentes de acordo com a linguagem natural do usuário, para permitir, no momento da especificidade, uma tradução dos conceitos em termos mais coerentes com o perfil dos usuários.
Treinamentos e orientações para os usuários	Para que saibam utilizar todos os mecanismos que estão interligados ao processo de indexação, como as linguagens de indexação.
Elaboração de Política de indexação	Com critérios claros e específicos, principalmente para os elementos de exaustividade e especificidade, a fim de que os níveis de revocação e precisão sejam altos e proporcionais na recuperação da informação; Com diretrizes gerais para todo o sistema e específicas para as bibliotecas de áreas específicas, uma vez que a indexação realizada com o mesmo padrão para áreas diferentes ocasiona resultados não eficientes na recuperação da informação.
Avaliações periódicas da indexação	Para a indexação ser moldada e aperfeiçoada, visando maior qualidade no seu processo.

Fonte: Elaborado pelas autoras, 2019.

Assim, não há uma fórmula padronizada para realizar a indexação, porém ela pode ser ajustada e aperfeiçoada para cada curso e área dos documentos de atuação da biblioteca, seguindo a filosofia do sistema e desejos de representação e recuperação da informação dos usuários, de modo que os resultados da recuperação da informação sejam satisfatórios para todas as pesquisas realizadas no catálogo, com mais documentos relevantes e menos irrelevantes.

Da mesma forma, para que haja uma maior qualidade no processo de indexação dos documentos analisados, torna-se indispensável a utilização de uma política de indexação com instruções gerais para todas as bibliotecas da UFPA e diretrizes específicas para o tratamento temático de documentos de áreas específicas, pois cada biblioteca possui uma realidade diferente de indexação para necessidades distintas de acesso à informação.

5 CONSIDERAÇÕES FINAIS

Diante do que foi exposto ao longo deste artigo, verifica-se que a indexação tem ganhado destaque dentro do contexto da organização do conhecimento e da informação como atividade reflexiva para a determinação dos assuntos do documento, cuja análise tem se expandido para além do conteúdo, agregando ao processo as demandas dos usuários concernentes as suas necessidades de informações.

Com a análise dos dados foram observados resultados promissores, em que se verificou uma indexação realizada com qualidade em muitos documentos, principalmente para o curso de Biologia nos quesitos de exaustividade e especificidade, pois foram recuperados documentos relevantes (alta revocação) com poucos irrelevantes (alta precisão), conceito de indexação bem-sucedida definida por Lancaster (2004). Já para os cursos de Psicologia e Engenharia civil, verificou-se uma qualidade da indexação em poucos documentos, pois na recuperação da informação obtiveram-se documentos relevantes, mas também muitos irrelevantes (baixa precisão), devido aos níveis variados de exaustividade e especificidade empregados na indexação que provocaram taxas variadas de revocação e de precisão na recuperação da informação.

Desta maneira, para uma maior qualidade na indexação dos documentos dos cursos de Psicologia e Engenharia civil, há necessidade de melhorias nos elementos de

exaustividade e especificidade, sobretudo neste último quesito, a fim de que haja maior precisão nos resultados da recuperação da informação, isto é, para que o sistema evite o máximo de documentos não relevantes.

Assim, com esta pesquisa, constata-se que avaliar a indexação para buscar e trazer melhorias para o processo, incluindo os sujeitos que fazem parte do ciclo documentário, é a forma mais benéfica para adequar o processo de indexação conforme o contexto do sistema de bibliotecas, de cada biblioteca integrante desse sistema e de seus usuários.

Da mesma forma, as opiniões dos usuários se mostram válidas para aprimorar a indexação, moldando os níveis adequados de exaustividade, especificidade, revocação e precisão, com vistas a uma recuperação da informação eficiente, uma vez que esse contato permite maior compreensão das suas necessidades quanto à representação da informação.

É oportuno frisar a importância da continuidade de estudos desta temática para a contribuição teórico-metodológica da área, a fim de aperfeiçoar a prática da indexação em bibliotecas e beneficiar os usuários com uma recuperação da informação com mais eficiência e qualidade em catálogos.

REFERÊNCIAS

ARAÚJO JÚNIOR, R. H. de. **Precisão no processo de busca e recuperação da informação**. Brasília, DF: Thesaurus, 2007.

CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 221-241, set. 1985. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/reb/>. Acesso em: 18 maio 2017.

DIAS, E. W.; NAVES, M. M. L. **Análise de assunto: teoria e prática**. 2. ed. rev. Brasília, DF: Briquet de Lemos, 2013.

FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2089>. Acesso em: 07 jan. 2018.

FUJITA, M. S. L. A política de indexação para representação e recuperação da informação. *In*: GIL LEIVA, I.; FUJITA, M. S. L. (ed.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. Capítulo 1, p. 17-28. Disponível em: https://www.marilia.unesp.br/Home/Publicacoes/politica-de-indexacao_ebook.pdf. Acesso em: 03 jan. 2017.

FUJITA, M. S. L. Política de indexação para bibliotecas: funções e finalidades. *In*: FUJITA, M. S. L. (org.). **Política de indexação para bibliotecas: elaboração, avaliação e implantação**.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2016. Capítulo 1, p. 15-19. Disponível em: https://www.marilia.unesp.br/Home/.../politicass-de-indexacao-para-bibliotecas_ebook. Acesso em: 22 mar. 2016.

FUJITA, M. S. L.; GIL LEIVA, I. Avaliação da indexação por meio da recuperação da informação. **Ciência da informação**, Brasília, DF, v. 41, n. 1, p. 50-66, jan./abr. 2014. Disponível em: revista.ibict.br/ciinf/article/view/1418. Acesso em: 21 jan. 2018.

FUJITA, M. S. L.; RUBI, M. P. Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. **DataGramZero: Revista de Ciência da Informação**, v. 7, n. 3, jun. 2006. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000003842/3bde87ebc04d88ac4ff43b566f71e33f>. Acesso em: 20 fev. 2018.

GIL LEIVA, I. Aspectos conceituais da indexação. In: GIL LEIVA, I.; FUJITA, M.S.L. (ed.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. Capítulo 2, p. 31-106. Disponível em: https://www.marilia.unesp.br/Home/Publicacoes/politica-de-indexacao_ebook.pdf. Acesso em: 03 jan. 2017.

GIL LEIVA, I. **Manual de indización**: teoría y práctica. Espanã: TREA, 2008.

GUIMARÃES, J. A. C. As políticas de indexação como elemento para a gestão do conhecimento nas organizações. In: VIDOTTI, S. A. B. G. (coord.). **Tecnologia e conteúdos informacionais**: abordagens teóricas e práticas. São Paulo: Polis, 2004. Cap. 3, p. 43-52. (Coleção palavra-chave, 15).

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. rev. atual. Brasília, DF: Briquet de Lemos, 2004.

LOPES, E. de F. Avaliação de serviços de indexação e resumo: critérios, medidas e metodologias. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 242-256, set. 1985. Disponível em: <http://www.brapci.inf.br/index.php/article/download/13786>. Acesso em: 11 jan. 2018.

NUNES, C. O. Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. **Biblos**: Revista do Instituto de Ciências Humanas e da Informação, Rio Grande, v. 16, p. 55-61, 2004. Disponível em: <http://www.seer.furg.br/index.php/biblos/article/view/411>. Acesso em: 28 abr. 2017.

PIECADE, Maria Antonieta Requião. Introdução à teoria da classificação. 2. ed. rev. e aum. Rio de Janeiro: Interciência, 1983.

RUBI, M. P. Os princípios da política de indexação na análise de assunto para catalogação: especificidade, exaustividade, revocação e precisão na perspectiva dos catalogadores e usuários. In: FUJITA, M. S. L. (org.). **A indexação de livros**: a percepção de catalogadores e usuários de bibliotecas universitárias. Um estudo de observação do contexto sociocognitivo com protocolos verbais. São Paulo: Editora UNESP: Cultura Acadêmica, 2009. Cap. 4, p. 81-

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

93. Disponível em: <https://repositorio.unesp.br/handle/11449/109109>. Acesso em: 12 nov. 2017.

RUBI, M. P.; FUJITA, M. S. L. Política de indexação na catalogação de assunto em bibliotecas universitárias: a visão sociocognitiva da atuação profissional com protocolo verbal. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 7, n. 2, p. 118-150, jan./jun. 2010. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1960>. Acesso em: 20 out. 2017.

SILVA, M. dos R. da; FUJITA, M. S. L. A prática da indexação: análise da evolução de tendências teóricas e metodológicas. **TransInformação**, Campinas, v. 16, n. 2, p. 133-161, maio/ago. 2004. Disponível em: <http://revistas.puccampinas.edu.br/transinfo/viewissue.php?id=7>. Acesso em: 10 jan. 2018.

VAN SLYPE, G. **Los lenguajes de indización**: concepción, construcción y utilización em los sistemas documentales. Tradução: Pedro Hípola e Félix de Moya (versión española). Madrid: Fundación Germán Sánchez Ruipérez: Pirámide, 1991. Disponível em: http://www.ugr.es/~phipola/Los_lenguajes_de_indizacion.pdf. Acesso em: 15 fev. 2018.

UNISIST. Princípios de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 10, n. 1, p. 83-94, 1981. Disponível em: <http://www.brapci.inf.br/v/a/2687>. Acesso em: 15 jan. 2018.