

ANÁLISE DA PROPORCIONALIDADE NA RELAÇÃO ENTRE O PROCESSO DE INDEXAÇÃO E O GRAU DE PRECISÃO DA BASE DE DADOS COMPENDEX

Josceline Lira

E-mail: liraxl.joy@gmail.com.

André Anderson Cavalcante Felipe

E-mail: andreandersonf@gmail.com.

Resumo: Este artigo versa sobre o processo de indexação em bases de dados com o objetivo de analisar o grau de precisão da recuperação da informação na base de dados especializada em engenharia *Compendex*, pertencente à plataforma do *Engineering Village*. O procedimento metodológico empregado é do tipo exploratório, iniciado por uma pesquisa bibliográfica e os conhecimentos discutidos são observados e experienciados por meio do método de estudo de caso. Os resultados mostram que a *Compendex* adota um processo de indexação realizada por intermédio do requisito da especificidade, o qual diminui a revocação e eleva o índice de precisão (unicidade). Além disso, utiliza o tesauro *Engineering Index Thesaurus*, que confere maior precisão na conversão de termos, e conta com a vantagem da especialização do público-alvo, o que facilita a comunicação entre a *Compendex* e os consulentes e também constitui mais um fator para o aumento do grau de precisão. Conclui que a temática pesquisada é importante para que a Ciência da Informação reafirme e reivindique seu papel de protagonista no trabalho com a informação para a eficácia da recuperação.

Palavras-Chave: Indexação. Precisão. *Compendex*. *Engineering Village*. Bases de dados especializadas.

1 INTRODUÇÃO

Este artigo foi apresentado no 21º Encontro Regional de Estudantes de Biblioteconomia, Documentação, Ciência e Gestão da Informação – EREBD – 2018, Recife, no Grupo de Trabalho 5 (GT5) – Organização e representação da

informação, na modalidade de artigo completo para comunicação oral.

A indexação é uma atividade compreendida no escopo da Ciência da Informação (CI) e é imprescindível para a inserção e recuperação eficazes de informações em coleções/acervos. Esta prática pode constituir um elo entre a coleção/acervo e os usuários por meio da comunicação entre estes com o uso de uma linguagem compreensível para os pares de cada área temática coberta por cada sistema de busca/base de dados.

Entretanto, sua aplicação em base de dados ainda é tímida, principalmente como fundamento para o alcance da precisão, definida por Hudon (2009, p. 259, tradução nossa) como uma “medida de desempenho de um sistema de informação que define a proporção de documentos relevantes recuperados em comparação com o conjunto que consiste de todos os documentos fornecidos em resposta a um pedido”.

A indexação possibilitaria aos sistemas de busca o desenvolvimento de critérios de precisão na recuperação da informação para favorecer o usuário no que tange à satisfação de receber apenas resultados precisos na realização de buscas em uma base de dados. Tal afirmação orienta a construção do seguinte questionamento: como utilizar a indexação como fundamento para otimizar o grau de precisão nas consultas realizadas em bases de dados?

A fim de delinear o caminho sugerido pela questão-problema proposta, este artigo objetiva analisar a existência de proporcionalidade na relação entre o processo de indexação e o grau de precisão da base de dados *Compendex*.

Este estudo justifica-se pelo realce da importância da indexação para o cumprimento da recuperação pretendida pelo sistema de busca e pelo usuário, além de mostrar os ganhos que se tem com a normalização da linguagem utilizada pelos sistemas que implica em sistemas mais precisos e especializados. A discussão desencadeada neste artigo também enaltece a CI enquanto pilar do desenvolvimento de produtos de recuperação de informação tecnológicos.

Como fundamentação metodológica, esta pesquisa utilizou a pesquisa de nível/tipo “exploratória”, que tem como “principal finalidade desenvolver, esclarecer e modificar conceitos e ideias” (GIL, 2008, p. 27), a qual se iniciou com

uma “pesquisa bibliográfica” em livros, teses, materiais de aula e artigos científicos disponíveis acerca dos assuntos tratados neste artigo.

Quanto aos fins, esta pesquisa é classificada como “pesquisa aplicada”, já que o conhecimento explorado neste estudo teve aplicação prática em um estudo de caso, método “caracterizado pelo estudo exaustivo de um ou de poucos objetos, de maneira a permitir o seu conhecimento amplo e detalhado” (GIL, 2008, p. 57-58).

A pesquisa também se caracteriza quanto à abordagem de análise dos dados oriundos do estudo de caso enquanto “qualitativa” por promover uma análise dirigida pelo olhar do pesquisador em relação à ocorrência do conhecimento (precisão em bases de dados consequentes da indexação) no contexto específico (base de dados *Compendex*) (GIL, 2008).

Vale ressaltar que este artigo foi apresentado e premiado no “GT5: ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO”, do 21º Encontro Regional de Estudantes de Biblioteconomia, Documentação, Ciência e Gestão da Informação – EREBD, em 2018 na cidade de Recife.

2 INDEXAÇÃO

A indexação constitui o tema principal deste trabalho, cuja intenção é apontar o importante papel de protagonista da CI (área na qual a indexação é componente) na inter-relação com outros campos de estudo, que no caso deste artigo, está representada pela conversa entre indexação e bases de dados especializadas.

A palavra “indexar” denota a atividade de ordenar em uma lista de assuntos (em um índice), por exemplo, tópicos referentes às informações contidas em um documento de maneira que este índice auxilie na localização rápida destas informações (FERREIRA, 2001). Esta afirmação traz o delineamento da teoria fulcral acerca da indexação que este artigo utiliza, a qual corresponde a Lancaster (2004, p. 6, **grifo nosso**), que entende que a “indexação de assuntos” compõe o processo de moldagem “[...] de uma **representação** do conteúdo temático dos documentos”.

A atividade de indexação é realizada pelo indexador que ao entrar em contato com o documento por intermédio da leitura e posterior

representação do documento mapeia os temas presentes neste. Sobre o papel da “pessoa” que faz a indexação, Ferneda (2003) exalta a vantagem da criação de índices pelo processo manual por este ser apoiado pela riqueza da mente humana, a qual é capaz de engendrar significados e concatenações ímpares a partir da interação com produtos de informação, por exemplo, ao afirmar que

[...] a vantagem óbvia está na utilização da insubstituível capacidade humana em julgar relevância e categorizar documentos, refletindo diretamente na qualidade do índice gerado e, consequentemente, na precisão da recuperação, desde que exista algum tipo de controle de vocabulário (FERNEDA, 2003, p. 96).

Assim, com a prática da indexação obtêm-se a síntese do conteúdo de um documento e a criação de “pontos de acesso” (os termos) “mediante os quais um item é localizado e recuperado, durante uma busca por assunto num índice publicado ou numa base de dados eletrônica” (LANCASTER, 2004, p. 6).

Enquanto atividade, o indexador desenvolve a indexação de assuntos em dois momentos: no primeiro ocorre a análise conceitual e, em seguida, a tradução (LANCASTER, 2004), os quais serão abordados nas subseções deste capítulo.

2.1 Análise Conceitual

Na análise conceitual, o indexador faz a leitura técnica do documento, que consiste numa leitura “direcionada para certas partes do documento onde vai-se encontrar elementos especialmente importantes para a identificação do assunto ou assuntos do documento” (DIAS; NAVES, 2007, p. 52). As partes mais importantes para a leitura técnica, segundo Lancaster (2004, p. 24) são:

- a) título;
- b) resumo, se houver;
- c) sumário;
- d) introdução, as frases e parágrafos de abertura de capítulos, e as conclusões;
- e) ilustrações, gráficos, tabelas e respectivas legendas;

f) palavras ou grupos de palavras que apareçam sublinhados ou grafados com tipos diferentes.

A leitura técnica desencadeia avaliações de termos que melhor representem o assunto do conteúdo que é assimilado pelo indexador/leitor (LANCASTER, 2004; DIAS; NAVES, 2007). Contudo, Lancaster (2004, p. 9) alerta para o fato de que uma indexação de assuntos eficiente “implica que se tome uma decisão não somente quanto ao que é tratado num documento, mas também por que ele se reveste de provável interesse para determinado grupo de usuário”. Este fato de o indexador pautar a escolha de termos também pela previsão das necessidades de informação do público-alvo depende diretamente da reunião deste público-alvo em um ambiente informacional específico no qual seja garantida a circulação de apenas informação do escopo determinado. Lancaster (2004, p. 10) complementa que quanto mais “especializada a clientela de um centro de informação maior a probabilidade de que a indexação possa e deva ser feita sob medida, ajustando-se com precisão aos interesses do grupo”.

Os termos selecionados para indicar os conceitos/assuntos do documento analisado na análise conceitual ainda são termos oriundos da linguagem natural, a qual “compreende os idiomas ou línguas criadas pelos seres humanos para fins de comunicação” que passarão por tratamento na próxima etapa da indexação, que será discutida na subseção seguinte (CORRÊA, 2016a, p. 7).

2.2 Tradução

A etapa subsequente à etapa de análise conceitual/análise de assunto é a tradução, a qual “envolve a conversão da análise conceitual de um documento num determinado conjunto de termos de indexação”. Esta conversão pode ser feita por meio da comparação entre os termos selecionados na análise conceitual e os termos que são equivalentes a estes em um vocabulário controlado, definido como “uma lista de termos autorizados” e padronizados, pois são termos que adotam a linguagem documentária. O indexador pode utilizar-se de termos encontrados no documento (termos extraídos do documento) ou de termos formulados a partir do seu

entendimento (termos inferidos pelo indexador), os quais compõem o grupo de termos representativos daquele conteúdo (LANCASTER, 2004, p. 18-19).

As linguagens documentárias são “construídas para indexação, armazenamento e recuperação da informação e correspondem a sistemas de símbolos, destinadas a traduzir os conteúdos dos documentos”, e por meio das linguagens documentárias é possível “representar, de maneira sintética, as informações materializadas nos textos” (CINTRA *et al*, 1994, p. 23-24). Entre os tipos de vocabulário controlado está o tesauro, o qual promove uma estrutura organizada de termos pertencentes ao universo que cada tesauro normaliza (LANCASTER, 2004), portanto, também se caracteriza como uma linguagem documentária (CINTRA *et al*, 1994).

Por fim, é importante atentar para o alerta de Cintra *et al* (1994) quanto a discernir o caráter ‘referencial’ que a representação dos conceitos/assuntos de um documento pela linguagem documentária tem e a completude insubstituível do próprio documento, pois a tarefa de indexação permite que se estruture uma maneira mais rápida de descobrir os assuntos que são tratados em documentos. Torna-se imprescindível que a indexação tenha alta qualidade, já que guiará a tomada de decisão na busca por assuntos, entretanto, os índices resultantes da indexação são apenas indícios do conteúdo dos documentos, ou seja, não os substituem.

2.3 Precisão e Revocação

O título deste artigo prevê a checagem da relação da indexação com a precisão de bases de dados, todavia, a literatura acerca da recuperação de informação versa sobre a precisão comumente associada à revocação. Por esse motivo este estudo entende como relevante abordar os dois conceitos.

Os conceitos de precisão e revocação são tratados na literatura pertinente à avaliação do desempenho de sistemas de recuperação de informação (SRI) e correspondem a índices desta avaliação que classificam o desempenho do SRI mediante a resposta que este retorna para as consultas realizadas no sistema.

Segundo Cardoso (2004, p. 34), “precisão é a fração dos documentos já examinados que são relevantes, e revocação é a fração dos documentos relevantes observada dentre os documentos examinados”. Nas palavras de Corrêa (2016b, p. 8), precisão é a “habilidade de recuperar somente itens relevantes (unicidade)” e revocação, também denominada por cobertura, é a “habilidade de recuperar todos os itens relevantes (completude)”. Diante das assertivas apresentadas, um sistema de recuperação de informação caracterizado pela precisão é aquele no qual são retornados apenas os documentos relevantes, ou seja, itens condizentes com o tema/assunto da consulta expressa pelo usuário, e um sistema caracterizado pela revocação é aquele no qual são retornados outros itens (irrelevantes) além dos documentos relevantes.

Corrêa (2016b), ao observar o conflito entre revocação e precisão, aponta uma relação importante entre estes índices e a indexação ao afirmar que

[...] a exaustividade e a especificidade na indexação são os dois fatores mais importantes que governam a revocação e a precisão; a exaustividade aumenta a revocação e diminui a precisão, enquanto que uma maior especificidade leva à menor revocação e maior precisão (CORRÊA, 2016b, p. 16).

Isto indica que uma representação de assuntos mais abrangente abre margem para a recuperação de muitos documentos e este alargamento de possibilidades acarreta na dispersão da consulta e dificulta o alcance e retorno dos documentos relevantes pelo sistema; ao mesmo passo em que uma representação concisa oferece mais probabilidade de os documentos relevantes e somente estes serem encontrados, pois há a delimitação na busca, ou seja, apenas aquele assunto ou tema será pesquisado na coleção.

3 A BASE DE DADOS ESPECIALIZADA COMPENDEX, DO ENGINEERING VILLAGE

De acordo com Rowley (2002, p. 125, **grifo nosso**), “uma base de dados é uma coleção geral e integrada de dados junto com a descrição deles, gerenciada de forma a atender a diferentes necessidades de seus usuários”. Ou seja, uma

base de dados compreende o ambiente (a plataforma) no qual são reunidos documentos representados enquanto ‘registros’, os quais correspondem às informações pertinentes a cada documento estruturadas em ‘campos’ (nome do autor, título, ano, entre outros) que apresentam os dados daquele documento, que compõem os ‘arquivos’, que são a coleção dos registros, integrados em uma base de dados (ROWLEY, 2002; ÁVILA, 2017a).

A base de dados especializada “é centralizada e focada em um determinado tema, gerenciada de forma a atender a diferentes necessidades de seus usuários específicos” (ÁVILA, 2017b, p. 9). De acordo com esta assertiva, o que caracteriza uma base de dados especializada é a especialização desta em um tema único, ou seja, o foco da base de dados em um tema único; por exemplo, uma base de dados especializada em ‘indexação’ terá seu escopo delimitado exclusivamente em documentos que versem sobre assuntos equivalentes ao tema indexação.

No tocante à linguagem, o sistema de busca de uma base de dados especializada pode utilizar linguagem documentária ou linguagem natural. No entanto, Rowley (2002), ao tecer comentários sobre a estrutura da informação e programas de computador, adverte que o uso de linguagem natural acarreta imprecisão ao explicar que

[...] há inúmeras palavras empregadas na linguagem natural, como ‘em média’, ‘um bocado de’, ‘poucos’ e ‘muitos’, que são imprecisas e difíceis para o sistema interpretar como um número específico ou determinada faixa de números (ROWLEY, 2002, p. 100).

A imprecisão também pode ser característica da ambiguidade das palavras da linguagem natural (ROWLEY, 2002), como, por exemplo, os significados imbricados na palavra ‘manga’, que podem se referir a um fruto ou a uma parte do vestuário, implicariam em dúvida na busca por documentos feita a partir deste termo em um sistema de busca que utiliza linguagem natural, pois seriam retornados itens de várias áreas e com vários olhares sobre esse objeto. Em contrapartida, a mesma autora afirma que o uso de linguagem controlada de indexação aumenta o grau de precisão de uma base de dados, ao comentar que

[...] em algumas bases há campos especiais onde se encontram termos da linguagem controlada de indexação, os quais são comumente empregados para representar o assunto do documento. As buscas feitas nesses termos costumam proporcionar maior precisão do que as feitas nos termos que aparecem, por exemplo, no texto integral ou no resumo do documento (ROWLEY, 2002, p. 227).

O *Engineering Village* é de propriedade da Elsevier¹, empresa integrante do RELX Group, sediado em Londres, o qual tem suas atividades voltadas para o fornecimento global de informações e análises para profissionais e empresários, principalmente, de quatro segmentos de mercado: científico, técnico e médico; análise de risco e negócios; legal; e exposições/eventos.

O *Engineering Village* é um serviço de busca em linha voltado para pesquisa em engenharia. Por serviço de busca em linha entende-se serviços os quais “instalam uma variedade de bases de dados num sistema de computador de grande porte e oferecem aos usuários o acesso, geralmente pago, a essas bases” (ROWLEY, 2002, p. 210). É o que o serviço em questão faz, este reúne bases de dados especializadas (focadas nos temas pertinentes à engenharia) essencialmente referenciais (que indicam ao usuário onde a informação pode ser encontrada e fornecem resumos, *proceedings*, citações, referências etc.), bases de dados de patentes, livros eletrônicos e notícias de várias áreas da engenharia (ELSEVIER, 2017).

A plataforma do *Engineering Village* é composta por doze bases de dados de pesquisa para a área de engenharia entre as quais está a *Compendex*, que fornece publicações revisadas – jornais, revistas, *proceeding* de conferências, séries de livros, artigos, dissertações, *papers* de conferências – e indexadas para engenheiros que abordam assuntos da literatura de engenharia nas áreas de elétrica, geral, civil, química, mecânica, mineração, física aplicada, óptica, bioengenharia, biotecnologia, tecnologia de alimentos, ciências de materiais, instrumentação, dispositivos médicos e nanotecnologia (ELSEVIER, 2017).

Todas as bases de dados do *Engineering Village* utilizam linguagem controlada e a indexação é feita de maneira específica/exaustiva para atender todos os termos possíveis (e apenas estes) dentro do vocabulário controlado da área de abrangência da base de dados para garantir maior precisão e relevância ao *corpus* (ELSEVIER, 2017). Sobre a linguagem controlada, as informações, no site da Elsevier, informam que na *Compendex*,

todos os registros são cuidadosamente selecionados e indexados usando o *Engineering Index Thesaurus* para garantir a descoberta e recuperação de literatura específica de engenharia em que estudantes de engenharia e profissionais possam confiar. Ao usar o *Ei Compendex*, os engenheiros podem ter certeza de que a informação é relevante, completa, precisa e de alta qualidade (ELSEVIER, 2017. Tradução nossa).

A *Compendex*, base de dados especializada do *Engineering Village*, tem como potencial usuário a comunidade cujas atividades são subsidiadas pela área do conhecimento da engenharia e concentra em um mesmo lugar informações similares e agiliza a recuperação de informação, pois a busca é centralizada em um só lugar e, com isto, o tempo de pesquisa é reduzido.

4 ANÁLISE E RESULTADOS

A seguir, será apresentada a análise advinda do estudo de caso realizado através da interação com a base de dados *Compendex* a fim de concatenar a experiência nesta base de dados com as teorias visitadas no referencial teórico e, assim, perceber se o grau de precisão nesta base de dados está atrelado às atividades de indexação.

Ao analisar o que foi supracitado nas discussões acerca da indexação e da precisão, juntamente com as informações coletadas no site da Elsevier, é possível afirmar que a *Compendex* é uma base de dados especializada com cobertura temática da área de engenharia que utiliza artifícios para garantir que apenas assuntos relacionados à engenharia sejam inseridos na sua coleção e que as informações nela englobadas

¹Fonte: <https://www.elsevier.com.br/>.

sejam recuperadas pelos usuários de maneira rápida e precisa. A *Compendex* centraliza e reúne de outras fontes todas as informações relevantes para seu público-alvo, poupando-o tempo e otimizando resultados assertivos.

Ao apresentar informações sobre a base de dados *Compendex*, a *Elsevier* deixa claro o artifício utilizado para a precisão na recuperação da informação, entregando ao usuário apenas itens relevantes, evitando que documentos irrelevantes estejam em meio aos documentos retornados. A base de dados adota um rigoroso processo de indexação realizado através do requisito da especificidade, o qual diminui a revocação e eleva o índice de precisão (unicidade), como dissera Corrêa (2016b), que se configura como pilar para a alta qualidade da *Compendex* enquanto base de dados que estabelece um caráter de precisão atinente ao tema engenharia.

O processo de indexação da *Compendex* é feito com a vantagem da especialização do público-alvo, que é composto por estudantes, pesquisadores e profissionais da área da engenharia. Isto facilita a comunicação entre a *Compendex* e os consultantes e também constitui mais um fator para o aumento do grau de precisão desta base de dados referente à indexação, pois os termos de indexação a serem escolhidos pela *Compendex* serão os pertencentes ao vocabulário daquele público.

Sobre a linguagem utilizada pela *Compendex*, ela é de cunho controlada, cuja discussão teórica deste tipo de linguagem foi guiada pelas ideias de Cintra *et al* (1994) de que esta linguagem melhor representa a síntese das informações contidas nos documentos, de Rowley (2002), que relaciona o uso da linguagem controlada ao aumento de precisão de bases de dados, e de Lancaster (2004), que afirma que o uso desta linguagem acarreta padronização e maior controle.

Na base de dados em análise, os termos, na etapa de tradução, são convertidos para termos do *Engineering Index Thesaurus*, o que confere maior precisão tanto na indexação de documentos que atendam à temática da engenharia (o que infere também na especialização) quanto na recuperação de documentos apenas relevantes. Resta apenas conhecer um pouco da interface da *Compendex*.

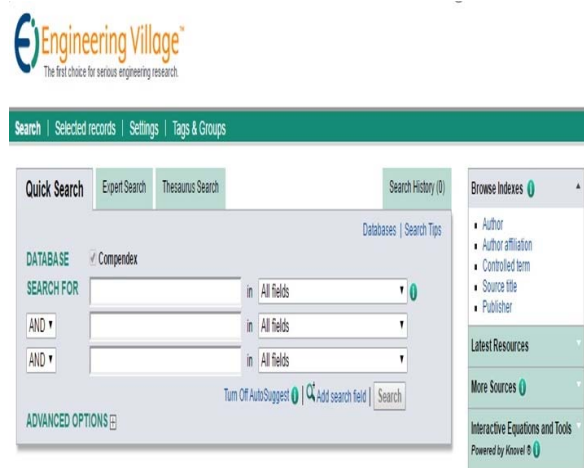


Figura 1. Representação da tela inicial do *Engineering Village* com permissão apenas para acesso à *Compendex*.
Fonte: <https://www-engineeringvillage-com.ez16.periodicos.capes.gov.br/search>.

Ao observar as abas de pesquisa (*quick search*, *expert search* e *thesaurus search*) na tela representada na figura 1 é percebido que a *Compendex* é destinada a especialistas que têm conhecimento da linguagem controlada de engenharia, o que proporcionará precisão para a busca, pois as três abas de opção de busca requerem conhecimento dos termos do vocabulário controlado.

Na opção de *quick search* é possível digitar os termos e combiná-los por operadores booleanos (AND, por exemplo). Mesmo tratando-se da opção de busca rápida, esta é feita por vocabulário controlado, mas o usuário conta com a facilidade da autossugestão, como demonstra a figura a seguir, que trata de uma simulação de busca com a digitação de letras aleatórias (l,a,g) realizada na base de dados *Compendex*.

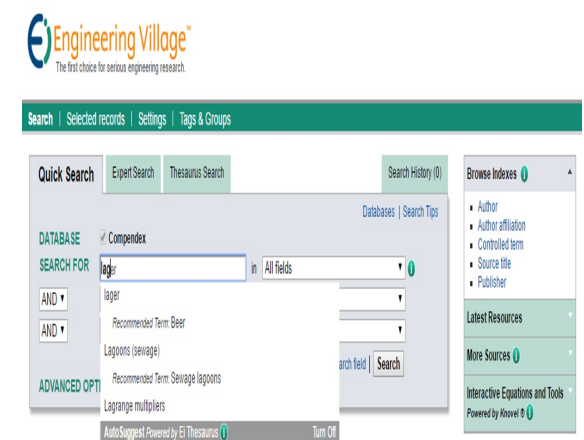


Figura 2 – Simulação de busca na opção *quick search* na *Compendex*.
Fonte: <https://www-engineeringvillage-com.ez16.periodicos.capes.gov.br/search>.

Há também a opção de busca mais especializada denominada *expert search*, em que o usuário cria uma expressão de busca em linguagem controlada, podendo usar códigos, como mostra a figura a seguir.



Figura 3 – Aba da opção de busca *expert search* selecionada na *Compendex*.

Fonte: <https://www-engineeringvillage-om.ez16.periodicos.capes.gov.br/search>

A última opção de busca é por meio da digitação do termo específico do tesauro. A figura a seguir mostra a aba *thesaurus search* selecionada.



Figura 4 – Aba da opção de *thesaurus search* selecionada na *Compendex*.

Fonte: <https://www-engineeringvillage-om.ez16.periodicos.capes.gov.br/search>.

Por fim, é possível inferir que a *Compendex* como base de dados apresenta um serviço de recuperação da informação altamente qualificado e sustentado pelo primor do processo de indexação e do caráter de precisão no acesso aos assuntos de engenharia.

5 CONSIDERAÇÕES FINAIS

Esta pesquisa partiu da problemática que questionava se utilizar a indexação como fundamento para a criação de bases de dados acarretava maior grau de precisão para elas, a qual suscitou a elaboração do objetivo do qual delineou o caminho a ser seguido para respondê-lo por meio da verificação da proporcionalidade na relação entre a atividade de indexação e o grau de precisão da base de dados especializada *Compendex*, a qual serviu como objeto de estudo desta pesquisa.

Por intermédio da coleta de informação para compor o referencial teórico e da interação com a *Compendex* foi possível responder positivamente a problemática de pesquisa e, consequentemente, o objetivo proposto.

Assim, foi descoberto que a utilização da indexação como fundamento na criação de bases de dados confere maior grau de precisão para elas. Fundamento no sentido de atentar para a atividade que permitirá que a base de dados cumpra seu papel principal de garantir a recuperação eficaz. A indexação é imprescindível para a precisão de bases de dados, configurando como pilar que controla e disponibiliza o acervo de maneira certa.

O estudo contido neste artigo é importante para que a CI, por meio do seu componente temático indexação, reafirme e reivindique seu papel de protagonista no trabalho com a informação para a eficaz recuperação. A importância deste artigo também está no fornecimento de literatura para a conscientização do uso correto da indexação e orientação para melhores resultados na criação de sistemas de busca.

REFERÊNCIAS

ÁVILA, B. T. **Introdução às bases de dados**. Recife: Bruno Tenório Ávila, 2017a. Slides de aula da disciplina Bases de Dados Especializadas do curso de Gestão da Informação da UFPE.

_____. **Bases de dados especializadas**. Recife: Bruno Tenório Ávila, 2017b. Slides de aula da disciplina Bases de Dados Especializadas do curso de Gestão da Informação da UFPE.

CARDOSO, O. N. P. Recuperação de Informação. **INFOCOMP Journal of Computer Science**, [S.l.], v.

2, n. 1, p. 33-38, nov. 2004. Disponível em: <<http://www.dcc.ufla.br/infocomp/index.php/INFOCOM-P/article/view/46>>. Acesso em: 24 jun. 2017.

CINTRA, A. M. *et al.* **Para entender as linguagens documentárias**. São Paulo: Polis, 1994.

CORRÊA, R. F. **Indexação e construção de índice**. Recife: Renato Fernandes Corrêa, 2016a. Slides de aula da disciplina Recuperação de Informação do curso de Gestão da Informação da UFPE.

_____. **Avaliação de Desempenho de Sistemas de Recuperação de Informação**. Recife: Renato Fernandes Corrêa, 2016b. Slides de aula da disciplina Recuperação de Informação do curso de Gestão da Informação da UFPE.

DIAS, E. W.; NAVES, M. M. L. **Análise de assunto: teoria e prática**. Brasília: Thesaurus, 2007.

ELSEVIER. **Content Compendex**. Disponível em: <<https://www.elsevier.com/solutions/engineering-village/content/compendex>>. Acesso em: 1 jul. 2017.

FERNEDA, E. **Recuperação de informação**: análise sobre a contribuição da ciência da computação para a ciência da informação. Tese (Doutorado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo, 2003. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/pt-br.php>>. Acesso em: 23 jun. 2017.

FERREIRA, A. B. de H. **Mini Aurélio século XXI: o minidicionário da língua portuguesa**. 5. ed. Rio de Janeiro: Nova Fronteira, 2001.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

HUDON, M. **Guide pratique pour l'élaboration d'un thésaurus documentaire**. Québec: Editions ASTED, 2009.

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

ROWLEY, J. **A biblioteca eletrônica**. Brasília: Briquet de Lemos, 2002. Disponível em: <https://pt.slideshare.net/SupperV/biblioteca-eletronica-jennifer-rowley>. Acesso em: 10 abr. 2017.

ANALYSIS OF PROPORTIONALITY IN THE RELATIONSHIP BETWEEN THE INDEXING

PROCESS AND PRECISION DEGREE THE COMPENDEX DATABASE

Abstract: This article deals with the database indexing process with the objective of analyzing the degree of accuracy of information retrieval in the specialized engineering database Compendex belonging to the Engineering Village platform. The methodological procedure employed is of the exploratory type, initiated by a bibliographical research and the knowledge discussed are observed and experienced through the case study method. The results show that Compendex adopts an indexing process performed through the requirement of specificity, which decreases recall and raises the precision index (uniqueness). In addition, it uses the Engineering Index Thesaurus, which gives greater precision in the conversion of terms and has the advantage of specialization of the target public that facilitates the communication between Compendex and the consultants, constituting as another factor for the increase the degree of precision. It concludes that the researched topic is important so that Information Science reaffirms and claims its protagonist role in the work with the information for the effectiveness of the recovery.

Keywords: *Indexing. Precision. Compendex. Engineering Village. Specialized databases.*