



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

ESTRATÉGIAS PARA SELEÇÃO DE TERMOS DO SISTEMA DE INDEXAÇÃO SISA

STRATEGIES FOR SELECTING SISA INDEXING SYSTEM TERMS

Isaque Katahira - Universidade Estadual Paulista Júlio de Mesquita Filho

Mariângela Spotti Lopes Fujita - Universidade Estadual Paulista Júlio de Mesquita Filho

Isidoro Gil Leiva - Universidad de Murcia

Roberta Caroline Vesu Alves - Universidade Estadual Paulista Júlio de Mesquita Filho

Modalidade: Trabalho Completo

Resumo: A organização e recuperação da informação são fundamentais no processo de divulgação do conhecimento científico. Contudo, para que o conhecimento produzido seja aproveitado, a organização e representação dos dados deve ser aprimorada com o processo de indexação para que a recuperação seja precisa. O objetivo deste artigo é analisar as estratégias de identificação e seleção de termos do Sistema de Indexação Semiautomático (SISA) para artigos científicos em português da área de Odontologia. A análise foi desenvolvida em duas etapas: a primeira realizou a análise das estratégias de identificação e seleção de termos das duas versões do SISA (Windows e Web) com base no Modelo de Leitura Documentária para textos científicos; a segunda identificou e comparou os termos selecionados nas duas versões do SISA. Na primeira etapa da análise, verificou-se que o SISA, nas duas versões, utiliza partes da estrutura textual para quantificar e qualificar os termos de indexação com base na frequência de aparecimento ponderados com a frequência em mais de uma parte da estrutura textual. Na segunda etapa, houve em média 17% de consistência entre as palavras-chave dos autores e do Modelo de Leitura Documentária; 15% em média do SISA Windows 2.0 e 30% em média do SISA Web 4.0. Conclui-se que a proposta de processo de indexação automática com base no uso do software SISA, atende de forma satisfatória quanto à representação na indexação dos artigos científicos selecionados.

Palavras-Chave: Sistema de Indexação Semiautomático (SISA); Modelo de Leitura Documentária; Indexação.

Abstract: The organization and retrieval of information are fundamental in the process of dissemination of scientific knowledge. However, for the knowledge produced to be harnessed, the organization and representation of the data must be enhanced with the indexing process for accurate retrieval. The aim of this paper is to analyze the term identification and selection strategies of the Semi-Automatic Indexing System (SISA) for Portuguese scientific articles in Dentistry. The analysis was carried out in two stages: the first performed the analysis of the identification and term selection strategies of the two versions of SISA (Windows and Web) based on the Documentary Reading Model for scientific texts; The second identified and compared the terms selected in the two versions of SISA. In the first stage of the analysis, it was found that SISA, in both versions, uses parts of the textual structure to quantify and qualify the frequency-weighted frequency-based indexing terms in more than one part of the textual structure. In the second stage, there was on average 17% consistency

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

between the keywords of the authors and the Documentary Reading Model; 15% on average for SISA Windows 2.0 and 30% on average for Web version 4.0 of SISA, respectively. It is concluded that the proposed automatic indexing process based on the use of SISA software satisfactorily meets the indexation representation of the selected scientific articles.

Keywords: Semi-Automatic Indexing System (SISA); Documentary Reading Model; Indexing.

1 INTRODUÇÃO

A organização e recuperação da informação são fundamentais no processo de divulgação do conhecimento científico. A divulgação do conhecimento produzido fundamenta e impulsiona novas descobertas. Nesse contexto, as bases de dados e os repositórios vinculados a instituições de pesquisa desempenham papel essencial, contudo, para que o conhecimento produzido seja efetivamente aproveitado, a organização dos dados deve ser feita de maneira que a recuperação seja realizada de forma rápida e precisa. Para tanto, o processo de indexação é fundamental a fim de representar e recuperar informações de modo a facilitar e agilizar o acesso, haja vista o grande e crescente número de publicações que requerem análise e catalogação.

Diante de tal necessidade, a Análise Documentária (AD), de acordo com Fujita e Nardi (1998, p. 21), realiza “um tratamento documentário de conteúdo com a finalidade de elaborar representações condensadas do que está contido no texto”. Assim, a indexação é feita por meio da descrição e identificação do conteúdo de um documento através de termos selecionados (VIEIRA, 1988). As ações requeridas pela indexação podem ser realizadas tanto pelo homem (indexação manual) quanto por programas computacionais (indexação automática) (VIEIRA, 1988; GIL-LEIVA, 2017).

Com o incremento crescente do número de publicações, a indexação manual tornou-se inviável para a organização de grandes bases de dados, seja pela morosidade de análise, seja pelo custo elevado ou pela falta de uniformidade na escolha dos termos representantes dos documentos. Nesse contexto, os sistemas de indexação automática se constituem como ferramentas fundamentais para a organização e recuperação de dados, especialmente no ambiente científico que requer sistematização e atualização precisa e ágil (ROBREDO, 1991; CORRÊA; BAZÍLIO, 2017).

Quanto à indexação automática, Lancaster (2004) destaca dois tipos, a saber: indexação por extração e indexação por atribuição. A indexação por extração seleciona automaticamente termos do texto dos documentos; já a indexação por atribuição relaciona termos da publicação analisada com termos de um vocabulário controlado também de forma automática.

Uma vez definidos os parâmetros de um método, é possível automatizar o processo de indexação de textos científicos por meio de termos significativos. Há diversos métodos utilizados para a indexação automática de termos: método de frequência; probabilístico; matemático; análise de *cluster*; associativo de palavras dentre outros (VIEIRA, 1988).

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

A partir da década de 1950, surgiram os primeiros sistemas automatizados de indexação. Posteriormente, a indexação automática foi aplicada a diversos contextos de análise (VIERIA, 1988; ROBREDO, 1991; GIL LEIVA, 1999; NARUKAWA; GIL-LEIVA; FUJITA, 2009; LAPA; CORREA, 2014; CORRÊA; BRAZILIO, 2017; BADIM; CORREA, 2019), fato que contribuiu para a organização do crescente volume de dados científicos publicados. Destacam-se, em especial, os softwares SISA, utilizado para a indexação automática por atribuição de artigos científicos (FUJITA, GIL-LEIVA, 2009), e OGMA, aplicado na indexação automática por extração de sintagmas nominais (CORRÊA; BRAZILIO, 2017).

Em 1999, quando poucas abordagens computacionais eram projetadas especificamente para a indexação automática ou semiautomática de textos científicos, o sistema desenvolvido pelo Prof. Dr. Isidoro Gil Leiva, da Universidade de Murcia, na Espanha, denominado SISA, buscou apresentar um método de indexação de textos científicos pautado em critérios de frequência (NARUKAWA; GIL LEIVA; FUJITA, 2009).

Especificamente o SISA, segundo Gil Leiva (2008), é um Sistema de Indexação automatizada para artigos científicos, elaborado para indexação semiautomática (se houver avaliação e edição de seus resultados pelo indexador) ou automática (sem edição dos seus resultados pelo indexador). O SISA foi desenvolvido para ambiente Windows na versão 2.0 e para a Web na versão 4.0 com diferenças entre as versões, principalmente, no que se refere à indexação de vários documentos ao mesmo tempo. A concepção do SISA é baseada na experiência e estratégias de indexadores humanos e, principalmente, no desenvolvimento de um processo de indexação com a vantagem de poder incluir os métodos utilizados para a indexação automática tal como a frequência de palavras.

Com base em estratégias metacognitivas de indexadores humanos, Fujita e Rubi (2006) elaboraram o Modelo de Leitura Documentária para indexação manual de textos. Esse modelo está fundamentado na estrutura básica de organização de artigos científicos para localização dos conceitos universais: objeto, ação, agente, métodos do agente, local ou ambiência, causa efeito, ponto de vista do autor e perspectiva. No processo de indexação de textos científicos, é possível utilizar o Modelo de Leitura Documentária para recuperação de dados, posto que “o texto científico possui uma estrutura argumentativa composta de justificativa e conclusão, acrescida da apresentação de um problema e sua solução.” (FUJITA, RUBI, 2006 apud VAN DIJK, 1992).

Diante do exposto, o problema desta pesquisa refere-se à identificação e validação de estratégias utilizadas no processo de indexação para identificação e seleção de termos mediante o uso das duas versões do SISA com relação ao Modelo de Leitura Documentária. Espera-se confirmar a hipótese de que as estratégias para identificação e seleção de termos de indexação são válidas. Também serão verificados os avanços entre as duas versões do SISA na indexação automática.

Nesse sentido, o objetivo deste artigo é analisar as estratégias de identificação e seleção de termos do SISA para artigos científicos em português da área de Odontologia. Justifica-se a realização do presente trabalho devido ao crescente volume de dados carentes de organização e à necessidade de estratégias que garantam a qualidade de indexação, bem como a qualidade de recuperação com maior agilidade. Assim, a partir da observação das convergências e divergências das duas formas de indexação, será possível comprovar a viabilidade do uso do software SISA para a indexação de artigos científicos.

2 INDEXAÇÃO: PROCESSO, MÉTODOS E SISTEMAS

A interpretação de um texto é suscetível a percepções subjetivas, já que evoca no interlocutor sua condição cognitiva no instante da leitura. Nesse aspecto, indexar é atribuir termos representativos de parte ou do todo de um documento (VIEIRA, 1988). A subjetividade é um fator latente na indexação manual, visto que esta é feita por pessoas diferentes, localizadas em tempos e espaços também distintos.

A partir da compreensão do indexador como leitor, o processo de representação do documento na indexação manual é altamente dependente da experiência do indexador, isso requer do profissional a capacidade de tradução da linguagem do usuário para uma linguagem indexada escolhida, por meio de instrumentos como tesouros, vocabulários controlados, listas de cabeçalhos de assunto dentre outros (VIEIRA, 1988).

O indexador é um leitor que interage com o texto para cumprir o objetivo da indexação. Nessa interação, o indexador lê o texto. Partindo dessa constatação, o estudo da leitura oferece uma outra perspectiva para a compreensão do processo de indexação e das dificuldades de um indexador frente a um documento: a de observação do processo de leitura do indexador para verificar estratégias de compreensão que visam à identificação de conceitos e de seu contexto para verificar os aspectos que influenciam e determinam a leitura documentária como leitura profissional (FUJITA; RUBI, 2006, p. 1).

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

A identificação e seleção de conceitos são considerados estágios do processo de indexação, segundo a Norma ABNT 12676 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992). A identificação de conceitos é realizada pela abordagem sistemática do conteúdo documental e tem como importante pressuposto basear-se em análise conceitual dos itens: ação, processo, instrumentos, local, métodos, objeto, parte do objeto e outros. A seleção de conceitos, por outro lado, é realizada por intermédio de um vocabulário controlado após abordagem do conteúdo documental. Portanto, a identificação de conceitos propiciará a representação com palavras significativas do conteúdo documental analisado, já a seleção de conceitos permitirá a representação das palavras significativas por termos de indexação. Nesse sentido, entre a identificação e a seleção é realizado um processo de controle de vocabulário que a Norma ABNT (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992) denomina de tradução.

Frente a complexidade e a variabilidade do processo de indexação de documentos, os autores Fujita e Rubi (2006), com base em resultados de processos de leitura de indexadores experientes, propõem um Modelo de Leitura Documentária baseado em estratégias de leitura e adoção de parâmetros de indexação, de modo a auxiliar o processo de indexação manual e torná-lo mais objetivo.

Para fins documentários, a leitura não é de caráter linear, uma vez que o leitor busca absorver a macroestrutura textual, de modo a evocar seus conhecimentos prévios para inferir significados e sugerir hipóteses para compreender a temática global do documento (FUJITA; RUBI, 2006). O Modelo de Leitura Documentária traz parâmetros para realizar a indexação manual, conforme disposto no Quadro 1.

Quadro 1: Componentes do Modelo de Leitura Documentária.

Conceitos	Questionamento para identificação de conceitos	Partes da estrutura textual
OBJETO e PARTE(S) DO OBJETO (algo ou alguém que está sob estudo do pesquisador)	O documento possui em seu contexto um objeto sob efeito desta ação?	Introdução (objetivos)
AÇÃO (processo sofrido por algo ou alguém)	O assunto contém uma ação (podendo significar uma operação, um processo etc.)?	Introdução (objetivos)
AGENTE (aquele ou algo que realizou a ação)	O documento possui um agente que praticou esta ação?	Introdução (objetivos)
MÉTODOS (métodos utilizados para realização da pesquisa)	Para estudo do objeto ou implementação da ação, o documento cita e/ou descreve modos específicos, por exemplo: instrumentos especiais, técnicas, métodos, materiais e equipamentos?	Metodologia

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

LOCAL OU AMBIÊNCIA (local físico onde foi realizada a pesquisa)	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	Metodologia
CAUSA E EFEITO Causa (ação+objeto)/Efeito	Considerando que a ação e o objeto identificam uma causa, qual é o efeito desta causa?	Resultados; discussão de resultados; Conclusões

Fonte: Extraído de Fujitae Rubi (2006).

Os dados explicitados nas três colunas do Quadro 1 proporcionam ao indexador ir diretamente às partes da estrutura textual para identificar conceitos, especialmente quanto à localização dos termos para agilizar o processo de indexação. Portanto, a identificação de termos é realizada a partir da estratégia de análise conceitual combinada com a estratégia de exploração da estrutura textual. Nesse sentido, podem ser utilizadas as estratégias de leitura interativa nos movimentos descendente (*bottom-up*) e ascendente (*top-down*); a primeira relacionada com a extração de significados contidos no texto e a segunda com a atribuição de significados pautados no conhecimento prévio do leitor (FUJITA; RUBI, 2006).

Exposto o Modelo de Leitura Documentária escolhido para a realização da indexação manual, dado o objetivo comparativo da presente pesquisa, é importante apresentar, ainda que brevemente, o histórico da indexação automática e as necessidades que ela buscou satisfazer com relação à organização e à recuperação de dados.

Com o incremento dos meios de comunicação, houve uma expansão significativa de informações científicas em diversas áreas do conhecimento, especificamente das áreas de ciências experimentais, a partir da década de 1950, com o auxílio do computador, a fim de atender cada vez mais a demanda dos cientistas. Nesse cenário de ampliação das formas de divulgação do conhecimento, os processos de organização de dados também foram impulsionados, inclusive, a indexação, nomeadamente, as estratégias para automatização desse processo (GIL LEIVA, 2017).

A abordagem da automatização da indexação deve ser realizada a partir de três perspectivas: a) Programas de computador que auxiliam no processo de armazenamento de termos de indexação, uma vez obtidos de modo intelectual (indexação assistida por computador durante o armazenamento); b) Sistemas que analisam documentos automaticamente, mas os termos de indexação propostos são validados e editados – se necessário – por um profissional (Indexação Semiautomática); e c) Programas sem qualquer validação, ou seja, os termos propostos são armazenados diretamente como descritores desse documento (Indexação Automática) (Gil LEIVA, 2008).

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

De acordo com Lapa e Corrêa:

A indexação automática pode ser definida como um conjunto de operações, basicamente matemáticas, linguísticas, de programação, destinadas a selecionar termos como elementos descritivos de um documento pelo processamento de seu conteúdo (LAPA; CORRÊA, 2014).

Na indexação automática, o processo ocorre por meio de algoritmos que buscam padrões estabelecidos para extrações significativas de termos que representem determinado documento sem a mediação do indexador. Segundo Corrêa e Bazílio (2017), as novas propostas de indexação automática possuem vantagens em relação à manual (intelectual):

As novas propostas de indexação, com grandes volumes de documentos digitais envolvem a indexação automática de documentos. A indexação automática apresenta algumas vantagens em relação à indexação intelectual: os descritores são homogêneos ou consistentes em suas regras de representação; e a rapidez na realização da indexação em grandes quantidades de documentos (CORRÊA; BAZÍLIO, 2017).

É importante destacar que a indexação automática, embora traga vantagens significativas, sobretudo com relação à rapidez com que a organização é feita, requer atenção especial nas modificações sofridas pela linguagem natural, de modo a manter a atualização dos termos representativos (GIL LEIVA, 2008). Dentre as abordagens de indexação automática, destacam-se os sistemas OGMA, que realiza a indexação por extração, e o SISA, que aplica a indexação por atribuição.

O sistema de indexação OGMA utiliza a extração por sintagmas nominais, isto é, a representação dos termos é realizada por nomes presentes no documento. Além da extração, o sistema permite cálculo de similaridade entre documentos e análise de textos (MAIA; SOUZA, 2010). Já a abordagem de indexação utilizada pelo sistema SISA realiza a indexação automática por meio de atribuição (GIL LEIVA, 2008), isto é, utiliza-se de vocabulários controlados e tesouros no idioma do texto que se deseja indexar (BADIM; CORREA, 2019).

Quanto à arquitetura, o SISA se estrutura inicialmente com uma etapa de pré-processamento para sinalizar as partes do texto a serem indexadas pelo software, a partir da linguagem de indexação e lista de palavras vazias. Assim, o processamento pode ser realizado a partir dos parâmetros selecionados e, posteriormente, a indexação é obtida (GIL LEIVA, 2008, p. 377, tradução nossa). Assim, é possível optar por duas versões do software SISA, a versão Windows 2.0 e a versão Web 4.0. A seguir, apresentam-se as principais características e parâmetros das duas versões.

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

A versão Windows 2.0 é um aplicativo para instalação monousuária e tem como pré-requisito a instalação do Sistema Operacional Windows. Quanto aos parâmetros de análise, o Quadro 2 expõe as codificações utilizadas para as marcações do texto nessa versão.

Quadro 2: Marcações para o pré-processamento do SISA.

Partes da Estrutura Textual	Marcação início	Marcação fim
Título	#CTI#	#FTI#
Resumo	#CR#	#FR#
Início e fim do texto	#CTE#	#FTE#

Fonte: Elaborado pelo autor.

Destaca-se, no SISA Windows 2.0, a atribuição dos termos pela frequência e recorrência nas partes do título, resumo e texto completo para seleção dos melhores candidatos a palavras indexadas. Assim, verificou-se que o SISA apresenta um modelo de leitura para embasar seus processos de indexação, isso ao delimitar as partes da estrutura textual que serão analisadas pelo software em relação aos aspectos de frequência e recorrência de palavras importantes para indexação.

A respeito da versão Web 4.0, trata-se de sistema online multiusuário que não requer instalação, nem o pré-requisito de instalação do Sistema Operacional Windows. Além disso, trata-se de um sistema multiplataformas, fato que amplia sua utilização devido ao acesso também por *smartphones* e *tablets*. Quanto aos parâmetros de análise, o Quadro 3 expõe as codificações utilizadas para as marcações do texto nessa versão.

Quadro 3: Subdivisões das estruturas textuais para indexação do SISA Web.

Estratégia de exploração da estrutura textual pelo SISA – versão 4.0
Título
Resumo
Palavras-chave
Título das tabelas
Título das figuras
Análise do 1º parágrafo da introdução
Análise de todo o texto da conclusão
Referências

Fonte: Elaborado pelo autor.

Além do maior número de subdivisões para identificação estrutural do documento, como se pode observar no Quadro 3, destaca-se que o usuário pode personalizar seus critérios

de seleção, selecionando todos os parâmetros disponíveis ou aqueles que atenderem a necessidade do conjunto de documentos a serem indexados. Ademais, há mais dois diferenciais importantes na versão Web 4.0: I) todos os dados são armazenados na nuvem, fato que minimiza o risco de perda de dados e II) há um mecanismo para recuperação de dados indexados a partir de campo de busca, no qual é possível procurar em todos os arquivos do SISA ou selecionar por parâmetro de indexação (título, resumo, palavras-chave ou outros descritores).

Dado o exposto, mostra-se oportuno salientar na literatura a aplicação do sistema SISA de textos em português na avaliação automática por atribuição em artigos científicos pelos autores:

- a) Nakamura, Gil Leiva e Fujita (2009): aplicaram o sistema SISA ao domínio de Odontologia na indexação automática por atribuição de artigos científicos, com o uso de Descritores em Ciências da Saúde (DeCS);
- b) Lima e Boccato (2009): aplicaram o sistema em resumos de teses e dissertações na área de Ciência da Informação com o vocabulário controlado do Sistema Integrado de Bibliotecas da Universidade de São Paulo (SIBiUSP);
- c) Souza e Gil Leiva (2016) aplicaram o sistema na área de Fruticultura com o vocabulário controlado em Agricultura (Thesagro).

No entanto, nenhum dos trabalhos relacionados tratam das estratégias de identificação e seleção de termos de indexação utilizadas pelas duas versões do SISA com base no modelo teórico de leitura documentária resultante da observação de estratégias metacognitivas de leitura documentária de indexadores experientes, fator que justifica a relevância deste trabalho de pesquisa.

3 MÉTODO

Em relação ao objetivo da pesquisa, após revisão bibliográfica, adotou-se uma abordagem qualitativa de análise das estratégias de identificação e seleção de termos de indexação das duas versões do SISA com base nas estratégias do Modelo de Leitura Documentária (FUJITA; RUBI, 2006). Para tanto, a análise foi realizada em duas etapas: I) verificação das estratégias que as versões do SISA utilizam, ou seja, exploração de estrutura textual, frequência de palavras, atribuição de palavras de vocabulário externo entre outros;

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

II) verificação e comparação dos termos selecionados nas duas versões do SISA em relação as palavras-chave dos autores.

Os artigos escolhidos para indexação foram selecionados a partir dos seguintes critérios: a) área de assunto compatível com vocabulário controlado a ser incluído no software SISA; b) publicações atuais; c) descritores utilizados na pesquisa na base de dados SCIELO: “odontologia” AND la:("pt") AND year_cluster:("2019"), conforme Quadro 4.

Quadro 4: Total de artigos encontrados em periódicos em julho de 2019, na base de dados SCIELO, com os parâmetros de busca.

Artigos	Quantidade
Total de artigos encontrados	24
Artigos relacionados a técnicas, metodologias e materiais	9
Artigos relacionados ao programa da saúde bucal	15
Total de artigos selecionados	17

Fonte: Elaborado pelo autor.

Da seleção de 24 artigos, conforme o Quadro 4, que envolve a área de Odontologia de periódicos atuais, foram selecionados 17 artigos na base de dados supracitada para aplicação neste trabalho pois, da totalidade dos artigos recuperados na base de dados com o termo de busca “Odontologia”, cinco não eram efetivamente da área e dois estavam duplicados e, portanto, foram excluídos da amostra. Os artigos selecionados para indexação se encontram dispostos no Quadro 5.

Quadro 5: Artigos selecionados para o estudo.

Nº	Artigos selecionados da área de Odontologia
1	CAMARGO, Fernanda Dalmolin de; BATISTA, Aline Krüger; UNFER, Beatriz. Ética e moral: reflexões de dentistas do serviço público. Revista Bioética , v. 27, n. 2, p. 297-303, 2019.
2	LORENZETTI, Camila Cruz et al. Influência de tratamento dentinário com EDTA sobre a resistência de união de sistemas adesivos autocondicionantes. Revista de Odontologia da UNESP , v. 48, p. 0-0, 2019.
3	MORAES, Daniela Abreu de et al. Precarização do trabalho odontológico na saúde suplementar: uma análise bioética. Ciência & Saúde Coletiva , v. 24, p. 705-714, 2019.
4	MUNIZ, Bruna Ariella Aguiar; FONTE, Diene Carvalho Belo; DOS SANTOS, Saulo Cabral. Percepção do portador de HIV/aids sobre o cirurgião-dentista. Revista Bioética , v. 27, n. 2, 2019.
5	MATICK, Anna Carolina Cenci et al. Avaliação da solubilidade e sorção em água de alguns materiais restauradores diretos. Revista de Odontologia da UNESP , v. 48, 2019.
6	PEREIRA, Flávia de Moura et al. Efeitos do laser de diodo de alta potência e da fotobiomodulação em dentes de ratos reimplantados tardiamente. Revista de Odontologia da UNESP , v. 48, 2019.
7	GARCEZ, Ramiro Heleno Mesquita et al. Caracterização de lesões bucomaxilofaciais decorrentes de agressão física: diferenças entre gênero. Ciência & Saúde Coletiva , v. 24, p. 1143-1152, 2019.

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

8	LIMA, Cássio de Almeida et al. Avaliação do processo de trabalho entre equipes de saúde da família de um município de Minas Gerais, Brasil. Trabalho, Educação e Saúde , v. 17, n. 1, 2019.
9	ESSVEIN, Gustavo et al. Atendimento odontológico na primeira infância no Brasil: da política pública à evidência. Revista de Saúde Pública , v. 53, p. 15, 2019.
10	NEVES, Matheus; GIORDANI, Jessye Melgarejo do Amaral; HUGO, Fernando Neves. Atenção primária à saúde bucal no Brasil: Processo de trabalho das equipes de saúde bucal. Ciência & Saúde Coletiva , v. 24, p. 1809-1820, 2019.
11	VALENTE PALMA, Pamella; GONÇALVES LEITE, Isabel Cristina; MARIA GRECO, Rosângela. Associação entre a qualidade de vida relacionada à saúde bucal e a capacidade para o trabalho de técnicos administrativos em educação: um estudo transversal. Cadernos Saúde Coletiva , v. 27, n. 1, 2019.
12	REGINA BORTOLI, Francieli; FRANCISCO KOVALESKI, Douglas; OTÁVIO MORETTI-PIRES, Rodrigo. Medicalização social e bucalidade: a busca pela superação da técnica. Cadernos Saúde Coletiva , v. 27, n. 1, 2019.
13	BARBOSA, Quésia Ferreira; RODRIGUES, Camila Serra; NOVAES, Maria Rita Carvalho Garbi. Integridade científica na educação de profissionais de saúde. Revista Bioética , v. 27, n. 1, 2019.
14	EDUARDO, DA-RÉ; DA SILVA GASQUE, Kellen Cristina; NETO, Rafael Tobias MORETTI. Rely XTM U200 versus Rely XTM ARC: uma comparação da resistência à microtração. Revista de Odontologia da UNESP , v. 48, p. 0-0, 2019.
15	DE ALMEIDA, Leonardo et al. Avaliação do manchamento e da rugosidade superficial de materiais restauradores diretos após diferentes sistemas de polimento: estudo in vitro. Revista de Odontologia da UNESP , v. 48, p. 0-0, 2019.
16	FORTE, Franklin Delano Soares et al. Olhar discente e a formação em Odontologia: interseções possíveis com a Estratégia Saúde da Família. Interface-Comunicação, Saúde, Educação , v. 23, p. e170407, 2019.
17	LEME, Pedro Augusto Thiene et al. A clínica do dentista na Estratégia Saúde da Família: entre a inovação e o conservadorismo. Physis: Revista de Saúde Coletiva , v. 29, p. e290111, 2019.

Fonte: Elaborado pelo autor.

Para realizar as duas etapas da análise, foi realizada a indexação automática por atribuição, por meio das duas versões do sistema SISA. Na sequência, foi aplicada a indexação manual, pautada no Modelo de Leitura Documentária adaptado por Fujita e Rubi (2006).

3 RESULTADOS E DISCUSSÃO

Conforme metodologia, a primeira etapa da análise consistiu na verificação das estratégias que uma e outra utilizam, ou seja, exploração de estrutura textual; frequência e recorrência de palavras e atribuição de palavras de vocabulário controlado. Por isso, conforme disposto no Quadro 6, observa-se que o SISA utiliza a estratégia de exploração da estrutura textual para identificar e selecionar termos frequentes que são relacionados aos locais título e resumo, potencialmente com mais visibilidade em buscas, enquanto que o Modelo de Leitura Documentária não realiza essa correlação dos termos identificados no corpo do texto com sua ocorrência no título e resumo.

Quadro 6: Comparativo de estratégias do SISA e do Modelo de Leitura Documentária.

Estratégia de exploração da estrutura textual pelo SISA Windows 2.0	Estratégia de exploração da estrutura textual pelo SISA Web 4.0	Estratégia de exploração da estrutura textual do Modelo de Leitura Documentária
<ul style="list-style-type: none">• Título;• Resumo;• Início e fim do texto.	<ul style="list-style-type: none">• Título;• Resumo;• Palavras-chave;• Título das tabelas;• Título das figuras;• Análise do 1º parágrafo da introdução;• Análise de todo o texto da conclusão;• Referências.	<ul style="list-style-type: none">• Introdução;• Metodologia;• Resultados e discussão;• Conclusão.

Fonte: Elaborado pelo autor.

Comparando o Modelo de Leitura Documentária proposto por Fujita e Rubi (2006) e o Modelo de Leitura do SISA, ambos utilizam parte da estrutura textual para quantificar e qualificar os termos para indexação. Nesse aspecto, os dois modelos obtêm os termos significativos pela localização das palavras em partes do texto, porém a versão Web 4.0 apresenta oito partes estruturais para identificação, enquanto o Modelo de Leitura Documentária apresenta quatro e a versão Windows 2.0 apenas três.

Destaca-se também que o Modelo de Leitura Documentária utiliza a estratégia de análise conceitual combinada com a exploração da estrutura textual para localização dos conceitos mais significativos, com base na relevância do tema presente no objetivo localizado no item de “Introdução” dos artigos. O Quadro 7 exemplifica a combinação das duas estratégias citadas, análise conceitual (primeira e segunda colunas) e exploração da estrutura textual (terceira coluna) no Modelo de Leitura Documentária, cujo resultado é a identificação de termos (quarta coluna) e a seleção de termos (quinta coluna) com uso do vocabulário controlado para tradução da linguagem natural.

Quadro 7: Modelo de Leitura Documentária aplicado ao artigo 1.

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

Conceitos	Questionamento para identificação de conceitos	Partes da estrutura textual	Identificação de conceitos	Seleção de conceitos (com DeCS)
OBJETO e PARTE(S) DO OBJETO (algo ou alguém que está sob estudo do pesquisador)	O documento possui em seu contexto um objeto sob efeito desta ação?	Introdução (objetivos)	Dentistas; serviço público	Odontólogos
AÇÃO (processo sofrido por algo ou alguém)	O assunto contém uma ação (podendo significar uma operação, um processo etc.)?	Introdução (objetivos)	Prática profissional	Prática profissional
AGENTE (aquele ou algo que realizou a ação)	O documento possui um agente que praticou esta ação?	Introdução (objetivos)	Ética;	Ética
MÉTODOS (métodos utilizados para realização da pesquisa)	Para estudo do objeto ou implementação da ação, o documento cita e/ou descreve modos específicos, por exemplo: instrumentos especiais, técnicas, métodos, materiais e equipamentos?	Metodologia	Análise de conteúdo; Código de Ética Odontológica	
LOCAL OU AMBIÊNCIA (local físico onde foi realizada a pesquisa)	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	Metodologia	Rio Grande do Sul; Sistema Único de Saúde – SUS;	Sistema Único de Saúde
CAUSA E EFEITO Causa (ação+objeto)/Efeito	Considerando que a ação e o objeto identificam uma causa, qual é o efeito desta causa?	Resultados; discussão de resultados; Conclusões	Causa: Prática profissional de dentistas Efeito: humanização; prevenção de saúde; Promoção de saúde	Humanização da Assistência

Fonte: Elaborado pelo autor.

Em relação à segunda parte da análise, que é constituída pela verificação e comparação dos termos selecionados nas duas versões do SISA (Windows 2.0, e Web 4.0), os 17 artigos foram indexados primeiro na versão 2.0 e depois na versão 4.0, como demonstrado no Quadro 7.

Quadro 7: Exemplo com os cinco primeiros artigos do comparativo dos termos indexados, ambos com o vocabulário controlado (DeCS)

Artigos	Palavras-chave dos autores	Modelo de Leitura Documentária	Descritores de SISA Windows 2.0	Descritores de SISA Web 4.0
----------------	-----------------------------------	---------------------------------------	--	------------------------------------

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

1	<ul style="list-style-type: none"> • Ética; • Odontologia; • Setor público. 	<ul style="list-style-type: none"> • Odontólogos; • Prática profissional; • Ética; • Sistema único de saúde; • Humanização da assistência. 	<ul style="list-style-type: none"> • Saúde; • Associações Profissionais; • Ética Odontológica; • Ensino; • Técnicas. 	<ul style="list-style-type: none"> • Odontologia; • Ética odontológica; • Setor público; • Saúde.
2	<ul style="list-style-type: none"> • Dentina; • Adesivo; • EDTA. 	<ul style="list-style-type: none"> • Dentina; • Reabilitação bucal; • Materiais dentários; • Ácido edético; • Smear layer; • Camada de esfregaço. 	<ul style="list-style-type: none"> • Adesivos; • Dentina; • Estatística; • Dente Molar. 	<ul style="list-style-type: none"> • Adesivos; • EDTA; • Sistema adesivo.
3	<ul style="list-style-type: none"> • Saúde suplementar; • Planos de saúde; • Odontologia suplementar; • Bioética; • Precarização do trabalho odontológico. 	<ul style="list-style-type: none"> • Odontólogos; • Trabalho; • Inquéritos e questionários; • Autogestão; • Planos de pré-pagamento em saúde. 	<ul style="list-style-type: none"> • Saúde; • Bioética; • Odontologia; • Associações Profissionais; • Mercado de Trabalho. 	<ul style="list-style-type: none"> • Bioética.
4	<ul style="list-style-type: none"> • HIV; • Síndrome de imunodeficiência adquirida; • Preconceito; • Estigma social; • Odontologia. 	<ul style="list-style-type: none"> • Síndrome da Imunodeficiência adquirida; • Odontólogos; • Percepção; • Inquéritos e questionários; • Preconceito; • Discriminação social. 	<ul style="list-style-type: none"> • Portador; • Saúde; • Síndrome. 	<ul style="list-style-type: none"> • Síndrome de imunodeficiência adquirida; • Doença; • HIV; • Odontologia; • Portador; • Saúde; • Brasil.
5	<ul style="list-style-type: none"> • Solubilidade; • Cimentos de ionômeros de vidro; • Resinas compostas; • Tratamento dentário estaurador sem trauma. 	<ul style="list-style-type: none"> • Solubilidade; • Odontólogos; • Cimentos de ionômeros de vidro. • Resinas compostas; • Materiais. 	<ul style="list-style-type: none"> • Solubilidade; • Materiais restauradores; • Materiais. 	<ul style="list-style-type: none"> • Resinas; • Compostas; • Avaliação.

Fonte: Elaborado pelo autor.

Em relação à segunda parte da análise, foi realizada o experimento utilizando a avaliação quantitativa em relação ao número de palavras-chave dos autores com as indexadas pelo Modelo de Leitura Documentária e as versões Web 4.0 e Windows 2.0 do SISA que visa medir o grau de consistência entre os modelos propostos.

Gil Leiva (2008 e 2017) observa que Hooper, em 1965, introduziu uma fórmula para encontrar consistência entre duas indexações e, posteriormente, Rolling, em 1981, introduziu uma variação dessa fórmula. Estas duas fórmulas têm sido amplamente utilizadas em

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

numerosos experimentos desde então. Uma variante da fórmula de Hooper é a que Gil Leiva vem usando desde 1999 em diferentes trabalhos, como em Narukawa, Gil Leiva e Fujita (2009) e mais recentemente em Gil Leiva (2017), que expressou da seguinte maneira:

$$C_i = \frac{T_{co}}{(A+B) - T_{co}}$$

Onde:

Tco = Número de termos comuns nas duas indexações;

A = Número de termos utilizados na indexação A;

B = Número de termos utilizados na indexação B.

Essa fórmula é aplicada neste trabalho para encontrar os índices de consistência entre as indexações, conforme o Quadro 9.

Quadro 9: Relação de consistência entre as palavras-chave dos autores com a versão Windows do SISA 2.0

Artigos	Palavras chave dos autores	Versão SISA 2.0	Termos comuns	Consistência
1	3	5	1	14%
2	3	4	2	40%
3	5	5	3	43%
4	4	3	1	17%
5	4	3	1	17%
6	4	2	0	0%
7	3	3	0	0%
8	5	5	1	11%
9	5	7	1	9%
10	5	4	1	13%
11	5	4	2	29%
12	3	4	2	40%
13	7	3	0	0%
14	3	2	0	0%
15	4	6	0	0%
16	4	4	1	14%
17	5	4	1	13%
Média	4	4	1	15%

Fonte: Elaborado pelo autor.

Na análise do Quadro 9, a amplitude de palavras-chave dos autores é de 3 a 7 termos, enquanto o mesmo critério, na versão SISA Windows 2.0, tem amplitude de 2 e 7 termos, com média de 4 termos representativos, perfazendo uma média geral de índice de consistência de 15% entre todos os documentos.

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

Quadro 10: Relação de consistência entre as palavras-chave dos autores com o SISA Web 4.0

Artigos	Palavras chave dos autores	SISA Web 4.0	Termos comuns	Consistência
1	3	4	2	40%
2	3	3	2	50%
3	5	1	1	20%
4	4	7	3	38%
5	4	3	2	40%
6	4	4	4	100%
7	3	4	1	17%
8	5	4	1	13%
9	5	6	3	38%
10	5	4	2	29%
11	5	2	1	17%
12	3	3	2	50%
13	7	2	1	13%
14	3	3	0	0%
15	4	4	1	14%
16	4	2	0	0%
17	5	4	2	29%
Média	4	4	2	30%

Fonte: Elaborado pelo autor.

Com relação aos termos propostos pelo SISA Web 4.0, a amplitude foi de 1 a 7 termos representativos, com média de 4 termos por artigo descrito. O número de termos comuns que representa a consistência dobrou em relação ao SISA Windows 2.0, com média geral de 30% entre as palavras-chave dos autores e os descritores atribuídos pelo SISA Web 4.0.

O Quadro 11 apresenta a relação de consistência entre palavras-chave atribuídas pelos autores e os termos atribuídos por meio do Modelo de Leitura Documentária.

Quadro 11: Relação de consistência entre as palavras-chave dos autores com o Modelo de Leitura Documentária

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

Artigos	Palavras chave dos autores	Modelo de Leitura Documentária	Termos comuns	Consistência
1	3	5	1	14%
2	3	6	1	13%
3	5	5	1	11%
4	4	6	2	25%
5	4	5	2	29%
6	4	6	1	11%
7	3	4	1	17%
8	5	6	2	22%
9	5	5	1	11%
10	5	4	0	0%
11	5	8	2	18%
12	3	5	2	33%
13	7	5	1	9%
14	3	5	1	14%
15	4	3	1	17%
16	4	8	1	9%
17	5	2	2	40%
Média	4	5%	1%	17%

Fonte: Elaborado pelo autor.

Observa-se a proximidade entre as médias dos resultados do Modelo de Leitura Documentária e do SISA Windows 2.0; enquanto o primeiro obteve 17% de consistência média, o segundo alcançou 15% de consistência em relação as palavras-chave atribuídas pelos autores. O Quadro 11 traz os resultados médios dos experimentos obtidos.

Quadro 11: Resumo das consistências obtidas.

Experimentos	
Palavras-chave dos autores e modelo de leitura documentária	17%
Palavras-chave dos autores e SISA Windows	15%
Palavras-chave dos autores e SISA Web	30%

Fonte: Elaborado pelo autor.

Diante dos índices apurados, verifica-se a possibilidade de que a consistência entre os termos atribuídos pela indexação com relação às palavras-chave escolhidas pelos autores tenha relação com a estratégia de exploração da estrutura textual, já que o SISA Windows 2.0, o Modelo de Leitura Documentária e o SISA Web 4.0 adotam três, quatro e oito partes estruturais estratégicas para a análise documental. Observa-se, portanto a similaridade e a proporcionalidade entre a quantidade de estruturas verificadas e as consistências médias obtidas.

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

Analisando-se os resultados deste trabalho, conclui-se que a proposta de processos de indexação automática por atribuição, com base no uso do software SISA e o Modelo de Leitura Documentária com o tesouro DeCS, atende de forma satisfatória quanto à qualidade na indexação de artigos científicos em português da área de Ciência da Odontologia.

4 CONSIDERAÇÕES FINAIS

Em meio a extensão de dados disponibilizados diariamente, as estratégias para seleção de informações relevantes para indexação têm ganhado destaque. No cenário de artigos científicos, essa realidade não é diferente, visto que o volume de dados proporcionados pela comunidade científica, ao passo que impulsiona avanços em todas as áreas do saber, também torna necessário o desenvolvimento de estratégias para indexar os dados produzidos. Nesse sentido, a indexação automática e os parâmetros do Modelo de Leitura Documentária são ferramentas fundamentais para seleção de informações significativas para representar os documentos.

Com relação ao corpus do trabalho, composto por 17 artigos selecionados da área de Odontologia na base de dados SCIELO, os resultados da análise dos dois modelos usam partes das estruturas textuais para atribuição dos termos representativos para indexação. Os resultados demonstram a proximidade entre as médias de consistência obtidas pelo Modelo de Leitura Documentária (17%) e o SISA Windows 2.0 (15%), enquanto a versão SISA Web 4.0 teve consistência bastante superior (30%), assim como a similaridade e a proporcionalidade entre a quantidade de estruturas textuais verificadas (quatro, três e oito, respectivamente) e as consistências médias obtidas.

Convém destacar que a indexação automática e as novas metodologias para indexação emergem no cenário atual como estratégias eficientes e essenciais para a representação e a recuperação da informação. Para trabalhos futuros, o método poderia ser testado em um corpus maior de artigos, bem como em outras áreas do conhecimento, de modo a ampliar a validação do comparativo, evidenciando a compreensão da indexação manual e automática e ampliando as observações dos parâmetros comparativos.

Por fim, é necessário continuar a investigação de modo a abranger um *corpus* maior de artigos para confirmar os dados obtidos, mas, principalmente, encontrar a relação entre a indexação do Modelo de Leitura Documentária e a indexação automática do SISA, com a

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

indexação obtida desses artigos por indexadores profissionais que realizam essa tarefa intelectualmente. Nesse sentido, também é necessário verificar as capacidades desses sistemas de indexação em relação à recuperação de informações em uma base de dados. Desse modo, abre-se as portas para novas pesquisas voltadas ao desafio da indexação de encontrar informações significativas em meio a avalanche de dados produzidos atualmente.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: métodos para análise de documentos – determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992. 4p.

BANDIM, M. A. S.; CORRÊA, R. F. Indexação automática por atribuição de artigos científicos em português da área de ciência da informação. **Transinformação**, v. 31, 2019. Disponível em: <http://www.brapci.inf.br/index.php/res/v/117217>. Acesso em: 2 de mar. 2019.

CORRÊA, Renato Fernandes; BAZÍLIO, Luiz Henrique Teixeira. Análise da extração de descritores como sintagmas nominais através do software OGMA. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, v. 22, n. 50, p. 44-58, 2017. Disponível em: <http://www.brapci.inf.br/index.php/res/v/34560>. Acesso em: 1 de jul. 2019.

FUJITA, Mariângela Spotti Lopes; LEIVA, Isidoro Gil. Políticas de indexação na América Latina. **Ibersid**: revista de sistemas de información y documentación, v. 3, p. 155-162, 2009. Disponível em: <https://www.iversid.eu/ojs/index.php/ibersid/article/view/3735>. Acesso em: 1 de jul. 2019.

FUJITA, Mariângela Spotti Lopes; GIL-LEIVA, Isidoro. Avaliação da indexação por meio da recuperação da informação. **Ciência da Informação**, v. 43, n. 1, 2014. Disponível em: <http://www.brapci.inf.br/index.php/res/v/21846>. Acesso em: 5 de jul. 2019.

FUJITA, M. S. L.; NARDI, M. I. A. A leitura em análise documentária. **Transinformação**, v. 10, n. 3, 1998. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/23924>. Acesso em: 31 mar. 2019.

FUJITA, Mariângela Spotti Lopes; RUBI, Milena Polsinelli. Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. **Datagramazero**, Rio de Janeiro, v. 7, n. 3, p. 1-18, 2006. Disponível em: <http://www.brapci.inf.br/index.php/res/v/5926>. Acesso em: 2 mar. 2019.

GIL-LEIVA, I. **La automatización de la indización de documentos**. Gijón (Asturias): Ediciones Trea, 1999. Disponível em: <https://www.trea.es/books/la-automatizacion-de-la-indizacion-de-documentos>. Acesso em: 01 ago. 2019.

GIL-LEIVA, Isidoro. **Manual de indización**: teoría y práctica. Trea, 2008. Disponível em: <https://webs.um.es/isgil/resources/Book2008%20Trea%20Isidoro%20Gil-Leiva.pdf>. Acesso em: 01 ago. 2019.

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

Gil-Leiva, Isidoro. La indización de artículos científicos con el sistema de indización automática SISA comparada con la indización en las bases de datos Agrícola, WoS y SCOPUS., 2017. *In: Congresso ISKO Espanha, 13.*, Coimbra, Portugal, 23-24 Noviembre 2017. **Atas [...]**. Coimbra: Universidade de Coimbra. Centro de Estudos Interdisciplinares do Século XX - CEIS20, 2017. p.509-24. Disponível em: <http://eprints.rclis.org/32645/>. Acesso em: 01 ago. 2019.

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

LAPA, Remi; CORREA, Renato. Indexação automática no âmbito da Ciência da Informação no Brasil. **Informação & Tecnologia**, v. 1, n. 2, p. 59-76, 2014. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/41624>. Acesso em: 04 fev. 2019.

LIMA, Vania Mara Alves; BOCCATO, Vera Regina Casari. O desempenho terminológico dos descritores em Ciência da Informação do Vocabulário Controlado do SIBi/USP nos processos de indexação manual, automática e semi-automática. **Perspectivas em Ciência da Informação**, v. 14, n. 1, p. 131-151, 2009. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/36563>. Acesso em: 01 ago. 2019.

MAIA, Luiz Cláudio; SOUZA, Renato Rocha. Uso de sintagmas nominais na classificação automática de documentos eletrônicos. **Perspectivas em Ciência Informação**, Belo Horizonte, v. 15, p. 154-172, 2010. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/38135>. Acesso em: 04 mar. 2019.

NARUKAWA, Cristina Miyuki; GIL-LEIVA, Isidoro; FUJITA, Mariângela Spotti Lopes. Indexação automatizada de artigos de periódicos científicos: análise da aplicação do software SISA com uso da terminologia DeCS na área de Odontologia. **Informação & Sociedade: Estudos**, p. 99-118, 2009. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/91939>. Acesso em: 01 ago. 2019.

SOUZA, Renato Rocha; GIL-LEIVA, Isidoro. Automatic indexing of scientific texts: A methodological Comparison. *In: International Society for Knowledge Organization Conference*. Rio de Janeiro, Brazil: **Proceedings**, 2016. p. 243-250. Disponível em: <https://doi.org/10.5771/0943-7444-2017-3-139>. Acesso em: 17 mar. 2019.

ROBREDO, Jaime. **Indexação automática de textos: uma abordagem otimizada e simples**. *Ciência da informação*, v. 20, n. 2, 1991. Disponível em: <http://www.brapci.inf.br/index.php/res/v/20479>. Acesso em: 01 mar. 2019.

VIEIRA, Simone Bastos. **Indexação automática e manual: revisão de literatura**. *Ciência da informação*, v. 17, n. 1, 1988. Disponível em: <http://www.brapci.inf.br/index.php/res/v/20414>. Acesso em: 05 mar. 2019.