



**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
MESTRADO EM CIÊNCIA DA INFORMAÇÃO**

MÁRCIO AÉRCIO SILVA BANDIM

**INDEXAÇÃO AUTOMÁTICA POR ATRIBUIÇÃO
DE ARTIGOS CIENTÍFICOS DA ÁREA DE CIÊNCIA DA INFORMAÇÃO**

**RECIFE
2017**

MÁRCIO AÉRCIO SILVA BANDIM

**INDEXAÇÃO AUTOMÁTICA POR ATRIBUIÇÃO
DE ARTIGOS CIENTÍFICOS DA ÁREA DE CIÊNCIA DA INFORMAÇÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco como requisito final para a obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Ciência da Informação.
Linha de Pesquisa: Comunicação e visualização da memória.

Orientador: Prof. Dr. Renato Fernandes Corrêa.

RECIFE
2017

Catálogo na fonte
Bibliotecário Jonas Lucas Vieira, CRB4-1204

B214i Bandim, Márcio Aécio Silva

Indexação automática por atribuição de artigos científicos da área de
Ciência da Informação / Márcio Aécio Silva Bandim. – Recife, 2017.

139 f.: il., fig.

Orientador: Renato Fernandes Corrêa.

Dissertação (Mestrado) – Universidade Federal de Pernambuco, Centro de
Artes e Comunicação. Ciência da Informação, 2017.

Inclui referências.

1. Indexação automática. 2. Indexação automática por atribuição. 3. Vocabulário controlado. 4. Sistemas de indexação automática. 5. Base de dados científicas I. Corrêa, Renato Fernandes (Orientador). II. Título.

020 CDD (22. ed.)

UFPE (CAC 2017-121)



SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO - PPGCI

MARCIO AERCIO SILVA BANDIM

**INDEXAÇÃO AUTOMÁTICA POR ATRIBUIÇÃO
DE ARTIGOS CIENTÍFICOS DA ÁREA DE CIÊNCIA DA INFORMAÇÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da
Universidade Federal de Pernambuco como requisito para obtenção do título de Mestre em
Ciência da Informação.

Aprovada em: 20/02/2017

Banca Examinadora

Prof. Dr. Renato Fernandes Corrêa (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Fábio Mascarenhas
Universidade Federal de Pernambuco (Examinador Interno)

Prof. Dr. Bruno Ávila (Examinador Externo)
Universidade Federal de Pernambuco

Recife, 20 de fevereiro de 2017.

Aos meus pais, Sebastião (*in memoriam*) e Camila, em reconhecimento ao que eles são e sempre serão para mim.

Meus eternos Mestres!!!!

AGRADECIMENTOS

A Deus.

À minha família e, em especial, aos meus filhos, Ellen, Marcia e Igor, pela amizade e apoio.

Aos amigos que sempre estão comigo na presença e na ausência.

Aos professores, colegas e funcionários do PPGCI-UFPE.

À Anna (*in memoriam*) pelo carinho, apoio e companheirismo e que sempre me deu forças para lutar e vencer os obstáculos da vida!

Especialmente aos Profs. Renato Fernandes Corrêa, Nadi Presser, André Fell, Fábio Pinho, Fábio Mascarenhas, Cely Brito, todos do PPGCI/UFPE pelo incentivo, confiança e apoio no dia a dia da caminhada ao Mestrado.

Só sei que nada sei,
e o fato de saber isso, me coloca em
vantagem sobre aqueles que acham
que sabem alguma coisa.

Sócrates
(Filósofo grego).

RESUMO

Investiga o processo de indexação automática por atribuição na representação de documentos no contexto da construção de base de dados científica na área de Ciência da Informação, visando a associação entre palavras-chaves e o conhecimento registrado em artigos científicos. O problema de pesquisa deste trabalho está configurado na análise do processo e da qualidade resultante na representação do assunto mediante a indexação automática por atribuição de artigos científicos na área da Ciência da Informação. Realiza levantamento das linguagens documentárias e tipo de indexação automática utilizadas em bases de dados científicas. Investiga o uso do Tesouro Brasileiro em Ciência da Informação (TBCI) conjuntamente com o Sistema de Indexação Semi-Automática (SISA), na indexação automática por atribuição de publicações científicas da área de Ciência da Informação. Utiliza como metodologia a pesquisa exploratória, com base em revisão da literatura brasileira em Ciência da Informação sobre indexação automática por atribuição e em estudo de caso via experimento. Utiliza o corpus composto por 60 artigos científicos referenciados na tese de doutorado de Souza (2005) para comparar as palavras-chave destes artigos com os termos atribuídos pelo sistema SISA usando o vocabulário controlado Tesouro TBCI. Na etapa pragmática desta pesquisa foi realizada a indexação automática por atribuição do corpus formado por estes 60 artigos científicos, bem como, análise e avaliação intrínseca qualitativa, via índices de consistência e, avaliação extrínseca quantitativa, via índices de precisão, revocação e medida F, dos termos propostos para indexação. Ao se avaliar os índices de consistência, adotando-se o critério de consistência relaxada, foi constatada uma média de 15 % nos índices de consistência, com uma variação de um mínimo de 0% a um máximo de 42 %. Quanto aos índices de revocação, precisão e medida F analisados neste experimento foram obtidos os seguintes resultados: Precisão média de 20 %; Revocação média de 42 % e Medida F média de 25 %. Os resultados da pesquisa permitem concluir sobre a importância do emprego do vocabulário controlado em um sistema indexação automática por atribuição uma vez que o Tesouro tem forte influência nos resultados da indexação automática. Também é possível afirmar que neste tipo de indexação feita pelo sistema SISA a habilitação do arquivo de termos gerais TCITG tem forte influência nos índices de qualidade os quais foram analisados pelos valores médios obtidos pelos indicadores de qualidade ou índices de consistência. Foi constatado ainda, que o uso de palavras-chave como termos de indexação precisa ser melhor estudado e pesquisado no sentido de se criar instrumentos que propiciem aos autores dos artigos científicos melhor definição conceitual destes termos e maior controle terminológico.

Palavras-chave: Indexação automática. Indexação automática por atribuição. Vocabulário Controlado. Sistemas de indexação automática. Base de Dados Científicas.

ABSTRACT

This work investigates the process of automatic indexing by attribution in the representation of documents within the context of scientific database construction in the area of Information Science, aiming to associate keywords to the knowledge recorded in scientific papers. This work's research problem is the analysis of the process and the resulting quality in the representation of the subject through automatic indexing by attribution of scientific papers in the area of Information Science. It appraises documentary language and type of automatic indexing used in scientific databases. Investigates the use of the Brazilian Thesaurus of Information Science (TBCI) along with the Semi-Automatic Indexing System (SISA) in automatic indexing by attribution of scientific publications in the area of Information Science. Its methodology is exploratory research, based on the revision of Brazilian Information Science literature on automatic indexing by attribution and a case study through an experiment. Utilizes the *corpus* made up of 60 scientific papers referenced in Souza's doctorate thesis to compare the keywords used in those papers with the terms attributed by SISA using the TBCI controlled vocabulary. In the pragmatic stage of this research, automatic indexing by attribution of the 60-paper *corpus* was used, as well as intrinsic qualitative analysis and evaluation, through consistency indices, and extrinsic quantitative evaluation, through precision, recall and F-measure indices of the terms proposed for indexing. Evaluation of the consistency indices, adopting the relaxed consistency criteria, showed an average of 15% in the consistency indices, with a variation going from a minimum of 0% to a maximum of 42%. As to recall, precision and F-measure indices analyzed in this experiment, the following results were obtained: average precision of 20%; average recall of 42%; and average F-measure of 25%. The research results allow us to conclude the importance of the use of controlled vocabulary in an automatic indexing by attribution system since the thesaurus has a strong influence in the results of automatic indexing. It is also possible to assert that in this type of indexing by the SISA system, the habilitation of the general terms archive has a strong influence in the quality indices, which were analyzed through the average values obtained by quality indicators as consistency indices. It was also evidenced that the use of keywords as indexing terms needs to be further studied and researched in order to create tools that provide the authors of scientific papers with a better conceptual definition of these terms and better terminology control.

Keywords: Automatic indexing. Automatic indexing by attribution. Controlled vocabulary. Automatic indexing systems. Scientific databases.

LISTA DE QUADROS

Quadro 1. Características das vertentes sobre tesouros.....	23
Quadro 2. Resumo dos Princípios dos cabeçalhos de assuntos.....	24
Quadro 3. Características dos cabeçalhos de assuntos.....	25
Quadro 4. Semelhanças e diferenças entre tesouros e ontologia.....	26
Quadro 5. As etapas da indexação.....	30
Quadro 6. Elementos que devem ser considerados na comparação de indexações.....	41
Quadro 7. Bases de Dados que usam indexação automática.....	51
Quadro 8. Abrangência dos Campos Semânticos no DOCUTES.....	57
Quadro 9. Índices médios de consistências na indexação e, de precisão, revocação e medida F nos cenários I e II.....	66
Quadro 10. Fatores intervenientes no processo de indexação por atribuição.....	67
Quadro 11. Os seis melhores casos de indexação feita pelo SISA com TCITG habilitado. (Cenário I).....	68
Quadro 12. Os seis piores casos de indexação feita pelo SISA com TCITG habilitado. (Cenário I).....	69
Quadro 13. Os sete melhores casos de indexação feita pelo SISA com TCITG não habilitado (Cenário II).....	69
Quadro 14. Os seis piores casos de indexação feita pelo SISA com TCITG não habilitado (Cenário II).....	69
Quadro 15. Análise comparativa dos melhores índices obtidos na indexação de um mesmo artigo nos Cenários I e II.....	70
Quadro 16. Análise comparativa dos piores índices obtidos na indexação de um mesmo artigo nos Cenários I e II.....	70
Quadro 17. Análise do índice de consistência dos melhores casos nos Cenários I e II.....	71
Quadro 18. Análise do índice de consistência dos piores casos nos Cenários I e II.....	72

LISTA DE FIGURAS

Figura 1. Relacionamentos entre termos estabelecidos nos tesauros.....	21
Figura 2. Fluxograma simplificado do processo de indexação automática.....	33
Figura 3. Algoritmo de indexação automática por atribuição com uso de Tesouro.....	34
Figura 4. Fatores que podem influenciar a qualidade da indexação.....	39
Figura 5. Âmbitos que intervêm na consistência da indexação.....	40
Figura 6. Modelo de arquitetura de um sistema de indexação automática com uso de vocabulário controlado (indexação automática por atribuição).....	44
Figura 7. Diagrama de fluxos do algoritmo SISA do módulo II. Modelo de indexação automática com uso de vocabulário controlado.....	49
Figura 8. Tela de configuração dos arquivos do SISA.....	59
Figura 9. Seleção dos arquivos para indexação no SISA.....	60
Figura 10. Tela apresentando um arquivo selecionado para indexação.....	60
Figura 11. Apresentação dos termos de indexação propostos pelo SISA.....	61

LISTA DE SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
BRAPCI	Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação.
CDD	Classificação Decimal de Dewey
CDU	Classificação Decimal Universal
CI	Ciência da Informação
DeCS	Descritores em Ciência da Saúde
IA	Indexação Automática
IBBD	Instituto Brasileiro de Bibliografia e Documentação
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
IBM	<i>International Business Machine (Corporation)</i>
KWIC	<i>Key Words in Context</i>
KWOC	<i>Key Word out of Context</i>
LILACS	Literatura Latino-americana e do Caribe em Ciências da Saúde
MeSH	<i>Medical Subject Headings</i>
OI/OC	Organização da Informação/Organização do Conhecimento
OWL	<i>Web Ontology Language</i>
SISA	<i>Sistema de Indizacion SemiAutomatica</i>
SRI	Sistema de Recuperação da Informação
TBCI	Tesouro Brasileiro de Ciência da Informação
TCI	Tesouro da Ciência da Informação
TG	Termo Geral
TICs	Tecnologias da Informação e Comunicação
TR	Termo relacionado
UNISIST	<i>World Information System for Science and Technology</i>
WEB (Teia)	<i>World Wide WEB</i>
WoS	<i>WEB of Science</i>

SUMÁRIO

1 INTRODUÇÃO	12
2 FUNDAMENTAÇÃO TEÓRICA	16
2.1 Indexação: definição e contexto	16
2.2 Linguagens de indexação	18
2.3 O processo de indexação	27
2.4 Indexação automática	32
2.5 Indexação automática por atribuição	37
2.5.1 Trabalhos, métodos e sistemas de indexação automática por atribuição	42
2.5.2 SISA	47
2.6 Uso da indexação automática em principais bases de dados científicas internacionais e da ciência da informação	50
2.7 Tesouros da área da ciência da informação	54
3 METODOLOGIA DA PESQUISA	57
4 ANÁLISE DOS RESULTADOS	64
4.1 Indexação automática com atribuição de termos gerais	64
4.2 Indexação automática sem atribuição de termos gerais	65
4.3 Discussões dos resultados	65
5 CONSIDERAÇÕES FINAIS	73
REFERÊNCIAS	77
APÊNDICE A - Corpus dos 60 artigos usados no experimento	83
APÊNDICE B - E-mails recebidos sobre uso de indexação usada	88
APÊNDICE C - Trabalhos publicados sobre Indexação automática por atribuição	93
APÊNDICE D - Índices de consistência, precisão, revocação e medida F com o TCITG habilitado	98
APÊNDICE E - Índices de consistência, precisão, revocação e medida F desabilitando o TCITG	100
APÊNDICE F - Termos de indexação atribuídos pelo SISA sem habilitar o TCITG	102
APÊNDICE G - Termos de indexação atribuídos pelo SISA com a habilitação do TCITG	115
APÊNDICE H - Modelo de um artigo científico formatado segundo critérios do SISA	129
ANEXO A - Lista de Descritores do TBCI	130
ANEXO B - Lista de palavras vazias	133
ANEXO C - Arquivo de Termos Gerais1	138

1 INTRODUÇÃO

Fazer Ciência é produzir conhecimento científico. Possuir este conhecimento é essencial para mudar os rumos da história da humanidade. Em todos os casos de descoberta e produção científica, existe a inevitável necessidade da divulgação deste novo saber, deste novo conhecimento que se dá através da Comunicação Científica.

Entre os instrumentos que potencializam a comunicação científica, se destacam as Bases de Dados de Publicações Científicas, que realizam os processos de indexação do conteúdo ou assunto, objetivando o acesso e recuperação da informação.

No entanto, a divulgação do conhecimento científico e sua recuperação através da informação registrada nos mais diversos meios ou suportes físicos - papel ou digital - sempre foi e, ainda é, uma tarefa difícil e complexa. Principalmente levando-se em consideração a forma pela qual o cientista ou pesquisador transmite, divulga ou comunica as informações, já que a eficácia desse processo depende do veículo ou meio empregado, da natureza destas informações e, não menos importante, do público-alvo (MEADOWS, 1999).

Esta questão, a da divulgação e comunicação do conhecimento científico é uma preocupação antiga fazendo parte do dia-a-dia dos profissionais da informação como bibliotecários, arquivistas, catalogadores e documentalistas.

Estes profissionais já estudavam e analisavam formas de organizar a informação disponibilizando o seu acesso e recuperação, criando sistemas de classificação como a Classificação Decimal de Dewey (CDD), a *Colon Classification* de Ranganathan, a Classificação Decimal Universal (CDU) de Otlet e La Fontaine, entre outros.

Entretanto, tal questão, ainda é investigada no sentido de se criar meios que disponibilizem de modo rápido e fácil o acesso a estes conhecimentos e informações.

Segundo Guimarães (2003), ao permitir o acesso à informação é preciso levar em consideração todas as questões que tratam dos processos de produção, coleta, organização e recuperação dessa informação.

Quando se trata da organização e recuperação da informação, a análise automática de textos é estudada e pesquisada desde a década de 1950, sendo também uma área muito importante na Ciência da Informação. Segundo Robredo (2005), para disponibilizar o acesso rápido às bases de dados científicas, o suporte do computador é de suma importância no processamento de dados e informações.

Assim, a organização e recuperação da informação se materializa, também, via indexação, que é um dos processos de análise documentária realizada com a finalidade de determinar, através do conteúdo dos documentos, um conjunto de palavras chave ou assuntos, facilitando sua armazenagem em bases de dados e, posterior recuperação, atendendo deste modo, as necessidades de recuperação da informação (FUJITA; GIL-LEIVA, 2010).

Também há que se considerar neste cenário a evolução das Tecnologias de Informação e Comunicação (TICs) que também disponibilizam instrumentos capazes de propiciar o aumento da produção científica como permitiram automatizar vários processos ligados a esta produção e posterior recuperação das informações.

Neste sentido, a indexação automática permite, segundo Robredo (2005), identificar e selecionar os termos que representam o conteúdo dos documentos sem a intervenção direta do profissional indexador, pois, é a que se configura como tendo mais condições de dar conta não só do volume crescente de produção de informação como também dispõe do suporte de programas já desenvolvidos e prontos para uso. Deste modo, tem-se a extração automática dos termos representativos do conteúdo dos documentos a serem indexados de modo mais rápido e, conseqüentemente, com menos custos uma vez que o fator tempo é consideravelmente racionalizado.

No contexto da indexação com o uso do computador, existem basicamente: a indexação automática, que faz a extração e seleção automática dos termos; a indexação semiautomática que, após a seleção automática dos termos, estes são avaliados por um profissional indexador e, a indexação automática por atribuição, na qual, para cada termo a ser atribuído, é formatado um conjunto de palavras ou expressões que ocorrem com frequência nos documentos (LANCASTER, 2004).

Nas palavras de Lancaster (2004), a indexação automática (IA) é definida como um processo que ocorre quando o computador é utilizado para substituir a indexação manual realizada por um indexador.

No Brasil, os estudos e a aplicação da indexação automática tem início no final dos anos 1960, com a utilização do programa *Keyword In Context* (KWIC) para elaborar os índices das bibliografias especializadas publicados pelo Instituto de Bibliografia e Documentação (IBBD), atual Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) (VIEIRA, 1988).

Aqui, vale lembrar a importância do uso de vocabulário controlado, um Tesouro, por exemplo, na indexação automática por atribuição. O processo de indexação não se encerra em *sim mesmo*, isto é: este processo objetiva, também, dar suporte à construção de Sistemas de

Recuperação da Informação (SRIs). Quando se trata de SRIs especializados, ou seja, aqueles construídos para o acesso a bases de dados científicas, o controle terminológico é muito importante, sendo que este controle é implementado através do uso das linguagens de indexação pós-coordenadas com controle de terminologia.

O tesauro permite o controle terminológico de sinônimos a partir de uma estrutura semântica, através de uma única forma padronizada com remissivas de todas as outras. Também é possível diferenciar homógrafos e reunir ou ligar termos cujos significados apresentam uma relação mais estrita entre si (LANCASTER, 2004).

Portanto, o uso do controle terminológico via vocabulário controlado permite ao indexador realizar e avaliar o processo de indexação, em termos de qualidade, analisando os índices de consistência, revocação, precisão e medida F.

Neste trabalho pesquisa-se a indexação automática por atribuição por descrever o uso do vocabulário controlado e, também, devido ao pouco conhecimento registrado na literatura da Ciência da Informação brasileira sobre este assunto, bem como, sua aplicação na construção de bases de dados científicas.

Assim, o problema desta pesquisa consiste na necessidade de se investigar não só o processo mas principalmente a qualidade resultante na representação do assunto mediante a indexação automática por atribuição de artigos científicos analisando-se os resultados da aplicação do Tesauro Brasileiro de Ciência da Informação (TBCI) no Sistema de Indexação SemiAutomática (SISA).

Em termos mais concisos a problemática a ser tratada nesta pesquisa pode ser expressa nas seguintes questões: Quais são os tipos de indexação automática e instrumentos de controle terminológicos - os vocabulários controlados - utilizados por bases de dados científicas na indexação de assuntos? Quais tesouros da área de Ciência da Informação contem termos em português e podem ser utilizados na composição do processo de indexação automática por atribuição? Como realizar o processo de indexação automática por atribuição utilizando o software SISA para artigos científicos em português da área de Ciência da Informação? No contexto de uma base de dados de artigos científicos em Ciência da Informação, a utilização da indexação automática por atribuição com base no TBCI aplicado ao SISA gera uma indexação de qualidade?

Este trabalho tem como objetivo geral, analisar a indexação automática por atribuição na representação de artigos científicos da área de Ciência da Informação, fazendo uso de descritores definidos em tesauro. Para tanto, será considerado o cenário de uma base de dados

de artigos de periódicos em Ciência da Informação, utilizando o Tesouro Brasileiro em Ciência da Informação e o software denominado SISA.

Assim, tem-se como objetivos específicos:

- a) Identificar e descrever os métodos, as técnicas e instrumentos de indexação automática por atribuição;
- b) Investigar as principais bases de dados que fazem uso de instrumentos de controle terminológico, como o uso de vocabulários controlados e de indexação automática por atribuição;
- c) Elaborar o processo de indexação automática por atribuição utilizando o software SISA e TBCI para artigos científicos da área de Ciência da Informação em português;
- d) Aplicar a indexação automática do SISA com vocabulário controlado TBCI em um corpus de 60 artigos científicos, comparando os termos propostos para indexação pelo SISA com a indexação manual, realizada pelos autores, através dos termos das palavras-chave;
- e) Analisar os resultados obtidos da indexação automática por atribuição feita pelo SISA considerando o cenário do experimento descrito acima, avaliando os índices de consistência, precisão, revocação e medida F.

Neste trabalho, foi formatado um corpus para uso com o software SISA juntamente com o vocabulário controlado Tesouro Brasileiro de Ciência da Informação, em amostras de artigos de periódicos em Ciência da Informação, o que trará como resultados, após efetuadas as análises, informações importantes para o uso da indexação automática por atribuição, na construção de Sistemas de Recuperação da Informação para bases de dados científicas nesta área.

A presente dissertação se encontra organizada da seguinte forma: O capítulo 1 consiste desta introdução; o capítulo 2 discute o referencial teórico e conceitual indispensáveis ao entendimento da pesquisa realizada; a metodologia da pesquisa é descrita; no capítulo 3 onde serão discutidos os índices de consistência em termos de percentuais esperados em uma boa indexação. No capítulo 4 será visto a Análise dos Resultados analisando e avaliando os percentuais dos índices de consistência, os fatores que influenciam os índices de consistência na indexação e os valores obtidos para os índices de precisão, revocação e medida F. No capítulo 5 tem-se as considerações finais sobre indexação automática por atribuição tanto em termos dos resultados obtidos nesta pesquisa como sugerindo a continuação dos estudos.

2 FUNDAMENTAÇÃO TEÓRICA

Serão tratados neste capítulo, as questões que dizem respeito ao referencial teórico e conceitual sobre indexação automática por atribuição com o uso de tesouros.

Em consonância com o exposto acima, na Subseção 2.1 serão discutidas as questões de cenários, contextos e definições do processo de indexação. Depois, na Subseção 2.2, serão vistas as diversas linguagens de indexação, como listas de cabeçalhos de assuntos, sistemas de classificação e tesouros. Na Subseção 2.3 será enfatizado o processo de indexação sob o ponto de vista histórico, conceitual e as atividades que são realizadas no ato de indexar.

A questão de usar a indexação automática ou manual levando em consideração as vantagens e desvantagens, serão discutidas na Subseção 2.4. Já o uso da indexação automática por atribuição com aplicação de vocabulários controlados é detalhado na Subseção 2.5 onde são discutidas também as questões relativas a trabalhos existentes na área, bem com os métodos e sistemas de indexação automática por atribuição. Na Subseção 2.6 será apresentado um panorama das principais bases de dados que fazem uso da indexação automática e vocabulário controlado e, por fim, na subseção 2.7 serão descritos alguns dos principais tesouros da área da Ciência da Informação.

2.1 Indexação: definição e contexto

Na Ciência da Informação discute-se o conceito de Organização e Representação do Conhecimento, que para Navarro e Marco (1995, p. 149):

É a disciplina dedicada ao estudo e desenvolvimento dos fundamentos e técnicas de planejamento, construção, gestão, uso e avaliação de sistemas de descrição, catalogação, ordenação, classificação, indexação, armazenamento, comunicação e recuperação dos documentos criados pelo homem para testemunhar, conservar e transmitir seu saber e seus atos, a partir de seu conteúdo, com a finalidade de garantir sua conversão em informação capaz de gerar novo conhecimento.

Subjacente a este conceito tem-se a necessidade constante de avaliação, análise e atualização do ferramental teórico-metodológico sobre a questão do tratamento e recuperação da informação e a indexação dos conteúdos dos documentos armazenados em bases de dados. Indexação esta que, segundo Pinto (2001):

É um conjunto de atividades que consiste em identificar, nos documentos, os seus traços descritivos (TD's) ou macro proposições e, em seguida, extrair os elementos/descriptores (sintagmas) indicadores do seu conteúdo, visando à sua recuperação posterior.

É preciso lembrar os elementos que definem a qualidade do produto resultante da indexação: a **consistência** que é o grau de coincidência quando da seleção e apresentação dos termos de indexação feito por dois sistemas de indexação ou por dois indexadores; a **especificidade** que é a precisão para detectar os conceitos mais importantes dos documentos; a **exaustividade** que diz respeito aos conceitos caracterizadores do conteúdo dos documentos e, a **correção** que é a ausência de erros óbvios ou a inclusão de termos desnecessários, afirma Leiva (1999). Tudo isto está diretamente associado a qualidade do atendimento das necessidades dos usuários da informação ao efetuar buscas via sistema de recuperação de informações (SRI).

Na indexação é usada uma linguagem própria também chamada de linguagem documentária ou linguagem de indexação, sendo definida, segundo Slype (1983), como um sistema de representação do conteúdo dos documentos e das perguntas, tendo como finalidade a recuperação dos documentos e dotada de estrutura própria, controlada, padronizada e hierarquizada.

Geralmente, as Instituições definem o tipo e método de indexação que devem ser aplicados na concepção de um processo de indexação através da formalização de uma política de indexação, tendo em mente que não existe uma política *certa* ou a *melhor* política.

Deste modo, se faz necessário explicar que uma política de indexação deve, necessariamente, atender aos objetivos e finalidades da instituição que pretende disponibilizar sua base de dados, tanto para o público interno como para o externo. Isto implica em toda uma formalização de processos, procedimentos, instrumentos e também toda uma filosofia profissional que permeia as atividades desta instituição.

Segundo Carneiro (*apud* NARUKAWA, 2011), esta política deve ser entendida como um guia para a tomada de decisões quando da definição do tipo de serviço que será oferecido. Mais ainda, é com base nesta política que serão identificados os usuários e que tipos ou necessidades de informações eles têm ou procuram. Assim, para atender com qualidade estes usuários se faz necessário definir os recursos humanos (capacitação, treinamento constante, etc.), materiais (equipamentos, softwares, suporte de fornecedores, etc.), e financeiros (doações, verbas orçamentárias, etc.) que possibilitem a construção de um SRI, eficaz e eficiente.

Neste cenário é que entra a questão do tempo gasto e do custo envolvidos na atividade de indexar. Ao avaliar tempo e custos envolvidos na indexação é que se leva em consideração a subjetividade do indexador. Apesar dos aspectos formais definidos na política de indexação, a subjetividade influencia sobremaneira a análise do documento atuando diretamente no tempo e no custo de indexação.

Neste ponto, é necessário considerar os argumentos dos defensores da indexação automática como uma alternativa à indexação manual. Nas palavras de Leiva (1999), sabe-se que uma análise do contexto dos indexadores tem mostrado uma carga excessiva de trabalho, não só nas bibliotecas como também em outras instituições de estudos e pesquisas científicas. Esta questão, da carga excessiva, aliada a evolução e disponibilização de recursos de informática (computadores, redes, Web, programas, etc.) têm servido de apoio para a adoção da indexação automática a qual permite o tratamento, a organização e a disponibilização da informação, de modo rápido e de fácil acesso, de qualquer lugar e a qualquer hora.

Portanto, segundo Gonzalez (2004), não se trata de justificar se é necessário ou não automatizar o processo de indexação ou se o trabalho do indexador (na indexação manual) envolve mais recursos financeiros e consome mais tempo. Mas, devem ser analisadas também: a) as variáveis de crescimento exponencial da produção de informações; b) se a indexação automática (IA), responde aos padrões determinados pela política de informações adotada pela Instituição e, c) se a IA atenderá e, em que grau, alguns aspectos de linguística que só podem ser obtidos apenas pela análise intelectual ou humana.

2.2 Linguagens de indexação

As linguagens de indexação são linguagens construídas com o propósito de servir como instrumento de representação temática da informação. Pressupõe-se que, em um processo de análise, conceitos representativos do documento são identificados e selecionados e, em seguida, os conceitos são traduzidos nos termos dessa linguagem com o objetivo de representarem e tornarem-se pontos de acesso entre a informação do documento e aqueles que buscam essa informação segundo palavras de Narukawa(*et al*, 2009).

Na indexação, quando se fala em análise intelectual ou humana, está se referindo aos vários tipos de linguagens de indexação, as quais diferem entre si, não só quanto à forma mas, também, quanto aos objetivos. As linguagens de indexação são: listas de cabeçalhos de assuntos, sistemas de classificação e tesauros (CAMPOS, 1986).

Todos estes tipos de linguagem de indexação fazem uso de descritores que são termos padronizados, definidos por especialistas ou indexadores, que servem para definir assuntos e recuperar a informação. Diferentemente de palavras-chaves que são termos simples ou expressões compostas, criadas pelo próprio autor, com o objetivo de definir os assuntos contidos nos textos da sua produção literária.

Como exemplos, tem-se os descritores definidos na linguagem documental da Medical Subject Headings (MeSH) usados na base de dados da MedLine e, os definidos na linguagem documental Descritores em Ciência da Saúde (DeCS) usados na base de dados Literatura Latino-Americana e do Caribe em Ciências da Saúde (LILACS).

Ainda há que se falar dos vocabulários controlados os quais incluem os cabeçalhos de assuntos e tesouros e que, na verdade, são linguagens de indexação nas quais a terminologia é controlada (LANCASTER, 2002). O controle de vocabulários objetiva tornar mais fácil a representação de assuntos de forma consistente evitando assim a dispersão de informações relacionadas entre si evitando que documentos que tem os mesmos temas ou assuntos apareçam sob indexações diversas (LEIVA, 2008).

O vocabulário controlado é definido por Naves (2004, p. 10-11) como um tipo de linguagem de indexação artificial. Esta linguagem é construída cuidadosamente, via processo minucioso da seleção dos termos a serem usados, levando-se consideração tanto o significado de cada termo como as relações que apresentam, segundo Borges (2009). Como exemplos de vocabulários controlados podem ser citados as Taxonomias, Tesouros, Listas de Cabeçalhos de Assuntos, entre outros.

Como neste trabalho está incluindo o uso do Tesouro Brasileiro de Ciência da Informação (TBCI) aplicado em um processo de indexação automática por atribuição, como linguagem de indexação, é necessário discorrer sobre este tipo de vocabulário controlado.

Segundo Narukawa (2011), o termo Thesaurus - expresso em latim como *thesaurus* - é de origem grega (*thesaurós*) e quer dizer tesouro no sentido de “armazenar” ou “repositório de palavras” (VICKERY, 1960). Sua aplicação como um tipo de linguagem de indexação acontece após a publicação do livro *Thesaurus of English Words and Phrases* por Peter M. Roget, em 1852. Neste tesouro existem duas partes: a primeira, é uma estrutura classificatória de ideias constituída por diversas categorias que são subdividas em tópicos e, na segunda parte, existe um índice alfabético, apresentando a associação entre os cabeçalhos (sob os quais ocorrem as palavras e frases) e os números, que representam as ideias na parte sistemática (CAMPOS, 2001).

A partir de 1951 os computadores começam a ser usados na organização e representação da informação embora que de forma bastante embrionária. Mas, um pouco antes, em 1950, segundo Foskett (1985), Hans Peter Luhn, especialista da IBM, pioneiro na aplicação da análise estatística de vocabulário para executar uma indexação automática, já utilizava o termo *thesaurus* para chamar o seu sistema de palavras, pois este já possuía uma estrutura de referências cruzadas entre palavras ou termos evidenciando noções ou conceitos que ligavam os termos uns aos outros, ou seja, estabelecendo relações entre eles.

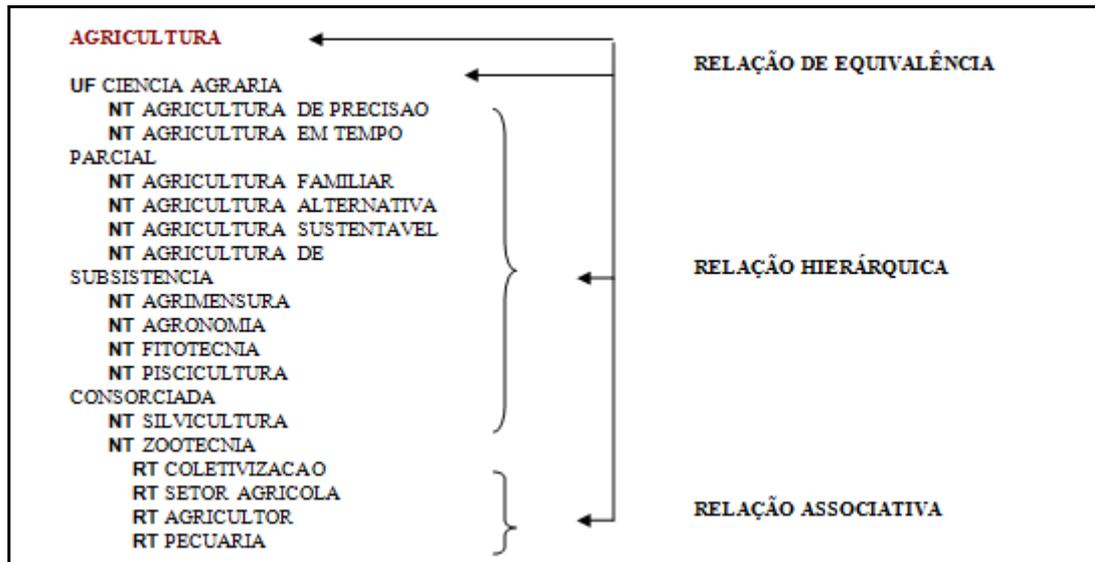
Também, neste período surge o sistema Unitermo, criado por Mortimer Taube e, que tem como característica “a representação do assunto por palavras únicas extraídas do texto dos documentos sem obedecer nenhum critério de controle” (LANCASTER, 1986). Vale lembrar que ainda neste período começa o uso das expressões “recuperação da informação” e “palavras-chaves” sendo que esta última viria a ser chamada de “descritores”. Estas expressões foram também introduzidas por Taube (CURRÁS, 2005).

Segundo Narukawa (2011), existem fatores que interferem nos processos de indexação e que levam à necessidade de estabelecer controle de vocabulários quando da construção de sistemas de recuperação da informação. Cesarino e Pinto (1978) lembram que entre estes fatores tem-se os fatores humanos relacionados às diferenças de cultura, de experiência dos autores e de domínio da terminologia entre indexadores, autores e usuários. Há também, os fatores característicos da linguagem natural associados aos fatores linguísticos de sinonímia, polissemia e sintaxe. Existem ainda, os denominados fatores hierárquicos, em que um conceito implica em outros mais amplos e ou mais restritos.

É preciso lembrar que quando se trata das relações semânticas definidas em um tesouro estas permitem expressar três tipos básicos de relações: equivalência, hierárquicas e associativas. Tais relações quando estruturadas nas formas alfabéticas e ou sistemática e ou gráfica e, ainda, sinalizadas por códigos especiais (USE, UP, TG, TE, TR), possibilitam uma organização adequada facilitando a sua aplicação tanto na representação quanto na recuperação da informação.

Na Figura 1 é reproduzido um exemplo, que mostra as relações semânticas usadas em um tesouro.

Figura 1. Relacionamentos entre termos estabelecidos nos tesauros



Fonte: (NARUKAWA, 2011).

Entre as principais características das relações de equivalências estão: o estabelecimento de associações entre os termos sinônimos para evitar incompatibilidade entre a linguagem do sistema e a do usuário; levar em consideração a coincidência de significado entre um termo antigo e um termo novo; um termo popular e seu correspondente científico, entre outras (GONZÁLEZ, 2004; AITCHISON; GILCHRIST, 1979; NARUKAWA, 2011).

As relações hierárquicas, afirma Currás (2005) e Cintra (*et al*, 2002) possibilita que se reúna os descritores definindo relações entre termos superiores-genéricos e termos subordinados-específicos, estabelecendo deste modo, relações genéricas que indicam que todo conceito que pertence a categoria do conceito específico - a espécie - é parte do conceito amplo - o gênero. Deste modo vale dizer que um conceito específico tem as características de um conceito mais amplo e, ao menos, uma característica distintiva adicional que permite diferenciar conceitos específicos dentro do mesmo nível de abstração.

Quando famílias ou grupos de termos afinssão analisadas no plano horizontal, considerando pontos de vistas diferentes, tem-se as relações associativas (CURRÁS, 2005).

Assim, considerando o exposto acima, vê-se a importância que tem a adequada definição das relações na estrutura dos tesauros, porque a partir dessa determinação de relações é que se tem uma ideia da organização de um domínio a partir do qual são representados novos conhecimentos e relações semânticas.

Outro ponto importante quando da definição de aportes teóricos para determinação das unidades que constituem os tesauros, é que as duas correntes, a americana e a europeia,

não haviam tratado *dos conceitos* que envolvem as unidades dos tesouros e, que só a partir da década de 1970, é quando surgem teorias a respeito, com a introdução da Teoria do Conceito, de Dahlberg.

Esta teoria disponibilizou aportes teórico-metodológicos para a pesquisa e construção dos tesouros conceituais onde se estabelece a identificação dos conceitos, dos termos, dos relacionamentos e da ordenação sistemática. É importante lembrar que, segundo Narukawa (2011, p. 36):

Tanto na vertente norte-americana quanto na europeia, a palavra ou o termo é considerado como unidade que constitui o tesouro, ao passo que, na Teoria do Conceito, considera-se que essa unidade é o conceito. Compreende-se que o conceito é constituído pelo conjunto de atributos que caracterizam um objeto e, nesse caso, o termo designa o conceito, sendo o que permite defini-lo em um sistema de conceitos são esses seus elementos, ou seja, as suas características.

Estas características (conceitos) podem ser consideradas como a matéria-prima do processo de indexação e, para Dahlberg (1978), conceito é a análise e a síntese de enunciados verdadeiros sobre um objeto, e pode ser entendido como uma “*unidade de conhecimento*”, representado por uma forma verbal.

Entende-se como tesouros conceituais aqueles baseados em conceitos, nos quais, são definidos princípios para determinação do termo e do conceito e suas relações. Assim a Teoria do Conceito e a Teoria da Classificação de Ranganathan deram suas contribuições para construção dos tesouros conceituais, determinando as bases para o reconhecimento dos conceitos, dos termos e das relações entre eles, bem como, para sua ordenação sistemática (CAMPOS; GOMES, 2006).

Existe uma grande dificuldade em definir fronteiras entre o controle de vocabulário e a forma de se expressar, relacionada com a representação da informação, em termos de disponibilizar ou comunicar uma dada informação de modo claro e eficiente.

No Quadro 1 estão as características das vertentes norte-americana e europeia, bem como, os atributos dos tesouros conceituais.

Quadro 1. Características das vertentes sobre tesouros

TESAUROS			
Características	Vertente norte-americana	Vertente europeia	Tesouro conceitual
	Unitermo	Thesaurofacet	Tesouro baseado em conceito
	Ruptura aos cabeçalhos de assuntos	Influência da Teoria da Classificação de Ranganathan	Terminologia e Teoria do Conceito de Dahlberg
	Uso de palavra única	Categorização	Categorização
	Ausência de controle de vocabulário	Controle de vocabulário	Definição do conceito
	Abordagem alfabética	Abordagem sistemática	Controle de vocabulário
	Sistema pós-coordenado	Sistema pós-coordenado	Abordagem sistemática
	Evolução pragmática	Estabelecimento de relações semânticas	Sistema pós-coordenado
	Levantamento do domínio	Levantamento do domínio	Estabelecimento de relações semânticas
			Levantamento do domínio

Fonte: (NARUKAWA, 2011).

A análise das características descritas no Quadro 1 permitem concluir que a principal diferença entre a vertente norte-americana e a europeia, é que naquela, verifica-se o uso de palavras simples, a não existência de controle de vocabulário e a estruturação seguindo uma ordem alfabética e, na europeia, há necessidade de controle de vocabulário e de organização sistemática objetivando estabelecer condições para se definirem relações semânticas não só de equivalência, mas, também, hierárquicas e associativas. Isto permite o uso de uma estrutura mais versátil.

No desenvolvimento da Teoria do Conceito, Dahlberg foi buscar aportes teóricos na Terminologia, o que resultou em um material teórico importante para pesquisa e construção dos tesouros terminológicos ou conceituais.

No próximo parágrafo são discutidas duas outras linguagens de indexação: os Cabeçalhos de Assuntos e as Ontologias.

Os cabeçalhos de assuntos surgiram nos Estados Unidos com o objetivo de servir de instrumento para catalogação de assuntos em bibliotecas. Constituem um tipo de linguagem pré coordenada, de estrutura associativa ou combinatória que consiste em uma lista alfabética de palavras ou expressões da linguagem natural, que, normalizadas, são capazes de representar os temas de que trata um documento e por meio dos quais se recuperam os documentos do acervo, afirma Narukawa (2011), ao mencionar Urdician(2004) e Leiva(2008).

Os cabeçalhos de assuntos foram fundamentados nos chamados princípios de Charles Ammi Cutter, publicados em seu livro “Rules for a Dictionary Catalog” em 1876, e que entre estes princípios estavam os de especificidade, o princípio de uso ou de garantia literária, o princípio sindético e o princípio de entrada direta. Os quadros 2 e 3 mostram o resumo destes princípios e as características dos cabeçalhos de assuntos.

Quadro 2. Resumo dos Princípios dos cabeçalhos de assuntos

PRINCÍPIOS DOS CABEÇALHOS DE ASSUNTOS	
Princípio de especificidade	Sugere que a entrada de um assunto deve ser realizada por um termo mais específico, em detrimento de um mais geral.
Princípio de uso	Indica que os assuntos serão definidos em função de como serão buscados por usuários daquela biblioteca.
Princípio sindético	Indica que deve ser elaborada uma rede bem construída de referências cruzadas para ajudar ou mesmo para superar o problema de aproximações absurdas de assunto e a separação de assuntos relacionados.
Princípio da entrada direta	Indica que, ao utilizar termos compostos, ou seja, termos formados por mais de uma unidade lexical, estes devem ser indicados na forma em que se apresentam na linguagem natural, sem a inversão das palavras.

Fonte: (NARUKAWA, 2011).

Quadro 3. Características dos cabeçalhos de assuntos

CABEÇALHOS DE ASSUNTOS
a) Linguagens estruturadas e pré-coordenadas que, de certa forma, apresentam limitações na pesquisa
b) Os termos do vocabulário controlado são selecionados de um dicionário existente, o que o caracteriza como sistema fechado
c) Os cabeçalhos de assuntos exercem função prescritiva
d) Linguagens não hierárquicas
e) São enumerativos, oferecendo poucas possibilidades de síntese
f) Arranjo alfabético
g) Linearidade, aplicável apenas a pesquisas unidimensionais.
h) Pouca sistemática na elaboração de cabeçalhos de assuntos e na elaboração de referências cruzadas

Fonte: (NARUKAWA, 2011).

Com o surgimento e evolução das Tecnologias de Informação e Comunicação (TICs) aconteceram grandes avanços em todos os campos do conhecimento humano - social, econômico e científico e, na Ciência da Informação não foi diferente, principalmente com a entrada em cena da Web.

Surgindo em 2001, o projeto Web Semântica do *World Wide Web Consortium* (W3C) é considerada uma extensão da Web. Segundo Narukawa (2001) a Web Semântica atribuiu à informação um sentido bem definido ao permitir que computadores e pessoas trabalhem em cooperação, ou seja, que a informação possa ser compreendida tanto por humanos como por máquinas.

Por outro lado, antes de 2001, no final da década de 1980 surge, também, uma nova forma ou instrumento de representação da informação, entra em cena as ontologias.

Segundo Ramalho (2010), ontologia, na Ciência da Informação, é um sistema capaz de representar o conhecimento, permitindo, não só, descrever formalmente, as características e os relacionamentos de um específico modelo conceitual, como também a realização de inferências automáticas nos processos de organização e recuperação da informação. Mais ainda, as ontologias, ao contrário das descrições textuais, disponibilizam a sua capacidade de representar uma visão de mundo com o uso de relações semânticas, indo além da representação dos aspectos descritivos e temáticos dos documentos, contribuindo, desta maneira, para contextualização das informações via representação dos próprios domínios.

Narukawa (2011, p. 43) afirma que as ontologias são constituídas por:

- 1) Classes e subclasses, que agrupam um conjunto de elementos de acordo com suas similaridades;
- 2) Propriedades descritivas, em que as características das classes são descritas;
- 3) Propriedades relacionais, que tratam dos relacionamentos entre classes de uma mesma hierarquia ou não, descrevendo os tipos de relações existentes;
- 4) Regras e axiomas, que são enunciados lógicos que impõem condições, possibilitando inferências automáticas;
- 5) Instâncias, que indicam os valores das classes e subclasses e;
- 6) Valores, que atribuem valores concretos às propriedades descritivas, indicando os formatos e os tipos de valores aceitos em cada classe.

Como, mais uma vez, tem-se que um dos objetivos deste trabalho é o estudo de vocabulário controlado - tesouro - como linguagem de indexação usada na indexação automática por atribuição de palavras chaves e ou descritores, o quadro 4 mostra algumas semelhanças e diferenças entre tesouros e ontologias.

Quadro 4. Semelhanças e diferenças entre tesouros e ontologia

	Tesouros	Ontologias
Objetivo	Representar e buscar informação	Organizar, explorar, compartilhar e reutilizar informação
Origem	Década de 1950	Década de 1980
Cobertura	Restrita a um campo do saber	Restrita a um âmbito do saber ou setor (econômico, sanitário, educativo, de mercado de trabalho, etc.)
Entorno	Analógico e digital	Digital
Fontes	Autorizadas (literatura científica e linguagem dos usuários)	Autorizadas (literatura científica, dados procedimentais, organogramas, causas-efeitos, sintomas-tratamentos, dados estatísticos, etc.)
Linguagem	Linguagem natural e linguagem controlada (terminologia consensual e normalizada)	Linguagem natural, linguagem controlada e linguagem formal
Estrutura	Sistemática ou macrotesauro, hierárquica, alfabética, índice (Kwic ou Kwoc)	Taxonomia, tabela com descritores, relações, atributos, valores, axiomas
Uso de taxonomias	Não	Sim
Custo de elaboração	Elevado	Muito elevado
Tipos de relações	Hierárquicas, associativas e de equivalência	Hierárquicas, associativas, de equivalência e qualquer outro tipo (temporais, familiares, causas-efeitos, sintomas-tratamento, etc.)
Inferências	Não	Sim
Definições	Contêm principalmente notas de como empregar um descritor na atividade de indexação e recuperação	Contêm definições universais e consensuais de cada um dos conceitos incluídos na ontologia
Axiomas	Não	Sim. Os axiomas permitem realizar inferências

continua

continuação

Reutilização	Toda ou parte da terminologia de um tesouro pode ser integrada em outro mais geral (por exemplo, um tesouro de urbanismo em outro de administração pública); ou também em outro mais específico (por exemplo, um tesouro sobre patrimônio histórico em outro de arqueologia). Custo elevado	As supraontologias (especificações formais do universo) podem ser reutilizadas nas ontologias de âmbito (especificações formais de um âmbito concreto); por exemplo, uma ontologia de medicina pode utilizar parte de uma linguística, como WordNet; ou, também, parte de uma ontologia de economia em outra de comércio eletrônico. Custo médio
Normas	ISO 25964 ANSI/NISO Z39.19 BS8723	W3C

Fonte: (SOLER MONREAL; GIL LEIVA, 2010, p. 374).

Portanto, as ontologias podem ser aplicadas aos sistemas de indexação como estruturas de apoio à organização, representação e recuperação da informação ajudando na contextualização de informações.

Nas seções a seguir - 2.3 a 2.7 - serão abordados os tipos, características e métodos de indexação, bem como, seu uso em bases de dados científicas.

2.3 O processo de indexação

Segundo Leiva e Muñoz (1996) a origem da indexação está nas tarefas realizadas pelos antigos escribas da Mesopotâmia. Naquela época, começaram a existir salas destinadas a execução de cópia das tábuas de argila, a elaboração de etiquetas e a conservação das placas. Os textos eram armazenados em prateleiras de madeira, colocados em nichos nas paredes ou eram dispostos em caixas de madeira. Para saber o que continham, colocavam uma pequena etiqueta anexada na lateral, onde escreviam o conteúdo dos documentos. Nessas tarefas rudimentares, há os primeiros passos do que hoje se conhece como, indexação (LEIVA; MUÑOZ, 1996, p. 53).

Com o passar dos tempos e, a partir do início do século XX, o acesso ao conhecimento científico já era preocupação de bibliotecários e documentalistas que criaram meios ou formas de organizar a informação como, por exemplo, o sistema de “Classificação Decimal Universal” (CDU) de Paul Otlet e Henry La Fontaine; o *Colon Classification* de Ranganathan e as Listas de Cabeçalhos de Assuntos.

Todos eles tendo como objetivo principal organizar e facilitar o acesso ao conhecimento produzido e disponibilizado por cientistas e estudiosos, via indexação.

Durante as últimas décadas, autores, pesquisadores e estudiosos produziram diversas definições e conceituações sobre o processo de indexação ou o ato de indexar. Segundo Lapa e Correa(2014, p. 15):

A indexação é um processo de tratamento temático essencial, pois consiste no ato de identificar e descrever um documento de acordo com o seu assunto, e cujo principal objetivo é orientar o usuário sobre esse conteúdo intelectual, permitindo, dessa forma, a sua recuperação de forma ágil e eficiente.

A indexação é definida como a forma de caracterizar os conceitos que estão presentes na escrita de certo tipo de documento, de maneira a permitir a sua recuperação. O propósito principal da indexação é o de representar tematicamente esses documentos, através de termos de indexação, organizados em índices, que podem também ser denominados descritores, palavras-chave ou metadados (NASCIMENTO, 2015).

O Sistema Mundial de Informação Científica (UNISIST) responsável pela conceituação e definição dos princípios gerais de indexação, contidos em um documento elaborado por esta entidade é, na verdade, uma das primeiras tentativas de normalização do processo de indexação. Assim, pela conceituação do UNISIST (1981, p. 83), a indexação é: “A atividade que descreve e identifica o conteúdo de um documento, através de termos. Os conceitos dos documentos podem ser representados por termos selecionados através da linguagem natural ou por símbolos”.

Ainda em termos de definição, Borges (2009), conceitua o processo de indexação como a realização de uma atividade que consiste em selecionar e definir termos ou expressões que serão usados para descrever (descritores) o conteúdo de um documento. Na seleção e definição destes termos ou descritores deve-se considerar também os possíveis usuários que necessitam acessar as informações contidas nestes documentos realizando o acesso e a busca via Sistemas de Recuperação da Informação.

Brito (1992), concorda com Navarro (1998), ao afirmar que a indexação, tal qual nós a vemos, é uma tradução lexical das unidades da língua, ou ainda uma tradução sintática, quando se trata de exprimir as relações entre as diferentes partes do discurso.

Passando das **definições** de indexação para o **processo** de indexação, pode-se afirmar que, pelos Princípios de indexação do UNISIST, de 1981, indexação é um processo que se constitui basicamente de dois estágios: no primeiro estágio são definidos os assuntos tratados no documento e, no segundo estágio, o da tradução, os conceitos são expressos em termos de uma linguagem de indexação.

Analisando o ponto de vista de vários autores sobre o processo de indexação, Fujita (2013) afirma que basicamente existem dois momentos ou etapas no ato de indexar. Um primeiro momento na análise do assunto, quando se identifica e se faz a seleção de conceitos via associação à conceitos universais, como ação, objeto que sofreu a ação e o agente que praticou a ação. No segundo momento, quando os termos são traduzidos segundo os termos de uma linguagem documental e, compatibilizando ou adequando os termos indexados com os termos que serão úteis ao atendimento das necessidades informacionais dos usuários.

Para Leiva (1997), o processo de indexação é composto de duas etapas:

1. Análise dos documentos e das perguntas para a seleção dos conceitos explícitos ou implícitos;
2. Conversão dos conceitos selecionados em linguagem natural para uma linguagem documental, com o uso de *ferramentas* como os vocabulários controlados.

Na literatura da Ciência da Informação, no quesito Organização e Representação da Informação, existem vários autores que indicam várias etapas no ato de indexar - cinco, seis etapas - no entanto é na análise de assunto que se verifica a etapa mais importante porque se identifica e se dá a seleção dos conceitos que representam a essência de um documento. (FUJITA, 2003). No Quadro 5 consta as etapas, básicas, de uma indexação.

Assim, no processo de indexação, quando envolve 2, 3 ou 4 etapas e, levando em consideração a visão de vários pesquisadores sobre o assunto, este processo de representação é resumido por Mai (2000, p. 269):

O procedimento em duas etapas consiste de uma etapa em que o assunto é determinado e outra etapa em que o assunto é traduzido e expresso na linguagem de indexação. O procedimento de três etapas insere mais uma etapa na qual o assunto é formulado em linguagem natural. No procedimento de quatro etapas, a tradução do assunto para a linguagem de indexação consiste de duas etapas. O indexador primeiro traduz o assunto de seu vocabulário no vocabulário utilizado na linguagem de indexação. Em seguida, o indexador constrói a entrada de assunto na linguagem de indexação na forma de termos de índice, em um código de classificação ou em um cabeçalho de assunto(Tradução nossa).

Aqui vale registrar que existe uma Norma chamada de “Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação”, publicada pela Associação Brasileira de Normas Técnicas - ABNT, em 1992, número 12676, estabelecendo normas sobre metodologia de análise e determinação de assuntos de documentos. De acordo com esta norma o processo de indexar “é identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos e que constituem

uma linguagem de indexação” (ABNT, 1992, p.2). Ainda, segundo a norma citada acima, a indexação se constitui de três estágios:

- a) Exame do documento e estabelecimento do assunto de seu documento;
- b) Identificação dos conceitos presentes no assunto; e
- c) Tradução desses conceitos nos termos de uma linguagem de indexação.

Quadro 5. As etapas da indexação

	ETAPAS	AUTOR/ES
Duas etapas	a) Determinação do assunto; b) Tradução dos conceitos nos termos da linguagem de indexação.	Unisist (1981)
	a) Reconhecimento e extração dos conceitos dos conceitos informativos; b) Tradução desses conceitos na linguagem documental.	Chaumier (1988)
	a) Análise dos conceitos e das perguntas para a seleção dos conceitos explícitos ou implícitos; b) Armazenamento das palavras chave tal como estão ou sua normalização por meio de um vocabulário controlado.	Gil Leiva (1997)
	a) Análise conceitual; b) Tradução.	Lancaster (2004)
	a) Análise de assunto; b) Tradução.	Fujita (2013)
Três etapas	a) Exame do documento e estabelecimento do assunto de seu conteúdo; b) Identificação dos conceitos presentes no assunto; c) Tradução desses conceitos nos termos de uma linguagem de indexação.	Norma 12675 (ABNT, 1972)
	a) Análise conceitual do conteúdo do documento; b) Expressão dessa análise por meio de códigos, palavras ou frases representativas do assunto; c) Tradução das descrições dos assuntos para a linguagem de indexação.	Robredo (2005)
	a) Análise do documento; b) Descrição do assunto; c) Entrada de assunto.	Mai (2001)
Quatro etapas	a) Contato com o documento; b) Identificação dos conceitos explícitos e implícitos do documento; c) Tradução dos conceitos expressados em linguagem natural por descritores; d) Estabelecimento de ligações sintáticas entre os descritores.	Van Slype (1977)
Cinco ou mais etapas	a) Lembrar os objetivos da operação, se necessário; b) Tomar conhecimento prévio do documento; c) Determinar o assunto principal do documento; d) Identificar os elementos do conteúdo que devem ser descritos e extrair os termos correspondentes; e) Verificar a pertinência dos termos selecionados; f) Traduzir os termos da linguagem natural nos termos correspondentes da linguagem documental, se for o caso; g) Verificar a pertinência da descrição; e h) Formalizar a descrição se o sistema prevê regras especiais de apresentação ou de escrita.	Guinchat e Menou (1994, p.177)
	a) Registro dos dados bibliográficos; b) Análise do conteúdo dos documentos a partir do título, resumo e texto completo; c) Determinação do assunto; d) Conversão dos conceitos extraídos em linguagem de indexação; e) Reexaminar a indexação.	Cleveland e Cleveland (1990 p. 104)

Fonte: (LAPA, 2014).

Deste modo, a indexação se constitui no principal meio de representação do documento objetivando a sua recuperação ou, em outras palavras, a recuperação da

informação (RI) está intimamente associada ao processo de indexação, uma vez que as linguagens de indexação compõem os SRIs. Portanto, a indexação é o componente mais importante para a eficácia de um SRI, pois, segundo Guinchat e Menou (1994, p. 175):

É a operação central do sistema para a armazenagem e pesquisa das informações. Situa-se, na maioria das vezes, no meio da cadeia documental, no momento da entrada dos documentos no subsistema de armazenamento ou no momento da pesquisa.

A indexação é um dos processos de OI/OC mais antigo e de maior subjetividade uma vez que o Indexador além de fazer uma análise do documento interpretando seu conteúdo tem que, ao mesmo tempo, conhecer os usuários ou comunidade de usuários que poderão querer acessar estes documentos. Daí o indexador faz inferências sobre as possíveis necessidades informacionais destes usuários procurando identificar e traduzir os conceitos contidos não só no documento como um todo, mas, também, em cada parágrafo e subparágrafo no sentido de identificar possíveis quebras de concatenação de ideias e sequências lógicas dos assuntos abordados. Também a indexação está associada de modo direto com a descrição (descritores) do objeto informacional e que constituem o que se chama de registro bibliográfico.

Enfim, de modo resumido, o processo de indexação é, basicamente, composto por duas etapas: a análise de assunto e a tradução desta análise quando então se “seleciona” os descritores que melhor representam o assunto analisado.

A análise de assunto é considerada a etapa mais importante do processo de indexação, segundo Cesarino (*apud* BORGES, 2009). A análise de assunto é composta por três estágios: 1 - Compreensão do conteúdo do documento como um todo, 2 - Identificação dos conceitos que representam esse conteúdo e 3 - Seleção dos conceitos válidos para recuperação.

Na compreensão do documento como um todo, a leitura integral do texto é “ideal” embora seja impraticável, assim “delimita-se” um roteiro das partes entendidas como mais importantes para que o indexador possa se nortear durante a leitura que são, segundo UNISIST (1981); FUJITA (2003):

- Título;
- Introdução e as primeiras frases de capítulos e parágrafos;
- Ilustrações;
- Tabelas;
- Diagrama e suas explicações;
- Conclusão;

- Palavras ou grupos de palavras sublinhadas ou impressas com tipo diferente.

Na análise do assunto o indexador faz a tradução, quando então procura representar o assunto do documento por meio da composição de um conjunto de termos, os quais, serão usados para indexação e, esta tradução é necessária mesmo em casos nos quais não houver o estabelecimento de regras formais. Estas regras podem ser definidas em função dos interesses da instituição ou do instrumento de controle terminológico. Por sua vez, este controle pode ser feito a partir do uso de um vocabulário controlado, sendo que, na maioria das vezes, essa tarefa é feita de forma intuitiva.

2.4 Indexação automática

A partir de meados do século XX com o surgimento de novas tecnologias aliado ao crescimento da produção científica, começam a ser discutidas questões ligadas a automação dos processos de tratamento dos documentos visando não só facilitar as tarefas de pesquisa/procura e acesso a base de dados, mas sobretudo, obter resultados mais eficazes nestes processos (NARUKAWA *et al*, 2009). Deste modo, surgem as primeiras pesquisas acerca da indexação automática.

Neste período, Stevens (1965), citado por Leiva (1997), já definia a indexação automática como a utilização de *ordenadores* (informática, computadores) para extrair ou assinalar (selecionar) termos de indexação sem interferência humana. Mais tarde, Vieira (*apud* SILVA; FUJITA, 2004, p. 145): “entende que é a operação que identifica palavras ou expressões significativas nos documentos para descrever seu conteúdo de forma condensada por meio de programas de computador”.

É interessante registrar que, mesmo existindo várias opiniões de pesquisadores e autores defendendo o uso da indexação manual, é preciso reconhecer que a indexação automática atende as necessidades de comunicação, disponibilização e acesso rápido a produção científica via uso do computador no processamento de dados e informações (ROBREDO, 2005).

Ainda, segundo Robredo (2005), a indexação automática é, “qualquer procedimento que permita identificar e selecionar os termos que representam o conteúdo dos documentos, sem a intervenção direta do indexador” ou, como afirma Lancaster (2004), é todo processo que utiliza o computador para substituir a indexação manual.

Na verdade, o processo de indexação automática guarda semelhanças com o processo de indexação manual ou, em outras palavras, é similar ao processo de leitura e memorização

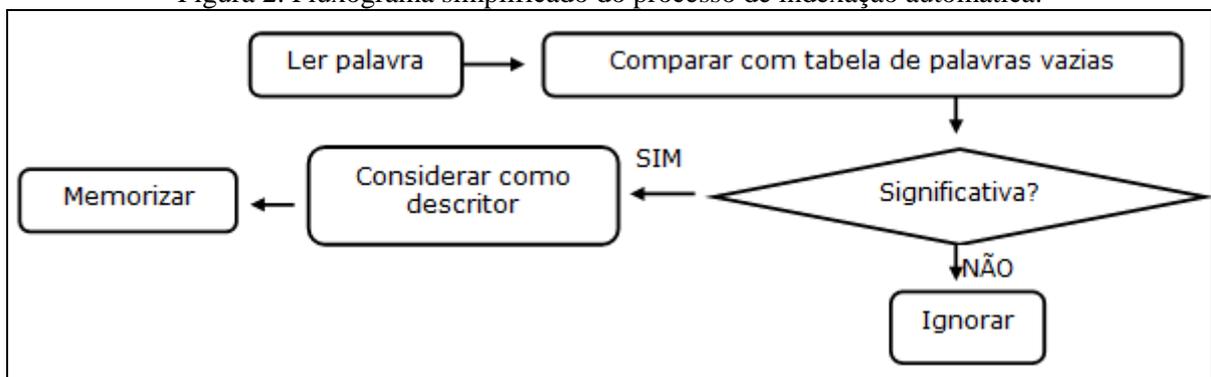
humano uma vez que está baseado na comparação de cada palavra do texto com uma lista de palavras vazias de significado segundo Robredo (*apud* BORGES, 2009). Esta lista de palavras é previamente definida, estabelecida, e o resultado desta comparação, leva, por eliminação, a considerar que as palavras restantes do texto são significativas, isto é, são usadas para representar o texto via indexação.

Nas palavras de Lapa (2014), ao mencionar Gomes (1989) e Santos e Ribeiro (2003), a indexação automática de documentos é compreendida como um processo onde o computador efetua a extração de palavras, expressões ou radicais de palavras utilizadas para representar o conteúdo do texto como um todo, sem a intervenção do documentalista. Para Araújo Júnior (2007), a indexação automática é qualquer procedimento que permita identificar e selecionar os termos que representam o assunto dos documentos sem a intervenção direta do homem.

O embasamento da indexação automática consiste, segundo Guinchat e Menou (1994, p. 182), “em fazer o computador reconhecer palavras que aparecem no título, no resumo do documento, ou no seu próprio texto. Os termos reconhecidos são incorporados em um arquivo de pesquisa e servem para recuperar o documento”.

A figura 2 representa o fluxo do processo básico de indexação automática de documentos, afirma Robredo (*apud* LAPA, 2014). Nesta figura é possível ver que a indexação automática está baseada na comparação de cada termo ou palavra do texto com uma *stoplist* (lista de palavras vazias), anteriormente estabelecida, e que após eliminação destas palavras vazias, as palavras restantes são consideradas como termos ou palavras significativas.

Figura 2. Fluxograma simplificado do processo de indexação automática.



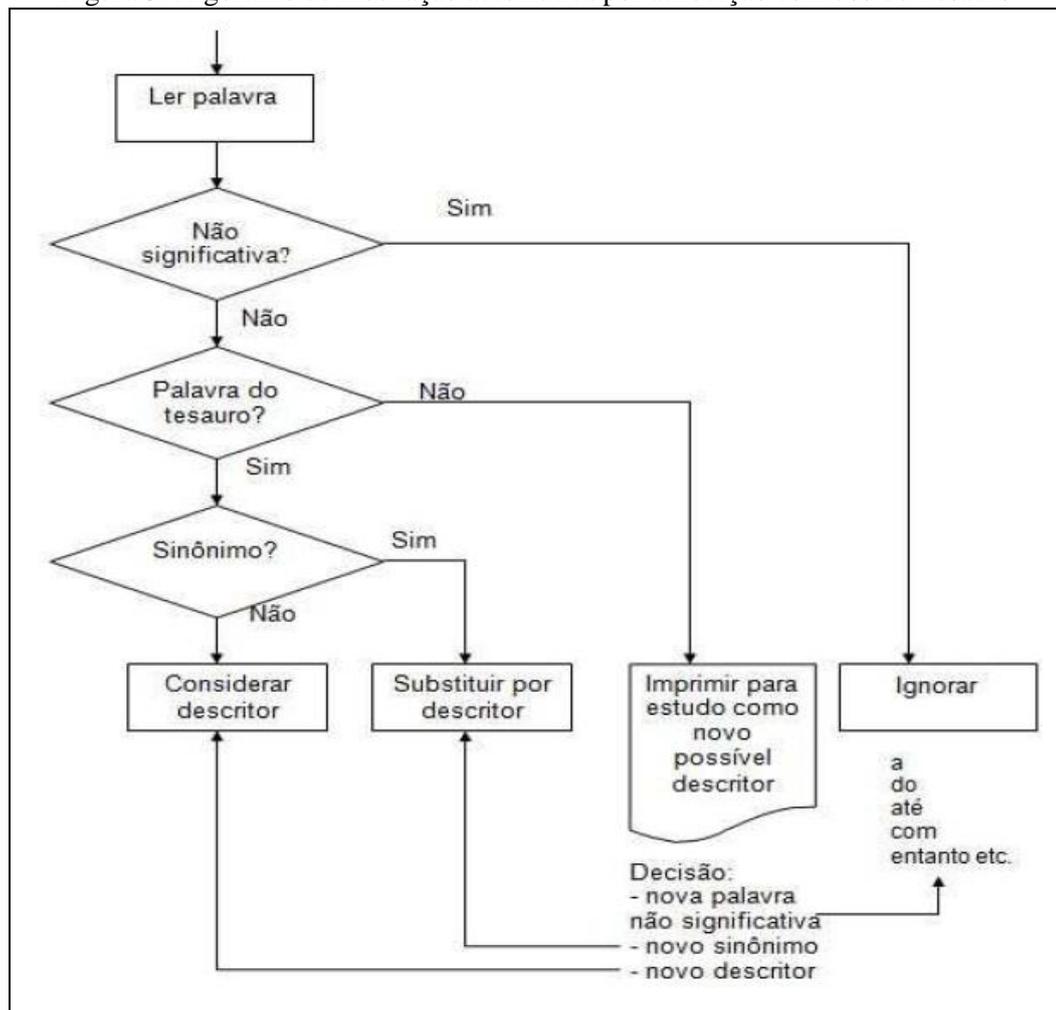
Fonte: (LAPA, 2014).

Ainda, segundo Robredo (2005), na indexação automática, vocabulários controlados também podem ser usados diminuindo assim as diferenças entre a indexação automática e a

manual e, afirma, que a aplicação desses dois tipos de indexação, em processos paralelos, é a proposta mais indicada para os casos de descrição temática e busca da informação.

Na figura 3 é mostrado um algoritmo de indexação automática com o uso de um vocabulário controlado, no caso, um tesouro.

Figura 3. Algoritmo de indexação automática por atribuição com uso de Tesouro



Fonte: (LAPA, 2014).

Aqui cabe uma explicação resumida, sobre o uso do computador para realizar a atividade de indexação. Existem basicamente três conceitos: indexação assistida por computador, indexação semiautomática e indexação automática (LEIVA, 1999; GONZÁLEZ, 2004).

Antes de descrever cada método é necessário fazer uma observação importante: na revisão da literatura sobre indexação automática, existe uma grande variedade terminológica para designar conceitos semelhantes. Expressões como: “indexação automatizada assistida”; “indexação automatizada”; “suporte automatizado para indexação”;

indexação auxiliada por computador”; indexação computadorizada”, entre muitas outras, sendo que “indexação automática” é sem dúvida a expressão mais utilizada. Todas estas expressões são importantes porque fazem referência a um processo de indexação *não* inteiramente manual, podem ser resumidas, segundo Leiva (*apud* LAPA; CORRÊA, 2014, p. 258), em apenas três conceitos diferentes:

1. Programas que dão suporte ao processo de armazenamento aos termos de indexação os quais são obtidos via processo intelectual. Estes sistemas proporcionam por meio de telas de ajuda, notas explicativas sobre a utilização de um termo, seus termos relacionados, permitem também a alocação de termos sem necessidade de digitá-los, a inclusão de termos e a consulta online de documentos que foram indexados para verificar outros itens de indexação. Este tipo de indexação é indexação assistida por computador pois usa o computador para armazenar os termos selecionados.
2. Sistema que analisam o documento de modo automático e os termos de indexação selecionados são analisados e validados por um profissional - Indexação semiautomática.
3. Programas que não necessitam de validação dos termos selecionados, isto é: os termos selecionados e propostos pelo programa, são armazenados diretamente como descritores do documento - Indexação automática.

Segundo Borges (*apud* LAPA, 2014), indexação assistida por computador é considerada inicialmente um modelo de extração com características estatísticas e probabilísticas e sua origem coincide com as tentativas iniciais de junção da informática e da estatística com a área de documentação.

Já a indexação semiautomática está relacionada ao processo em que um sistema computacional realiza a atividade de análise do conteúdo do documento e, posteriormente, um indexador humano avalia os termos para indexação propostos pelo sistema, decidindo quais termos serão efetivamente usados para representar o documento.

No processo de indexação automática ocorre a atividade de análise do conteúdo do documento por um sistema computacional sem que haja uma avaliação posterior. Isto é, os termos de indexação são definidos apenas pela análise realizada pelo sistema (LEIVA, 1999).

Como já dito acima os primeiros processos de indexação automática foram desenvolvidos, basicamente, via métodos estatísticos e probabilísticos e, só a partir da década de 1980, começaram a ser incorporados nestes processos os métodos linguísticos.

Vale lembrar que os métodos estatísticos foram desenvolvidos aplicando-se o princípio do “mínimo esforço” criado por Zipf (1949), o qual diz que “a relação entre a frequência das palavras e a posição que estas ocupam na ordem frequencial tem valor constante”. A partir dessas ideias, Luhn sugeriu, em 1957, que a frequência das palavras em um texto está ligada à utilidade que teriam na indexação e, a partir da ideia de frequência das

palavras como critério para indexação, outras concepções começaram a surgir (NARUKAWA, 2011).

Luhn desenvolveu o processo de indexação extraindo palavras do título do documento para construir o chamado índice *Key-Word In Context* (KWIC). Nesse tipo de índice, a palavra considerada como ponto de entrada é situada no centro, com o restante do título de ambos os lados (incluindo as palavras vazias). Esse método foi utilizado, segundo NARUKAWA (2011), por Crestadoro na compilação do catálogo da Biblioteca Pública de Manchester, no século XIX; contudo, o seu valor no processamento por computador foi determinado por Hans Peter Luhn (FOSKETT, 1985). O método de indexação automática empregado por Luhn na construção de índices consistia em confrontar uma lista de palavras vazias com o texto do documento e, dessa forma, eliminar as palavras insignificantes, tais como artigos, preposições e conjunções, restando, assim, as palavras que figurariam como termos de indexação.

É preciso registrar que, segundo Rodriguez e Gonzalez (1999) os primeiros modelos de indexação automática baseados em critérios estatísticos e ou probabilísticos, deram contribuição importante ao processo de indexação automática. Isto aconteceu porque foram os primeiros métodos que surgiram como alternativa à indexação manual. Estes critérios, foram aproveitados juntos, com os avanços da informática e, ainda, continuam sendo usados integrados tanto aos métodos linguísticos, como na extração de palavras para elaboração de linguagens controladas, como os tesouros.

Na década de 1980 dá-se início ao processamento de indexação em linguagem natural com emprego do sentidosemântico dos termoscom o empregoda análise semântica que objetiva descobrir o significado das palavras (semântica léxica), reconhecer sinônimos, situar o significado das palavras dentro das orações (semântica gramatical), estabelecer o conjunto de palavras que se relacionam com um mesmo campo semântico (semântica contextual), determinar os termos gerais e específicos e estabelecer enlaces com os antônimos” (NARUKAWA *et al*, 2009).

Os métodos mistos ou híbridos de indexação automática são caracterizados pelo uso da estatística, linguística textual e os tesouros como meios ou instrumentos para controle de vocabulário “contribuindo para eliminar problemas como os causados pela sinonímia e pela não identificação das funções sintáticas dos termos, proporcionando benefícios à revocação na recuperação da informação (NARUKAWA *et al*, 2009).

Afirma Lancaster (2004) que a indexação pode ser feita por extração automática ou por atribuição automática dos termos. Na indexação automática por extração, as palavras são

extraídas do texto por um programa que adota o critério de frequência com que estas palavras aparecem no texto, então são usadas como descritores para representar o documento isto é: a indexação é realizada em linguagem natural. Vale lembrar que estes descritores são *comparados* com uma lista de palavras vazias (*stopwords*) e os que estão nesta lista são desconsiderados como descritores.

A indexação automática por atribuição será discutida, em detalhes, na próxima seção (2.5), por se tratar do foco da presente pesquisa.

2.5 Indexação automática por atribuição

A indexação automática por atribuição consiste no processo de representação temática do conteúdo do documento por meio de termos selecionados de algum vocabulário controlado que pode ser um tesouro ou lista alfabética. O vocabulário controlado, segundo Lancaster (2004), é uma lista de termos autorizados que constitui uma linguagem de indexação. Afirma ainda Lancaster (2004), que na indexação automática por atribuição é necessário desenvolver para cada termo a ser atribuído, um perfil de palavras ou expressões que costumam ocorrer frequentemente nos documentos.

Uma metodologia nesse sentido é utilizada no *Sistema de Indización Semi Automática* (SISA), no qual se efetua a indexação por comparação entre o documento - constituído por título, resumo e texto - e um vocabulário controlado, partindo de critérios de frequência preestabelecidos pelo sistema para propor os termos de indexação (LOPES, 2002).

A indexação automática por atribuição é um processo complexo e delegar a atividade de atribuição de termos aos sistemas automáticos requer análise tanto dos métodos quanto das características dos instrumentos de indexação (NARUKAWA *et al*, 2011).

Basicamente a questão que se apresenta na indexação automática por atribuição é a seleção dos termos que irão representar os documentos e que serão também essenciais sob o ponto de vista dos SRIs em termos de revocação e precisão.

Assim, diferindo do processo da indexação por extração automática, o processo de indexação por atribuição automática é mais complexo de ser realizado uma vez que é necessário um controle terminológico dos termos que irão representar o conteúdo temático do documento, fazendo com que este tipo de indexação seja mais difícil de realizar.

Como já dito, acima, na indexação automática por atribuição, deve-se desenvolver, para cada termo atribuído, um “perfil” de palavras ou expressões que costumam ocorrer nos documentos, como por exemplo, para o termo ‘chuva ácida’, poder-se-ia incluir, entre outras,

as seguintes expressões: ‘precipitação ácida’, ‘poluição atmosférica’ e ‘dióxido de enxofre’, etc. Outro exemplo que mostra este problema é dificuldade do computador em atribuir o descritor “TOXICIDADE” ao texto, uma vez que a frase: “dois dias depois de a substância haver sido ingerida surgiram diversos sintomas” pode ser legitimamente indexada por uma pessoa sob o assunto ‘intoxicação’, entretanto, para um programa computacional, essa tarefa é verdadeiramente difícil (LANCASTER, 2004).

A indexação automática por atribuição é, das duas citadas por Lancaster, a que mais se identifica com o processo de indexação humana. Os analistas de software encontram dificuldades ao elaborar programas (algoritmos) que executem as tarefas de atribuição, principalmente porque neste tipo de indexação aos termos selecionados são atribuídos outros termos autorizados. Ou seja, este tipo de indexação está, basicamente, relacionado a um vocabulário controlado e que quando da análise do conteúdo do documento para extração dos termos mais frequentes, aos termos “selecionados pelo sistema” são atribuídos novos termos equivalentes definindo novos valores para indexação automática feita no primeiro momento.

É o caso, por exemplo, que existe um texto para indexar e a palavra “indexação automática” aparece com uma razoável frequência. Então se faz uma associação deste termo com um vocabulário controlado e, podem ser atribuídos os seguintes termos: termos relacionados como indexação semiautomática, indexação manual, indexação de documentos não textuais, etc. ou, termos equivalentes como *automatic indexing*, *mechanized indexing*, etc.

No processo de indexação automática por atribuição, os vocabulários controlados atuam no próprio processo de análise automática do documento e na representação, ou seja, condicionam os resultados na atribuição de descritores, assim, há sempre a necessidade de se levar em consideração todos os seus atributos, além da necessidade de considerar a sua atuação associada aos métodos de indexação automática pelos quais se realiza a análise do conteúdo temático dos documentos, segundo afirma Narukawa (*et al*, 2009).

Na indexação automática por atribuição, um dos quesitos importantes é a avaliação desta indexação.

O ato de avaliar ou de fazer uma avaliação não é uma tarefa fácil pois, envolve a subjetividade de um julgamento, de um juízo de valor sobre um processo e seus resultados.

No sentido de se verificar até que ponto os requisitos de qualidade estão sendo cumpridos de acordo com o planejado ou esperado, é que se faz a avaliação dos resultados alcançados quando se executa uma atividade ou processo.

No processo de indexação não é diferente. Sendo uma atividade importante na organização e recuperação da informação, se faz necessário avaliar se sua atividade fim ou

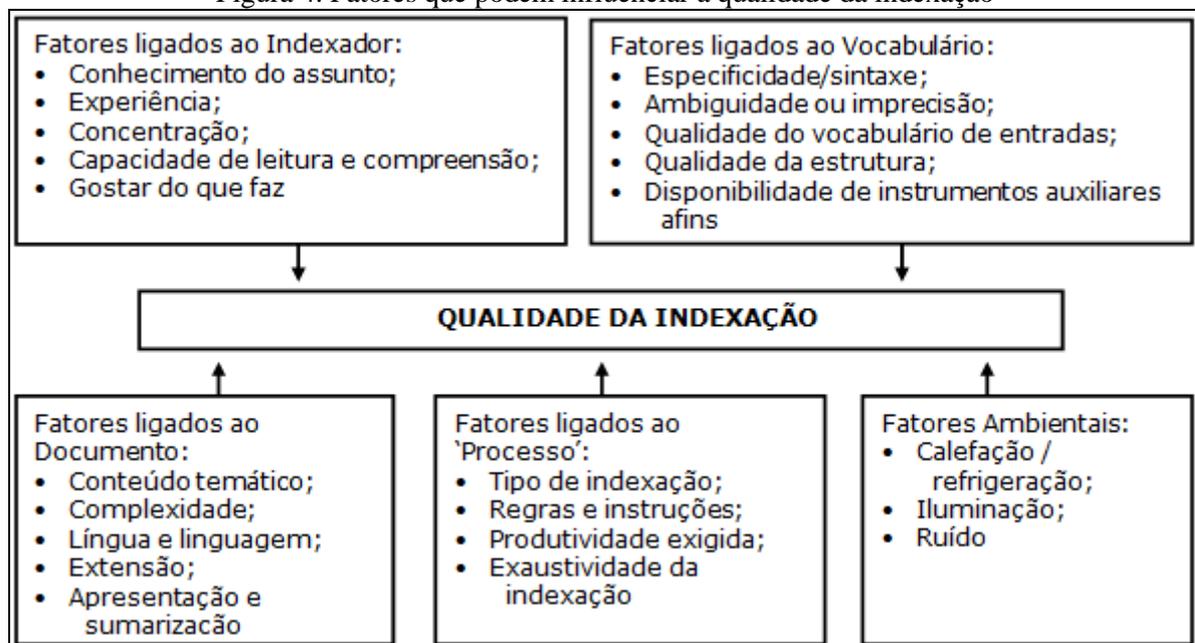
seu objetivo principal está sendo atingido, qual seja, atender as necessidades de informação daqueles que acessam as bases de dados, via sistemas de recuperação da informação.

No processo de indexação tem-se, basicamente, a avaliação intrínseca e a avaliação extrínseca. Na avaliação intrínseca é medido o grau de consistência da indexação, que de acordo com Leiva (1997, p. 22) é definida como o grau de concordância entre indexadores de um mesmo grupo ou entre indexadores de grupos diferentes, quando da representação da informação essencial de um documento, por meio de um conjunto de termos de indexação selecionados por estes indexadores.

Segundo Lapa (2014) ao mencionar Lancaster (2004) quando há concordância entre dois ou mais indexadores, ocorre à coerência inter-indexadorese, no momento em que a concordância for entre o mesmo indexador em diferentes momentos ocorre à coerência intra-indexador. Deste modo, para que haja qualidade na indexação, é necessário que se tenha o mesmo grau de consistência inter e intra indexadores.

Mas não é só a consistência que garante a qualidade no processo de indexação. Lancaster (2004) relaciona outros fatores que eventualmente podem afetar a qualidade da indexação e que estão listados na figura 4.

Figura 4. Fatores que podem influenciar a qualidade da indexação



Fonte: (LANCASTER, 2004).

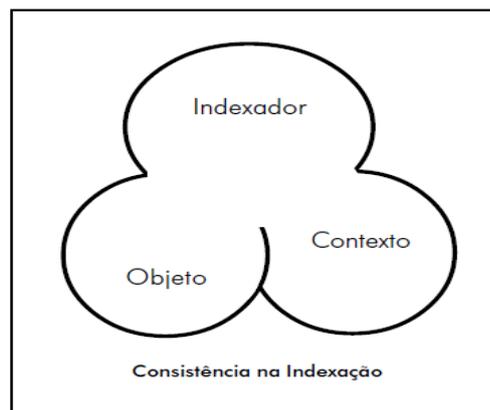
Como afirma Lancaster (2004), a indexação não é uma atividade-fim em si mesma, ou seja, os parâmetros de qualidade devem ser aplicados aos resultados deste processo como um todo, levando-se em consideração, também, o cenário de uma específica base de dados

tanto em formato eletrônico como formato impresso, bem como, as necessidades informacionais dos seus usuários. Portanto, é preciso fazer a avaliação extrínseca do processo de indexação para se analisar os parâmetros de revocação, precisão e medida F quando da recuperação da informação.

A indexação e o seu resultado estão associados, impreterivelmente, a um conjunto de elementos ao quais são, obrigatoriamente, considerados na hora de analisar os índices de consistência. Estes elementos, estão descritos na Figura 5 e, são:

- 1) A formação, os conhecimentos na matéria, a profissão e a motivação do indexador;
- 2) As características do objeto indexado e,
- 3) As condições em que se dá a indexação.

Figura 5. Âmbitos que intervêm na consistência da indexação.



Fonte: (LEIVA, 2008).

Enfim, quando se tem como um dos objetivos da indexação comparar o grau de consistência, tanto intraconsistência ou consistência intraindexador e/ou interconsistência ou consistência interindexador é preciso manter o controle sobre o número de elementos ou fatores que influenciam diretamente os índices de consistência (quadro 6).

Quadro 6. Elementos que devem ser considerados na comparação de indexações

INDEXADOR	<ul style="list-style-type: none"> - Formação e experiência em indexação: indexador experiente contra principiante - Conhecimento da matéria - Domínio das ferramentas da indexação (linguagens de indexação) - Profissionalismo
CONTEXTO	<ul style="list-style-type: none"> - Políticas de indexação marcadas pela instituição - Objetivo da indexação: temas principais contra especificidade - Tipos e necessidades dos usuários - Carga de trabalho e tempo dedicado
OBJETO	<ul style="list-style-type: none"> - Complexidade do objeto indexado: livro infantil contra patentes - Características e propriedades do objeto indexado: material textual contra material gráfico ou audiovisual - Tamanho: indexação de textos curtos contra textos compridos
MOMENTO	<ul style="list-style-type: none"> - A comparação executa-se com as palavras-chave pinçadas diretamente do texto ou também, uma vez convertidas essas palavras-chave em descritores, após sua filtragem com o vocabulário controlado.
FÓRMULA	<ul style="list-style-type: none"> - São variadas as fórmulas matemáticas utilizadas para conseguir os índices de consistência.

Fonte: (LEIVA, 2008).

Os elementos do Quadro 6 estão, segundo Leiva (2008), relacionados com o indexador, seja automático ou um profissional da indexação, os quais têm critérios próprios definidos. Também devem ser considerados não só a política e os objetivos da indexação, mas, também, as características dos conteúdos dos artigos a serem indexados. Mais ainda, deve ser considerado a comparação que se faz entre as palavras-chave que o autor do artigo retira do texto e aquelas que foram convertidas via vocabulário controlado.

Como salienta Leiva (1999), até mesmo a exploração dos índices de consistência como indicador de uma indexação correta pode ser problemática, e isso se deve ao fato de que pode ser encontrada uma indexação consistente, porém com algumas correções a fazer, o que pode ocorrer por erros de análises - os dois indexadores cometem o mesmo erro -, ou pela utilização de vocabulários controlados diferentes (NARUKAWA, 2011).

Lancaster (2004) afirma que as pesquisas de análise de consistência apontam a dificuldade de se alcançar altos índices de consistência em indexações e, por isto, é preciso “ver” a inconsistência na indexação como um “fato” inerente à atividade de indexação.

2.5.1 Trabalhos, métodos e sistemas de indexação automática por atribuição

Nesta seção serão apresentados e discutidos alguns dos trabalhos publicados na literatura brasileira em Ciência da Informação, que discutem ou usam a indexação automática por atribuição, bem como, os métodos e tipos empregados na realização deste tipo de indexação.

Também, procura-se analisar e avaliar o emprego deste tipo de indexação associada à aplicação dos diversos programas ou sistemas de extração automática de termos e palavras-chaves, vinculados aos métodos empregados, bem como, os diversos tipos de vocabulários controlados com ênfase nos tesouros.

O trabalho de produção de pesquisa acadêmica: “A indexação automática no Brasil no âmbito da ciência da informação (1973-2012)” realizado por Lapa (2014), além de preencher uma lacuna existente na literatura pertinente, apresenta e discute entre outros pontos: o mapeamento e análise da produção científica no Brasil sobre indexação automática, abrangendo um período de quase quarenta anos (1973-2012); a realização de estudos de análise bibliométrica, analisando as publicações por autoria, instituições publicadoras, ano, fonte de publicação e instituições acadêmicas; a realização de análise de conteúdo observando as publicações pelas categorias de objetivos e de aspectos metodológicos.

Como resultados alcançados na realização da pesquisa, apresenta o processo de **atribuição** como o meio mais adotado para identificar os termos, a indexação semiautomática como procedimento mais aplicado na validação dos termos, sendo ultrapassada, nos estudos dos últimos dez anos, pela indexação automática.

Neste trabalho, Lapa (2014) discorre sobre a questão dos métodos utilizados na indexação automática, onde apresenta um quadro descrevendo os métodos básicos e, entre estes métodos, comenta o método comparativo composto das seguintes etapas:

1) Compara, através do computador, os termos do resumo com os descritores da linguagem documentária (vocabulário controlado - tesouro, por exemplo);

2) Cada vez que um descritor aparece no resumo o documento é indexado com o termo correspondente.

Este método é eficaz somente em campos muito específicos do conhecimento onde a linguagem natural e a linguagem documentária são muito próximas, afirma Guinchat e Menou (*apud*LAPA, 2014).

Na questão de sistemas utilizados em indexação automática, após analisar os resultados, foi verificado que entre os mais citados e usados estão:

1) O sistema PRECIS que é usado na construção automática de índices de assunto (FUJITA, 1989 *apud* LAPA, 2014);

2) Já Robredo (1982) propõe a utilização do sistema de indexação automática, nomeado como AUTOMINDEX/II, para gerar uma tabela de frequências de aparecimento dos descritores encontrados em cada arquivo;

3) A proposta de Narukawa (2011) é a de investigar o uso e a adequação de linguagens documentárias alfabéticas e vocabulários controlados (tesauros) em indexação automatizada e indexação automática por atribuição na aplicação do (SISA).

No capítulo Conclusões, Lapa (2014) afirma que, com relação à pequena diferença que há, em termos percentuais, entre os trabalhos que pesquisaram a identificação dos termos por meio da extração (47%) e os que optaram pelo da atribuição (53%), é explicado, em virtude da evolução e disponibilização de novas tecnologias de processamento da linguagem natural. Também, o aumento de pesquisas sobre aplicação de tesauros e vocabulários controlados motivaram, paralelamente, estudos e pesquisas sobre a atribuição.

Outra produção acadêmica sobre indexação automática é o trabalho realizado por Narukawa (2011) que discorre sobre indexação automática por atribuição, fazendo uso de um conjunto de instrumentos teórico-metodológicos, como a aplicação de um vocabulário controlado tipo tesouro e da atribuição automática de termos via sistema SISA.

A indexação automática por atribuição, objeto deste trabalho, consiste no processo de representação do conteúdo do documento, mediante termos selecionados de algum vocabulário controlado.

Metodologia nesse sentido é utilizada no (SISA), no qual se efetua a indexação por comparação entre o documento, constituído por título, resumo e texto e um vocabulário controlado, partindo de critérios de frequência preestabelecidos pelo sistema para propor os termos de indexação (NARUKAWA, 2011).

Quanto aos sistemas de indexação automática com o uso de vocabulários controlados, Narukawa (2011) descreve e discute vários sistemas, entre os quais, estão: FAIRS, Autoindex, Concept Assigner, HEPIIndexer, Cadis, Smart, Precis e SISA.

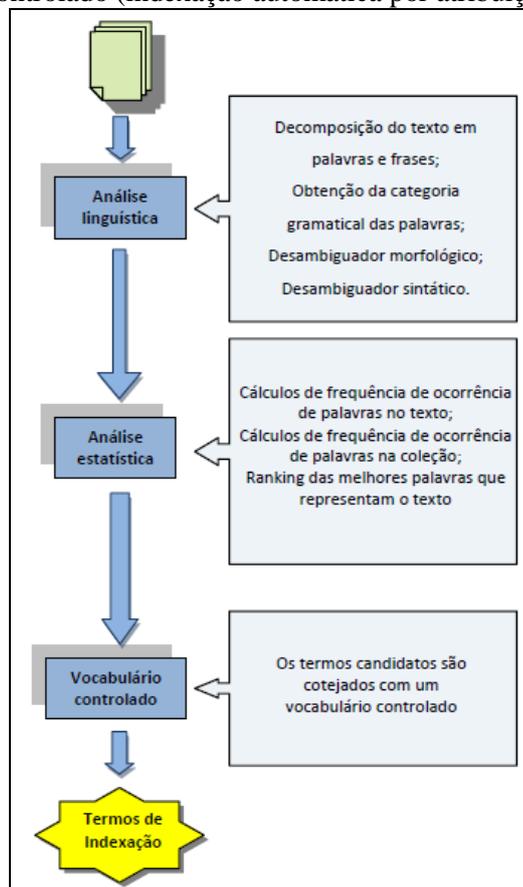
No quesito, qualidade de indexação, a autora lembra que, em pesquisa realizada por Narukawa (*et al*, 2009), foi proposta a investigação da indexação automática por atribuição por meio da análise dos índices de consistência, revocação e precisão, estes dois últimos, avaliados na recuperação da informação. Para isso, realizou-se análise comparativa entre a indexação automática com o uso da linguagem do Descritores em Ciências da Saúde (DeCS) e a indexação manual do Centro Latino-Americano e do Caribe de Informação em Ciências da

Saúde (BIREME), buscando contribuir com o aperfeiçoamento do SISA e com o desenvolvimento teórico da indexação automática.

Em termos resumidos, Narukawa (2011) afirma que, os resultados da pesquisa mostraram a necessidade de adaptação do vocabulário controlado, principalmente no que diz respeito ao estabelecimento de relações entre os termos, uma vez que as relações semânticas entre conceitos devem ser mantidas durante a indexação automática para garantir qualidade na recuperação da informação.

Leiva (2008) sugere um modelo de arquitetura de um sistema de indexação automática, no qual, é realizada, numa primeira etapa, a análise linguística via tratamentos morfológico, sintático e semântico. Na etapa seguinte, opera o analisador estatístico, calculando a frequência de ocorrência dos termos que foram tratados linguisticamente na etapa anterior. E, na etapa final, aplica-se um vocabulário controlado, cujos termos são comparados com os termos candidatos à indexação, determinando, deste modo, a *atribuição definitiva* dos termos de indexação(Figura 6).

Figura 6. Modelo de arquitetura de um sistema de indexação automática com uso de vocabulário controlado (indexação automática por atribuição)



Fonte: (LEIVA, 2008).

Resumidamente, o uso do vocabulário controlado na indexação automática por atribuição, possibilita definir o significado das palavras, selecionar as palavras candidatas à indexação no contexto de uma área do conhecimento específica e ajuda na determinação das palavras ou termos que serão *atribuídos* na indexação. Tudo isto em consonância com a política de indexação adotada cujas diretrizes devem “atender as necessidades e particularidades de determinado sistema de informação”, não esquecendo, “o domínio do conhecimento em que o sistema de indexação automática atuará, uma vez que as particularidades da literatura, das fontes de informação e da terminologia de cada área podem influir no desempenho do sistema e merecem análise e adaptação” lembra (NARUKAWA, 2011).

Após mostrar as aplicações dos vocabulários controlados no uso do sistema de indexação automática SISA bem como analisar e avaliar os resultados, Narukawa (2011, p. 178), aponta entre outras constatações, que:

Sob uma perspectiva mais ampla dos problemas enfrentados na indexação automática, tivemos a comprovação de que o vocabulário controlado atua como instrumento que condiciona os resultados de indexação automática e, portanto, é extremamente importante compreendê-lo dentro de um processo em que estão envolvidos aspectos relacionados aos métodos de identificação das unidades representativas da informação, à aplicação de tratamento linguístico, à relação entre o vocabulário controlado e os instrumentos de indexação automática, às relações semânticas no vocabulário controlado, às características da área de conhecimento, ao idioma do vocabulário controlado, à atualização do vocabulário controlado e à possibilidade de uso conjunto de vários vocabulários controlados.

Sobre o uso de ontologias na indexação automática existe o trabalho intitulado, “Um método para utilização de ontologias na indexação automática”, produzido por Nicolino e Ferneda (2014), ambos, pesquisadores da Universidade Estadual Paulista (Unesp).

Nesta produção acadêmica, Nicolino e Ferneda (2014) afirmam que as ontologias, por disponibilizarem uma estrutura conceitual e terminológica restrita a um domínio específico e que, originalmente, são representadas em linguagens processáveis por computador, introduzem novas perspectivas para o desenvolvimento das pesquisas em indexação automática. Deste modo, continuam os autores, o uso de ontologias no processo de indexação automática possibilita adicionar a esse processo, uma linguagem de um domínio específico e uma estrutura lógica e conceitual que pode ser utilizada para realizar inferências, proporcionando uma expansão dos termos diretamente extraídos do texto do documento.

Uma ontologia é composta de um conjunto de conceitos dentro de um determinado domínio, organizados em uma taxonomia. Taxonomia é a classificação de informações no

formato de uma hierarquia, de acordo com os relacionamentos que estabelecem com entidades do mundo real que representam. Assim, servem justamente para classificar informação em uma hierarquia, utilizando o relacionamento de generalização (“tipo-de” ou “pai-filho”) (NICOLINO; FERNEDA, 2014). Na Ciência da Computação, Guimarães (2003, p.53), apresenta algumas vantagens no uso das ontologias, quais sejam:

- 1) Ontologias fornecem um vocabulário para representação do conhecimento. Vocabulário esse que traz uma conceitualização que o sustenta, evitando ambiguidades.
- 2) Permitem o compartilhamento de conhecimento. Sendo assim, caso exista uma ontologia que modele adequadamente certo domínio do conhecimento, essa pode ser compartilhada e usada por pessoas que desenvolvem aplicações para o mesmo domínio.
- 3) Fornecem uma descrição exata do conhecimento. Diferentemente da linguagem natural, em que as palavras podem ter semântica diferente conforme o contexto, a ontologia é escrita em linguagem formal, ou seja, estabelecendo formalmente, então, as definições de um termo, eliminando ambiguidades.
- 4) Há a possibilidade de mapeamento da linguagem da ontologia sem que com isso seja alterada a sua conceitualização, isto é, uma mesma conceitualização pode ser expressa em várias línguas.
- 5) É possível estender o uso de uma ontologia genérica de tal forma que ela se adeque a um domínio específico.

Em seguida, neste trabalho de Nicolino e Ferneda (2014), estes autores na seção 4, fazem abordagens sobre a *Web Ontology Language* (OWL) que é a linguagem recomendada pelo Consórcio W3C para construções de ontologias. Os elementos básicos para a construção de uma ontologia OWL são as classes, as instâncias das classes (indivíduos), propriedades e relacionamentos entre classes e instâncias. Neste trabalho são apresentados apenas alguns elementos da linguagem OWL considerados essenciais para o entendimento do método proposto (NICOLINO; FERNEDA, 2014).

Na sequência, na seção 5, é apresentado o método, proposto pelos autores, para a utilização de ontologias na indexação automática. Foi utilizada uma ontologia de Termos de Pediatria (PedTerm) que apresenta informações relacionadas à saúde e ao desenvolvimento infantil desde o pré-natal até os 21 anos de idade. Essa ontologia está disponível no BioPortal, um grande repositório de ontologias na área biomédica. Como foi originalmente criada no idioma Inglês, para cada classe foi incluída propriedades *label* em três idiomas: inglês, espanhol e português (NICOLINO; FERNEDA, 2014). Para extração dos termos, atribuição de conceitos, termos sinônimos e indexação multilíngue, há descrições de como fazer, orientações sobre construção dos algoritmos mostrando os exemplos.

Por fim, Nicolino e Ferneda (2014) afirmam a exequibilidade e o potencial do uso de ontologias nos sistemas de indexação automática. Também, acreditam eles, que o método exposto neste trabalho pode ser aplicado em diversas ideias, sendo possível imaginar diversas funcionalidades que podem ser desenvolvidas em um sistema de recuperação de informação baseado em ontologia.

Aqui, há concordância quanto a importância de se realizar investigações, estudos e análises sobre domínios específicos, levando em consideração suas peculiaridades pois, os mesmos, podem dar indícios que ajudam na adaptação de sistemas de indexação permitindo assim a aplicação de métodos mais eficazes.

2.5.2 SISA

O Sistema de Indexação SemiAutomática (SISA) é um sistema de indexação desenvolvido pelo Professor Doutor Isidoro Gil Leiva durante seu doutorado, concluído em 1999, na Universidade de Murcia, Espanha. A proposta é aplicá-lo a artigos científicos, haja vista as estruturas do documento analisadas pelo sistema, delimitadas por marcadores específicos que indicam o título (#CTI# e #FTI#), o resumo (#CR# e #FR#) e o texto (#CTE# e #FTE#).

O SISA exige para sua aplicação, como fonte de entrada, um documento em texto completo, um vocabulário controlado e uma lista de palavras vazias. A metodologia empregada pelo software SISA na indexação automática é descrita a seguir via aplicação do processo de análise automática composto de três módulos ou fases (NARUKAWA, 2011).

Módulo I: fase de pré-processamento do documento a indexar no formato TXT sinalizado com os marcadores exigidos pelo SISA que são: #CTI# e #FTI# para o Título, #CR# e #FR# para o Resumo e #CTE# e #FTE# para o Texto. Tudo isto é preparado para que, posteriormente, os cálculos de ponderação sejam realizados a partir da identificação da frequência nessas estruturas denominadas fontes. Neste módulo são realizadas duas etapas: a eliminação das palavras vazias por meio do confronto do documento com a lista de palavras vazias (*stopwords*) e a horizontalização, em que frases e orações compreendidas entre os sinais de pontuação (. , ; :) são dispostas em forma horizontal, isto é, são separadas em cada linha do texto.

Módulo II: fase de processamento, na qual o documento é analisado e os termos de indexação são identificados e selecionados. Na Fig. 7 é possível verificar em detalhes como ocorre o processo (LEIVA, 2008):

1º) Extrai-se o primeiro termo do vocabulário controlado (tesauro);

2º) Extrai-se o primeiro termo da fonte (texto completo);

3º) Verifica-se se os termos extraídos são iguais;

4º) Se não são iguais, verifica-se se existem mais termos nas fontes. Se não existem mais palavras, verifica-se se existem mais termos no vocabulário controlado. Se não houver, o processo é finalizado. No caso de haver mais palavras em alguma fonte, extrai-se a seguinte, que é novamente comparada com o termo do vocabulário controlado. Esse processo se repete até serem iguais ou não existirem mais palavras nas fontes.

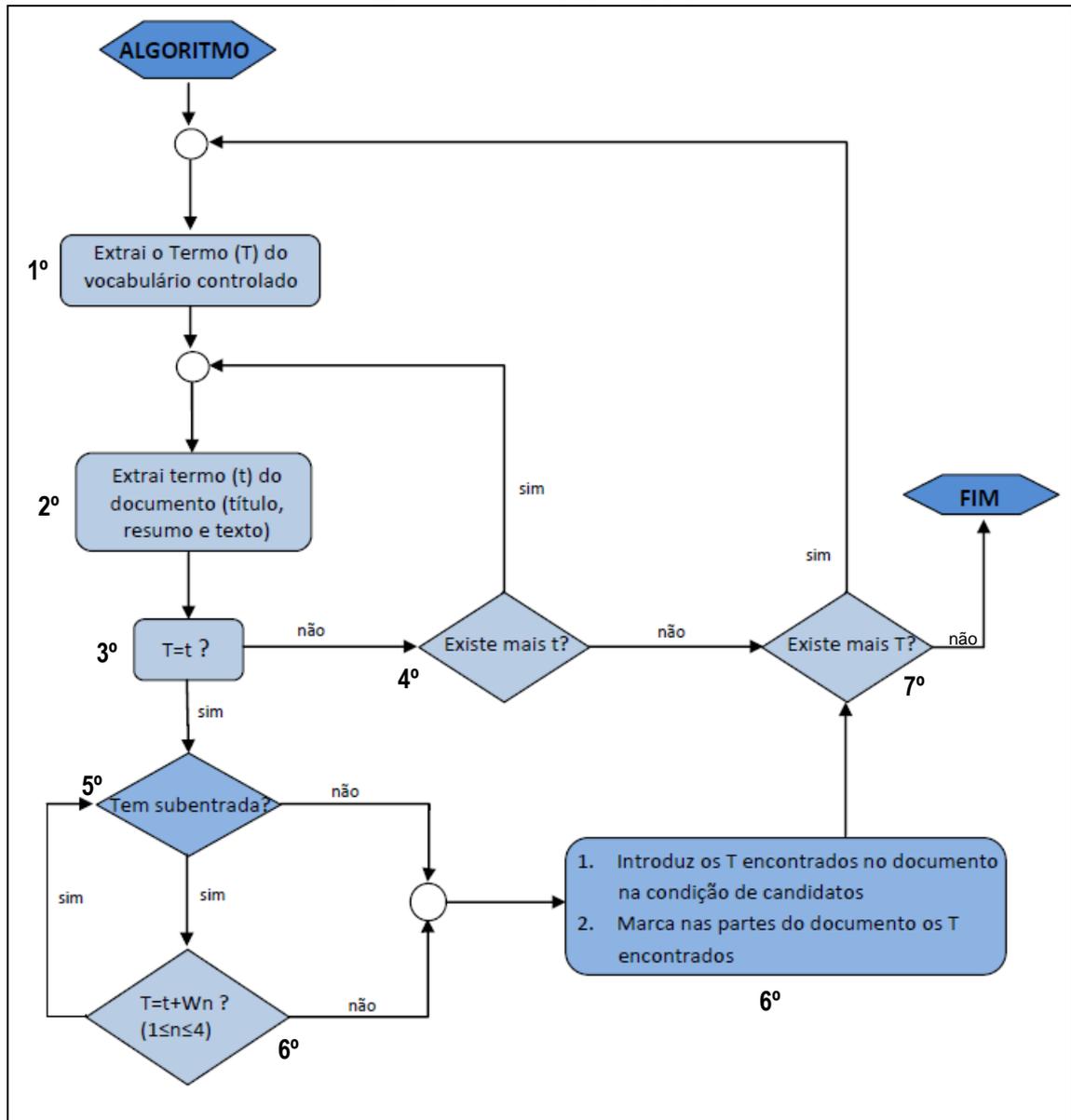
5º) Se a palavra da fonte e o termo do vocabulário controlado são iguais, verifica-se se a entrada do vocabulário controlado tem subentradas, ou seja, se tem termos que estão relacionados entre si: Se não tem subentradas, verifica-se se há relação de equivalência da entrada principal, sendo o termo autorizado enviado ao módulo de candidato a termo de indexação. Esse termo é marcado para que não seja necessário localizá-lo novamente. Se tiver subentradas, verifica-se se coincide com alguma das quatro palavras seguintes à direita da última palavra localizada na fonte.

6º) Se alguma das quatro palavras coincide com o termo do vocabulário controlado, comprova-se se este tem subentradas e repete-se o mesmo processo. No caso de não ter mais subentradas ou nenhuma das quatro palavras da fonte coincidir com a do vocabulário, comprova-se se existem relações de equivalência e transfere-se (transferem-se) o(s) termo(s) autorizado(s) ao módulo de candidatos e marca-se (marcam-se) o(s) localizado(s).

7º) Confirma-se se há mais termos (**T**) no vocabulário controlado: Se não existem mais palavras, finaliza-se o processo. Se houver mais palavras, continua-se o processo desde o 2º. passo.

A Figura 7 representa o MODELO usado pelo *software* SISA na indexação automática usando vocabulário controlado.

Figura 7. Diagrama de fluxos do algoritmo SISA do módulo II. Modelo de indexação automática com uso de vocabulário controlado



Fonte: (LEIVA, 2008).

Módulo III: nesta fase, os termos que foram analisados pelo sistema são ponderados pois, do contrário, todos os termos do vocabulário controlado que coincidem com os das fontes do documento seriam selecionados e propostos como termos de indexação. Assim para evitar isto, o sistema considera os seguintes critérios para apresentar os termos de indexação:

- 1) Se um termo autorizado aparece na fonte-título e na fonte-resumo;
- 2) Se um termo autorizado aparece na fonte-título e na fonte-texto;
- 3) Se um termo autorizado aparece na fonte-resumo e na fonte-texto;
- 4) Se o termo candidato a descritor aparece no título, no resumo e no texto;

5) Se um termo candidato a descritor aparece no texto dez vezes ou mais, além de aparecer em oito parágrafos diferentes ou mais, e não está incluído em nenhum dos termos propostos, apresenta-se como termo candidato a termo de indexação.

2.6 Uso da indexação automática em principais bases de dados científicas internacionais e da ciência da informação

As bases de dados científicas foram criadas com o propósito de disponibilizar, em um único local, milhares de artigos de periódicos científicos evitando a perda de tempo por parte do pesquisador. Em outras palavras, as bases de dados científicas são um local onde se encontram centenas de milhares de artigos evitando, assim, a busca por informação em diversos sites por parte do usuário.

Vale lembrar que há bases que oferecem o texto completo dos artigos e há bases de dados referenciais, ou seja, as que indicam somente a existência do artigo, o seu resumo e em qual revista encontra-se este material.

Na literatura da Ciência da Informação não se tem, claramente, quais bases de dados, pelo menos as mais acessadas, usam indexação automática e que tipo de indexação automática, se assistida por computador, semiautomática ou automática e, mais ainda, se usam algum tipo de vocabulário controlado para organizar e disponibilizar as suas informações. Assim, foi realizada uma pesquisa no sentido de se analisar algumas bases de dados científicas, descritas a seguir, procurando-se identificar quais usam indexação automática e, caso afirmativo, que tipo de indexação é utilizada e, quais usam vocabulários controlados.

De acordo com Porter e Cunningham (2005), as principais bases de dados em Ciência e Tecnologia (Science & Technology - S&T) são, **Engineering Village**, incluindo *Engineering Index Compendex*, **Chem Abstracts (CAS)** com o seu software especial, SciFinder, que disponibiliza o acesso a área de Ciências Exatas, cobrindo, inclusive, literatura química e patentes (<http://www.cas.org>) e, **MEDLINE** (<http://www.ncbi.nlm.nih.gov/PubMed/>).

Além das bases acima citadas, serão analisadas as bases *Web of Science* (WoS), Inspec, Scopus, ERIC e, em Ciência da Informação, as bases LISA e BRAPCI.

Após consulta aos gestores das bases de dados acima referenciadas, tem-se o Quadro 7. Vide Apêndice B onde constam as respostas recebidas. As bases: *Web of Science*, Lisa, *Engineering Index Compendex*, Scopus e *Chem Abstracts* não responderam aos emails.

Quadro 7. Bases de Dados que usam indexação automática

Base de Dados	Indexação Automática	Tipo de Indexação Automática	Vocabulário Controlado
BRAPCI	Sim	Automática no processo de coleta de termos, comparados com vocabulário controlado	Lista de Cabeçalhos de Assuntos (remissiva).
PUBMED - MEDLINE	Sim	Por atribuição automática, assistida por computador.	Tesouro MeSH Medical Subject Heading
ERIC	Não	Indexação manual e indexação assistida por computador.	Vocabulário controlado ERIC Thesaurus.

Fonte: (Elaborado pelo autor, 2016).

A **BRAPCI** é um produto de informação, fruto de projeto de pesquisa criado pelo Departamento de Ciência e Gestão da Informação da Universidade Federal do Paraná (UFPR). A base de dados é constituída pelos artigos dos periódicos da área de Ciência da Informação (CI).

A BRAPCI tem em seu banco de dados mais de 25.000 termos somente no idioma Português, destes, muitos são remetidos para termos equivalentes (remissivas). Desta forma, a indexação é automática quando se faz o processo de coleta, porém, os termos selecionados são comparados com um vocabulário controlado que está em constante validação. Caso o termo “indexado” tenha uma remissiva, utiliza-se esta para indexação, ou seja, a Brapci utiliza o processo de indexação automática por atribuição pós-coordenado, através de vocabulário controlado.

A **PUBMED - MEDLINE** é uma base de dados composta por MedLine e PubMed. PUBMED é um recurso gratuito (motor de busca) desenvolvido e mantido pela Biblioteca Nacional de Medicina (NLM®) dos Estados Unidos. Além de oferecer acesso aos recursos relacionados ao Medline, o PubMed também contém: registros de artigos em fase de indexação, informações sobre os editores das revistas, Mesh® (Vocabulário controlado da NLM), OLDMEDLINE, registros de livros disponíveis no NCBI Bookshelf, links para sites

que possuam artigos com texto completo e outros assuntos relacionados. Já o MEDLINE (*Medical Literature Analysis and Retrieval System Online*) é uma base de dados online que oferece acesso gratuito a referências e resumos de revistas científicas da área Biomédica. São indexados nesta base aproximadamente 5.400 periódicos dos Estados Unidos e de mais 80 países. O Medline é o principal componente do PubMed. Além de estar disponível no portal PubMed, o Medline pode ser acessado em outras interfaces como na Biblioteca Virtual em Saúde. A PUBMED-MEDLINE usa indexação automática com os indexadores comparando os termos selecionados (citação bibliográfica) com os termos de um vocabulário controlado *Medical Subject Headings* (MeSH).

A WEB OF SCIENCE (WoS) - Designação comum que é dada a um conjunto de bases de dados também conhecidas como *Science Citation Indexes* (*Science Citation Index*, *Social Science Citation Index*, *Arts and Humanities Citation Index*), compiladas pelo ISI (*Institute for Scientific Information*). Além destas, estão também incluídas na *Web of Science* as bases de dados *Current Chemical Reactions* e *Index Chemicus*. Trata-se de bases de referência bibliográfica, que não contêm o texto integral dos documentos, mas que possuem uma característica especial: é possível ver quais os artigos citados por determinado artigo, ou verificar quantas vezes um artigo foi citado e por quem. É a partir destas bases de dados que é calculado o fator de impacto das publicações periódicas, bem como outros indicadores bibliométricos, presentes no *Journal Citation Reports*. Não usa indexação automática para texto completo.

A LISA - *Library information science & technology abstracts* é uma base de dados internacional de resumos e indexação, elaborada para profissionais da área de biblioteconomia e outros especialistas da área da informação. A LISA conta atualmente com resumos de mais de 440 periódicos, publicados em mais de 68 países, num número superior a 20 idiomas. Gestor da base de dados não respondeu ao email sobre que tipo de indexação é usada. Vide Apêndice B.

A INSPEC é uma base de dados editada pelo *The Institution of Engineering and Technology* (IET). A edição de 2012 contém 18.755 termos dos quais 9.722 termos são termos chaves (chamado de indexação controlada, descritores ou cabeçalhos de assunto em diversas plataformas). 9.033 termos são lead-in (sinônimos) - termos sugeridos ao pesquisador os quais eles devem usar para a pesquisa. Os registros INSPEC contêm, em média, de três a seis termos do *Thesaurus* que refletem os principais conceitos descritos pelo título e resumo. Isso permite que o usuário possa recuperar todos os principais conceitos abordados pelos

autores sem ter que saber todos os seus sinônimos. Gestor da base de dados não respondeu ao email sobre que tipo de indexação é usada. Vide Apêndice B.

A ERIC-O *Educational Research Information Center* ou Centro de Informações de Recursos Educativos (ERIC) é uma biblioteca digital online de pesquisa em educação e informação. O ERIC é patrocinado pelo Instituto de Ciências da Educação do Departamento de Educação dos Estados Unidos. É um banco de dados de textos completos que inclui bibliografias, documentos de investigação e pesquisa em educação e informação para educadores, pesquisadores e público em geral. Fornece acesso a 1,5 milhões de registros bibliográficos (citações, resumos e outros dados pertinentes) de artigos de jornal e outros materiais relacionados com a educação, com centenas de novos registros adicionados a cada semana. Um componente chave do ERIC é a sua coleção de literatura cinzenta na educação, que é largamente disponível em texto completo. A ERIC usa indexação manual e indexação assistida por computador. Indexadores profissionais “atribuem” manualmente termos de assunto do ERIC Thesaurus (vocabulário controlado) para cada registro, e existe um mapeamento de termos que são correlacionados uns com os outros para aparecer automaticamente quando um é usado (a indexação assistida por computador).

A ENGINEERING INDEX COMPENDEX é uma plataforma que proporciona acesso integrado a diversas fontes de informação especializadas, para atender as demandas de diversas áreas da Engenharia, e também outras áreas do conhecimento como Química, Física, GeoCiências, dentre outras. Engloba as bases Compendex, Inspec, PaperChem, entre outras, que formam um conjunto de fontes referenciais que, integradas, são uma valiosa fonte para apoiar o processo de pesquisa, desenvolvimento e inovação.

A CHEM ABSTRACTS é um departamento da Sociedade Americana de Química (*American Chemical Society*) que produz os *Chemical Abstracts*, um índice da literatura científica sobre a química e os ramos coligados. Os *Chemical Abstracts* são publicados desde 1907. CAS está sediada em Columbus, Ohio. A CAS pode ser acessada pela plataforma SciFinder. A CAPES, por meio do Portal de Periódicos (www.periodicos.capes.gov.br) disponibiliza à comunidade acadêmica, o acesso online à base SciFinder (<http://scifinder.cas.org>). SciFinder é uma plataforma projetada para dar suporte aos alunos, docentes, pesquisadores e profissionais da área de Química, em suas pesquisas e trabalhos acadêmicos. Isto se dá pela disponibilização de coleções de excelente qualidade: livros, dissertações, periódicos, patentes, substâncias químicas, entre outros.

A SCOPUS é uma base referencial da Editora Elsevier, disponibilizada pelo Portal de Periódicos da CAPES. Essa base indexa títulos acadêmicos revisados por pares, títulos de

acesso livre, anais de conferências, publicações comerciais, séries de livros, páginas web de conteúdo científico (reunidos no Scirus) e patentes de escritórios. Dispõe de funcionalidades de apoio à análise de resultados (bibliometria) como identificação de autores e filiações, análise de citações, análise de publicações e índice H. Cobre as áreas de Ciências Biológicas, Ciências da Saúde, Ciências Físicas e Ciências Sociais.

2.7 Tesouros da área da ciência da informação

Como já mencionado anteriormente, é proposto neste trabalho, o uso de um vocabulário controlado, o Tesouro Brasileiro de Ciência da Informação (TBCI) como linguagem de indexação no processo de indexação automática por atribuição nos 60 artigos selecionados e utilizados por Souza (2005) em sua Tese de Doutorado, listados no Apêndice A. O sistema SISA será utilizado para selecionar os termos candidatos a indexação.

No entanto, além do Tesouro TBCI, existem outros Tesouros na área de Ciência da Informação os quais, devem ser mencionados neste trabalho no sentido descrever tais instrumentos disponíveis para a construção de novos instrumentos de controle terminológicos, sendo fontes de sugestões de termos que podem ser associados à uma base de dados, nos diversos processos de representação e recuperação dos documentos.

Assim, de acordo Santos (2015), para a elaboração de um instrumento de controle terminológico, verifica-se a importância e necessidade de se considerar as garantias literárias, estrutural e de uso. O princípio de garantia literária para elaboração de um instrumento de controle terminológico está associado à seleção de um vocabulário/linguagem proveniente das produções científicas da área de Ciência da Informação. A garantia estrutural está relacionada à aplicação de termos no instrumento de representação da informação, cujas relações entre eles devem estar ligadas à estrutura de um tesouro no intuito de apresentar uma função estrutural útil. Com a garantia de uso, os termos selecionados para a elaboração e atualização do instrumento de controle devem estar de acordo com descritores empregados pelos usuários no momento das buscas das produções científicas disponíveis no sistema.

a) Tesouro Brasileiro de Ciência da Informação

O Tesouro Brasileiro de Ciência da Informação - TBCI foi lançado em sua última versão em 2014 pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) que engloba, em média, 1.800 termos os quais (em sua maioria) tem versão em língua inglesa e

espanhola e são complementados por definições, no intuito de assumir um papel central na recuperação da informação no Brasil e em países lusófonos. Esse instrumento de controle terminológico é direcionado para indexadores, pesquisadores, professores e profissionais da informação em geral. Disponível em: <<http://www.ibict.br/publicacoes-e-institucionais/tesauro-brasileiro-de-Ciência-da-informacao-1>>. Acesso em: 25 jun. 2016.

b) Tesouro em Ciência da Informação - TCI

Tesouro disponibilizado pela Pontifícia Universidade Católica de Minas Gerais (PUC- Minas) e Universidade Federal de Minas Gerais (UFMG) e elaborado com base nos tesouros existentes na área de Ciência da Informação, em 2012. Esse instrumento se configura como parte da pesquisa de Doutorado de Manoel Palhares (2005), da Escola de Ciência da Informação, da UFMG.

O Tesouro em Ciência da Informação (TCI) contém 1.891 termos, sendo que 1.694 são descritores e 197 não-descritores (MOREIRA, 2003).

c) Tesouro de Ciência da Informação da ASIS&T

Tesouro da Associação para a Ciência da Informação e Tecnologia (*Association for Information Science and Technology - ASIS&T*) lançado nos Estados. Esse instrumento abrange os campos da Ciência da Informação e da Biblioteconomia, bem como de áreas afins como a linguística, Ciência da Computação, ciências comportamentais e cognitivas, entre outras áreas relacionadas. O Tesouro de Ciência da Informação da ASIS&T inclui 1.316 termos preferenciais (descritores), 691 descritores não preferenciais e 37 indicadores de facetas (SANTOS, 2015).

d) Tesouro de Ciências da Documentação - DOCUTES

O Tesouro de Ciências da Documentação é um instrumento normalizador e de apoio para o processo de indexação e recuperação do conteúdo de documentos. É administrado pelo Departamento de Biblioteconomia e Documentação da Faculdade de Filosofia e Letras da Universidade de Leon, Espanha, com a última versão em 2002/2003. Este tesouro está estruturado em seis campos (domínios) temáticos:

- 1) Ciências da Documentação: história, teoria e sistemas.
- 2) Informação, documentos, fontes de informação.

- 3) Investigação e Metodologia documental.
- 4) Representação e Recuperação da Informação.
- 5) Sistemas de Informação.
- 6) Tecnologias da Informação.

Estes domínios temáticos transcendem a divisão tradicional por disciplinas, tais como, Biblioteconomia, Arquivística e Documentação. Refletem as facetas que as compõem, ou seja, métodos e técnicas de investigação e organização, ferramentas tecnológicas e sistemas.

O tesouro DOCUTES tem um índice KWOC que permite localizar os descritores sintagmáticos ou pré coordenados pelo segundo e terceiro termos significativos. Este tesouro tem um total de 2.196 termos dos quais 1.612 são descritores e 584 não descritores. O Quadro 8 mostra os campos semânticos e a quantidade de descritores existentes no DOCUTES.

Quadro 8. Abrangência dos Campos Semânticos no DOCUTES

CAMPOS SEMÂNTICOS	DESCRITORES
Ciências da Documentação	191
Informação, Documentos e Fontes.	216
Investigação e metodologia documental	155
Representação e recuperação da informação	187
Sistemas de informação	471
Tecnologia da Informação	392
TOTAL	1612

Fonte: (BRAVO, [s/d])

3 METODOLOGIA DA PESQUISA

Utiliza como metodologia a pesquisa exploratória. Foi realizada revisão da literatura em Ciência da Informação sobre indexação automática e indexação automática por atribuição, a qual, configura o caráter teórico deste trabalho. Em termos pragmáticos, realiza estudo de caso ao usar o corpus composto por 60 artigos científicos referenciados na Tese de Doutorado de Souza (2005) ao comparar as palavras-chave destes artigos com os termos atribuídos pelo sistema SISA usando o vocabulário controlado Tesouro TBCI.

Vale lembrar que a pesquisa exploratória de senso qualitativo e cunho bibliográfico tem como objetivo, “proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a construir hipóteses” (GIL, 1996). Também, segundo Marconi e Lakatos (2010, p. 190), esse tipo de pesquisa, exploratória, tem uma tripla finalidade: “Desenvolver hipóteses, aumentar a familiaridade do pesquisador com o ambiente, fato ou fenômeno, para realização de uma pesquisa futura mais precisa ou modificar e clarificar conceitos”.

Por outro lado, a realização da pesquisa empírica tem sua justificativa concretizada no valor da informação obtida via análise do resultado da aplicação de método de indexação automática por atribuição no corpus de documentos compilado, contribuindo, assim, com pesquisas futuras em indexação automática.

O objeto de estudo desta pesquisa é a indexação automática por atribuição com o uso do tesouro TBCI e aplicação do sistema SISA na indexação de artigos de periódicos em Ciência da Informação.

Esta pesquisa consiste das seguintes etapas: 1) revisão de literatura sobre indexação automática por atribuição, etapa esta que dará aportes teóricos-metodológicos à pesquisa; 2) análise do uso de vocabulários controlados e métodos de indexação automática em bases de dados científicas na WEB. Foram acessadas e analisadas, as bases de dados: BRAPCI, PUBMED / MEDLINE, LISA, INSPEC, ENGINEERING INDEX COMPENDEX, SCOPUS, CHEM ABSTRACTS e ERIC; 3) Planejamento de experimento que permita a análise da qualidade da indexação automática por atribuição realizada pelo SISA através da avaliação do grau de consistência, revocação, precisão e medida F na indexação automática; 4) Construção do corpus, execução do experimento e análise dos resultados do experimento.

A seguir é apresentada a metodologia de aplicação do sistema SISA com o uso do vocabulário controlado TBCI, metodologia esta, adaptada do Capítulo 4 - Procedimentos Metodológicos, da Dissertação de Mestrado (NARUKAWA 2011).

Primeiro prepara-se os arquivos utilizados pelo sistema SISA que basicamente são três tipos de arquivos:

a) Linguagem de indexação: constando de dois arquivos: 1) lista de descritores elaborada a partir do tesauro TBCI, organizada em ordem alfabética, enumerada sequencialmente, onde os descritores estabelecem apenas a relação de equivalência por meio da indicação USE, arquivo criado através do TBCI convertido para o formato TXT (Anexo A) e 2) lista de Termos Gerais (TCITG) que é uma listagem de termos que permitem categorizar os termos incluídos no vocabulário controlado, neste arquivo representa-se a hierarquização dos termos do vocabulário controlado através das relações de superordenação, atendendo à uma eventual necessidade de agrupamento de termos, tendo sido criado através do TBCI convertido para o formato TXT, recortando as relações do tipo TG entre termos. (Anexo C).

b) Lista de palavras vazias: A lista de palavras vazias, em língua portuguesa, cedida pelo projeto de pesquisa Mapeador Temático de Teses e Dissertações (MTTD), coordenado pelo orientador da dissertação, foi adaptada a partir da conferência das palavras, removendo as que faziam parte da lista de descritores do TBCI. Vide Anexo B;

c) Artigos científicos: nesta dissertação foi compilado um corpus composto por 60 artigos científicos selecionados e utilizados por Renato Rocha Souza em sua Tese de Doutorado (2005) o qual é reproduzido no Apêndice A. No experimento serão utilizados os 60 artigos, formatados, para entrada no software SISA da seguinte forma: os arquivos são convertidos para o formato TXT; as referências dos artigos científicos são removidas; o título, o resumo e o texto completo são extraídos e delimitados por marcadores #CTI# (começo do título), #FTI# (fim do título), #CR# (começo do resumo), #FR# (fim do resumo), #CTE# (começo do texto) e #FTE# (fim do texto) estabelecidos pelo SISA. Vide Apêndice H um exemplo de um artigo formatado segundo critérios do SISA.

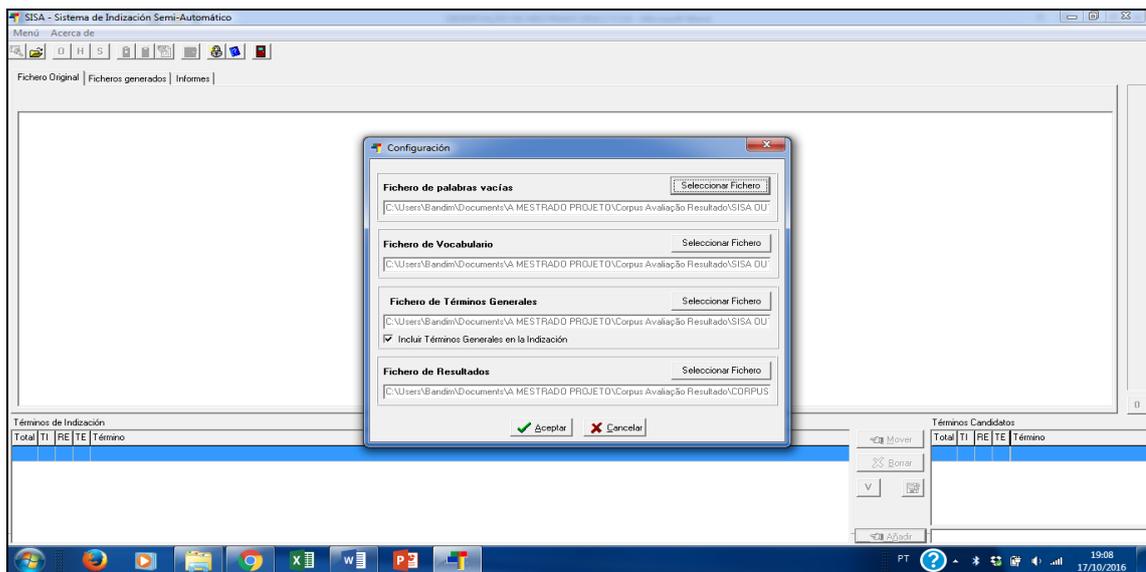
Os artigos referenciados acima são documentos textuais de língua portuguesa, publicados nos anos de 2002 e 2003 em duas revistas científicas eletrônicas da área da Ciência da Informação. As revistas são: DataGramZero e Ciência da Informação do IBICT.

A escolha destes artigos se deve ao fato de que, além de terem sido publicados em revistas renomadas que divulgam artigos na área da Ciência da Informação reconhecidas pelo programa QUALIS da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, apresentam os requisitos necessários à formatação exigida pelo SISA para executar a indexação automática, ou seja, contém título, resumo e texto completo em formato padronizado. Outro fator motivador é a possibilidade de realizar comparações entre os resultados alcançados neste trabalho com os de Souza (2005).

Após a formatação dos três arquivos de configuração do SISA, inicia-se a indexação automática dos artigos. Esta atividade pode ser dividida em três etapas, descritas a seguir.

Na primeira etapa é feita a configuração do SISA habilitando os arquivos: Lista de palavras vazias, Lista de descritores e Lista de Termos Gerais.Figura 8.

Figura 8. Tela de configuração dos arquivos do SISA



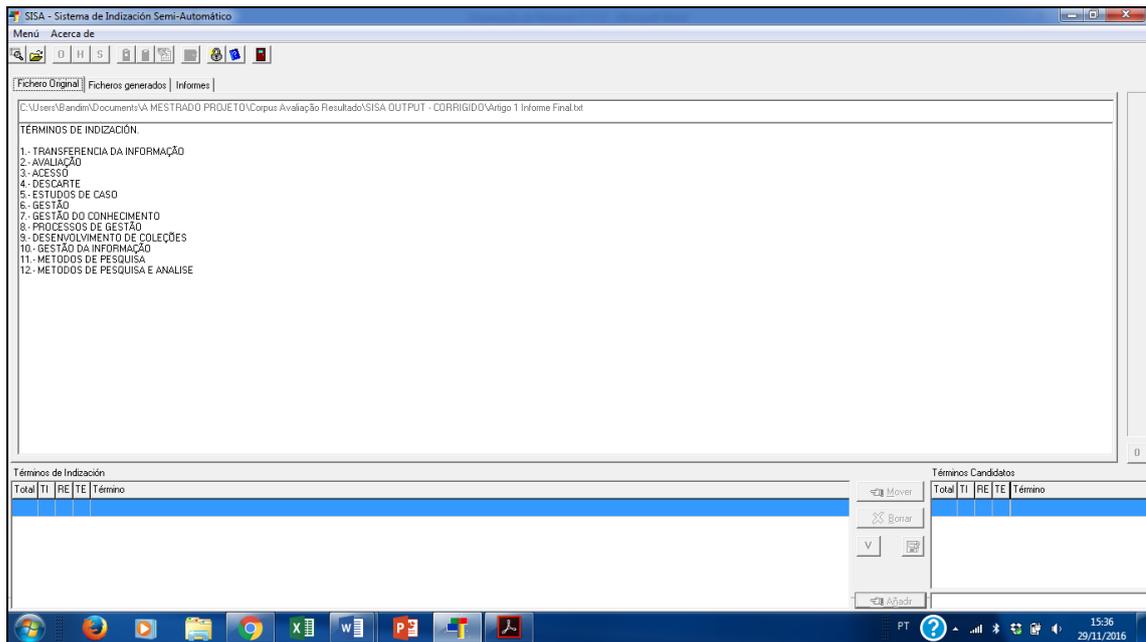
Fonte: Sistema de Indexação Semi-Automática (SISA).

Depois, na segunda etapa, os artigos (arquivos) são selecionados para indexação. Na Figura 9 tem-se a seleção dos arquivos para indexação pelo SISA. Na Figura 10 é mostrado um arquivo selecionado para indexação. E na Figura 11 tem-se a apresentação dos termos de indexação propostos pelo SISA.

Na terceira etapa é solicitada ao SISA a indexação do arquivo. Após a indexação de cada arquivo (artigo), o SISA mostra a tela onde são apresentados os termos propostos para indexação. Figura 11.

Por fim, repete-se todo o processo, a partir da segunda etapa, para cada artigo (arquivo) a ser indexado.

Figura 11. Apresentação dos termos de indexação propostos pelo SISA



Fonte: Sistema de Indexação Semi-Automática (SISA).

Após a indexação automática de cada artigo, os termos de indexação propostos pelo SISA foram dispostos num quadro onde constam também as palavras-chave dos respectivos artigos. Assim torna-se possível comparar os dois conjuntos de termos e avaliar a qualidade na indexação. Vide Apêndice F.

Em seguida, é realizada a análise de consistência, onde são comparados os termos de indexação atribuídos pelas palavras-chave, definidas pelos autores dos artigos científicos, com os termos de indexação atribuídos e propostos pelo SISA.

Neste trabalho são feitos dois tipos de avaliação dos resultados: avaliação intrínseca e extrínseca.

Baseado nos procedimentos de avaliação de indexação definidos por Gil Leiva (1999), a avaliação intrínseca, medirá o grau de consistência da indexação e, a avaliação extrínseca, medirá o grau de revocação, precisão e medida F (média harmônica ponderada entre a precisão e revocação) na indexação automática.

O índice de revocação é obtido através da relação entre os termos relevantes recuperados e o total de termos relevantes existentes para cada artigo. Equação (1).

$$\text{Revocação} = \frac{\text{Número de termos relevantes recuperados}}{\text{Número total de termos relevantes}} \quad \text{Equação (1)}$$

Já para se obter os valores para o índice de precisão calcula-se a relação entre os termos relevantes recuperados e o total de termos recuperados. Equação (2).

$$\text{Precisão} = \frac{\text{Número de termos relevantes recuperados}}{\text{Número total de termos recuperados}} \quad \text{Equação (2)}$$

A medida F é a média harmônica ponderada entre o índice de precisão (P) e o índice de revocação (R), sendo uma maneira de combinar a precisão e revocação em um único número. Se nenhum termo relevante foi recuperado a medida F assume valor 0 e, quando todos os termos recuperados são relevantes e foram exaustivamente recuperados todos os termos relevantes, assume valor 1.

É preciso lembrar que a medida F só assume valores altos quando os índices de revocação e precisão são altos também e, a procura em determinar um valor máximo para medida F é visto como uma tentativa de se encontrar um *equilíbrio* entre estes dois índices. A equação (3) é usada para o cálculo da medida F.

$$\text{Medida F} = \frac{2 \times (\text{Precisão} \times \text{Revocação})}{\text{Precisão} + \text{Revocação}} \quad \text{Equação (3)}$$

As métricas de qualidade na indexação serão avaliadas fazendo-se a indexação automática por atribuição dos 60 artigos científicos já citados anteriormente. Para tanto o SISA será configurado habilitando-se os arquivos: lista de descritores TBCI, lista de Palavras Vazias (stopwords), lista de termos gerais os artigos formatados em TXT, sendo estes devidamente sinalizados com os marcadores exigidos pelo SISA. Após feita as indexações, os termos das palavras-chave dos artigos e os termos de indexação atribuídos pelo SISA serão organizados em um quadro comparativo, Apêndices F e G, o que possibilitará a aplicação da fórmula matemática de consistência.

A indexação dos artigos será feita em dois cenários: Cenário I – com a lista de termos gerais (TCITG) habilitada na configuração do SISA e Cenário II – sem habilitar esta lista. Isto se deve ao fato de que, como se verá adiante, a habilitação ou não desta lista tem influência nos índices de consistência, revocação, precisão e, conseqüentemente, na medida F. A equação (4) é usada para o cálculo dos índices de consistência. $Tco =$ Número de termos

comuns nas duas indexações; A = Número de termos usados na indexação A e B = Número de termos usados na indexação B.

$$C_i = \frac{Tco}{(A + B) - Tco} \text{ Equação (4)}$$

Vale lembrar que, segundo Narukawa (2011), esta fórmula com algumas variações, foi utilizada por Salton e McGill (1983), Lancaster (1991), Silvester Genuardi y Klingbiel (1994) e Gil Leiva (1999) para verificar a consistência entre a indexação manual e a automática. Mais recentemente, foi aplicada nas pesquisas de Soler Monreal (2009); Oliveira (2009); Narukawa, Gil Leiva e Fujita (2009); Gil Leiva, Rubi e Fujita (2008); e Gil Leiva (2001; 2002).

No próximo capítulo será apresentada a análise dos resultados.

4 ANÁLISE DOS RESULTADOS

Será apresentada neste capítulo a análise dos resultados do experimento, envolvendo a avaliação da qualidade da indexação automática por atribuição por meio dos índices de consistência, revocação, precisão e medida F.

4.1 Indexação automática com atribuição de termos gerais

Os resultados do experimento no Cenário I (Apêndice G), isto é, com a lista de Termos Gerais habilitada na configuração do SISA, são descritos a seguir.

Na avaliação dos termos presentes no campo palavras-chave, houve um mínimo de 2 e um máximo de 9 termos atribuídos para indexação, dando uma média de 4 a 5 descritores atribuídos por documento. Já com relação aos termos propostos pelo SISA houve um mínimo de 3 e um máximo de 24 descritores atribuídos, o que dá uma média de 9 a 10 descritores ou termos de indexação por documento.

O número de termos comuns variou de um mínimo de 0 (zero) termos comuns a um máximo de 5 termos comuns.

Assim, foi constatada uma média de 15 % nos índices de consistência, com uma variação de um mínimo de 0% a um máximo de 42 %. Vide Apêndice G.

Estas médias influenciam diretamente os índices de consistência uma vez que a quantidade de descritores que são atribuídos em cada tipo de indexação, é considerada nos cálculos de consistência de indexação.

Segundo Narukawa (2011), os diversos estudos realizados sobre a indexação automática por atribuição, mostram que os índices de consistência obtidos, estão, aproximadamente, entre 10% e 60% (LANCASTER, 1968; LEONARD, 1975; FUNK e REID, 1983; MARKEY, 1984; MIDDLETON, 1984; TONTA, 1991; SIEVER; ANDREWS, 1991; IIVONEN; KIVIMÄKI, 1998; LEININGER, 2000; LEIVA, 2001 e 2002; SAARTI, 2002; NESHAT e HORRI, 2006; LEIVA; RUBI; FUJITA, 2008 e KIPP, 2009).

Portanto, a média de 15% obtida no índice de consistência para o corpus dos 60 artigos indexados nesse cenário do experimento, está dentro do intervalo verificado nos estudos realizados sobre índices de consistência na indexação automática por atribuição, estudos estes, cujos autores foram citados no parágrafo anterior.

Existem alguns fatores que influenciam a atribuição de termos de indexação pelo SISA, lembrando que para entender estes fatores se faz necessário considerar que o SISA somente atribui os termos que constam no vocabulário controlado por um processo de comparação com os termos do artigo, considerando a presença destes termos no título, resumo e texto do artigo analisado. Entre tais fatores podem ser citados: termos no singular e plural; dificuldades em atribuir termos compostos; termos escritos entre aspas, hifens e parênteses, entre outros (NARUKAWA, 2011).

Quanto aos índices de revocação, precisão e medida F obtidos neste experimento foram obtidos os seguintes resultados: Precisão média de 20 %; Revocação média de 42 % e Medida F média de 25 %. Vide Apêndice D.

4.2 Indexação automática sem atribuição de termos gerais

A seguir são mostrados os resultados obtidos para os índices de consistência, precisão, revocação e medida F no Cenário II, isto é, sem a atribuição de Termos Gerais (TCITG) pelo SISA (Apêndice F).

Após ser desabilitado o arquivo de Termos Gerais e se executar novamente a indexação dos 60 artigos, os termos propostos pelo SISA mudaram em quantidade, afetando diretamente os índices de consistência na indexação, bem como, os de precisão, revocação e medida F.

Conforme pode ser verificado, os índices de consistência foram otimizados quando comparados com os índices obtidos na indexação com o SISA configurado com a habilitação dos Termos Gerais (Cenário I).

Neste segundo cenário a média dos índices de consistência foi de 19%, o mínimo de 0% e máximo de 71%, ficando o índice médio, mais uma vez, dentro dos valores registrados nos estudos e pesquisas, anteriormente citados.

As médias dos outros índices foram: Revocação média de 37 %; Precisão média de 30% e Medida F média de 30%. Vide Apêndice E. Na próxima subseção 4.3, discutiremos os resultados globais do experimento.

4.3 Discussões dos resultados

Avaliando-se os resultados globais do experimento pode-se afirmar que, comparando os termos de indexação automática atribuídos pelo SISA com os termos atribuídos pelas

palavras-chave e, levando em consideração os dois cenários citados, verificou-se que o SISA atribuiu uma grande quantidade de termos de indexação, influenciando diretamente os cálculos dos índices de consistência na indexação e, os índices de precisão, revocação e medida F. Este elevado número de termos é explicado no Cenário I porque o SISA atribui além dos termos específicos, os termos gerais também.

Quadro 9. Índices médios de consistências na indexação e, de precisão, revocação e medida F nos cenários I e II.

	Cenário I TCITG habilitado	Cenário II TCITG desabilitado
Consistência na Indexação	15 %	19 %
Precisão	20 %	30 %
Revocação	42 %	37 %
Medida F	25 %	30 %

Fonte: (Dados da pesquisa, 2016).

No Quadro9 pode ser visto a influência direta nos cálculos dos índices da habilitação ou não do TCITG.

Neste experimento, foi obtido para o índice de precisão, uma média de 20% com o TCITG habilitado e uma média de 30% desabilitado. A não habilitação da atribuição do termo geral TCITG teve influência direta no índice de precisão, passando de 20% para 30%, melhorando a precisão na atribuição de termos relevantes, evitando que termos gerais diferentes das palavras-chaves fossem atribuídos.

Quanto ao índice de revocação, a não habilitação do TCITG resultou em uma média menor para este índice, o qual passou de 42% (TCITG habilitado) para 37% (TCITG desabilitado), uma perda de cerca de 5% de termos relevantes em média para todo o corpus.

Estes valores influíram diretamente nos percentuais médios da medida F, cuja fórmula nos permite calcular a média harmônica entre o índice de precisão e o índice de revocação, sendo uma maneira de combinar a precisão e revocação em um único número. Para se conseguir valores médios altos para F é necessário que se tenha valores altos para revocação e precisão.

Neste contexto, o aumento da média do índice de precisão influenciou diretamente na média percentual da medida F que passou de 25% para 30%. Como o objetivo da medida F é conseguir um equilíbrio entre precisão e revocação, neste experimento este objetivo foi

atingido, mesmo considerando o decréscimo da média do índice de revocação que passou de 42% para 37% que foi compensado pelo aumento, em termos percentuais médios, do índice de precisão que passou de 20% para 30%.

É preciso lembrar que existem alguns fatores chamados de fatores intervenientes, que agem diretamente nos processos de indexação. Segundo Gil Leiva (2008) esses fatores estão diretamente associados ao indexador, uma vez que, tanto o indexador automático como o humano têm alguns critérios definidos como: o contexto da indexação, ou seja, a política e os objetivos da indexação; à diversidade e às características do objeto informacional a ser indexado e, ainda, os fatores relacionados ao momento da indexação, isto é, na comparação entre palavras-chave retiradas diretamente do texto e aquelas retiradas de um vocabulário controlado.

Desta maneira, segundo Narukawa (2011), o SISA somente atribui termos do artigo científico que estão no vocabulário controlado. Assim aplicando-se um processo de “comparação de padrões” os termos de indexação são selecionados levando-se em consideração sua frequência (aparecimento) no título, resumo e texto, que constituem os artigos (vide Subseção 2.5.2 para mais detalhes).

No experimento desta pesquisa foram identificados alguns destes fatores intervenientes, os quais, influenciaram os índices de consistência. (Quadro 10).

Quadro 10. Fatores intervenientes no processo de indexação por atribuição

Fatores Intervenientes	Termos Palavras-chave	Termos SISA
Termos no singular e no plural	- <i>Biblioteca virtual</i> - <i>Estratégia de busca</i> - <i>Biblioteca híbrida</i>	- <i>Bibliotecas virtuais</i> - <i>Estratégias de busca</i> - <i>Bibliotecas híbridas</i>
Dificuldade em atribuir termos compostos	- <i>Compressão semântica</i> - <i>Ciência da informação no Brasil</i> - <i>Internet e produção científica</i>	- <i>Semântica</i> - <i>Ciência da informação</i> - <i>Internet</i>
Diferenças entre as estruturas dos termos de indexação	- <i>Recuperação de informação</i> - <i>Profissional da informação</i> - <i>Transferência de informação</i>	- <i>Recuperação da informação</i> - <i>Profissional de informação</i> - <i>Transferência da informação</i>
Dificuldades em atribuir conceitos implícitos	- <i>Arquivos-abertos</i> - <i>Sistema de publicação</i> - <i>Interação humano-computador (IHC)</i>	- <i>Acesso</i> - <i>Tipos de documentos</i> - <i>Buscas de informação</i>
Atribuição automática de termo geral e de termo específico	- <i>Informação</i> - <i>Ciência</i> - <i>Caos</i>	- <i>Direito a informação</i> - <i>Ciência da Informação</i> - <i>Teoria do caos</i>

Fonte: (DO AUTOR, 2016).

No caso de termos no singular e no plural foram constatadas variações de um mesmo conceito como por exemplo *biblioteca virtual* e *bibliotecas virtuais*, isto é: o termo do artigo não foi atribuído porque no vocabulário controlado este termo está no singular.

Outro exemplo de fator interveniente é o caso do SISA ter dificuldades em atribuir termos compostos como aconteceu na indexação automática de um documento onde o autor usou o termo composto – *compressão semântica* - para delimitar o assunto que será discutido no artigo e, no decorrer do artigo usou somente a primeira palavra – *semântica* – que compõe o termo composto. O SISA atribuiu o termo *semântica*. (Quadro 10).

Enfim, é preciso lembrar que esta questão, a dos chamados fatores intervenientes não constituem erros do sistema SISA e, sim, estão a sugerir melhorias e/ou aprimoramentos ou ainda adaptações nos algoritmos do SISA para que possa selecionar e atribuir mais termos importantes na indexação automática por atribuição com o uso de vocabulário controlado.

Portanto, na indexação automática por atribuição, os índices de consistência são influenciados pelos ditos fatores intervenientes, os quais, estão associados às questões que caracterizam as diferenças morfológicas, linguísticas e semânticas, considerando a aplicação dos vocabulários controlados que no nosso caso é o tesouro TBCI. Isto acontece porque no processo de análise documentária automática do conteúdo do documento pelo sistema SISA, estes vocabulários condicionam a seleção automática dos termos que serão atribuídos para indexação.

Na indexação automática por atribuição feita pelo SISA com o arquivo TCITG habilitado é mostrado no Quadro 11 os seis melhores casos e no Quadro 12 os seis piores casos. (Cenário D). Mais uma vez verifica-se claramente a influência direta nos valores médios dos índices quando a indexação é feita habilitando-se este arquivo.

Quadro 11. Os seis melhores casos de indexação feita pelo SISA com TCITG habilitado. (Cenário D).

Artigo Científico	No. de Palavras-chave	No. de termos do SISA	No. de Termos comuns	Consistência Relaxada	Precisão	Revocação	Medida F
Artigo 9	5	4	2	29%	50%	40%	44%
Artigo 15	2	5	1,5	27%	30%	75%	43%
Artigo 16	7	7	3,5	33%	50%	50%	50%
Artigo 29	4	9	3	30%	33%	75%	46%
Artigo 44	3	6	2	29%	33%	67%	44%
Artigo 58	7	10	5	42%	50%	71%	59%

Fonte: (DO AUTOR, 2016).

Quadro 12. Os seis piores casos de indexação feita pelo SISA com TCITG habilitado. (Cenário I).

Artigo Científico	No. de Palavras-chave	No. de termos do SISA	No. de Termos comuns	Consistência Relaxada	Precisão	Revocação	Medida F
Artigo 5	5	4	0,5	6%	13%	10%	11%
Artigo 10	5	5	0,5	5%	10%	10%	10%
Artigo 23	3	3	0	0%	0%	0%	0%
Artigo 31	6	7	0,5	4%	7%	8%	8%
Artigo 34	3	4	0	0%	0%	0%	0%
Artigo 49	3	11	0,5	4%	5%	17%	7%

Fonte: (DO AUTOR, 2016).

Já com o arquivo TCITG não habilitado é mostrado no Quadro 13 os sete melhores casos Cenário I e no Quadro 14 os seis piores casos. (Cenário II).

Quadro 13. Os sete melhores casos de indexação feita pelo SISA com TCITG não habilitado (Cenário II).

Artigo Científico	No. de Palavras-chave	No. de termos do SISA	No. de Termos comuns	Consistência Relaxada	Precisão	Revocação	Medida F
Artigo 15	2	3	1,5	43%	50%	75%	60%
Artigo 16	7	4	3,5	47%	88%	50%	64%
Artigo 29	4	5	3	50%	60%	75%	67%
Artigo 35	5	8	3	30%	38%	60%	46%
Artigo 39	5	8	3	30%	38%	60%	46%
Artigo 58	7	5	5	71%	100%	71%	83%
Artigo 60	4	8	3,5	41%	44%	88%	58%

Fonte: (DO AUTOR, 2016).

Quadro 14. Os seis piores casos de indexação feita pelo SISA com TCITG não habilitado (Cenário II).

Artigo Científico	No. de Palavras-chave	No. de termos do SISA	No. de Termos comuns	Consistência Relaxada	Precisão	Revocação	Medida F
Artigo 5	5	3	0	0%	0%	0%	0%
Artigo 10	5	2	0	0%	0%	0%	0%
Artigo 19	5	5	0	0%	0%	0%	0%
Artigo 23	3	1	0	0%	0%	0%	0%
Artigo 31	6	3	0	0%	0%	0%	0%
Artigo 34	3	3	0	0%	0%	0%	0%

Fonte: (DO AUTOR, 2016).

Comparando os percentuais obtidos nos dois cenários, verifica-se claramente, que não só os índices de consistência como os índices de precisão, revocação e, conseqüentemente, a medida F foram fortemente alterados.

Por exemplo, analisando o artigo 15 que aparece entre “os melhores casos” nos dois cenários vê-se que o índice de consistência passou de 27% no Cenário I para 43% no Cenário II isto porque o número de termos atribuídos pelo SISA foi menor no Cenário II (três termos) do que no Cenário I (cinco termos). (Quadro 15).

Quadro 15. Análise comparativa dos melhores índices obtidos na indexação de um mesmo artigo nos Cenários I e II.

Artigo Científico 15	No. de Palavras-chave	No. de termos do SISA	No. de Termos comuns	Consistência Relaxada	Precisão	Revocação	Medida F
Cenário I	2	5	1,5	27%	30%	75%	43%
Cenário II	2	3	1,5	43%	50%	75%	60%

Fonte: (DO AUTOR, 2016).

Já analisando os índices obtidos na indexação dos artigos que são entre os piores casos, por exemplo para o artigo 10 (Quadro 16), o índice de consistência obtido no Cenário I foi de 5% e no cenário II foi de 0%.

Quadro 16. Análise comparativa dos piores índices obtidos na indexação de um mesmo artigo nos Cenários I e II.

Artigo Científico 10	No. de Palavras-chave	No. de termos do SISA	No. de Termos comuns	Consistência Relaxada	Precisão	Revocação	Medida F
Cenário I	5	5	0,5	5%	10%	10%	10%
Cenário II	5	2	0	0%	0%	0%	0%

Fonte: (DO AUTOR, 2016).

Estes valores do Quadro 16 indicam que a habilitação do TCITG produziu uma pequena melhora no índice de consistência para a maioria dos piores casos, isto porque a atribuição dos termos gerais possibilitou um número de termos comuns com as palavras-chaves igual a 0,5 usando o critério de consistência relaxada. Entretanto, para que este 0,5 termo comum fosse encontrado, foram atribuídos pelo SISA muitos termos gerais aos documentos que não tinham nenhuma semelhança com as palavras-chaves. De qualquer forma, estes valores de consistência estão muito abaixo dos valores percentuais médios dos

índices de consistência obtidos na indexação dos 60 artigos os quais foram de 15% no Cenário I, com TCTTG habilitado e, 19% no Cenário II sem TCITG.

É interessante perceber que analisando os melhores casos nos dois cenários fica claro que a inclusão do TCITG diminui em muito os valores percentuais dos índices de consistência, isto porque o número de termos comuns permaneceu o mesmo, mas o número de termos atribuídos pelo SISA aumentou. No Quadro 17 tem-se esta conclusão, inclusive, considerando o exemplo dos mesmos artigos nos dois cenários. Só lembrando: o número de termos comuns (Tco) é parâmetro fundamental na fórmula de cálculo do índice de consistência. (Vide Seção 3).

Quadro 17. Análise do índice de consistência dos melhores casos nos Cenários I e II

Artigos	Índices de Consistência Cenário I	Índices de Consistência Cenário II	No. de Termos comuns Cenários I e II	No. de Termos SISA Cenário I	No. de Termos SISA Cenário II
15	27 %	43 %	1,5	5	3
16	33 %	47 %	3,5	7	4
29	30 %	50 %	3	9	5

Fonte: (DO AUTOR, 2016).

Fazendo-se a mesma análise realizada no parágrafo anterior, mas desta vez, com os piores casos verifica-se nos dois cenários que há discretas variações nos valores percentuais dos índices de consistência, mas ainda assim, estes valores estão muito abaixo dos valores percentuais médios dos índices de consistência obtidos na indexação dos 60 artigos os quais foram de 15% no Cenário I, com TCTTG habilitado e, 19% no Cenário II sem TCITG. (Quadro 18).

Aqui, mais uma vez, é verificado a influência do parâmetro de termos comuns (Tco) no cálculo dos índices de consistência.

Quadro 18. Análise do índice de consistência dos piores casos nos Cenários I e II

Artigos	Índices de Consistência Cenário I	Índices de Consistência Cenário II	No. de Termos comuns Cenário I	No. de Termos comuns Cenário II	No. de Termos SISA Cenário I	No. de Termos SISA Cenário II
5	6%	0%	0,5	0	4	3
10	5%	0%	0,5	0	5	2
23	0%	0%	0	0	3	1
31	4%	0%	0,5	0	7	3
34	0%	0%	0	0	4	3

Fonte: (DO AUTOR, 2016).

Enfim, a habilitação do arquivo TCITG não gera bons resultados em termos de qualidade para indexação automática por atribuição.

5 CONSIDERAÇÕES FINAIS

Considerando os estudos realizados neste trabalho de pesquisa e tendo por base os aspectos teóricos-metodológicos bem como os resultados das experiências pragmáticas que foram realizadas, pode-se afirmar que os objetivos foram atingidos e as questões da problemática foram respondidas. Nos próximos parágrafos são discutidos estes pontos.

Começando com os objetivos específicos temos:

a) Identificar e descrever os métodos, as técnicas e instrumentos de indexação automática por atribuição.

Os métodos e instrumentos de indexação automática são discutidos nas seções 2.1 a 2.6 quando, entre outros pontos, descrevemos o contexto e definição da indexação automática e indexação automática por atribuição com o uso de vocabulários controlados como a escolha do vocabulário controlado (TBCI) que melhor atende os requisitos dos termos que irão representar o conteúdo dos documentos armazenados nas bases de dados e o sistema SISA usado para indexação automática por atribuição.

b) Investigar as principais bases de dados que fazem uso de instrumentos de controle terminológico, como o uso de vocabulários controlados e de indexação automática por atribuição.

A análise e identificação das principais linguagens de indexação – cabeçalhos de assuntos e vocabulários controlados - usadas na indexação automática por atribuição de artigos científicos em bases de dados científicas tratadas nas seções 2.5 e 2.6 como por exemplo a base de dados BRAPCI, que usa a indexação automática no processo de coleta de termos comparados com Lista de Cabeçalhos de Assuntos (remissiva); a PUBMED/MEDLINE que usa a indexação automática por atribuição e, também, indexação assistida por computador aplicando o tesouro *Medical Subject Heading* (MeSH), entre outras.

c) Elaborar o processo de indexação automática por atribuição utilizando o software SISA e TBCI para artigos científicos da área de Ciência da Informação em português.

Foi elaborado processo de indexação automática por atribuição, discutido principalmente na Seção 3, fazendo-se uma revisão do instrumental teórico-metodológico sobre indexação automática por atribuição e considerando o cenário composto pelo sistema de indexação SISA, o tesouro TBCI, o corpus dos artigos científicos objeto da indexação e os três arquivos formatados em TXT: lista de palavras vazias, lista de descritores TBCI e lista de termos gerais (TCITG).

d) Aplicar a indexação automática do SISA com vocabulário controlado TBCI num corpus de 60 artigos científicos, comparando os termos propostos para indexação com a indexação manual realizada pelos autores através dos termos das palavras-chave.

Foi realizada a indexação automática por atribuição e feita a comparação entre os termos das palavras-chave e os atribuídos pelo SISA.

e) Analisar os resultados obtidos da indexação automática por atribuição feita pelo SISA considerando o cenário do experimento descrito acima, avaliando os índices de consistência, precisão, revocação e medida F.

A Análise dos Resultados, descrita em detalhes na Seção 4, mostrou a influência direta da habilitação do TCITG nas médias de consistência, precisão, revocação e medida F e permitiu comprovar que a indexação automática por atribuição gera qualidade nos resultados do processo de indexação. Isto está demonstrado não só na avaliação intrínseca via parâmetros de consistência como também na avaliação extrínseca via parâmetros de revocação, precisão e medida F.

Quanto as questões da Problemática :

1) Quais são os tipos de indexação automática e instrumentos de controle terminológicos - os vocabulários controlados - utilizados por bases de dados científicas na indexação de assuntos?

Neste trabalho foi constatado que os tipos de indexação usados são a indexação automática, semiautomática e automática por atribuição e, também, em algumas bases de dados, a indexação assistida por computador como por exemplo a base de dados ERIC. Quanto aos instrumentos de controle terminológico, verificou-se que o mais usado é o Tesouro como por exemplo nas bases BRAPCI, PubMed/MedLine, entre outras.

2) Quais tesouros da área de Ciência da Informação contem termos em português e podem ser utilizados na composição do processo de indexação automática por atribuição?

Nesta questão identificou-se e foram investigados os tesouros Tesouro Brasileiro da Ciência da Informação (TBCI/IBICT) e Tesouro em Ciência da Informação (PUC-MINAS), tendo optado pelo primeiro devido a atualidade.

3) Como realizar o processo de indexação automática por atribuição utilizando o software SISA para artigos científicos em português da área de Ciência da Informação?

Todo o processo foi detalhado na Seção 3 - Metodologia da Pesquisa. Processo este aplicado no experimento, no qual, foi realizada a indexação automática por atribuição do corpus formado por 60 artigos científicos, utilizados por Souza (2005) como também,

utilizando o tesauro TBCI, o sistema de indexação SISA e os termos das palavras-chave dos mesmos artigos acima referenciados, como padrão de referência.

4) No contexto de uma base de dados de artigos científicos em Ciência da Informação, a utilização da indexação automática por atribuição com base no tesauro TBCI e o uso do software SISA gera qualidade na indexação?

Como resultados deste experimento - Vide Seção 4 - é mostrado, considerando os dois cenários: Cenário I, com o arquivo de Termos Gerais (TCITG) e, Cenário II, sem habilitar este arquivo, que o uso do tesauro TBCI junto com o SISA, pode, sim, gerar qualidade na indexação automática, principalmente levando em consideração os índices de consistência obtidos neste experimento. Nas Seções 4.1 a 4.3 foi detalhada a obtenção destes índices e os significados dos valores obtidos.

Foi visto que a utilização da indexação automática por atribuição com base no tesauro TBCI e o uso do software SISA atende as necessidades de indexação de artigos científicos uma vez que os percentuais médios dos índices de consistência, que ficaram entre 15% e 19%, estão dentro de valores obtidos nos estudos e pesquisas de vários autores, conforme mostrados na seção 4.1.

Deve ser registrado que no sentido de melhorar os índices de consistência, revocação, precisão e medida F obtidos neste experimento e em outros processos de indexação automática por atribuição, devem ser consideradas algumas sugestões, dadas a seguir, resguardadas as políticas, os objetivos e os ambientes de indexação de cada caso.

Uma das limitações deste trabalho é que nos 60 artigos analisados, existem alguns com apenas duas palavras-chave e, outros, com mais de 8 palavras-chave. Tal limitação pode ser resolvida dedicando-se mais tempo na análise documentária e se indexando os trabalhos com o uso do TBCI.

Em termos de trabalhos futuros é muito importante continuar os estudos e pesquisas sobre indexação automática por atribuição com a aplicação de outras linguagens de indexação como ontologias, objetivando atender as necessidades semânticas quando da seleção automática dos termos de indexação. Isto porque como visto nesta pesquisa, o SISA via vocabulário controlado tesauro, faz a seleção automática dos termos de indexação usando a frequência relativa dos termos existentes nos documentos, não levando em consideração os aspectos morfológicos, sintáticos e semânticos relativos a equivalência e especificidade dos termos atribuídos para indexação.

Portanto, conclui-se que, mesmo tendo atingido todos os objetivos e respondida todas as questões referentes ao processo de indexação automática por atribuição de artigos

científicos propostos neste trabalho, concorda-se com a opinião de alguns autores e pesquisadores que defendem o uso tanto da indexação automática como da indexação manual. Isto se deve ao fato de que sempre será necessária a intervenção do profissional indexador no sentido de “entender” e “decidir” qual o melhor termo ou termos que representarão determinado objeto informacional seguindo, claro, as diretrizes e política informacional da instituição.

Por fim, também é possível concluir que neste tipo de indexação feita pelo sistema SISA, a habilitação do arquivo de termos gerais TCITG tem forte influência nos índices de qualidade os quais foram analisados pelos valores médios obtidos pelos indicadores de qualidade ou índices de consistência. Foi constatado ainda, que o uso de palavras-chave como termos de indexação precisa ser melhor estudado e pesquisado no sentido de se criar instrumentos que propiciem aos autores dos artigos científicos melhor definição conceitual destes termos e maior controle terminológico.

REFERÊNCIAS

- ABNT - Associação Brasileira de Normas Técnicas. **NBR 12676/1992**. Disponível em: <<http://docs11.minhateca.com.br/387786954,BR,0,0,NBR-12676.pdf>>. Acesso em: 20 jun. 2016.
- AITCHISON, J.; GILCHRIST, A. G. **Manual para construção de tesouros**. Rio de Janeiro: BNG / Brasilart, 1979.
- ARAÚJO J. R. H. **Precisão no processo de busca e recuperação da informação**. Brasília: Thesaurus, 2007.
- BORGES, G. S. B. **Indexação automática de documentos textuais**: proposta de critérios essenciais. Dissertação de Mestrado em Ciência da Informação. Programa de Pós-Graduação em Ciência da Informação, Escola de Ciência da Informação, Universidade Federal de Minas Gerais - UFMG, Belo Horizonte, 2009.
- BRAPCI. **Base de dados referenciais de artigos de periódicos em ciência da informação**. Disponível em: <<http://www.brapci.ufpr.br/brapci/>>. Acesso em 18 jun. 2016.
- BRAVO, B. R. **Docutes**: tesouro de ciencias de la documentación. [s/d]. Disponível em: <http://eprints.rclis.org/5875/1/TESAURO_DOCUTES.pdf>. Acesso em: 19 jan. 2017.
- BRITO, M. de. Sistemas de informação em linguagem natural: em busca de uma indexação automática. **Ciência da Informação**. Brasília, v. 21, n. 3, p. 223-232, set/dez. 1992.
- CAMPOS, A. T. Linguagens documentárias. **Revista Biblioteconomia de Brasília**. Brasília, v. 14, n. 1, p. 85-88, 1986.
- CAMPOS, M. L. de A. **Linguagens documentárias**: teorias que fundamentam sua elaboração. Niterói, RJ: EDUFF, 2001.
- CAMPOS, M. L. de A.; GOMES, H. E.; MOTTA, D. F. da. **Elaboração de tesouro documentário**: Tutorial. 2006. Disponível em: <<http://www.conexao rio.com/bit/tesouro/>>. Acesso em: 18 jan. 2017.
- CESARINO, M. A.; PINTO, M. C. M. F. Cabeçalhos de assunto como linguagem de indexação. **Revista da Escola de Biblioteconomia UFMG**. Belo Horizonte, v. 7, n. 2, p. 268-288, set. 1978.
- CINTRA, A. M. M. *et al.* **Para entender as linguagens documentárias**. 2. ed. rev. atual. São Paulo: Polis, 2002. 96 p.
- CURRÁS, E. **Ontologias, taxonomias y tesouros**: manual de construcción y uso. 3. ed. Gijón: Trea, 2005. 337 p.
- DAHLBERG, I. Teoria do conceito. **Ciência da informação**. Rio de Janeiro, v. 7, n. 2, p. 101-107, jul./dez. 1978.

- FOSKETT, D. T. **Subject and information analysis**. New York: M. Dekker, 1985.
- 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.
- FUJITA, M. S. L. **A representação documentária no processo de indexação com o modelo de leitura documentária para textos científicos e livros: uma abordagem cognitiva com protocolo verbal**. Ponto de Acesso, Salvador, v. 7, n. 1, p. 42-66, abr. 2013.
- FUJITA, M. S. L.; LEIVA, I. G. As linguagens de indexação em bibliotecas nacionais, arquivos nacionais e sistemas de informação na América Latina, 2010. **XVI Seminário Nacional de Bibliotecas Universitárias y II Seminário Internacional de Bibliotecas Digitais**. [Conference paper]. Rio de Janeiro, Brazil, october 2010.
- FUNK, M. E.; REID, C. A. Indexing consistency. *In: MEDLINE. Bulletin of the Medical Library Association*, v. 71, p. 176-183, 1983.
- GOMES, H. E. O indexador face às novas tecnologias de informação. **Transinformação**, v. 1, n. 2, p. 161-171, maio/ago. 1989.
- GONZÁLEZ, J. A. M. **El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural**. Gijón (Astúrias): Trea, 2004.
- GUIMARÃES, J.A.C. **A análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais**. In: RODRIGUES, J.M.; LOPES, I.L. (org.) **Organização e representação do conhecimento na perspectiva da Ciência da informação**. Brasília: Thesaurus, 2003. (Estudos avançados em Ciência da informação; 2). p. 100-117.
- GUINCHAT, C.; MENOU, M. **Introdução geral às ciências e técnicas da informação e documentação**. 2. ed. corr. aum. Brasília: IBICT, 1994.
- IIVONEN, M.; KIVIMAKI, K. **Common entities and missing properties: similarities and differences in the indexing of concepts**. Knowledge Organization, v. 25, n. 3, p. 90-102, 1998.
- KIPP, M. E. I. Exploring Measures of Inter-Tagger Consistency. **SIG-CR Workshop Poster, Annual Meeting of the American Society for Information Science and Technology**, 2009, Vancouver, British Columbia, Canada. Disponível em: <<http://eprints.rclis.org/17218/>>. Acesso em: 03 jan. 2017.
- LAKATOS, E. M.; MARCONI, M. A. **Fundamentos de metodologia científica**. São Paulo: Atlas, 2010.
- LANCASTER, F. W. **El control del vocabulário en la recuperación de información**. 2. ed. Saragoza: Universidad de València; 2002. 281 p.
- LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos Livros, 2004. 452p.
- LANCASTER, F. W. **Vocabulary control for information retrieval**. 2. ed. Virginia : IRP, 1986.

LANCASTER, F.W. **Evaluation of the MEDLARS demand search service.** Washington, D.C. National Library of Medicine, 1968.

LAPA, R. **Indexação Automática no Brasil no âmbito da Ciência da Informação (1973-2012).** 2014. 287 f. Dissertação (Mestrado) - Curso de Ciência da Informação, Departamento de Ciência da Informação, Universidade Federal de Pernambuco, Recife, 2014.

LAPA, R.; CORRÊA, R. F. Indexação Automática no Âmbito da Ciência da Informação no Brasil. **Informação & Tecnologia (itec).** Marília/João Pessoa, v. 2, n. 1, p.1-18, 1 jul. 2014. Semestral.

LEININGER, K. Interindexer consistency in PsycINFO. **Journal of Librarianship and Information Science**, v. 32, n. 1, p. 4-8, 2000.

LEIVA, I. G. **Consistencia en la asignación de materias en bibliotecas públicas del Estado.** Boletín de la Asociación Andaluza de Bibliotecarios, n. 63, p. 69-96, 2001. Disponível em: <<http://webs.um.es/isgil>>. Acesso em: 13 jun. 2016.

LEIVA, I. G. **Consistencia en la indización de documentos entre indizadores noveles.** Anales de Documentación, v. 5, p. 99-111, 2002. Disponível em: <<http://webs.um.es/isgil>>. Acesso em 14 nov. 2016.

LEIVA, I. G. **La automatización de la indización de documentos.** Gijón: Trea, 1999.

LEIVA, I. G. **La automatización de la indización, propuesta teórico-metodológica: aplicación al área de Biblioteconomía y Documentación.** 1997. 268f. Tese – Universidad de Murcia, Murcia, España, 1997.

LEIVA, I. G. **Manual de indización: teoría y práctica.** Gijón: Trea, 2008. 429 p.

LEIVA, I. G.; RUBI, M. P.; FUJITA, M. S. L.. **Consistência na indexação em Bibliotecas Universitárias Brasileiras.** Transinformação, Campinas, v. 20, n. 3, p.233-253 set/dez., 2008. Disponível em: <<http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/523>>. Acesso em: 12 jan. 2016.

LEIVA, I. G; MUÑOZ, J. V. R. **Los Orígenes del almacenamiento y recuperación de información.** Boletín de la Asociación Andaluza de Bibliotecarios, Málaga, n. 42, p. 51-66, 1996.

LEONARD, L. E. **Inter-indexer consistency and retrieval effectiveness: measurement of relations.** Champaign: University of Illinois, 1975 (PhD Thesis).

LOPES, I. L.. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. **Ciência da Informação**, Brasília, v. 31, n. 1, p.41-52, 1 jan-abril. 2002.

MAI, J. E. **Deconstructing the indexing process.** Advances in Librarianship, v. 23, p. 269-298, 2000.

MARKEY, K. Interindexer consistency tests: A literature review and report of a test of consistency in indexing visual materials. **Library and Information Science Research**, v. 6, n. 2, p. 155-177, 1984.

- MEADOWS, A. J. **A comunicação científica**. Brasília, DF: Briquet de Lemos Livros, 1999.
- MIDDLETON, M.R. A comparison of indexing consistency and coverage in the AEI, ERIC, and APAIS databases. **Behavioral & Social Sciences Librarian**, v. 3, n. 4, p. 33-43, 1984.
- MOREIRA, A. **Tesauros e ontologias**: estudo de definições presentes na literatura das áreas das Ciências da Computação e da Informação, utilizando-se o método analítico-sintético. Belo Horizonte, 2003. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, 2003.
- NARUKAWA, C. M. **Estudo de Vocabulário Controlado na Indexação Automática**: Aplicação no Processo de Indexação do Sistema de Indización Semiautomática (SISA). 2011. 222 f. Dissertação (Mestrado) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2011.
- NARUKAWA, C. M.; LEIVA, I. G.; FUJITA, M. S. L. **Indexação automatizada de artigos de periódicos científicos**: análise da aplicação do software SISA com uso da terminologia DeCS na área de odontologia. *Informação. & Sociedade: Est.*, João Pessoa, v. 19, n. 2, p. 99-118, maio/ago. 2009.
- NASCIMENTO, G. D. **Dos Sintagmas Nominais aos Descritores Documentais**: estudo de caso na indexação de teses e dissertações da área de Direito. 2015. 200 f. Dissertação (Mestrado) - Curso de Ciência da Informação, Departamento de Ciência da Informação, Universidade Federal de Pernambuco, Recife, 2015.
- NAVARRO, M. A. N.; MARCO, F. J. G. **Las primeras jornadas sobre organización del conocimiento**: organización del conocimiento e información científica. Scire, Zaragoza, v.1,n.1,p.149-157,1995. Disponível em:<<http://ibersid.eu/ojs/index.php/scire/article/view/1038/1020>>. Acesso em: 15 jan. 2017.
- NAVARRO, S. Interface entre lingüística e indexação: uma revisão da literatura. **Revista Brasileira de Biblioteconomia e Documentação**. São Paulo, v.21, n. 1/2, p. 46-62, jan/jun. 1988.
- NAVES, M. M. L. **Curso de indexação**: princípios e técnicas de indexação, com vistas à recuperação da informação. Material didático. Belo Horizonte: UFMG, Biblioteca Universitária, 2004. 23p.
- NESHAT, N.; HORRI, A. **A study of subject indexing consistency between the National Library of Iran and Humanities Libraries in the Area of Iranian Studies**. *Cataloging & Classification Quarterly*, v. 43, n. 1, p. 67-76, 2006.
- NICOLINO, M. E. V. P.; FERNEDA, E. Um Método para a Utilização de Ontologias na Indexação Automática. **Informação & Tecnologia (ITEC)**. Marília/João Pessoa, v. 2, n. 1, p.13-33, 1 jul. 2014.
- OLIVEIRA, M. I. **Estudo do contexto de bibliotecas universitárias pelas abordagens de indexação e recuperação em domínios específicos**. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2009.

PINTO, V. B. Indexação documentária: uma forma de representação do conhecimento registrado. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.6, n.2, p.223-234, jul/dez, 2001.

PORTER, A. L.; CUNNINGHAM, S. W. **TECH MINING: exploiting new technologies for competitive advantage**. New Jersey: A John Wiley & Sons, Inc., 2005. 384 p.

RAMALHO, R. A. S. **Desenvolvimento e utilização de ontologias em Bibliotecas Digitais**: uma proposta de aplicação. 2010. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2010. Disponível em: <http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho_ras_do_mar.pdf>. Acesso em: 02 jan. 2017.

ROBREDO, J. **Documentação de hoje e de amanhã**: uma abordagem revisitada e contemporânea da Ciência da Informação e de suas aplicações biblioteconômicas, documentárias, arquivísticas e museológicas. 4. ed. rev. e ampl. Brasília DF: Edição de autor, 2005. 410 p.

ROBREDO, J. Otimização dos processos de indexação dos documentos e de recuperação da informação mediante o uso de instrumentos de controle terminológico. **Ciência da Informação**. Brasília, v. 11, n. 1, p. 3-18, 1982.

SAARTI, J. Consistency of subject indexing of novels by public library professionals and patrons. **Journal of Documentation**. v. 58, n. 1, p. 49-65, 2002.

SALTON, G.; MCGILL, M. J. **Introduction to modern information retrieval**. Nova York: McGraw-Hill, 1983.

SANTOS, G. C.; RIBEIRO, C. M. **Acrônimos, siglas e termos técnicos: arquivística, biblioteconomia, documentação, informática**. Campinas, SP: Editora Átomo, 2003

SANTOS, R. F. **Modelos colaborativos de indexação social e a sua aplicabilidade na base de dados referencial de artigos de periódicos em ciência da informação (BRAPCI)**. 200 f. Dissertação (Mestrado) - Curso de Ciência da Informação, Dci/ufpe, Universidade Federal de Pernambuco, Recife, 2015.

SIEVERT, M. E.; ANDREWS, M. J. Indexing consistency in Information Science Abstracts. **Journal of the American Society for Information Science**, v. 42, n. 1, p. 1-6, 1991.

SILVA, M. R.; 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.puc-campinas.edu.br/transinfo/viewarticle.php?id=65>>. Acesso em: 12/01/2017.

SILVESTER, J. P.; GENUARDI, M. T.; KLINGBIEL, P. H. Machine-aided indexing at NASA. **Information Processing & Management**, v. 30, n. 5, p. 631-45, 1994.

SLYPE. Van G. **Linguagem documentária e linguística**. Trad. Cordélia R. Cavalcanti. Brasília: UnB; Departamento de Biblioteconomia, 1983.

SOLER MONREAL, C.; GIL LEIVA, I. Posibilidades y límites de los tesauros frente a otros sistemas de organización del conocimiento: folksonomías, taxonomías y ontologías. **Revista Inter**. SILVA, W. T.; 2010.

SOLER MONREAL, M. C. **Evaluacion de vocabulários controlados en la indizacion dedocumentos mediante índices de consistência entre indizadores**. 2009. Tese (Doutorado) – Universidad Politecnica de Valencia, Valência, 2009.

SOUZA, R. R. **Uma proposta de metodologia para escolha automática de descritores utilizando sintagmas nominais**. 2005. 215 f. Tese (Doutorado) - Curso de Ciência da Informação, Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.

STEVENS, M. E. **Indexação automática**: um estado da arte, relatório, Monografia 91, National Bureau of Standards, Washington, D.C. 1965.

TESAURO BRASILEIRO DE CIÊNCIA DA INFORMAÇÃO – TBCI. **Labcom - Laboratório de Pesquisas em Comunicação Científica - IBICT**, Rio de Janeiro: 2014. 384 p. Disponível em: <<http://www.ibict.br/publicacoes-e-institucionais/tesauro-brasileiro-de-Ciência-da-informacao-1>>. Acesso em: 12 jan. 2017.

TESAURO EM CIÊNCIA DA INFORMAÇÃO -TCI. Disponível em: <http://icei.pucminas.br/ci/tci/index.php?option=com_weblinks&catid=1&Itemid=4>. Acesso em: 21 jun 2016.

TONTA, Y. A Study of indexing consistency between Library of Congress and BritishLibrary Catalogers. *Library Resources and Technical Services*, v. 35, n. 2, p. 177-185, 1991.

UNISIST. Princípios de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v.10, n.1, p.83-94, mar. 1981.

URDICIAIN, B. G. **Manual de lenguajes documentales**. 2. ed. rev. e ampl. Gijón: Trea, 2004.

VICKERY, B. C. Thesaurus: a new word in documentation. **Journal of Documentation**, v. 16, n. 4, dec. 1960.

VIEIRA, S. B. Indexação automática e manual: revisão de literatura. **Ciência da Informação**. v. 17, n. 1, p. 43-57, 1988.

ZIPF, G. K. **Comportamento humano e o princípio do menor esforço**. Cambridge: Harvard, 1949.

APÊNDICE A - Corpus dos 60 artigos usados no experimento

Corpus utilizado para validação da extração automática de termos, isto é; comparação da indexação automática com a indexação manual.

Referências bibliográficas dos artigos selecionados.

1) SANTOS, P. C. L. V. A. C.; SANT'ANA, R. C. S. G. A. Transferência da informação: análise para valoração de unidades de conhecimento, Datagramazero: Revista de Ciência da Informação , v. 3, n. 2, abr.2002.
2) MUELLER, Suzana P. M.. Popularização do Conhecimento Científico. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 2, abr. 2002.
3) CASTRO, Ana Lúcia Siaines de. O Valor da Informação: um desafio permanente. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 3, jun. 2002.
4) CAFÉ, Lígia; LAGE, Marcia Basílio. Auto-arquivamento: uma opção inovadora para a produção científica. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 3, jun. 2002.
5) BURNHAM, Teresinha Fróes. Análise contrastiva: memória da construção de uma metodologia para investigar a tradução de conhecimento científico em conhecimento público. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 3, jun. 2002.
6) LEVACOV, Marília et al. O Tesouro Eletrônico do Mundo do Trabalho: produto de um esforço interdisciplinar. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 4, ago. 2002.
7) VALENTIM, Marta Lígia Pomim. Inteligência competitiva em organizações: dado, informação e conhecimento. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 4, ago. 2002.
8) MIRANDA, Antonio; SIMEÃO, Elmira. A conceituação de massa documental e o ciclo de interação entre tecnologia e o registro do conhecimento. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 4, ago. 2002.
9) STAREC, C. Informação e universidade: os pecados informacionais e barreiras na comunicação da informação para a tomada de decisão na universidade. DataGramazero: Revista de Ciência da Informação , v. 3, n. 4, p. 0-0, 2002.
10) PORCARO, Rosa Maria. Implicações da 'nova economia' para a mensuração estatística: desajustes conceituais e metodológicos. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 4, ago. 2002.

11) BRITO, Luiz Carlos. Por uma nova Ciência da Informação: ensino, pesquisa e formação. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 5, out. 2002.
12) DIAS, Eduardo Wense. Ensino e pesquisa em Ciência da Informação. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 5, out. 2002.
13) CARVALHO, Kátia de. O Profissional da Informação: o humano multifacetado. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 5, out. 2002.
14) TARAPANOFF, Kira; SUAIDEN, Emir; OLIVEIRA, Cecília Leite. Funções Sociais e Oportunidades para Profissionais da Informação. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 5, out. 2002.
15) RODRIGUES, Mara Eliane Fonseca. Relação Ensino-Pesquisa: em discussão a formação do Profissional da Informação. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 5, out. 2002.
16) CARDOSO, Ana Maria Pereira. Educação para a Informação:: desafios contemporâneos para a Ciência da Informação. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 5, out. 2002.
17) GRAÇAS, Maria das. Novas tecnologias e produção científica: uma relação de causa e efeito ou uma relação de muitos efeitos?. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 6, dez. 2002.
18) DAGNINO, Renato. Enfoques sobre a relação Ciência, Tecnologia e Sociedade: Neutralidade e Determinismo. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 6, dez. 2002.
19) RODRIGUES, Ricardo. Inteligência empresarial: uma avaliação de fontes de informação sobre o ambiente organizacional externo. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 6, dez. 2002.
20) SMIT, Johanna W.; DIAS, Eduardo Wense; SOUZA, Rosali Fernandez de. Contribuição da Pós-graduação para a Ciência da Informação no Brasil: uma visão. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 6, dez. 2002.
21) DUMONT, Lígia Maria Moreira. Os múltiplos aspectos e interfaces da leitura. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 6, dez. 2002.
22) SANTOS, Nilton Bahlis dos. A Informação e o Paradigma Holográfico: a Utopia de Vannevar Bush. Ciência da Informação , Brasília, v. 3, n. 6, p.15-25, dez. 2002.
23) COSTA, Icléia Thiesen Magalhães. Informação, Memória e Espaço Prisional no Rio de Janeiro. Datagramazero: Revista de Ciência da Informação , Brasília, v. 4, n. 1, fev. 2003.
24) GÓMEZ, González de; NÉLIDA, Maria. O Contrato Social da Pesquisa:: em busca de uma nova equação entre a autonomia epistêmica e autonomia política. Datagramazero: Revista da Ciência da Informação , Brasília, v. 4, n. 1, fev. 2003.

25) MUELLER, Suzana Pinheiro Machado; SANTANA, Maria Gorette. A Ciência da Informação no CNPq: fomento à formação de recursos humanos e à pesquisa entre 1994-2002. <i>DataGramaZero: Revista de Ciência da Informação</i> , v. 4, n. 1, fev. 2003.
26) BARRETO, Aldo de Albuquerque. Políticas de Monitoramento da Informação por Compressão Semântica dos seus Estoques. <i>Datagramazero: Revista da Ciência da Informação</i> , Brasília, v. 4, n. 2, abr. 2003.
27) OLINTO, Gilda. Bolsas de Pesquisador do CNPq: informações sobre política de C&T a partir da base que contém os dados cadastrais dos bolsistas. <i>Ciência da Informação</i> , Brasília, v. 4, n. 2, p.23-39, abr. 2003.
28) PACHECO, Roberto; KERN, Vinícius. Arquitetura conceitual e resultados da integração de sistemas de informação e gestão da Ciência e tecnologia. <i>Ciência da Informação</i> , Brasília, v. 4, n. 2, p.56-71, abr. 2003.
29) MARCONDES, Carlos Henrique; JARDIM, José Maria. Políticas de Informação Governamental:: a construção de Governo Eletrônico na Administração Federal do Brasil. <i>Datagramazero: Revista da Ciência da Informação</i> , Brasília, v. 4, n. 2, abr. 2003.
30) DIAS, Guilherme Ataíde. Avaliação do acesso a periódicos eletrônicos na web pela análise do arquivo de log de acesso. <i>Ciência da Informação</i> , Brasília, v. 31, n. 1, jan. abr. 2002.
31) GÓMEZ, González de; NÉLIDA, Maria. Novos cenários políticos para a informação. <i>Ciência da Informação</i> , Brasília, v. 31, n. 1, jan. 2002.
32) LOPES, Ilza Leite. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. <i>Ciência da Informação</i> , Brasília, v. 31, n. 1, p.41-52, 1 jan-abril. 2002.
33) OHIRA, Maria Lourdes Blatt. Bibliotecas virtuais e digitais: análise de artigos de periódicos brasileiros (1995/2000). <i>Ciência da Informação</i> , Brasília, v. 31, n. 1, p.61-74, jan-abril. 2002.
34) PRYSTHON, Cecília; SCHMIDT, Susana. Experiência do Leaal/UFPE na produção e transferência de tecnologia. <i>Ciência da Informação</i> , Brasília, v. 31, n. 1, p.84-90, jan-abril, 2002.
35) ALMEIDA, Maurício Barcellos. Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares. <i>Ciência da Informação</i> , Brasília, v. 31, n. 2, p.5-13, maio-agosto, 2002.
36) ALVARADO, Rubén Urbizagástegui. A Lei de Lotka na bibliometria brasileira. <i>Ciência da Informação</i> , Brasília, v. 31, n. 2, p.14-20, maio-agosto, 2002.
37) CENDÓN, Beatriz Valadares. Bases de dados de informação para negócios. <i>Ciência da Informação</i> , Brasília, v. 31, n. 2, p.30-43, maio-agosto, 2002.
38) GARCEZ, Eliane Maria Stuart. Biblioteca híbrida: um novo enfoque no suporte à educação a distância. <i>Ciência da Informação</i> , Brasília, v. 31, n. 2, p.44-51, maio-agosto, 2002.

39) LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. **Ciência da Informação**, Brasília, v. 31, n. 2, p.60-71, maio-agosto, 2002.

40) MARCHIORI, Patricia Zeni. A Ciência e a gestão da informação: compatibilidades no espaço profissional. **Ciência da Informação**, Brasília, v. 31, n. 2, p.72-79, maio-agosto, 2002.

41) POBLACIÓN, Dinah Aguiar; NORONHA, Daisy Pires. Produção das literaturas “branca” e “cinzenta” pelos docentes/doutores dos programas de pós-graduação em Ciência da informação no Brasil. **Ciência da Informação**, Brasília, v. 31, n. 2, maio 2002.

42) REZENDE, Yara. Informação para negócios: os novos agentes do conhecimento e a gestão do capital intelectual. **Ciência da Informação**, Brasília, v. 31, n. 1, p.75-83, jan-abril 2002.

43) SILVA, Janete Fernandes; FERREIRA, Marta Araújo Tavares; BORGES, Mônica Erichsen Nassif. Análise metodológica dos estudos de necessidades de informação sobre setores industriais brasileiros: proposições: proposições. **Ciência da Informação**, Brasília, v. 31, n. 2, p.129-141, maio-agosto 2002.

44) SILVA, Sergio Luis da. Informação e competitividade: a contextualização da gestão do conhecimento nos processos organizacionais. **Ciência da Informação**, Brasília, v. 31, n. 2, p.142-151, maio-agosto, 2002.

45) VANTI, Nadia Aurora Peres. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, Brasília, v. 31, n. 2, p.152-162, maio-agosto, 2002.

46) BORGES, Paulo César Rodrigues. Métodos quantitativos de apoio à bibliometria: a pesquisa operacional pode ser uma alternativa?. **Ciência da Informação**, Brasília, v. 31, n. 3, set. - dez., 2002.

47) DIAS, Guilherme Ataíde. Periódicos eletrônicos: considerações relativas à aceitação deste recurso pelos usuários. **Ciência da Informação**, Brasília, v. 31, n. 3, set - dez., 2002.

48) COHEN, Max F.. Alguns aspectos do uso da informação na economia da informação. **Ciência da Informação**, Brasília, v. 31, n. 3, set. - dez., 2002.

49) ORTIZ, Lúcia Cunha; ORTIZ, Wilson Aires; SILVA, Sergio Luis da. Ferramentas alternativas para monitoramento e mapeamento automatizado do conhecimento. **Ciência da Informação**, Brasília, v. 31, n. 3, p.66-76, set. dez.,2002.

50) SILVA, Edna Lúcia da; CUNHA, Riam Vieira da. A formação profissional no século XXI: desafios e dilemas. **Ciência da Informação**, Brasília, v. 31, n. 3, p.77-82, set. dez., 2002.

51) TORRES, Elisabeth Fátima; MAZZONI, Alberto Angel; ALVES, João Bosco da Mota. A acessibilidade à informação no espaço digital. **Ciência da Informação**, Brasília, v. 31, n. 3, p.83-91, set. dez., 2002.

52) FREIRE, Isa Maria et al. Estudos de usuários: o padrão que une três abordagens. **Ciência da Informação**, Brasília, v. 31, n. 3, set. dez., 2002.

53) AIRES, R. V. N. X.; ALUÍSIO, S. M. Como incrementar a qualidade dos resultados das máquinas de busca: da análise de logs à interação em português. **Ciência da Informação**, v. 32, n. 1, p. 5-16, 2003.

54) DUDZIAK, Elisabeth Adriana. Information literacy:: princípios, filosofia e prática. **Ciência da Informação**, Brasília, v. 32, n. 1, jan. 2003.

55) FERREIRA, Danielle Thiago. Profissional da informação:: perfil de habilidades demandadas pelo mercado de trabalho. **Ciência da Informação**, Brasília, v. 32, n. 1, jan. abr., 2003.

56) FREIRE, Isa Maria. O olhar da consCiência possível sobre o campo científico. **Ciência da Informação**, Brasília, v. 32, n. 1, jan. abr., 2003.

57) GÓMEZ, Maria Nélide González de. As relações entre Ciência, Estado e sociedade:: um domínio de visibilidade para as questões da informação. **Ciência da Informação**, Brasília, v. 32, n. 1, jan. abr., 2003.

58) LIMA, G. N. B. M. O. Interfaces entre a Ciência da informação e a Ciência cognitiva. **Ciência da Informação**, v. 32, n. 1, jan. abr., 2003.

59) MOSTAFA, Solange Puntel; MÁXIMO, Luis Fernando. A produção científica da Anped e da Intercom no GT da Educação e Comunicação. **Ciência da Informação**, Brasília, v. 32, n. 1, jan. abr., 2003.

60) FREIRE, A Maria. SILVA, Helena Pereira da. Inteligência competitiva na Internet:: um processo otimizado por agentes inteligentes. **Ciência da Informação**, Brasília, v. 32, n. 1, jan. abr., 2003.

APÊNDICE B- E-mails recebidos sobre tipo de indexação usada

E-mails recebidos com as respostas sobre consultas feitas as bases de dados, acerca do tipo de indexação usada, descritas na Subseção 2.6 - **USO DA INDEXAÇÃO AUTOMÁTICA EM PRINCIPAIS BASES DE DADOS CIENTÍFICAS INTERNACIONAIS E DA CIÊNCIA DA INFORMAÇÃO.**

BRAPCI=====

=====Bom dia, Rene.

Você sabe me informar, por gentileza, se a BRAPCI usa indexação automática para representar os documentos armazenados nas bases de dados.

grato,
abçs,
Marcio Bandim

=====

Ola Marcio,

Desculpe-me, seu e-mail acabou ficando perdido em minha caixa de e-mail.

A Brapci tem em seu banco de dados mais de 25.000 termos somente no idioma Português, destes muitos são remetidos para termos equivalentes (remissivas). Desta forma a indexação é automática quando no processo de coleta, porém é comparado com um vocabulário que está em constante validação. Caso o termo "indexado" tenha uma remissiva, utiliza-se este para indexação.

A Brapci utiliza o processo pós-coordenado do vocabulário. A dificuldade está no volume de informações, pois foram revisados somente 15% (aproximadamente) dos 25.000 termos.

Qualquer dúvida, entre em contato!

Sds,

Rene F. Gabriel Junior
Acessado em 28junho2016.

PUBMED - MEDLINE

Hi there,

I want to know if you use automatic indexing to represents your documents stored in all data base. Automatic indexing is one in several type of methods to bulding index of documents to make possible retrieve or recovery information.

thanks a lot,
Marcio Bandim

=====

Dear Marcio Bandim,

This is in response to your message about how we index the citations in PubMed. Please see the FAQ

about Indexing for MEDLINE at <https://www.nlm.nih.gov/bsd/indexfaq.html>

Thanks for your interest in NLM products and services.

Sincerely,

Cathy Sorge, MSLS

NIH Contractor Librarian

National Library of Medicine

National Library of Medicine

O site <www.nlm.nih.gov/bsd/indexfaq.html> tem as informações sobre como são indexados os termos nesta base de dados. Como são vários textos explicativos e textos longos então, em termos resumidos, esta base de dados usa indexação automática assistida por computador com os indexadores comparando os termos selecionados (citação bibliográfica) com os termos de um vocabulário controlado o Medical Subject Headings (MeSH). Vide link acima.

Acessado em 26junho2016.

Web of Science (ISI)

Dear Marcio,

I wanted to let you know that Web of Science does not actually index the Full Text Documents for the articles it index. We only index the bibliographic information on an article. The actual article is only available through the publisher. Your webform mentions INSPEC which is actually a hosted database. If you have any questions with INSPEC data then please reach out to them directly at <inspec.corrections@theiet.org>. We do have a system where all of our data is loaded on our servers but not the actual articles themselves.

William Sickler

Customer Technical Support Representative

Thomson Reuters

<http://ip-science.thomsonreuters.com/support> Acessado em 27junho2016.

Hi, William,

A time ago you said me that WoS:

"I wanted to let you know that Web of Science does not actually index the Full Text Documents for the articles it index. We **only index the bibliographic information** on an article."

How you indexe the bibliographic information? What type of indexing: automatic, semiautomatic or indexing aided by computer?.

thank you,

Marcio Bandim

(12 julho 2016).

Hi, William.

In addition to the last email please let me know if you use "manual indexing" too.

thank you very much,

Marcio Bandim

Gestor não mais respondeu. 01/11/ 2016.

LISA

Hello Marcio,

Thank you for contacting technical support. I have forwarded your question to my colleagues in content and will let you know when they provide responses. Please let us know if you need anything else in the meantime.

Regards,

Kristen Stroehmann, Technical Product Specialist

ProQuest | 789 E. Eisenhower Parkway, P.O. Box 1346 | Ann Arbor MI 48106-1346 USA | 800 889 3358

ProQuest... Start here. 2013 InformationWeek 500 Top Innovator

Email History:

Subject Indexing documents.

Descriptin Hi There,
 What type of indexing is used by LISA - Library and Information Science Abstrats to represents the contents of documents stored in your database ? Automatic indexing, semiautomatic indexing or indexing aided by computer? It's used controlled vocabulary like thesauros or subject head list?
 thank ypu a lot,
 Marcio.

Gestor não mais respondeu. 01/11/ 2016.

A INSPEC - IET

Hi, there.

I want to know what type of indexing you use to represents the contents of documents stored in your databases. Automatic indexing, semiautomatic indexing or automatic indexing aided by computer. Do you use some controlled vocabulary like thesauros?

thanks,
 Marcio Bandim

Gestor não mais respondeu. 01/11/ 2016.

ERIC

=====

Hi, there.

Please I want to know how the documents are indexing in ERIC data bases.
Automatic indexing, semi automatic indexing or indexing aided by computer to select the terms witch represents the docs stored in data base?
ERIC uses controlled vocabulary like Thesaurus?

thanks a lot,
Marcio Bandim
Mastering of Information Science

=====

Dear Marcio,

Eric uses manual and computer-assisting indexing to process ERIC records. Professional indexers manually assign subject terms from the ERIC Thesaurus (our controlled vocabulary) to each record, and we have mapped terms that are correlated with each other to automatically appear when one is used (computer-assisted indexing).

Best regards,

Nancy Cawley
ERIC Communications
ERIC Customer Service Desk
Acessado em 27jun2016.
Gestor não mais respondeu. 01/11/ 2016.

ENGINEERING INDEX COMPENDEX:

=====

Subject
Indexing documents.
Description
Hi There,

What type of indexing is used by Engineering Village to represents the contents of documents stored in your database ? Automatic indexing, semiautomatic indexing or indexing aided by computer? It's used controlled vocabulary like thesaurus or subject head list?

thank you a lot,
Marcio.

=====

Dear Marcio Bandim,

Thank you for contacting Elsevier's Online Customer Support Department with regards to the Indexing Process for Engineering Village.

Please be advised that these are internal processes and information is not readily available for the public. May I know the purpose why you need this information?

Thank you for your patience on this matter. I will wait for your reply.

Kind regards,

Charlie Sun

Elsevier Online Solutions.

=====

Hi,

I'm in Information of Science Master degree at Federal University of Pernambuco - Brazil. My advisor asked me to analyze the main bases of international scientific data on how the indexing terms representing the contents of the documents stored in these databases is made.

If you use automatic indexing, semi-automatic indexing or indexing computer assisted with the application of a controlled vocabulary like thesaurus. That's all I want to know.

thank you so much,

Marcio Bandim

Mastering in Science of Information.

=====

Dear Marcio Bandim,

Thank you for your patience and on the information provided.

I consulted with our content team and indexing process for Engineering Village is automatic. They did not provide information if articles are being sorted via controlled vocabulary.

I hope this information will be helpful to you, please let me know if you need any further assistance.

Kind regards,

Charlie Sun.

Gestor não mais respondeu. 01/11/ 2016.

SCOPUS

=====

Olá,

Por favor quero saber qual tipo de indexação é usada para selecionar os termos que representam o conteúdo dos documentos/informações armazenadas nos bancos de dados da SCOPUS.

Indexação automática, indexação semiautomática, indexação assistida por computador ou indexação manual. É usado algum tipo de vocabulário controlado como Tesouro, por exemplo?

muito grato,
Marcio Bandim.

http://www.americalatina.elsevier.com/corporate/solicitar_teste_e_informacoes.php

Aguardando resposta.
Informações no site:

http://periodicos.capes.gov.br/?option=com_pcollection&mn=70&smn=79&cid=63

Gestor não mais respondeu. 01/11/ 2016.

CHEM ABSTRACTS

Hi, there.

Please let me know how the terms that represents the contents of documents stored in your database are indexed: automatic indexing, semiautomatic indexing, automated indexing aided by computer or manual indexing. Do you use controlled vocabulary like thesaurus?

thank you a lot,

Marcio Bandim
University Federal of Pernambuco / Brasil.
Aguardando resposta. 13 julho 2016.
Gestor não mais respondeu. 01/11/ 2016.

APÊNDICE C - Trabalhos publicados sobre Indexação automática por atribuição

Trabalhos publicados, no Brasil, sobre indexação automática por atribuição, os quais, foram identificados no trabalho de pesquisa e registrados na Tese de Remi Lapa (2014), INDEXAÇÃO AUTOMÁTICA NO BRASIL NO ÂMBITO DA Ciência DA INFORMAÇÃO (1973-2012).

LEGENDA

ID: número de identificação do documento
 AUT: nome do(s) responsável(is) intelectual(is) da obra
 TIT: título da obra
 IPUB: nome da fonte publicadora da obra
 ANO: ano em que a obra foi publicada
 FONT: tipo de fonte onde a obra foi publicada
 INST: instituição(ões) responsável(is) pelo(s) desenvolvimento(s) da obra
 CIT: citação dos r responsável(is) intelectual (is) da obra
 OOBJ: objetivo da obra
 OPRO: quando propõe determinar se trata de um sistema, método ou fórmula
 OAPL: quando aplica determinar se trata de um sistema, método ou fórmula
 OSMF: nome do sistema, método ou fórmula
 MAV: como ocorreu a avaliação do sistema, método ou fórmula
 MNC: natureza da obra
 MTC: tipologia da obra
 MVAL: tipo de indexação utilizada para validar os termos
 MIT: se identificação dos termos ocorreu por extração ou atribuição
 MED: tratamento do texto na entrada de dados
 MNLI: tipologia da natureza da linguagem de indexação
 MNTE: tipologia da natureza dos termos

MMAE: método de Identificação/Ponderação/Seleção de termos da análise estatística
 MMAL: método de Identificação/Ponderação/Seleção de termos da análise linguística
 MMAM: método de Identificação/Ponderação/Seleção de termos do aprendizado de máquina
 MMRC: método de Identificação/Ponderação/Seleção de termos da representação do conhecimento

TRABALHOS PUBLICADOS

AUT: Jean Habran
 TIT: MISTRAL: sistema de recuperação documentária
 IPUB: Revista de Biblioteconomia de Brasília
 ANO: 1985
 FONT: Periódico
 INST: ABC BULL S.A. Telematic
 CIT: HABRAN, J.
 OOBJ: Propõe
 OPRO: Sistema
 OAPL:
 OSMF: MISTRAL
 MAV:
 MNC:
 MTC:
 MVAL: semiautomática; automática
 MIT: Atribuição
 MED: Estruturado
 MNLI: Natural; controlada
 MNTE: Palavras isoladas; termos compostos
 MMAE: Radicalização; stopwords; frequência
 MMAL: Morfológica; sintática; semântica
 MMAM:
 MMRC: vocabulário controlado

=====
 AUT: Wagner Teixeira da Silva; Ruy Luiz Milidui
 TIT: Indexação e recuperação da informação com função de crença,
 IPUB: Ciência da Informacao
 ANO: 1991
 FONT: Periodico
 INST: UNB ; PUC-Rio
 CIT: SILVA, W. T. ; MILIDIU, R. L.
 OOBJ: Propoe
 OPRO: Metodo
 OAPL:
 OSMF: Funcaodecrenca
 MAV:
 MNC:
 MTC:
 MVAL: Automatico

MIT: Atribuicao
 MED: Naoestruturado
 MNLI: Natural; controlada
 MNTE: Palavrasisoladas; termoscompostos
 MMAE: Frequencia; Pesonumerico
 MMAL: semântica
 MMAM:
 MMRC: vocabulariocontrolado

=====

AUT: Marisa Brascher
 TIT: Ambiguidade na Recuperacao da Informacao
 IPUB: DataGramZero - Revista de Ciência da Informacao
 ANO: 2002
 FONT: Periodico
 INST: IBICT
 CIT: BRASCHER, M.
 OOBJ: Aplica
 OPRO:
 OAPL: Sistema
 OSMF: Zstation
 MAV:
 MNC: Textocompleto
 MTC: Documentosoficiais
 MVAL: Automatica
 MIT: Atribuicao
 MED: Naoestruturado
 MNLI: Natural ; controlada
 MNTE: Palavrasisoladas ; termoscompostos
 MMAE:
 MMAL: Morfologica; sintatica; sintagmasnominais; semantica
 MMAM:
 MMRC: Graficosconceituais; vocabulariocontrolado

=====

AUT: Graciane Silva Bruzanga; Benildes Coura Moreira dos Santos Maculan; Gercina
 Angela Borem de Oliveira Lima
 TIT: Indexacao Automatica e semantica: estudo da analise do conteudo de teses e
 dissertacoes
 IPUB: ENCONTRO NACIONAL DE PESQUISA EM Ciência DA INFORMACAO,
 ANO: 2007
 FONT: Anais...
 INST: UFMG
 CIT: BRUZINGA, G. S.; MACULAN, B. C. M. S.; LIMA, G. A. B. O.
 OOBJ: Propoe
 OPRO: Metodo
 OAPL:
 OSMF: Sintagmasnominais
 MAV:
 MNC: Textocompleto

MTC: Trabalhocientifico
 MVAL: Automatica
 MIT: Atribuicao
 MED: Naoestruturado
 MNLI: Natural ; controlada
 MNTE: Palavraisoladas
 MMAE:
 MMAL: Morfologica ; sintatica ; semantica ; sintagmasnominais
 MMAM:
 MMRC: Vocabulariocontrolado

=====
 AUT: Auto Tavares da Camara Junior
 TIT: Indexacao automatica de acordaos por meio de processamento de linguagem
 natural
 IPUB: ENCONTRO NACIONAL DE PESQUISA EM CiênciA DA INFORMACAO
 ANO: 2007
 FONT: Anais...
 INST: UNB
 CIT: CAMARA JUNIOR, A. T.
 OOBJ: Aplicaepropoe
 OPRO:
 OAPL: Metodo
 OSMF: Sintagmasnominais
 MAV: Indexacaomanual
 MNC: Textocompleto
 MTC: Acordao
 MVAL: Automatica
 MIT: Atribuicao
 MED: Naoestruturado
 MNLI: Natural ; controlada
 MNTE: Palavraisoladas ; termoscompostos
 MMAE: pesonumerico
 MMAL: Morfologica; sintatica; semantica; sintagmasnominais
 MMAM: Cluster
 MMRC: vocabulariocontrolado

=====
 AUT: Auto Tavares da Camara Junior
 TIT: Indexacao automática de acórdãos por meio de processamento de linguagem
 natural
 IPUB:
 ANO: 2007
 FONT: Dissertação
 INST: UNB
 CIT: CAMARA JUNIOR, A. T.
 OOBJ: Aplicaepropoe
 OPRO:
 OAPL: Metodo
 OSMF: Sintagmasnominais

MAV: Indexacaomanual
 MNC: Textocompleto
 MTC: Acordao
 MVAL: Automatica
 MIT: Atribuicao
 MED: Naoestruturado
 MNLI: Natural ; controlada
 MNTE: Palavrasisoladas ; termoscompostos
 MMAE: pesonumerico
 MMAL: Morfologica ; sintatica ; semantica ; sintagmasnominais
 MMAM:
 MMRC: vocabulariocontrolado

=====
 AUT: Graciane Silva Bruzinga Borges ; Benildes Coura Moreira dos Santos Maculan ;
 Gercina Angela Borem de Oliveira Lima

TIT: Indexacao Automatica e semantica: estudo de analise do conteúdo de teses e
 dissertações

IPUB: Informacao & Sociedade: Estudos

ANO: 2008

FONT: Periodico

INST: UFMG

CIT: BORGES, G. S. B. ; MACULAN, B. C. M. S. ; LIMA, G. A. B. O.

OOBJ: Propoe

OPRO: Metodo

OAPL:

OSMF: Sintagmasnominais

MAV:

MNC: Textocompleto

MTC: Trabalho científico

MVAL: Automatica

MIT: Atribuicao

MED: Naoestruturado

MNLI: Natural ; controlada

MNTE: Palavrasisoladas

MMAE:

MMAL: Morfologica ; sintatica ; semantica ; sintagmasnominais

MMAM:

MMRC: vocabulariocontrolado

=====
 AUT: Edberto Ferneda; Maria Cristiane Barbosa Galvao; Joeli Espirito Santo Rocha

TIT: Um Metodo de Indexacao automatica de Documentos: aplicacao em laudos de
 exames radiologicos

IPUB: ENCONTRO NACIONAL DE PESQUISA EM Ciência DA INFORMACAO

ANO: 2010

FONT: Anais...

INST: UNESP ; USP

CIT: FERNEDA, E. ; GALVAO, M. C. B. ; ROCHA, J. E. S.

OOBJ: Aplicaeproe

OPRO:
OAPL: Sistema
OSMF: sintagMed
MAV: Indexacaomanual
MNC: Textocompleto
MTC: Laudo
MVAL: Automatica
MIT: Atribuicao
MED: Naoestruturado
MNLI: Natural
MNTE: Palavrasisoladas ; termoscompostos
MMAE: Stopwords; matrizbinaria; frequencia
MMAL: Sintagmasnominais
MMAM:
MMRC: vocabulariocontrolado

APÊNDICE D - Índices de consistência, precisão, evocação e medida F com o TCITG habilitado.

Índices de consistência, precisão, evocação e medida F entre os termos de Indexação proposto pelo SISA e os termos de Palavras-chave habilitando a lista de termos gerais TCITG.

Artigo Científico	No. de Palavras-chave	No. de termos do SISA	No. de Termos comuns	Consistência Relaxada	Precisão	Revocação	Medida F
Artigo 1	3	12	2	15%	17%	67%	27%
Artigo 2	2	11	1,5	13%	14%	75%	23%
Artigo 3	5	9	1,5	12%	17%	30%	21%
Artigo 4	5	10	1,5	11%	15%	30%	20%
Artigo 5	5	4	0,5	6%	13%	10%	11%
Artigo 6	7	9	1	7%	11%	14%	13%
Artigo 7	5	4	1,5	20%	38%	30%	33%
Artigo 8	5	4	1	13%	25%	20%	22%
Artigo 9	5	4	2	29%	50%	40%	44%
Artigo 10	5	5	0,5	5%	10%	10%	10%
Artigo 11	4	9	1,5	13%	17%	38%	23%
Artigo 12	6	6	2,5	26%	42%	42%	42%
Artigo 13	3	7	1,5	18%	21%	50%	30%
Artigo 14	6	4	1	11%	25%	17%	20%
Artigo 15	2	5	1,5	27%	30%	75%	43%
Artigo 16	7	7	3,5	33%	50%	50%	50%
Artigo 17	3	8	1,5	16%	19%	50%	27%
Artigo 18	4	4	0,5	7%	13%	13%	13%
Artigo 19	5	8	1	8%	13%	20%	15%
Artigo 20	4	17	2	11%	12%	50%	19%
Artigo 21	4	4	0,5	7%	13%	13%	13%
Artigo 22	9	6	2,5	20%	42%	28%	33%
Artigo 23	3	3	0	0%	0%	0%	0%
Artigo 24	6	6	2	20%	33%	33%	33%
Artigo 25	2	10	1	9%	10%	50%	17%
Artigo 26	3	15	1,5	9%	10%	50%	17%
Artigo 27	3	10	1	8%	10%	33%	15%
Artigo 28	7	12	3	19%	25%	43%	32%
Artigo 29	4	9	3	30%	33%	75%	46%
Artigo 30	3	17	2	11%	12%	67%	20%
Artigo 31	6	7	0,5	4%	7%	8%	8%

Artigo 32	6	19	2,5	11%	13%	42%	20%	
Artigo 33	5	24	3,5	14%	15%	70%	24%	
Artigo 34	3	4	0	0%	0%	0%	0%	
Artigo 35	5	16	4	24%	25%	80%	38%	
Artigo 36	4	15	2	12%	13%	50%	21%	
Artigo 37	2	14	2	14%	14%	100%	25%	
Artigo 38	3	14	1	6%	7%	33%	12%	
Artigo 39	5	19	3	14%	16%	60%	25%	
Artigo 40	2	10	2	20%	20%	100%	33%	
Artigo 41	4	20	2	9%	10%	50%	17%	
Artigo 42	5	12	2,5	17%	21%	50%	29%	
Artigo 43	4	7	1,5	16%	21%	38%	27%	
Artigo 44	3	6	2	29%	33%	67%	44%	
Artigo 45	5	16	2,5	14%	16%	50%	24%	
Artigo 46	6	15	3	17%	20%	50%	29%	
Artigo 47	3	14	1	6%	7%	33%	12%	
Artigo 48	3	7	1	11%	14%	33%	20%	
Artigo 49	3	11	0,5	4%	5%	17%	7%	
Artigo 50	2	10	1	9%	10%	50%	17%	
Artigo 51	5	5	2	25%	40%	40%	40%	
Artigo 52	5	12	2,5	17%	21%	50%	29%	
Artigo 53	5	4	1,5	20%	38%	30%	33%	
Artigo 54	7	14	2	11%	14%	29%	19%	
Artigo 55	4	7	2	22%	29%	50%	36%	
Artigo 56	5	12	2	13%	17%	40%	24%	
Artigo 57	7	6	1,5	13%	25%	21%	23%	
Artigo 58	7	10	5	42%	50%	71%	59%	
Artigo 59	5	8	1	8%	13%	20%	15%	
Artigo 60	4	17	3,5	20%	21%	88%	33%	
Média	4,5	9,9	1,8	15%	20%	42%	25%	
Desvio	1,6	4,9	1,0	8%	12%	23%	12%	
Mínimo	2	3	0	0%	0%	0%	0%	
Máximo	9	24	5	42%	50%	100%	59%	
No. de documentos com nenhuma palavra-chave atribuída							2	
No. de documentos com meia palavra-chave atribuída							6	
No. de documentos com uma ou mais de uma palavras-chave atribuída							52	
Total de documentos							60	

APÊNDICE E - Índices de consistência, precisão, revocação e medida F desabilitando o TCITG

Índices de consistência, precisão, revocação e medida F entre os termos de Indexação proposto pelo SISA e os termos de Palavras-chave *desabilitando* o arquivo de Termos Gerais (TCITG).

Artigo Científico	No. de Palavras-chave	No. de termos do SISA	No. de termos comuns	Consistência (relaxada)	Precisão	Revocação	Medida F
Artigo 1	3	7	2	25%	29%	67%	40%
Artigo 2	2	5	1	17%	20%	50%	29%
Artigo 3	5	4	1,5	20%	38%	30%	33%
Artigo 4	5	5	1,5	18%	30%	30%	30%
Artigo 5	5	3	0	0%	0%	0%	0%
Artigo 6	7	4	0,5	5%	13%	7%	9%
Artigo 7	5	3	1	14%	33%	20%	25%
Artigo 8	5	3	0,5	7%	17%	10%	13%
Artigo 9	5	3	1,5	23%	50%	30%	38%
Artigo 10	5	2	0	0%	0%	0%	0%
Artigo 11	4	5	1	13%	20%	25%	22%
Artigo 12	6	3	2	29%	67%	33%	44%
Artigo 13	3	4	1	17%	25%	33%	29%
Artigo 14	6	4	1	11%	25%	17%	20%
Artigo 15	2	3	1,5	43%	50%	75%	60%
Artigo 16	7	4	3,5	47%	88%	50%	64%
Artigo 17	3	5	0,5	7%	10%	17%	13%
Artigo 18	4	2	0,5	9%	25%	13%	17%
Artigo 19	5	5	0	0%	0%	0%	0%
Artigo 20	4	7	2	22%	29%	50%	36%
Artigo 21	4	2	0,5	9%	25%	13%	17%
Artigo 22	9	3	2	20%	67%	22%	33%
Artigo 23	3	1	0	0%	0%	0%	0%
Artigo 24	6	3	1	13%	33%	17%	22%
Artigo 25	2	4	1	20%	25%	50%	33%
Artigo 26	3	6	1,5	20%	25%	50%	33%
Artigo 27	3	5	0,5	7%	10%	17%	13%
Artigo 28	7	8	3	25%	38%	43%	40%
Artigo 29	4	5	3	50%	60%	75%	67%
Artigo 30	3	6	1,5	20%	25%	50%	33%
Artigo 31	6	3	0	0%	0%	0%	0%

Artigo 32	6	8	2	17%	25%	33%	29%	
Artigo 33	5	11	3,5	28%	32%	70%	44%	
Artigo 34	3	3	0	0%	0%	0%	0%	
Artigo 35	5	8	3	30%	38%	60%	46%	
Artigo 36	4	6	2	25%	33%	50%	40%	
Artigo 37	2	9	2	22%	22%	100%	36%	
Artigo 38	3	7	1	11%	14%	33%	20%	
Artigo 39	5	8	3	30%	38%	60%	46%	
Artigo 40	2	7	2	29%	29%	100%	44%	
Artigo 41	4	10	2	17%	20%	50%	29%	
Artigo 42	5	9	2,5	22%	28%	50%	36%	
Artigo 43	4	5	1,5	20%	30%	38%	33%	
Artigo 44	3	4	1,5	27%	38%	50%	43%	
Artigo 45	5	6	2	22%	33%	40%	36%	
Artigo 46	6	5	2,5	29%	50%	42%	45%	
Artigo 47	3	5	1	14%	20%	33%	25%	
Artigo 48	3	5	1	14%	20%	33%	25%	
Artigo 49	3	4	0,5	8%	13%	17%	14%	
Artigo 50	2	5	1	17%	20%	50%	29%	
Artigo 51	5	4	2	29%	50%	40%	44%	
Artigo 52	5	6	2,5	29%	42%	50%	45%	
Artigo 53	5	3	1,5	23%	50%	30%	38%	
Artigo 54	7	8	2	15%	25%	29%	27%	
Artigo 55	4	5	2	29%	40%	50%	44%	
Artigo 56	5	6	1,5	16%	25%	30%	27%	
Artigo 57	7	3	0,5	5%	17%	7%	10%	
Artigo 58	7	5	5	71%	100%	71%	83%	
Artigo 59	5	4	0,5	6%	13%	10%	11%	
Artigo 60	4	8	3,5	41%	44%	88%	58%	
Média	4,5	5,1	1,5	19%	30%	37%	30%	
Desvio	1,6	2,1	1,0	13%	20%	25%	18%	
Mínimo	2	1	0	0%	0%	0%	0%	
Máximo	9	11	5	71%	100%	100%	83%	
No. de documentos com nenhuma palavras-chave atribuída							6	
No. de documentos com meia palavras-chave atribuída							9	
No. de documentos com uma ou mais de uma palavras-chave atribuída							45	
Total de documentos							60	

APÊNDICE F – Termos de indexação atribuídos pelo SISA sem habilitar o TCITG

Artigo Científico	Termos de Indexação Palavras-chave	Termos de Indexação SISA
<p>Artigo 1</p> <p>SANTOS, P. C. L. V. A. C.; SANT'ANA, R. C. S. G. A. Transferência da informação: análise para valoração de unidades de conhecimento, Datagramazero: Revista de Ciência da Informação, v. 3, n. 2, abr.2002.</p>	<p>1 - TRANSFERÊNCIA DE INFORMAÇÃO 2- GESTÃO DO CONHECIMENTO 3 - VALOR DE UNIDADES DE CONHECIMENTO.</p>	<p>1-TRANSFERENCIA DA INFORMAÇÃO 2- AVALIAÇÃO 3 – ACESSO 4 – DESCARTE 5 - ESTUDOS DE CASO 6 – GESTÃO 7 - GESTÃO DO CONHECIMENTO</p>
<p>Artigo 2</p> <p>MUELLER, Suzana P. M. Popularização do Conhecimento Científico. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 2, abr. 2002.</p>	<p>1 - POPULARIZAÇÃO DA <i>CIÊNCIA</i> 2 - COMUNICAÇÃO CIENTÍFICA.</p>	<p>1 - COMUNICAÇÃO CIENTÍFICA 2 - <i>CIENCIA</i> DA INFORMAÇÃO 3 - ESTUDOS DE CASO 4 - EDUCAÇÃO 5 - NOTÍCIAS</p>
<p>Artigo 3</p> <p>CASTRO, Ana Lúcia Siaines de. O Valor da Informação: um desafio permanente. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 3, jun. 2002.</p>	<p>1 - <i>INFORMAÇÃO</i> 2 - VALOR INFORMACIONAL 3 - DIREITO À INFORMAÇÃO 4 - MEMÓRIA SOCIAL 5 - ESTOQUE INFORMACIONAL.</p>	<p>1 – AVALIAÇÃO 2 - DIREITO 3 - DIREITO A INFORMAÇÃO 4 - RECUPERAÇÃO DA <i>INFORMAÇÃO</i></p>
<p>Artigo 4</p> <p>CAFÉ, Lígia; LAGE, Marcia Basílio. Auto-arquivamento: uma opção inovadora para a produção científica. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 3, jun. 2002.</p>	<p>1 - ARQUIVOS-ABERTOS 2 - SISTEMA DE <i>PUBLICAÇÃO</i> 3 - BUDAPEST OPEN ACCESS INITIATIVE 4 - ACESSO LIVRE 5 - AUTO-ARQUIVAMENTO.</p>	<p>1 - ACESSO 2 - ACESSO LIVRE 3 - CONCEITOS 4 - FILOSOFIA 5 - <i>PUBLICAÇÕES</i></p>
<p>Artigo 5</p> <p>BURNHAM, Teresinha Fróes. Análise contrastiva: memória da construção de uma metodologia para</p>	<p>1 - CONHECIMENTO CIENTÍFICO 2 - CONHECIMENTO PRIVADO 3 - CONHECIMENTO ESCOLAR; 4 - DEMOCRATIZAÇÃO DA CIÊNCIA</p>	<p>1 – ACESSO 2 - ESTUDANTES 3 - PESQUISA</p>

investigar a tradução de conhecimento científico em conhecimento público. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 3, jun. 2002.		
	5 - COMUNICAÇÃO CIENTÍFICA.	
Artigo 6 LEVACOV, Marília et al. O Tesouro Eletrônico do Mundo do Trabalho: produto de um esforço interdisciplinar. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 4, ago. 2002.	1 - TESAURO ELETRÔNICO; 2 - MUNDO DO TRABALHO; 3 - RECUPERAÇÃO DA INFORMAÇÃO ; 4 - INTERFACE DE CONSULTA; 5 - SISTEMA DE INFORMAÇÃO; 6 - INTERDISCIPLINARIDADE 7 - INTERAÇÃO HUMANO-COMPUTADOR (IHC).	1 - CIENCIA DA INFORMAÇÃO 2- CIENCIA DA COMPUTAÇÃO 3 - NAVEGAÇÃO 4 - UNIVERSIDADES
Artigo 7 VALENTIM, Marta Lúcia Pomim. Inteligência competitiva em organizações: dado, informação e conhecimento. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 4, ago. 2002.	1 - INTELIGÊNCIA COMPETITIVA; 2 - GESTÃO DO CONHECIMENTO; 3 - GESTÃO DA INFORMAÇÃO; 4 - FLUXOS INFORMACIONAIS; 5 - TRANSFERÊNCIA DA INFORMAÇÃO.	1-INTELIGENCIA COMPETITIVA 2-COMPETITIVIDADE 3-DADOS
Artigo 8 MIRANDA, Antonio; SIMEÃO, Elmira. A conceituação de massa documental e o ciclo de interação entre tecnologia e o registro do conhecimento. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 4, ago. 2002.	1 - INFORMAÇÃO ; 2 - MASSA DOCUMENTAL; 3 - CONCEITO DE INFORMAÇÃO; 4 - TECNOLOGIA; 5 - REGISTRO DO CONHECIMENTO.	1- AREAS DO CONHECIMENTO 2- CIENCIA DA INFORMAÇÃO 3- PESQUISA
Artigo 9 STAREC, C. Informação e universidade: os pecados informacionais e barreiras na comunicação da informação para a tomada de decisão na universidade. DataGramZero: Revista de Ciência da Informação , v. 3, n. 4, p. 0-0, 2002.	1 - UNIVERSIDADE; 2 - GESTÃO DO FLUXO DE INFORMAÇÃO NA UNIVERSIDADE; 3 - INTELIGÊNCIA COMPETITIVA; 4 - BARREIRAS NA COMUNICAÇÃO DA INFORMAÇÃO ; 5 - PECADOS INFORMACIONAIS.	1- FLUXO DA INFORMAÇÃO 2- INTELIGENCIA COMPETITIVA 3- RELEVANCIA
Artigo 10 PORCARO, Rosa Maria. Implicações da 'nova economia' para a mensuração estatística: desajustes conceituais e metodológicos. Datagramazero: Revista de Ciência da Informação , Brasília, v. 3, n. 4, ago. 2002.	1 - INFORMAÇÃO ESTATÍSTICA; 2 - NOVA ECONOMIA; 3 - MENSURAÇÃO ESTATÍSTICA; 4 - DESAJUSTE CONCEITUAL; 5 - METODOLOGIA ESTATÍSTICA	1- LEVANTAMENTOS 2- PERTINENCIA
Artigo 11 BRITO, Luiz Carlos. Por uma nova	1 - CIÊNCIA DA INFORMAÇÃO , 2 - ARMAZENAMENTO E RECUPERAÇÃO, 3 - CURSO EM INFORMAÇÃO	1- CIENCIA DA INFORMAÇÃO 2- PESQUISA 3- CLASSIFICAÇÃO

<p>Ciência da Informação: ensino, pesquisa e formação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.</p>	<p>4 - UNIDADE E ESPECIFICIDADE DA INFORMAÇÃO</p>	<p>4- DADOS 5- PROPRIEDADE INTELECTUAL</p>
<p>Artigo 12 DIAS, Eduardo Wense. Ensino e pesquisa em Ciência da Informação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.</p>	<p>1 - CIÊNCIA DA INFORMAÇÃO; 2 - BIBLIOTECONOMIA; 3 - SISTEMA DE INFORMAÇÃO; 4 - ARQUIVOLOGIA; 5 - ENSINO; 6 - PESQUISA.</p>	<p>1- CIENCIA DA INFORMAÇÃO 2- PESQUISA 3- ACESSO</p>
<p>Artigo 13 CARVALHO, Kátia de. O Profissional da Informação: o humano multifacetado. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.</p>	<p>1 - <i>PROFISSIONAL DA INFORMAÇÃO</i>, 2 - INFORMAÇÃO ORGANIZACIONAL, 3 - <i>FORMAÇÃO E PROFISSIONAL DA INFORMAÇÃO</i></p>	<p>1 - <i>FORMAÇÃO PROFISSIONAL</i> 2 - ECONOMIA 3 - PAPEL 4 - <i>SISTEMAS DE INFORMAÇÃO</i></p>
<p>Artigo 14 TARAPANOFF, Kira; SUAIDEN, Emir; OLIVEIRA, Cecília Leite. Funções Sociais e Oportunidades para Profissionais da Informação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.</p>	<p>1 - <i>PROFISSIONAIS DA INFORMAÇÃO</i>; 2 - FUNÇÕES SOCIAIS; 3 - <i>PERFIS DE PROFISSIONAIS DA INFORMAÇÃO</i>; 4 - INCLUSÃO DIGITAL; 5 - GESTÃO DA INFORMAÇÃO; 6 - GESTÃO DO CONHECIMENTO.</p>	<p>1 - ACESSO 2 - GESTORES 3 - <i>FORMAÇÃO PROFISSIONAL</i> 4. <i>PROFISSIONAIS DE INFORMAÇÃO</i></p>
<p>Artigo 15 RODRIGUES, Mara Eliane Fonseca. Relação Ensino-Pesquisa: em discussão a formação do Profissional da Informação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.</p>	<p>1 - <i>FORMAÇÃO PROFISSIONAL</i>; 2 - <i>ENSINO E PESQUISA</i>.</p>	<p>1 - <i>FORMAÇÃO PROFISSIONAL</i> 2 - EDUCAÇÃO 3 - <i>PESQUISA</i></p>
<p>Artigo 16 CARDOSO, Ana Maria Pereira. Educação para a Informação: desafios contemporâneos para a Ciência da Informação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.</p>	<p>1 - CIÊNCIA DA INFORMAÇÃO 2 - <i>FORMAÇÃO PROFISSIONAL</i>; 3 - EDUCAÇÃO SUPERIOR NO BRASIL; 4 - SOCIEDADE DA INFORMAÇÃO 5 - EDUCAÇÃO; 6 - <i>CIÊNCIA DA INFORMAÇÃO E BIBLIOTECONOMIA</i>; 7 - CIÊNCIA DA INFORMAÇÃO - CURSO DE GRADUAÇÃO.</p>	<p>1 - CIENCIA DA INFORMAÇÃO 2 - EDUCAÇÃO 3 - <i>BIBLIOTECONOMIA</i> 4 - <i>FORMAÇÃO PROFISSIONAL</i></p>
<p>Artigo 17 GRAÇAS, Maria das. Novas tecnologias e produção científica: uma relação de causa e efeito ou uma</p>	<p>1 - <i>INTERNET E PRODUÇÃO CIENTÍFICA</i>; 2 - NOVAS TECNOLOGIAS DE INFORMAÇÃO E DE COMUNICAÇÃO; 3 - PRODUÇÃO CIENTÍFICA E NOVAS TECNOLOGIAS.</p>	<p>1 - AUTORIA 2 - COMPLEXIDADE 3 - CONTROLE BIBLIOGRAFICO 4 - <i>INTERNET</i> 5 - PROPRIEDADE INTELECTUAL</p>

relação de muitos efeitos? Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.		
Artigo 18 DAGNINO, Renato. Enfoques sobre a relação Ciência, Tecnologia e Sociedade: Neutralidade e Determinismo. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.	1 - ESTUDOS SOCIAIS DA CIÊNCIA; 2 - SOCIOLOGIA DA CIÊNCIA; 3 - CIÊNCIA E SOCIEDADE; 4 - TECNOLOGIA E SOCIEDADE.	1 - CATEGORIAS 2 - ESTUDOS DE CASO
Artigo 19 RODRIGUES, Ricardo. Inteligência empresarial: uma avaliação de fontes de informação sobre o ambiente organizacional externo. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.	1 - INTELIGÊNCIA EMPRESARIAL; 2 - MONITORAÇÃO AMBIENTAL; 3 - FONTES DE INFORMAÇÃO; 4 - GESTÃO DO CONHECIMENTO; 5 - GESTÃO DA INFORMAÇÃO.	1 - AVALIAÇÃO 2 - AMBIENTE ORGANIZACIONAL 3 - BIBLIOTECAS 4 - PESQUISA 5 - RELEVANCIA
Artigo 20 SMIT, Johanna W.; DIAS, Eduardo Wense; SOUZA, Rosali Fernandez de. Contribuição da Pós-graduação para a Ciência da Informação no Brasil: uma visão. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.	1 - CIÊNCIA DA INFORMAÇÃO NO BRASIL; 2 - AVALIAÇÃO 2001 CAPES; 3 - PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO; 4 - PESQUISA EM CIÊNCIA DA INFORMAÇÃO NO BRASIL.	1 - CIENCIA DA INFORMAÇÃO 2 - AVALIAÇÃO 3 - DISSERTAÇÕES E TESES 4 - PROGRAMAS DE POS-GRADUAÇÃO 5 - PROGRAMAS DE POS-GRADUAÇÃO EM CIENCIA DA INFORMAÇÃO 6 - PUBLICAÇÕES 7 - PESQUISA
Artigo 21 DUMONT, Lígia Maria Moreira. Os múltiplos aspectos e interfaces da leitura. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.	1 - LEITURA-TEORIA; 2 - COGNÓSCIO, CONHECIMENTO-INTROJEÇÃO; 3 - LEITURA E SOCIEDADE; 4 - INFORMAÇÃO E SOCIEDADE.	1 - AREAS DO CONHECIMENTO 2 - ESTUDOS DE CASO
Artigo 22 SANTOS, Nilton Bahlis dos. A Informação e o Paradigma Holográfico: a Utopia de Vannevar Bush. Ciência da Informação, Brasília,	1 - PARADIGMA; 2 - HOLOGRAFIA; 3 - CIÊNCIA DA INFORMAÇÃO; 4 - TECNOLOGIA DA INFORMAÇÃO; 5 - HIPERTEXTO; 6 - COMPLEXIDADE;	1 - CIENCIA DA INFORMAÇÃO 2 - COMPLEXIDADE 3 - TERMOS

v. 3, n. 6, p.15-25, dez. 2002.	7 - INTERATIVIDADE; 8 - VIRTUAL; 9 - TOTALIDADE.	
Artigo 23 COSTA, Icléia Thiesen Magalhães. Informação, Memória e Espaço Prisional no Rio de Janeiro. Datagramazero: Revista de Ciência da Informação, Brasília, v. 4, n. 1, fev. 2003.	1 - INFORMAÇÃO; 2 - MEMÓRIA SOCIAL; 3 - ESPAÇO PRISIONAL.	1 - HISTORIA
Artigo 24 GÓMEZ, González de; NÉLIDA, Maria. O Contrato Social da Pesquisa: em busca de uma nova equação entre a autonomia epistêmica e autonomia política. Datagramazero: Revista da Ciência da Informação, Brasília, v. 4, n. 1, fev. 2003.	1 - CONTRATO SOCIAL; 2 - CIÊNCIA; 3 - PESQUISA; 4 - PESQUISADORES; 5 - AUTONOMIA; 6 - ECOLOGIA DOS CONHECIMENTOS.	1 - PESQUISA 2 - ECOLOGIA 3 - VALIDADE
Artigo 25 MUELLER, Suzana Pinheiro Machado; SANTANA, Maria Gorette. A Ciência da Informação no CNPq: fomento à formação de recursos humanos e à pesquisa entre 1994- 2002. DataGramaZero: Revista de Ciência da Informação, v. 4, n. 1, fev. 2003.	1 - FOMENTO À PESQUISA - CIÊNCIA DA INFORMAÇÃO; 2 - CNPQ - FOMENTO À PESQUISA EM CIÊNCIA DA INFORMAÇÃO.	1 - CIÊNCIA DA INFORMAÇÃO 2 - PESQUISA 3 - DADOS 4 - ORÇAMENTO
Artigo 26 BARRETO, Aldo de Albuquerque. Políticas de Monitoramento da Informação por Compressão Semântica dos seus Estoques. Datagramazero: Revista da Ciência da Informação, Brasília, v. 4, n. 2, abr. 2003.	1 - COMPRESSÃO SEMÂNTICA; 2 - MONITORAMENTO DA INFORMAÇÃO; 3 - ESTOQUES DE INFORMAÇÃO.	1 - MONITORAMENTO 2 - SEMANTICA 3 - CIENCIA DA INFORMAÇÃO 4 - ESTUDOS DE CASO 5 - GESTÃO 6 - SUBSIDIOS
Artigo 27	1 - INDICADORES CIENTÍFICOS;	1 - DADOS

<p>OLINTO, Gilda. Bolsas de Pesquisador do CNPq: informações sobre política de C&T a partir da base que contém os dados cadastrais dos bolsistas. <i>Ciência da Informação</i>, Brasília, v. 4, n. 2, p.23-39, abr. 2003.</p>	<p>2 - POLÍTICA CIENTÍFICA E TECNOLÓGICA; 3 - GESTÃO DE CIÊNCIA E TECNOLOGIA.</p>	<p>2 - BASES DE DADOS 3 - INDICADORES 4 - MONITORAMENTO 5 - RELEVANCIA</p>
<p>Artigo 28</p> <p>PACHECO, Roberto; KERN, Vinícius. Arquitetura conceitual e resultados da integração de sistemas de informação e gestão da Ciência e tecnologia. <i>Ciência da Informação</i>, Brasília, v. 4, n. 2, p.56-71, abr. 2003.</p>	<p>1 - GOVERNO ELETRÔNICO; 2 - ARQUITETURA DE SISTEMAS DE INFORMAÇÃO; 3 - INTEGRAÇÃO DE INFORMAÇÕES; 4 - GESTÃO DE C&T; 5 - BIBLIOTECAS DIGITAIS; 6 - PLATAFORMA LATTES; 7 - REDE SCenTI.</p>	<p>1 - GESTÃO 2 - SISTEMAS DE INFORMAÇÃO 3 - BIBLIOTECAS 4 - BIBLIOTECAS DIGITAIS 5 - GESTÃO DA INFORMAÇÃO 6 - GOVERNO ELETRONICO 7 - INOVAÇÃO 8 - PAPEL</p>
<p>Artigo 29</p> <p>MARCONDES, Carlos Henrique; JARDIM, José Maria. Políticas de Informação Governamental: a construção de Governo Eletrônico na Administração Federal do Brasil. <i>Datagramazero: Revista da Ciência da Informação</i>, Brasília, v. 4, n. 2, abr. 2003.</p>	<p>1 - GOVERNO ELETRÔNICO; 2 - POLÍTICAS DE INFORMAÇÃO; 4 - INFORMAÇÃO GOVERNAMENTAL.</p>	<p>1 - GOVERNO ELETRONICO 2 - INFORMAÇÃO GOVERNAMENTAL 3 - POLITICAS DE INFORMAÇÃO 4 - ACESSO 5 - GESTÃO</p>
<p>Artigo 30</p> <p>DIAS, Guilherme Ataíde. Avaliação do acesso a periódicos eletrônicos na web pela análise do arquivo de log de acesso. <i>Ciência da Informação</i>, Brasília, v. 31, n. 1, jan. abr. 2002.</p>	<p>1 - PERIÓDICOS ELETRÔNICOS; 2 - AVALIAÇÃO DE ACESSO; 3 - ARQUIVO DE LOG DE ACESSO.</p>	<p>1 - AVALIAÇÃO 2 - ACESSO 3 - PERIODICOS 4 - PERIODICOS ELETRONICOS 5 - ESTUDOS DE CASO 6 - WORLD WIDE WEB</p>
<p>Artigo 31</p> <p>GÓMEZ, González de; NÉLIDA, Maria. Novos cenários políticos para a informação. <i>Ciência da Informação</i>,</p>	<p>1 - POLÍTICA DE INFORMAÇÃO; 2 - SOCIEDADE DA INFORMAÇÃO; 4 - INTERNET; 5 - INSTITUCIONALIZAÇÃO DA INFORMAÇÃO; 6 - ESTADO.</p>	<p>1 - CENARIOS 2 - ECONOMIA 3 - PAPEL</p>

Brasília, v. 31, n. 1, jan. 2002.		
<p>Artigo 32</p> <p>LOPES, Ilza Leite. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. Ciência da Informação, Brasília, v. 31, n. 1, p.41-52, 1 jan-abril. 2002.</p>	<p>1 - BASES DE DADOS; 2 - ESTRATÉGIA DE BUSCA; 3 - LINGUAGEM CONTROLADA; 4 - LINGUAGEM NATURAL. 5 - RECUPERAÇÃO DA INFORMAÇÃO; 6 - ARTIGO DE REVISÃO.</p>	<p>1 - BASES DE DADOS 2 - DADOS 3 - CD-ROM 4 - COMPLEXIDADE 5 - ESTUDOS DE CASO 6 - ESTRATEGIAS DE BUSCA 7 - PLANEJAMENTO 8 - TERMOS</p>
<p>Artigo 33</p> <p>OHIRA, Maria Lourdes Blatt. Bibliotecas virtuais e digitais: análise de artigos de periódicos brasileiros (1995/2000). Ciência da Informação, Brasília, v. 31, n. 1, p.61-74, jan-abril. 2002.</p>	<p>1 - BIBLIOTECA DIGITAL; 2 - BIBLIOTECA VIRTUAL; 3 - PRODUÇÃO CIENTÍFICA; 4 - PRODUÇÃO BIBLIOGRÁFICA; 5 - PERIÓDICOS.</p>	<p>1 - BIBLIOTECAS 2 - BIBLIOTECAS VIRTUAIS 3 - PERIODICOS 4 - AUTORIA 5 - BIBLIOTECAS DIGITAIS 6 - COLEÇÕES 7 - DESENVOLVIMENTO DE COLEÇÕES 8 - INTERNET 9 - PLANEJAMENTO 10 - PROFISSIONAIS DE INFORMAÇÃO 11 - REFERENCIAS BIBLIOGRAFICAS</p>
<p>Artigo 34</p> <p>PRYSTHON, Cecília; SCHMIDT, Susana. Experiência do Leaal/UFPE na produção e transferência de tecnologia. Ciência da Informação, Brasília, v. 31, n. 1, p.84-90, jan-abril, 2002.</p>	<p>1 - INFORMAÇÃO TECNOLÓGICA; 2 - TRANSFERÊNCIA DE INFORMAÇÃO; 3 - TRANSFERÊNCIA TECNOLÓGICA.</p>	<p>1 - ACESSO 2 - CIENCIAS DA SAUDE 3 - NUTRIÇÃO</p>
Artigo 35 *	1 - XML;	1 - CONCEITOS

<p>ALMEIDA, Maurício Barcellos. Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares. Ciência da Informação, Brasília, v. 31, n. 2, p.5-13, maio-agosto, 2002.</p>	<p>2 - HTML; 3 - LINGUAGENS DE MARCAÇÃO; 4 - INTERNET; 5 - INTRANET.</p>	<p>2 - INTERNET 3 - XML 4 - COMPLEXIDADE 5 - CIENCIA DA INFORMAÇÃO 6 - DADOS 7 - HTML 8 - USUARIOS</p>
<p>Artigo 36</p> <p>ALVARADO, Rubén Urbizagástegui. A Lei de Lotka na bibliometria brasileira. Ciência da Informação, Brasília, v. 31, n. 2, p.14-20, maio-agosto, 2002.</p>	<p>1 - BIBLIOMETRIA; 2 - LEI DE LOTKA; 3 - PRODUTIVIDADE DE AUTORES; 4 - BRASIL.</p>	<p>1 - BIBLIOMETRIA 2 - LEI DE LOTKA 3 - BIBLIOTECONOMIA 4 - CARTAS 5 - DADOS 6 - MEDICINA</p>
<p>Artigo 37</p> <p>CENDÓN, Beatriz Valadares. Bases de dados de informação para negócios. Ciência da Informação, Brasília, v. 31, n. 2, p.30-43, maio-agosto, 2002.</p>	<p>1 - INFORMAÇÃO PARA NEGÓCIOS; 2 - BASES DE DADOS.</p>	<p>1 - BASES DE DADOS 2 - DADOS 3 - CENARIOS 4 - CATEGORIAS 5 - CLASSIFICAÇÃO 6 - EMPRESAS 7 - INFORMAÇÃO PARA NEGOCIOS 8 - INDUSTRIA DA INFORMAÇÃO 9 - NOTICIAS</p>
<p>Artigo 38</p> <p>GARCEZ, Eliane Maria Stuart. Biblioteca híbrida: um novo enfoque no suporte à educação a distância. Ciência da Informação, Brasília, v. 31, n. 2, p.44-51, maio-agosto, 2002.</p>	<p>1 - BIBLIOTECA HÍBRIDA; 2 - TIPOS DE USUÁRIOS; 3 - BENS E SERVIÇOS.</p>	<p>1 - EDUCAÇÃO 2 - BIBLIOTECAS 3 - BIBLIOTECAS HIBRIDAS 4 - INTERNET 5 - PAPEL 6 - UNIVERSIDADES 7 - USUARIOS</p>

<p>Artigo 39</p> <p>LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. Ciência da Informação, Brasília, v. 31, n. 2, p.60-71, maio-agosto, 2002.</p>	<p>1 - ESTRATÉGIA DE BUSCA; 2 - RECUPERAÇÃO DA INFORMAÇÃO; 3 - TÉCNICAS DE ESTRATÉGIA DE BUSCA; 4 - BASES DE DADOS; 5 - ARTIGO DE REVISÃO.</p>	<p>1 - RECUPERAÇÃO DA INFORMAÇÃO 2 - ACESSO 3 - BASES DE DADOS 4 - COMPLEXIDADE 5 - CONCEITOS 6 - DADOS 7 - ESTRATEGIAS DE BUSCA 8 - PLANEJAMENTO</p>
<p>Artigo 40</p> <p>MARCHIORI, Patricia Zeni. A Ciência e a gestão da informação: compatibilidades no espaço profissional. Ciência da Informação, Brasília, v. 31, n. 2, p.72-79, maio-agosto, 2002.</p>	<p>1 - CIÊNCIA DA INFORMAÇÃO; 2 - GESTÃO DA INFORMAÇÃO.</p>	<p>1 - GESTÃO 2 - GESTÃO DA INFORMAÇÃO 3 - FORMAÇÃO PROFISSIONAL 4 - AVALIAÇÃO 5 - CIENCIA DA INFORMAÇÃO 6 - PROFISSIONAIS DE INFORMAÇÃO 7 - RECURSOS DE INFORMAÇÃO</p>
<p>Artigo 41</p> <p>POBLACIÓN, Dinah Aguiar; NORONHA, Daisy Pires. Produção das literaturas “branca” e “cinzenta” pelos docentes/doutores dos programas de pós-graduação em Ciência da informação no Brasil. Ciência da Informação, Brasília, v. 31, n. 2, maio 2002.</p>	<p>1 - PRODUÇÃO CIENTÍFICA; 2 - LITERATURA BRANCA; 3 - LITERATURA CINZENTA; 4 - CIÊNCIA DA INFORMAÇÃO.</p>	<p>1 - CIENCIA DA INFORMAÇÃO 2 - PROGRAMAS DE POS-GRADUAÇÃO 3 - PROGRAMAS DE POS-GRADUAÇÃO EM CIENCIA DA INFORMAÇÃO 4 - AUTORIA 5 - DADOS 6 - INDICE 7 - LITERATURA CINZENTA 8 - PESQUISA 9 - PUBLICAÇÕES 10 - PERIODICOS</p>

<p>Artigo 42</p> <p>REZENDE, Yara. Informação para negócios: os novos agentes do conhecimento e a gestão do capital intelectual. Ciência da Informação, Brasília, v. 31, n. 1, p.75-83, jan-abril 2002.</p>	<p>1 - GESTÃO DO CONHECIMENTO; 2 - CAPITAL INTELECTUAL; 3 - INFORMAÇÃO PARA NEGÓCIOS; 4 - SISTEMAS DE INFORMAÇÃO PARA NEGÓCIOS; 5 - AGENTES DO CONHECIMENTO.</p>	<p>1 - GESTÃO 2 - INFORMAÇÃO PARA NEGÓCIOS 3 - BIBLIOTECAS 4 - BIBLIOTECAS VIRTUAIS 5 - EMPRESAS 6 - GESTÃO DO CONHECIMENTO 7 - GESTÃO DA INFORMAÇÃO 8 - PROFISSIONAIS DE INFORMAÇÃO 9 - SISTEMAS DE INFORMAÇÃO</p>
<p>Artigo 43</p> <p>SILVA, Janete Fernandes; FERREIRA, Marta Araújo Tavares; BORGES, Mônica Erichsen Nassif. Análise metodológica dos estudos de necessidades de informação sobre setores industriais brasileiros: proposições. Ciência da Informação, Brasília, v. 31, n. 2, p.129-141, maio-agosto 2002.</p>	<p>1 - NECESSIDADE DE INFORMAÇÃO TECNOLÓGICA; 2 - INFORMAÇÃO TECNOLÓGICA; 3 - SETOR INDUSTRIAL; 4 - INOVAÇÃO.</p>	<p>1 - ESTUDOS DE CASO 2 - NECESSIDADES DE INFORMAÇÃO 3 - EMPRESAS 4 - INOVAÇÃO 5 - PESQUISA</p>
<p>Artigo 44</p> <p>SILVA, Sergio Luis da. Informação e competitividade: a contextualização da gestão do conhecimento nos processos organizacionais. Ciência da Informação, Brasília, v. 31, n. 2, p.142-151, maio-agosto, 2002.</p>	<p>1 - GESTÃO DO CONHECIMENTO; 2 - INFORMAÇÃO E COMPETITIVIDADE; 3 - PROCESSOS ORGANIZACIONAIS.</p>	<p>1 - COMPETITIVIDADE 2 - GESTÃO 3 - GESTÃO DO CONHECIMENTO 4 - PAPEL</p>
<p>Artigo 45</p> <p>VANTI, Nadia Aurora Peres. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. Ciência da Informação, Brasília, v. 31, n. 2, p.152-162, maio-agosto, 2002.</p>	<p>1 - BIBLIOMETRIA; 2 - CIENCIOMETRIA; 3 - INFORMETRIA; 4 - WEBOMETRIA; 5 - MÉTODOS QUANTITATIVOS DE AVALIAÇÃO.</p>	<p>1 - BIBLIOMETRIA 2 - CIENCIA DA INFORMAÇÃO 3 - CONCEITOS 4 - INFORMETRIA 5 - INTERNET 6 - PESQUISA</p>
<p>Artigo 46</p> <p>BORGES, Paulo César Rodrigues. Métodos quantitativos de apoio à bibliometria: a pesquisa operacional</p>	<p>1 - BIBLIOMETRIA; 2 -LEI DE BRADFORD; 3 - PESQUISA OPERACIONAL; 4 - CIÊNCIA DA INFORMAÇÃO; 5 - INFERÊNCIA BAYESIANA.</p>	<p>1 - BIBLIOMETRIA 2 - PESQUISA 3 - CIENCIA DA INFORMAÇÃO 4 - PERIODICOS 5 - TEORIA DO CAOS</p>

<p>pode ser uma alternativa?. Ciência da Informação, Brasília, v. 31, n. 3, set. - dez., 2002.</p>	<p>6 - CAOS</p>	
<p>Artigo 47</p> <p>DIAS, Guilherme Ataíde. Periódicos eletrônicos: considerações relativas à aceitação deste recurso pelos usuários. Ciência da Informação, Brasília, v. 31, n. 3, set. - dez., 2002.</p>	<p>1 - PERIÓDICOS ELETRÔNICOS; 2 - USABILIDADE; 3 - NOVAS TECNOLOGIAS.</p>	<p>1 - PERIODICOS 2 - PERIODICOS ELETRONICOS 3 - CIENCIA DA INFORMACÃO 4 - PERIODICOS CIENTIFICOS 5 - WORLD WIDE WEB</p>
<p>Artigo 48</p> <p>COHEN, Max F.. Alguns aspectos do uso da informação na economia da informação. Ciência da Informação, Brasília, v. 31, n. 3, set. - dez., 2002.</p>	<p>1 - USO DA INFORMACÃO; 2 - ECONOMIA DA INFORMACÃO; 3 - MODELO GENÉRICO.</p>	<p>1 - ECONOMIA 2 - ECONOMIA DA INFORMACÃO 3 - CUSTOS 4 - EMPRESAS 5 - INOVAÇÃO</p>
<p>Artigo 49</p> <p>ORTIZ, Lúcia Cunha; ORTIZ, Wilson Aires; SILVA, Sergio Luis da. Ferramentas alternativas para monitoramento e mapeamento automatizado do conhecimento. Ciência da Informação, Brasília, v. 31, n. 3, p.66-76, set. dez.,2002.</p>	<p>1 - MONITORAMENTO DA INFORMACÃO; 2 - BIBLIOTECONOMIA; 3 - CIÊNCIA DA INFORMACÃO.</p>	<p>1 - MONITORAMENTO 2 - DESCRITORES 3 - PESQUISA 4 - PERIODICOS</p>
<p>Artigo 50</p> <p>SILVA, Edna Lúcia da; CUNHA, Riam Vieira da. A formação profissional no século XXI: desafios e dilemas. Ciência da Informação, Brasília, v. 31, n. 3, p.77-82, set. dez., 2002.</p>	<p>1 - EDUCACÃO DOS BIBLIOTECÁRIOS; 2 - PROFISSIONAL DA INFORMACÃO.</p>	<p>1 - FORMAÇÃO PROFISSIONAL 2 - BIBLIOTECARIOS 3 - COOPERACÃO 4 - EDUCACÃO 5 - PAPEL</p>
<p>Artigo 51</p>	<p>1 - ACESSIBILIDADE;</p>	<p>1 - ACESSIBILIDADE</p>

<p>TORRES, Elisabeth Fátima; MAZZONI, Alberto Angel; ALVES, João Bosco da Mota. A acessibilidade à informação no espaço digital. Ciência da Informação, Brasília, v. 31, n. 3, p.83-91, set. dez., 2002.</p>	<p>2 - ESPAÇO DIGITAL;</p> <p>3 - BIBLIOTECAS;</p> <p>4 - PESSOAS PORTADORAS DE DEFICIÊNCIA;</p> <p>5 - AJUDAS TÉCNICAS.</p>	<p>2 - BIBLIOTECAS</p> <p>3 - CATEGORIAS</p> <p>4 - USUARIOS</p>
<p>Artigo 52</p> <p>FREIRE, Isa Maria et al. Estudos de usuários: o padrão que une três abordagens. Ciência da Informação, Brasília, v. 31, n. 3, set. dez., 2002.</p>	<p>1 - <i>ESTUDOS DE USUÁRIOS</i>;</p> <p>2 - <i>EDUCAÇÃO AMBIENTAL</i>;</p> <p>3 - <i>INTERNET</i>;</p> <p>4 - <i>HIPERTEXTO</i>;</p> <p>5 - <i>PESQUISA PARTICIPANTE</i>.</p>	<p>1 - <i>USUARIOS</i></p> <p>2 - CIENCIA DA INFORMAÇÃO</p> <p>3 - <i>EDUCAÇÃO</i></p> <p>4 - <i>INTERNET</i></p> <p>5 - <i>PESQUISA</i></p> <p>6 - PAPEL</p>
<p>Artigo 53</p> <p>AIRES, R. V. N. X.; ALUÍSIO, S. M. Como incrementar a qualidade dos resultados das máquinas de busca: da análise de logs à interação em português. <i>Ciência da Informação</i>, v. 32, n. 1, p. 5-16, 2003.</p>	<p>1 - ANÁLISE DE LOGS;</p> <p>2 - MÁQUINAS DE BUSCA;</p> <p>3 - RECUPERAÇÃO DE INFORMAÇÃO;</p> <p>4 - <i>COMPORTEAMENTO DE USUÁRIOS</i>;</p> <p>5 - ESTRATÉGIAS DE BUSCA.</p>	<p>1 - PROFESSORES</p> <p>2 - RECUPERAÇÃO DA INFORMAÇÃO</p> <p>3 - <i>USUARIOS</i></p>
<p>Artigo 54</p> <p>DUDZIAK, Elisabeth Adriana. Information literacy: princípios, filosofia e prática. Ciência da Informação, Brasília, v. 32, n. 1, jan. 2003.</p>	<p>1 - INFORMATION LITERACY;</p> <p>2 - COMPETÊNCIA EM INFORMAÇÃO;</p> <p>3 - ALFABETIZAÇÃO INFORMACIONAL;</p> <p>4 - <i>BIBLIOTECA APRENDENTE</i>;</p> <p>5 - <i>BIBLIOTECÁRIO EDUCADOR</i>;</p> <p>6 - SOCIEDADE DE APRENDIZAGEM;</p> <p>7 - HABILIDADES INFORMACIONAIS.</p>	<p>1 - FILOSOFIA</p> <p>2 - ACESSO</p> <p>3 - AVALIAÇÃO</p> <p>4 - BIBLIOTECAS</p> <p>5 - <i>BIBLIOTECARIOS</i></p> <p>6 - COMPETENCIA EM INFORMAÇÃO</p> <p>7 - EDUCAÇÃO</p> <p>8 - PAPEL</p>
<p>Artigo 55</p> <p>FERREIRA, Danielle Thiago. Profissional da informação: perfil de habilidades demandadas pelo mercado de trabalho. Ciência da Informação, Brasília, v. 32, n. 1, jan. abr., 2003.</p>	<p>1 - <i>PROFISSIONAL DA INFORMAÇÃO</i>;</p> <p>2 - <i>PROFISSIONAL DA INFORMAÇÃO – HABILIDADES</i>;</p> <p>3 - <i>PERFIL E ATUAÇÃO PROFISSIONAL</i>;</p> <p>4 - MERCADO DE TRABALHO.</p>	<p>1 - MERCADO DE TRABALHO</p> <p>2 - <i>FORMAÇÃO PROFISSIONAL</i></p> <p>3 - CIENCIA DA INFORMAÇÃO</p> <p>4 - EMPRESAS</p> <p>5 - <i>PROFISSIONAIS DE INFORMAÇÃO</i></p>
<p>Artigo 56</p> <p>FREIRE, Isa Maria. O olhar da consciência possível sobre o campo</p>	<p>1 - <i>TEORIA DA CIÊNCIA DA INFORMAÇÃO</i>;</p> <p>2 - <i>SOCIOLOGIA DA INFORMAÇÃO</i>;</p> <p>3 - HISTÓRIA DA CIÊNCIA DA INFORMAÇÃO;</p> <p>4 - <i>COMUNICAÇÃO CIENTÍFICA</i>;</p>	<p>1 - <i>CIENCIA DA INFORMAÇÃO</i></p> <p>2 - CIENTISTAS</p> <p>3 - <i>CIENTISTAS DA INFORMAÇÃO</i></p> <p>4 - <i>INFORMAÇÃO CIENTIFICA E</i></p>

científico. Ciência da Informação , Brasília, v. 32, n. 1, jan. abr., 2003.	5 - RESPONSABILIDADE SOCIAL.	TECNOLOGICA 5 - PESQUISA 6 - RELEVANCIA
Artigo 57 GÓMEZ, Maria Néida González de. As relações entre Ciência, Estado e sociedade:: um domínio de visibilidade para as questões da informação. Ciência da Informação , Brasília, v. 32, n. 1, jan. abr.,2003.	1 - RECUPERAÇÃO DA INFORMAÇÃO; 2 - INTELIGÊNCIA CIENTÍFICA; 3 - INTEGRAÇÃO DOS CONHECIMENTOS; 4 - ESTADO; 5 - CIÊNCIA; 6 - SOCIEDADE; 7 - INFORMAÇÃO.	1 - AVALIAÇÃO 2 - CIENCIA DA INFORMAÇÃO 3 - PESQUISA
Artigo 58 LIMA, G. N. B. M. O. Interfaces entre a Ciência da informação e a Ciência cognitiva. Ciência da Informação , v. 32, n. 1, jan. abr., 2003.	1 - CIÊNCIA DA INFORMAÇÃO; 2 - CIÊNCIA COGNITIVA; 3 - PROCESSAMENTO DA INFORMAÇÃO; 4 - CATEGORIZAÇÃO; 5 - INDEXAÇÃO; 6 - RECUPERAÇÃO DA INFORMAÇÃO; 7 - INTERAÇÃO HOMEM-COMPUTADOR.	1 - CIENCIA DA INFORMAÇÃO 2 - CIENCIA COGNITIVA 3 - INDEXAÇÃO 4 - INTERAÇÃO HOMEM-COMPUTADOR 5 - RECUPERAÇÃO DA INFORMAÇÃO
Artigo 59 MOSTAFA, Solange Puntel; MÁXIMO, Luis Fernando. A produção científica da Anped e da Intercom no GT da Educação e Comunicação. Ciência da Informação , Brasília, v. 32, n. 1, jan. abr., 2003.	1 - COMUNICAÇÃO CIENTÍFICA; 2 - BIBLIOMETRIA; 3 - COMUNICAÇÃO E EDUCAÇÃO; 4 - ESTUDO DE CITAÇÕES; 5 - CIENTOMETRIA.	1 - EDUCAÇÃO 2 - ESTUDOS DE CASO 3 - FRENTE DE PESQUISA 4 - PESQUISA
Artigo 60 FREIRE, A Maria. SILVA, Helena Pereira da. Inteligência competitiva na Internet:: um processo otimizado por agentes inteligentes. Ciência da Informação , Brasília, v. 32, n. 1, jan. abr., 2003.	1 - INTELIGÊNCIA COMPETITIVA; 2 - INTERNET; 3 - MONITORAMENTO DE FONTES DE INFORMAÇÃO; 4 - AGENTES INTELIGENTES.	1 - INTELIGENCIA COMPETITIVA 2 - INTERNET 3 - AGENTES INTELIGENTES 4 - ESTUDOS DE CASO 5 - GESTÃO 6 - GESTÃO DA INFORMAÇÃO 7 - INOVAÇÃO 8 - MONITORAMENTO

APÊNDICE G – Termos de indexação atribuídos pelo SISA com a habilitação do TCITG.

Artigo Científico	Termos de Indexação Palavras-chave	Termos de Indexação SISA
<p>Artigo 1</p> <p>SANTOS, P. C. L. V. A. C.; SANT'ANA, R. C. S. G. A. Transferência da informação: análise para valoração de unidades de conhecimento, Datagramazero: Revista de Ciência da Informação, v. 3, n. 2, abr.2002.</p>	<p>1 - TRANSFERÊNCIA DE INFORMAÇÃO 2- GESTÃO DO CONHECIMENTO 3 - VALOR DE UNIDADES DE CONHECIMENTO.</p>	<p>1-TRANSFERENCIA DA INFORMAÇÃO 2- AVALIAÇÃO 3 – ACESSO 4 – DESCARTE 5 - ESTUDOS DE CASO 6 – GESTÃO 7 - GESTÃO DO CONHECIMENTO 8 - PROCESSOS DE GESTÃO 9 - DESENVOLVIMENTO DE COLEÇÕES 10 - GESTÃO DA INFORMAÇÃO 11 - METODOS DE PESQUISA 12 - METODOS DE PESQUISA E ANALISE</p>
<p>Artigo 2</p> <p>MUELLR, Suzana P. M. Popularização do Conhecimento Científico. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 2, abr. 2002.</p>	<p>1 - POPULARIZAÇÃO DA CIÊNCIA 2 - COMUNICAÇÃO CIENTÍFICA.</p>	<p>1 - COMUNICAÇÃO CIENTIFICA 2 - CIENCIA DA INFORMAÇÃO 3 - ESTUDOS DE CASO 4 – EDUCAÇÃO 5 – NOTICIAS 6 - CIENCIAS SOCIAIS APLICADAS 7 - AREAS DO CONHECIMENTO 8 - METODOS DE PESQUISA 9 - METODOS DE PESQUISA E ANALISE 10 - CIENCIAS HUMANAS 11 - CONTEUDOS DA INFORMAÇÃO</p>
<p>Artigo 3</p> <p>CASTRO, Ana Lúcia Siaines de. O Valor da Informação: um desafio permanente. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 3, jun. 2002.</p>	<p>1 – INFORMAÇÃO 2 - VALOR INFORMACIONAL 3 - DIREITO À INFORMAÇÃO 4 - MEMÓRIA SOCIAL 5 - ESTOQUE INFORMACIONAL.</p>	<p>1 – AVALIAÇÃO 2 – DIREITO 3 - DIREITO A INFORMAÇÃO 4 - RECUPERAÇÃO DA INFORMAÇÃO 5 - PROCESSOS DE GESTÃO 6 - CIENCIAS SOCIAIS APLICADAS 7 - AREAS DO CONHECIMENTO 8 - DIREITOS CIVIS 9 - DIREITOS HUMANOS</p>
<p>Artigo 4</p> <p>CAFÉ, Lúgia; LAGE, Marcia Basílio. Auto-arquivamento: uma opção inovadora para a produção científica. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 3, jun. 2002.</p>	<p>1 - ARQUIVOS-ABERTOS 2 - SISTEMA DE PUBLICAÇÃO 3 - BUDAPEST OPEN ACCESS INITIATIVE 4 - ACESSO LIVRE 5 - AUTO-ARQUIVAMENTO .</p>	<p>1 – ACESSO 2 - ACESSO LIVRE 3 - CONCEITOS 4 - FILOSOFIA 5 - PUBLICAÇÕES 6 - CONTEUDOS DA INFORMAÇÃO 7 - CIENCIAS HUMANAS 8 - AREAS DO CONHECIMENTO 9 - DOCUMENTOS POR CONTEUDO, DISPONIBILIDADE E ACESSO</p>

		10 - TIPOS DE DOCUMENTO
<p>Artigo 5</p> <p>BURNHAM, Teresinha Fróes. Análise contrastiva: memória da construção de uma metodologia para investigar a tradução de conhecimento científico em conhecimento público. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 3, jun. 2002.</p>	<p>1 - CONHECIMENTO CIENTÍFICO</p> <p>2 - CONHECIMENTO PRIVADO</p> <p>3 - CONHECIMENTO ESCOLAR;</p> <p>4 - DEMOCRATIZAÇÃO DA CIÊNCIA</p> <p>5 - COMUNICAÇÃO <i>CIENTÍFICA</i>.</p>	<p>1 - ACESSO</p> <p>2 - ESTUDANTES</p> <p>3 - PESQUISA</p> <p>4 - COMUNIDADES <i>CIENTÍFICAS</i></p>
<p>Artigo 6</p> <p>LEVACOV, Marília et al. O Tesouro Eletrônico do Mundo do Trabalho: produto de um esforço interdisciplinar. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 4, ago. 2002.</p>	<p>1 - TESAURO ELETRÔNICO;</p> <p>2 - MUNDO DO TRABALHO;</p> <p>3 - RECUPERAÇÃO <i>DA INFORMAÇÃO</i>;</p> <p>4 - INTERFACE DE CONSULTA;</p> <p>5 - SISTEMA <i>DE INFORMAÇÃO</i>;</p> <p>6 - INTERDISCIPLINARIDADE</p> <p>7 - INTERAÇÃO HUMANO-COMPUTADOR (IHC).</p>	<p>1 - CIENCIA <i>DA INFORMAÇÃO</i></p> <p>2- CIENCIA DA COMPUTAÇÃO</p> <p>3 - NAVEGAÇÃO</p> <p>4 - UNIVERSIDADES</p> <p>5 - CIENCIAS SOCIAIS APLICADAS</p> <p>6 - AREAS DO CONHECIMENTO</p> <p>7 - CIENCIAS EXATAS E DA TERRA</p> <p>8 - BUSCAS <i>DE INFORMAÇÃO</i></p> <p>9 - INSTITUIÇÕES DE ENSINO E PESQUISA</p>
<p>Artigo 7</p> <p>VALENTIM, Marta Lígia Pomim. Inteligência competitiva em organizações: dado, informação e conhecimento. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 4, ago. 2002.</p>	<p>1 - INTELIGÊNCIA COMPETITIVA;</p> <p>2 - GESTÃO DO CONHECIMENTO;</p> <p>3 - GESTÃO <i>DA INFORMAÇÃO</i>;</p> <p>4 - FLUXOS INFORMACIONAIS;</p> <p>5 - TRANSFERÊNCIA DA INFORMAÇÃO.</p>	<p>1-INTELIGENCIA COMPETITIVA</p> <p>2-COMPETITIVIDADE</p> <p>3-DADOS</p> <p>4-CONTEUDOS <i>DA INFORMAÇÃO</i></p>
<p>Artigo 8</p> <p>MIRANDA, Antonio; SIMEÃO, Elmira. A conceituação de massa documental e o ciclo de interação entre tecnologia e o registro do conhecimento. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 4, ago. 2002.</p>	<p>1 - <i>INFORMAÇÃO</i>;</p> <p>2 - MASSA DOCUMENTAL;</p> <p>3 - CONCEITO DE INFORMAÇÃO;</p> <p>4 - TECNOLOGIA;</p> <p>5 - REGISTRO <i>DO CONHECIMENTO</i>.</p>	<p>1- AREAS <i>DO CONHECIMENTO</i></p> <p>2- CIENCIA <i>DA INFORMAÇÃO</i></p> <p>3- PESQUISA</p> <p>4- CIENCIAS SOCIAIS APLICADAS</p>
<p>Artigo 9</p> <p>STAREC, C. Informação e universidade: os pecados informacionais e barreiras na comunicação da informação para a tomada de decisão na universidade. DataGramaZero: Revista de Ciência da Informação, v. 3, n. 4, p. 0-0, 2002.</p>	<p>1 - UNIVERSIDADE;</p> <p>2 - GESTÃO DO <i>FLUXO DE INFORMAÇÃO</i> NA UNIVERSIDADE;</p> <p>3 - INTELIGÊNCIA COMPETITIVA;</p> <p>4 - BARREIRAS NA COMUNICAÇÃO <i>DA INFORMAÇÃO</i>;</p> <p>5 - PECADOS INFORMACIONAIS.</p>	<p>1- <i>FLUXO DA INFORMAÇÃO</i></p> <p>2- INTELIGENCIA COMPETITIVA</p> <p>3- RELEVANCIA</p> <p>4- ATRIBUTOS <i>DA INFORMAÇÃO</i></p>
<p>Artigo 10</p> <p>PORCARO, Rosa Maria. Implicações da 'nova economia' para a mensuração estatística: desajustes conceituais e metodológicos. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3,</p>	<p>1 - <i>INFORMAÇÃO</i> ESTATÍSTICA;</p> <p>2 - NOVA ECONOMIA;</p> <p>3 - MENSURAÇÃO ESTATÍSTICA;</p> <p>4 - DESAJUSTE CONCEITUAL;</p>	<p>1- LEVANTAMENTOS</p> <p>2- PERTINENCIA</p> <p>3- DOCUMENTOS POR ESTRUTURA</p> <p>4- TIPOS DE DOCUMENTO</p>

n. 4, ago. 2002.	5 - METODOLOGIA ESTATÍSTICA	5- ATRIBUTOS DA <i>INFORMAÇÃO</i>
Artigo 11 BRITO, Luiz Carlos. Por uma nova Ciência da Informação: ensino, pesquisa e formação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.	1 - CIÊNCIA DA INFORMAÇÃO , 2 - ARMAZENAMENTO E RECUPERAÇÃO, 3 - CURSO EM INFORMAÇÃO 4 - UNIDADE E ESPECIFICIDADE DA INFORMAÇÃO	1- CIENCIA DA INFORMAÇÃO 2- PESQUISA 3- CLASSIFICAÇÃO 4- DADOS 5- PROPRIEDADE INTELECTUAL 6- CIENCIAS SOCIAIS APLICADAS 7- AREAS DO CONHECIMENTO 8- ORGANIZAÇÃO DO CONHECIMENTO 9- CONTEUDOS DA INFORMAÇÃO
Artigo 12 DIAS, Eduardo Wense. Ensino e pesquisa em Ciência da Informação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.	1 - CIÊNCIA DA INFORMAÇÃO ; 2 - BIBLIOTECONOMIA; 3 - SISTEMA DE INFORMAÇÃO ; 4 - ARQUIVOLOGIA; 5 - ENSINO; 6 - PESQUISA.	1- CIENCIA DA INFORMAÇÃO 2- PESQUISA 3- ACESSO 4- ACESSO A INFORMAÇÃO 5- CIENCIAS SOCIAIS APLICADAS 6- AREAS DO CONHECIMENTO
Artigo 13 CARVALHO, Kátia de. O Profissional da Informação: o humano multifacetado. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.	1 - PROFISSIONAL DA INFORMAÇÃO , 2 - INFORMAÇÃO ORGANIZACIONAL, 3 - FORMAÇÃO E PROFISSIONAL DA INFORMAÇÃO	1 - FORMAÇÃO PROFISSIONAL 2 - ECONOMIA 3 - PAPEL 4 - SISTEMAS DE INFORMAÇÃO 5 - CIENCIAS SOCIAIS APLICADAS 6 - AREAS DO CONHECIMENTO 7 - SUPORTES DE INFORMAÇÃO
Artigo 14 TARAPANOFF, Kira; SUAIDEN, Emir; OLIVEIRA, Cecília Leite. Funções Sociais e Oportunidades para Profissionais da Informação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.	1 - PROFISSIONAIS DA INFORMAÇÃO ; 2 - FUNÇÕES SOCIAIS; 3 - PERFIS DE PROFISSIONAIS DA INFORMAÇÃO ; 4 - INCLUSÃO DIGITAL; 5 - GESTÃO DA INFORMAÇÃO; 6 - GESTÃO DO CONHECIMENTO.	1 - ACESSO 2 - GESTORES 3 - FORMAÇÃO PROFISSIONAL 4. PROFISSIONAIS DE INFORMAÇÃO
Artigo 15 RODRIGUES, Mara Eliane Fonseca. Relação Ensino-Pesquisa: em discussão a formação do Profissional da Informação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.	1 - FORMAÇÃO PROFISSIONAL ; 2 - ENSINO E PESQUISA.	1 - FORMAÇÃO PROFISSIONAL 2 - EDUCAÇÃO 3 - PESQUISA 4 - CIENCIAS HUMANAS 5 - AREAS DO CONHECIMENTO
Artigo 16 CARDOSO, Ana Maria Pereira. Educação para a Informação: desafios contemporâneos para a Ciência da Informação. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 5, out. 2002.	1 - CIÊNCIA DA INFORMAÇÃO 2 - FORMAÇÃO PROFISSIONAL ; 3 - EDUCAÇÃO SUPERIOR NO BRASIL; 4 - SOCIEDADE DA INFORMAÇÃO 5 - EDUCAÇÃO ; 6 - CIÊNCIA DA INFORMAÇÃO E BIBLIOTECONOMIA ; 7 - CIÊNCIA DA INFORMAÇÃO - CURSO DE GRADUAÇÃO.	1 - CIENCIA DA INFORMAÇÃO 2 - EDUCAÇÃO 3 - BIBLIOTECONOMIA 4 - FORMAÇÃO PROFISSIONAL 5 - CIENCIAS SOCIAIS APLICADAS 6 - AREAS DO CONHECIMENTO 7 - CIENCIAS HUMANAS

<p>Artigo 17</p> <p>GRAÇAS, Maria das. Novas tecnologias e produção científica: uma relação de causa e efeito ou uma relação de muitos efeitos?. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.</p>	<p>1 - <i>INTERNET</i> E PRODUÇÃO CIENTÍFICA; 2 - NOVAS TECNOLOGIAS DE <i>INFORMAÇÃO E DE COMUNICAÇÃO</i>; 3 - PRODUÇÃO CIENTÍFICA E NOVAS TECNOLOGIAS.</p>	<p>1 - AUTORIA 2 - COMPLEXIDADE 3 - CONTROLE BIBLIOGRÁFICO 4 - INTERNET 5 - PROPRIEDADE INTELECTUAL 6 - ATRIBUTOS DA <i>INFORMAÇÃO</i> 7 - REDES DE TELECOMUNICAÇÕES 8 - REDES DE <i>COMUNICAÇÃO</i></p>
<p>Artigo 18</p> <p>DAGNINO, Renato. Enfoques sobre a relação Ciência, Tecnologia e Sociedade: Neutralidade e Determinismo. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.</p>	<p>1 - <i>ESTUDOS</i> SOCIAIS DA CIÊNCIA; 2 - SOCIOLOGIA DA CIÊNCIA; 3 - CIÊNCIA E SOCIEDADE; 4 - TECNOLOGIA E SOCIEDADE.</p>	<p>1 - CATEGORIAS 2 - <i>ESTUDOS</i> DE CASO 3 - METODOS DE PESQUISA 4 - METODOS DE PESQUISA E ANÁLISE</p>
<p>Artigo 19</p> <p>RODRIGUES, Ricardo. Inteligência empresarial: uma avaliação de fontes de informação sobre o ambiente organizacional externo. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.</p>	<p>1 - INTELIGÊNCIA EMPRESARIAL; 2 - MONITORAÇÃO AMBIENTAL; 3 - FONTES <i>DE INFORMAÇÃO</i>; 4 - GESTÃO DO CONHECIMENTO; 5 - GESTÃO <i>DA INFORMAÇÃO</i>.</p>	<p>1 - AVALIAÇÃO 2 - AMBIENTE ORGANIZACIONAL 3 - BIBLIOTECAS 4 - PESQUISA 5 - RELEVANCIA 6 - PROCESSOS DE GESTÃO 7 - UNIDADES <i>DE INFORMAÇÃO</i> 8 - ATRIBUTOS <i>DA INFORMAÇÃO</i></p>
<p>Artigo 20</p> <p>SMIT, Johanna W.; DIAS, Eduardo Wense; SOUZA, Rosali Fernandez de. Contribuição da Pós-graduação para a Ciência da Informação no Brasil: uma visão. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.</p>	<p>1 - <i>CIÊNCIA DA INFORMAÇÃO</i> NO BRASIL; 2 - AVALIAÇÃO 2001 CAPES; 3 - <i>PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO</i>; 4 - <i>PESQUISA</i> EM <i>CIÊNCIA DA INFORMAÇÃO</i> NO BRASIL.</p>	<p>1 - <i>CIENCIA DA INFORMAÇÃO</i> 2 - AVALIAÇÃO 3 - DISSERTAÇÕES E TESES 4 - PROGRAMAS DE POS-GRADUAÇÃO 5 - PROGRAMAS <i>DE POS-GRADUAÇÃO EM CIENCIA DA INFORMAÇÃO</i> 6 - PUBLICAÇÕES 7 - <i>PESQUISA</i> 8 - CIENCIAS SOCIAIS APLICADAS 9 - AREAS DO CONHECIMENTO 10 - PROCESSOS DE GESTÃO 11 - DOCUMENTOS POR ESTRUTURA 12 - TIPOS DE DOCUMENTO 13 - EDUCAÇÃO SUPERIOR 14 - EDUCAÇÃO 15 - CIENCIAS HUMANAS 16 - ENSINO DE <i>CIENCIA DA INFORMAÇÃO</i> 17 - DOCUMENTOS POR CONTEUDO, DISPONIBILIDADE E ACESSO</p>
<p>Artigo 21</p> <p>DUMONT, Lúgia Maria Moreira. Os múltiplos aspectos e interfaces da leitura. Datagramazero: Revista de Ciência da Informação, Brasília, v. 3, n. 6, dez. 2002.</p>	<p>1 - LEITURA-TEORIA; 2 - COGNÓSCIO, <i>CONHECIMENTO</i>-INTROJEÇÃO; 3 - LEITURA E SOCIEDADE; 4 - INFORMAÇÃO E SOCIEDADE.</p>	<p>1 - AREAS DO <i>CONHECIMENTO</i> 2 - ESTUDOS DE CASO 3 - METODOS DE PESQUISA 4 - METODOS DE PESQUISA E ANÁLISE</p>

<p>Artigo 22</p> <p>SANTOS, Nilton Bahlis dos. A Informação e o Paradigma Holográfico: a Utopia de Vannevar Bush. <i>Ciência da Informação</i>, Brasília, v. 3, n. 6, p.15-25, dez. 2002.</p>	<p>1 - PARADIGMA; 2 - HOLOGRAFIA; 3 - CIÊNCIA DA INFORMAÇÃO; 4 - TECNOLOGIA <i>DA INFORMAÇÃO</i>; 5 - HIPERTEXTO; 6 - COMPLEXIDADE; 7 - INTERATIVIDADE; 8 - VIRTUAL; 9 - TOTALIDADE.</p>	<p>1 - CIENCIA DA INFORMAÇÃO 2 - COMPLEXIDADE 3 - TERMOS 4 - CIENCIAS SOCIAIS APLICADAS 5 - AREAS DO CONHECIMENTO 6 - ATRIBUTOS <i>DA INFORMAÇÃO</i></p>
<p>Artigo 23</p> <p>COSTA, Icléia Thiesen Magalhães. Informação, Memória e Espaço Prisional no Rio de Janeiro. <i>Datagramazero: Revista de Ciência da Informação</i>, Brasília, v. 4, n. 1, fev. 2003.</p>	<p>1 - INFORMAÇÃO; 2 - MEMÓRIA SOCIAL; 3 - ESPAÇO PRISIONAL.</p>	<p>1 - HISTORIA 2 - CIENCIAS HUMANAS 3 - AREAS DO CONHECIMENTO</p>
<p>Artigo 24</p> <p>GÓMEZ, González de; NÉLIDA, Maria. O Contrato Social da Pesquisa: em busca de uma nova equação entre a autonomia epistêmica e autonomia política. <i>Datagramazero: Revista da Ciência da Informação</i>, Brasília, v. 4, n. 1, fev. 2003.</p>	<p>1 - CONTRATO SOCIAL; 2 - CIÊNCIA; 3 - PESQUISA; 4 - PESQUISADORES; 5 - AUTONOMIA; 6 - ECOLOGIA DOS CONHECIMENTOS.</p>	<p>1 - PESQUISA 2 - ECOLOGIA 3 - VALIDADE 4 - CIENCIAS BIOLÓGICAS 5 - AREAS DO CONHECIMENTO 6 - ATRIBUTOS DA INFORMAÇÃO</p>
<p>Artigo 25</p> <p>MUELLER, Suzana Pinheiro Machado; SANTANA, Maria Gorette. A Ciência da Informação no CNPq: fomento à formação de recursos humanos e à pesquisa entre 1994-2002. <i>DataGramZero: Revista de Ciência da Informação</i>, v. 4, n. 1, fev. 2003.</p>	<p>1 - FOMENTO À PESQUISA - CIÊNCIA DA INFORMAÇÃO; 2 - CNPQ - FOMENTO À PESQUISA EM CIÊNCIA DA INFORMAÇÃO.</p>	<p>1.- CIENCIA DA INFORMAÇÃO 2.- PESQUISA 3.- DADOS 4.- ORÇAMENTO 5.- CIENCIAS SOCIAIS APLICADAS 6.- AREAS DO CONHECIMENTO 7.- CONTEUDOS DA INFORMAÇÃO</p>
<p>Artigo 26</p> <p>BARRETO, Aldo de Albuquerque. Políticas de Monitoramento da Informação por Compressão Semântica dos seus Estoques. <i>Datagramazero: Revista da Ciência da Informação</i>, Brasília, v. 4, n. 2, abr. 2003.</p>	<p>1 - COMPRESSÃO SEMÂNTICA; 2 - MONITORAMENTO DA INFORMAÇÃO; 3 - ESTOQUES DE INFORMAÇÃO.</p>	<p>1 - MONITORAMENTO 2 - SEMANTICA 3 - CIENCIA DA INFORMAÇÃO 4 - ESTUDOS DE CASO 5 - GESTÃO 6 - SUBSIDIOS 7 - PROCESSOS DE GESTÃO 8 - SEMIOTICA 9 - LINGUISTICA 10 - LINGUISTICA, LETRAS E ARTES 11 - AREAS DO CONHECIMENTO</p>

		12 - CIENCIAS SOCIAIS APLICADAS 13 - METODOS DE PESQUISA 14 - METODOS DE PESQUISA E ANALISE 15 - CAPTAÇÃO DE RECURSOS
Artigo 27 OLINTO, Gilda. Bolsas de Pesquisador do CNPq: informações sobre política de C&T a partir da base que contém os dados cadastrais dos bolsistas. <i>Ciência da Informação</i> , Brasília, v. 4, n. 2, p.23-39, abr. 2003.	1 - INDICADORES CIENTÍFICOS; 2 - POLÍTICA CIENTÍFICA E TECNOLÓGICA; 3 - GESTÃO DE CIÊNCIA E TECNOLOGIA.	1 - DADOS 2 - BASES DE DADOS 3 - INDICADORES 4 - MONITORAMENTO 5 - RELEVANCIA 6 - CONTEUDOS DA INFORMAÇÃO 7 - DOCUMENTOS POR ESTRUTURA 8 - TIPOS DE DOCUMENTO 9 - PROCESSOS DE GESTÃO 10 - ATRIBUTOS DA INFORMAÇÃO
Artigo 28 PACHECO, Roberto; KERN, Vinícius. Arquitetura conceitual e resultados da integração de sistemas de informação e gestão da Ciência e tecnologia. <i>Ciência da Informação</i> , Brasília, v. 4, n. 2, p.56-71, abr. 2003.	1 - GOVERNO ELETRÔNICO ; 2 - ARQUITETURA DE SISTEMAS DE INFORMAÇÃO ; 3 - INTEGRAÇÃO DE INFORMAÇÕES; 4 - GESTÃO DE C&T; 5 - BIBLIOTECAS DIGITAIS ; 6 - PLATAFORMA LATTES; 7 - REDE SCenTI.	1 - GESTÃO 2 - SISTEMAS DE INFORMAÇÃO 3 - BIBLIOTECAS 4 - BIBLIOTECAS DIGITAIS 5 - GESTÃO DA INFORMAÇÃO 6 - GOVERNO ELETRONICO 7 - INOVAÇÃO 8 - PAPEL 9 - PROCESSOS DE GESTÃO 10 - UNIDADES DE INFORMAÇÃO 11 - SERVIÇOS DE INFORMAÇÃO 12 - SUPORTES DE INFORMAÇÃO
Artigo 29 MARCONDES, Carlos Henrique; JARDIM, José Maria. Políticas de Informação Governamental: a construção de Governo Eletrônico na Administração Federal do Brasil. <i>Datagramazero: Revista da Ciência da Informação</i> , Brasília, v. 4, n. 2, abr. 2003.	1 - GOVERNO ELETRÔNICO ; 2 - POLÍTICAS DE INFORMAÇÃO ; 4 - INFORMAÇÃO GOVERNAMENTAL .	1 - GOVERNO ELETRONICO 2 - INFORMAÇÃO GOVERNAMENTAL 3 - POLITICAS DE INFORMAÇÃO 4 - ACESSO 5 - GESTÃO 6 - SERVIÇOS DE INFORMAÇÃO 7 - CONTEUDOS DA INFORMAÇÃO 8 - POLITICAS PUBLICAS 9 - PROCESSOS DE GESTÃO
Artigo 30 DIAS, Guilherme Ataíde. Avaliação do acesso a periódicos eletrônicos na web pela análise do arquivo de log de acesso. <i>Ciência da Informação</i> , Brasília, v. 31, n. 1, jan. abr. 2002.	1 - PERIÓDICOS ELETRÔNICOS ; 2 - AVALIAÇÃO DE ACESSO ; 3 - ARQUIVO DE LOG DE ACESSO.	1 - AVALIAÇÃO 2 - ACESSO 3 - PERIODICOS 4 - PERIODICOS ELETRONICOS 5 - ESTUDOS DE CASO 6 - WORLD WIDE WEB 7 - PROCESSOS DE GESTÃO 8 - PUBLICAÇÕES SERIADAS 9 - PUBLICAÇÕES 10 - DOCUMENTOS POR CONTEUDO, DISPONIBILIDADE E ACESSO 11 - TIPOS DE DOCUMENTO 12 - PERIODICOS CIENTIFICOS 13 - METODOS DE PESQUISA 14 - METODOS DE PESQUISA E ANALISE 15 - INTERNET

		16 - REDES DE TELECOMUNICAÇÕES 17 - REDES DE COMUNICAÇÃO
Artigo 31 GÓMEZ, González de; NÉLIDA, Maria. Novos cenários políticos para a informação. Ciência da Informação , Brasília, v. 31, n. 1, jan. 2002.	1 - POLÍTICA DE INFORMAÇÃO ; 2 - SOCIEDADE DA INFORMAÇÃO; 4 - INTERNET; 5 - INSTITUCIONALIZAÇÃO DA INFORMAÇÃO; 6 - ESTADO.	1 - CENARIOS 2 - ECONOMIA 3 - PAPEL 4 - METODOS DE ANALISE NA INTELIGENCIA COMPETITIVA 5 - CIENCIAS SOCIAIS APLICADAS 6 - AREAS DO CONHECIMENTO 7 - SUPORTES DE INFORMAÇÃO
Artigo 32 LOPES, Ilza Leite. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. Ciência da Informação , Brasília, v. 31, n. 1, p.41-52, 1 jan-abril. 2002.	1 - BASES DE DADOS; 2 - ESTRATÉGIA DE BUSCA; 3 - LINGUAGEM CONTROLADA; 4 - LINGUAGEM NATURAL. 5 - RECUPERAÇÃO DA INFORMAÇÃO ; 6 - ARTIGO DE REVISÃO.	1 - BASES DE DADOS 2 - DADOS 3 - CD-ROM 4 - COMPLEXIDADE 5 - ESTUDOS DE CASO 6 - ESTRATEGIAS DE BUSCA 7 - PLANEJAMENTO 8 - TERMOS 9 - DOCUMENTOS POR ESTRUTURA 10 - TIPOS DE DOCUMENTO 11 - CONTEUDOS DA INFORMAÇÃO 12 - DISCOS COMPACTOS 13 - DISCOS OTICOS 14 - SUPORTES OTICOS 15 - SUPORTES DE INFORMAÇÃO 16 - ATRIBUTOS DA INFORMAÇÃO 17 - METODOS DE PESQUISA 18 - METODOS DE PESQUISA E ANALISE 19 - PROCESSOS DE GESTÃO
Artigo 33 OHIRA, Maria Lourdes Blatt. Bibliotecas virtuais e digitais: análise de artigos de periódicos brasileiros (1995/2000). Ciência da Informação , Brasília, v. 31, n. 1, p.61-74, jan-abril. 2002.	1 - BIBLIOTECA DIGITAL; 2 - BIBLIOTECA VIRTUAL; 3 - PRODUÇÃO CIENTÍFICA; 4 - PRODUÇÃO BIBLIOGRÁFICA ; 5 - PERIÓDICOS.	1 - BIBLIOTECAS 2 - BIBLIOTECAS VIRTUAIS 3 - PERIODICOS 4 - AUTORIA 5 - BIBLIOTECAS DIGITAIS 6 - COLEÇÕES 7 - DESENVOLVIMENTO DE COLEÇÕES 8 - INTERNET 9 - PLANEJAMENTO 10 - PROFISSIONAIS DE INFORMAÇÃO 11 - REFERENCIAS BIBLIOGRAFICAS 12 - UNIDADES DE INFORMAÇÃO 13 - PUBLICAÇÕES SERIADAS 14 - PUBLICAÇÕES 15 - DOCUMENTOS POR CONTEUDO, DISPONIBILIDADE E ACESSO 16 - TIPOS DE DOCUMENTO 17 - GESTÃO DA INFORMAÇÃO 18 - GESTÃO 19 - PROCESSOS DE GESTÃO 20 - REDES DE TELECOMUNICAÇÕES 21 - REDES DE COMUNICAÇÃO 22 - REGISTROS BIBLIOGRAFICOS 23 - REGISTROS 24 - DOCUMENTOS POR ESTRUTURA

<p>Artigo 34</p> <p>PRYSTHON, Cecília; SCHMIDT, Susana. Experiência do Leaal/UFPE na produção e transferência de tecnologia. Ciência da Informação, Brasília, v. 31, n. 1, p.84-90, jan-abril, 2002.</p>	<p>1 - INFORMAÇÃO TECNOLÓGICA; 2 - TRANSFERÊNCIA DE INFORMAÇÃO; 3 - TRANSFERÊNCIA TECNOLÓGICA.</p>	<p>1 - ACESSO 2 - CIENCIAS DA SAUDE 3 - NUTRIÇÃO 4 - AREAS DO CONHECIMENTO</p>
<p>Artigo 35</p> <p>ALMEIDA, Maurício Barcellos. Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares. Ciência da Informação, Brasília, v. 31, n. 2, p.5-13, maio-agosto, 2002.</p>	<p>1 - XML; 2 - HTML; 3 - LINGUAGENS DE MARCAÇÃO; 4 - INTERNET; 5 - INTRANET.</p>	<p>1 - CONCEITOS 2 - INTERNET 3 - XML 4 - COMPLEXIDADE 5 - CIENCIA DA INFORMAÇÃO 6 - DADOS 7 - HTML 8 - USUARIOS 9 - CONTEUDOS DA INFORMAÇÃO 10 - REDES DE ITELECOMUNICAÇÕES 11 - REDES DE COMUNICAÇÃO 12 - LINGUAGENS DE MARCAÇÃO 13 - NORMAS E PROTOCOLOS 14 - ATRIBUTOS DA INFORMAÇÃO 15 - CIENCIAS SOCIAIS APLICADAS 16 - AREAS DO CONHECIMENTO</p>
<p>Artigo 36</p> <p>ALVARADO, Rubén Urbizagástegui. A Lei de Lotka na bibliometria brasileira. Ciência da Informação, Brasília, v. 31, n. 2, p.14-20, maio-agosto, 2002.</p>	<p>1 - BIBLIOMETRIA; 2 - LEI DE LOTKA; 3 - PRODUTIVIDADE DE AUTORES; 4 - BRASIL.</p>	<p>1 - BIBLIOMETRIA 2 - LEI DE LOTKA 3 - BIBLIOTECONOMIA 4 - CARTAS 5 - DADOS 6 - MEDICINA 7 - METODOS DE ANALISE NA INTELIGENCIA COMPETITIVA 8 - LEIS BIBLIOMETRICAS 9 - CIENCIA DA INFORMAÇÃO 10 - CIENCIAS SOCIAIS APLICADAS 11 - AREAS DO CONHECIMENTO 12 - DOCUMENTOS POR ESTRUTURA 13 - TIPOS DE DOCUMENTO 14 - CONTEUDOS DA INFORMAÇÃO 15 - CIENCIAS DA SAUDE</p>
<p>Artigo 37</p> <p>CENDÓN, Beatriz Valadares. Bases de dados de informação para negócios. Ciência da Informação, Brasília, v. 31, n. 2, p.30-43, maio-agosto, 2002.</p>	<p>1 - INFORMAÇÃO PARA NEGÓCIOS; 2 - BASES DE DADOS.</p>	<p>1 - BASES DE DADOS 2 - DADOS 3 - CENARIOS 4 - CATEGORIAS 5 - CLASSIFICAÇÃO 6 - EMPRESAS 7 - INFORMAÇÃO PARA NEGOCIOS 8 - INDUSTRIA DA INFORMAÇÃO 9 - NOTICIAS 10 - DOCUMENTOS POR ESTRUTURA 11 - TIPOS DE DOCUMENTO 12 - CONTEUDOS DA INFORMAÇÃO</p>

		13 - METODOS DE ANALISE NA INTELIGENCIA COMPETITIVA 14 - ORGANIZAÇÃO DO CONHECIMENTO
Artigo 38 GARCEZ, Eliane Maria Stuart. Biblioteca híbrida: um novo enfoque no suporte à educação a distância. Ciência da Informação , Brasília, v. 31, n. 2, p.44-51, maio-agosto, 2002.	1 - BIBLIOTECA HÍBRIDA ; 2 - TIPOS DE USUÁRIOS ; 3 - BENS E SERVIÇOS.	1 - EDUCAÇÃO 2 - BIBLIOTECAS 3 - BIBLIOTECAS HÍBRIDAS 4 - INTERNET 5 - PAPEL 6 - UNIVERSIDADES 7 - USUARIOS 8 - CIENCIAS HUMANAS 9 - AREAS DO CONHECIMENTO 10 - UNIDADES DE INFORMAÇÃO 11 - REDES DE TELECOMUNICAÇÕES 12 - REDES DE COMUNICAÇÃO 13 - SUPORTES DE INFORMAÇÃO 14 - INSTITUIÇÕES DE ENSINO E PESQUISA
Artigo 39 LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. Ciência da Informação , Brasília, v. 31, n. 2, p.60-71, maio-agosto, 2002.	1 - ESTRATÉGIA DE BUSCA ; 2 - RECUPERAÇÃO DA INFORMAÇÃO ; 3 - TÉCNICAS DE ESTRATÉGIA DE BUSCA; 4 - BASES DE DADOS ; 5 - ARTIGO DE REVISÃO.	1 - RECUPERAÇÃO DA INFORMAÇÃO 2 - ACESSO 3 - BASES DE DADOS 4 - COMPLEXIDADE 5 - CONCEITOS 6 - DADOS 7 - ESTRATEGIAS DE BUSCA 8 - PLANEJAMENTO 9 - SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO 10 - USUARIOS 11 - DOCUMENTOS POR ESTRUTURA 12 - TIPOS DE DOCUMENTO 13 - ATRIBUTOS DA INFORMAÇÃO 14 - CONTEUDOS DA INFORMAÇÃO 15 - PROCESSOS DE GESTÃO 16 - PROGRAMAS DE RECUPERAÇÃO DA INFORMAÇÃO 17 - APLICATIVOS 18 - PROGRAMAS DE COMPUTADOR QUANTO AO USO 19 - PROGRAMAS DE COMPUTADOR
Artigo 40 MARCHIORI, Patricia Zeni. A Ciência e a gestão da informação: compatibilidades no espaço profissional. Ciência da Informação , Brasília, v. 31, n. 2, p.72-79, maio-agosto, 2002.	1 - CIÊNCIA DA INFORMAÇÃO ; 2 - GESTÃO DA INFORMAÇÃO .	1 - GESTÃO 2 - GESTÃO DA INFORMAÇÃO 3 - FORMAÇÃO PROFISSIONAL 4 - AVALIAÇÃO 5 - CIENCIA DA INFORMAÇÃO 6 - PROFISSIONAIS DE INFORMAÇÃO 7 - RECURSOS DE INFORMAÇÃO 8 - PROCESSOS DE GESTÃO 9 - CIENCIAS SOCIAIS APLICADAS 10 - AREAS DO CONHECIMENTO
Artigo 41 POBLACIÓN, Dinah Aguiar; NORONHA, Daisy Pires. Produção das literaturas “branca” e “cinzenta” pelos docentes/doutores dos programas de	1 - PRODUÇÃO CIENTÍFICA; 2 - LITERATURA BRANCA; 3 - LITERATURA CINZENTA ; 4 - CIÊNCIA DA INFORMAÇÃO .	1 - CIENCIA DA INFORMAÇÃO 2 - PROGRAMAS DE POS-GRADUAÇÃO 3 - PROGRAMAS DE POS-GRADUAÇÃO EM CIENCIA DA INFORMAÇÃO 4 - AUTORIA 5 - DADOS 6 - INDICE

<p>pós-graduação em Ciência da informação no Brasil. Ciência da Informação, Brasília, v. 31, n. 2, maio 2002.</p>		<p>7 - LITERATURA CINZENTA 8 - PESQUISA 9 - PUBLICAÇÕES 10 - PERIODICOS 11 - CIENCIAS SOCIAIS APLICADAS 12 - AREAS DO CONHECIMENTO 13 - EDUCAÇÃO SUPERIOR 14 - EDUCAÇÃO 15 - CIENCIAS HUMANAS 16 - ENSINO DE CIENCIA DA INFORMAÇÃO 17 - CONTEUDOS DA INFORMAÇÃO 18 - DOCUMENTOS POR CONTEUDO, DISPONIBILIDADE E ACESSO 19 - TIPOS DE DOCUMENTO 20 - PUBLICAÇÕES SERIADAS</p>
<p>Artigo 42</p> <p>REZENDE, Yara. Informação para negócios: os novos agentes do conhecimento e a gestão do capital intelectual. Ciência da Informação, Brasília, v. 31, n. 1, p.75-83, jan-abril 2002.</p>	<p>1 - GESTÃO DO CONHECIMENTO; 2 - CAPITAL INTELECTUAL; 3 - INFORMAÇÃO PARA NEGÓCIOS; 4 - SISTEMAS DE INFORMAÇÃO PARA NEGÓCIOS; 5 - AGENTES DO CONHECIMENTO.</p>	<p>1 – GESTÃO 2 - INFORMAÇÃO PARA NEGOCIOS 3 - BIBLIOTECAS 4 - BIBLIOTECAS VIRTUAIS 5 - EMPRESAS 6 - GESTÃO DO CONHECIMENTO 7 - GESTÃO DA INFORMAÇÃO 8 - PROFISSIONAIS DE INFORMAÇÃO 9 - SISTEMAS DE INFORMAÇÃO 10 - PROCESSOS DE GESTÃO 11 - CONTEUDOS DA INFORMAÇÃO 12 - UNIDADES DE INFORMAÇÃO</p>
<p>Artigo 43</p> <p>SILVA, Janete Fernandes; FERREIRA, Marta Araújo Tavares; BORGES, Mônica Erichsen Nassif. Análise metodológica dos estudos de necessidades de informação sobre setores industriais brasileiros: proposições. Ciência da Informação, Brasília, v. 31, n. 2, p.129-141, maio-agosto 2002.</p>	<p>1 - NECESSIDADE DE INFORMAÇÃO TECNOLÓGICA; 2 - INFORMAÇÃO TECNOLÓGICA; 3 - SETOR INDUSTRIAL; 4 - INOVAÇÃO.</p>	<p>1 - ESTUDOS DE CASO 2 - NECESSIDADES DE INFORMAÇÃO 3 - EMPRESAS 4 - INOVAÇÃO 5 - PESQUISA 6 - METODOS DE PESQUISA 7 - METODOS DE PESQUISA E ANALISE</p>
<p>Artigo 44</p> <p>SILVA, Sergio Luis da. Informação e competitividade: a contextualização da gestão do conhecimento nos processos organizacionais. Ciência da Informação, Brasília, v. 31, n. 2, p.142-151, maio-agosto, 2002.</p>	<p>1 - GESTÃO DO CONHECIMENTO; 2 - INFORMAÇÃO E COMPETITIVIDADE; 3 - PROCESSOS ORGANIZACIONAIS.</p>	<p>1 – COMPETITIVIDADE 2 - GESTÃO 3 - GESTÃO DO CONHECIMENTO 4 - PAPEL 5 - PROCESSOS DE GESTÃO 6 - SUPORTES DE INFORMAÇÃO</p>
<p>Artigo 45</p> <p>VANTI, Nadia Aurora Peres. Da bibliometria à webometria: uma exploração conceitual dos mecanismos</p>	<p>1 - BIBLIOMETRIA; 2 - CIENCIOMETRIA; 3 - INFORMETRIA; 4 - WEBOMETRIA; 5 - MÉTODOS QUANTITATIVOS DE AVALIAÇÃO.</p>	<p>1 - BIBLIOMETRIA 2 - CIENCIA DA INFORMAÇÃO 3 - CONCEITOS 4 - INFORMETRIA 5 - INTERNET</p>

<p>utilizados para medir o registro da informação e a difusão do conhecimento. Ciência da Informação, Brasília, v. 31, n. 2, p.152-162, maio-agosto, 2002.</p>		<p>6 - PESQUISA 7 - METODOS DE ANALISE NA INTELIGENCIA COMPETITIVA 8 - CIENCIAS SOCIAIS APLICADAS 9 - AREAS DO CONHECIMENTO 10 - CONTEUDOS DA INFORMAÇÃO 11 - METRIAS DA INFORMAÇÃO E COMUNICAÇÃO 12 - METODOS MATEMATICOS E ESTATISTICOS 13 - ANALISE <i>QUANTITATIVA</i> 14 - METODOS DE PESQUISA E ANALISE 15 - REDES DE TELECOMUNICAÇÕES 16 - REDES DE COMUNICAÇÃO</p>
<p>Artigo 46</p> <p>BORGES, Paulo César Rodrigues. Métodos quantitativos de apoio à bibliometria: a pesquisa operacional pode ser uma alternativa?. Ciência da Informação, Brasília, v. 31, n. 3, set. - dez., 2002.</p>	<p>1 - BIBLIOMETRIA; 2 -LEI DE BRADFORD; 3 - <i>PESQUISA</i> OPERACIONAL 4 - <i>CAOS</i> 5 - CIÊNCIA DA INFORMAÇÃO; 6 - INFERÊNCIA BAYESIANA.</p>	<p>1 - BIBLIOMETRIA 2 - <i>PESQUISA</i> 3 - CIENCIA DA INFORMAÇÃO 4 - PERIODICOS 5 - TEORIA DO <i>CAOS</i> 6 - METODOS DE ANALISE NA INTELIGENCIA COMPETITIVA 7 - CIENCIAS SOCIAIS APLICADAS 8 - AREAS DO CONHECIMENTO 9 - PUBLICAÇÕES SERIADAS 10 - PUBLICAÇÕES 11 - DOCUMENTOS POR CONTEUDO, DISPONIBILIDADE E ACESSO 12 - TIPOS DE DOCUMENTO 13 - MODELOS MATEMATICOS 14 - MODELOS DE ANALISE 15 - METODOS DE PESQUISA E ANALISE</p>
<p>Artigo 47</p> <p>DIAS, Guilherme Ataíde. Periódicos eletrônicos: considerações relativas à aceitação deste recurso pelos usuários. Ciência da Informação, Brasília, v. 31, n. 3, set. - dez., 2002.</p>	<p>1 - PERIÓDICOS ELETRÔNICOS; 2 - USABILIDADE; 3 - NOVAS TECNOLOGIAS.</p>	<p>1 - PERIODICOS 2 - PERIODICOS ELETRONICOS 3 - CIENCIA DA INFORMAÇÃO 4 - PERIODICOS CIENTIFICOS 5 - WORLD WIDE WEB 6 - PUBLICAÇÕES SERIADAS 7 - PUBLICAÇÕES 8 - DOCUMENTOS POR CONTEUDO, DISPONIBILIDADE E ACESSO 9 - TIPOS DE DOCUMENTO 10 - CIENCIAS SOCIAIS APLICADAS 11 - AREAS DO CONHECIMENTO 12 - INTERNET 13 - REDES DE TELECOMUNICAÇÕES 14 - REDES DE COMUNICAÇÃO</p>
<p>Artigo 48</p> <p>COHEN, Max F.. Alguns aspectos do uso da informação na economia da informação. Ciência da Informação, Brasília, v. 31, n. 3, set. - dez., 2002.</p>	<p>1 - USO DA INFORMAÇÃO; 2 - ECONOMIA DA INFORMAÇÃO; 3 - MODELO GENÉRICO.</p>	<p>1 - ECONOMIA 2 - ECONOMIA DA INFORMAÇÃO 3 - CUSTOS 4 - EMPRESAS 5 - INOVAÇÃO 6 - CIENCIAS SOCIAIS APLICADAS 7 - AREAS DO CONHECIMENTO</p>
<p>Artigo 49</p>	<p>1 - <i>MONITORAMENTO</i> DA INFORMAÇÃO; 2 - BIBLIOTECONOMIA;</p>	<p>1 - <i>MONITORAMENTO</i> 2 - DESCRITORES</p>

<p>ORTIZ, Lúcia Cunha; ORTIZ, Wilson Aires; SILVA, Sergio Luis da. Ferramentas alternativas para monitoramento e mapeamento automatizado do conhecimento. Ciência da Informação, Brasília, v. 31, n. 3, p.66-76, set. dez.,2002.</p>	<p>3 - CIÊNCIA DA INFORMAÇÃO.</p>	<p>3 - PESQUISA 4 - PERIODICOS 5 - PROCESSOS DE GESTÃO 6 - TERMOS DE INDEXAÇÃO 7 - TERMOS 8 - PUBLICAÇÕES SERIADAS 9 - PUBLICAÇÕES 10 - DOCUMENTOS POR CONTEUDO, DISPONIBILIDADE E ACESSO 11 - TIPOS DE DOCUMENTO</p>
<p>Artigo 50</p> <p>SILVA, Edna Lúcia da; CUNHA, Riam Vieira da. A formação profissional no século XXI: desafios e dilemas. Ciência da Informação, Brasília, v. 31, n. 3, p.77-82, set. dez., 2002.</p>	<p>1 - EDUCAÇÃO DOS BIBLIOTECÁRIOS; 2 - PROFISSIONAL DA INFORMAÇÃO.</p>	<p>1 - FORMAÇÃO PROFISSIONAL 2 - BIBLIOTECARIOS 3 - COOPERAÇÃO 4 - EDUCAÇÃO 5 - PAPEL 6 - PESSOAL DA BIBLIOTECA 7 - PROCESSOS DE GESTÃO 8 - CIENCIAS HUMANAS 9 - AREAS DO CONHECIMENTO 10 - SUPORTES DE INFORMAÇÃO</p>
<p>Artigo 51</p> <p>TORRES, Elisabeth Fátima; MAZZONI, Alberto Angel; ALVES, João Bosco da Mota. A acessibilidade à informação no espaço digital. Ciência da Informação, Brasília, v. 31, n. 3, p.83-91, set. dez., 2002.</p>	<p>1 - ACESSIBILIDADE; 2 - ESPAÇO DIGITAL; 3 - BIBLIOTECAS; 4 - PESSOAS PORTADORAS DE DEFICIÊNCIA; 5 - AJUDAS TÉCNICAS.</p>	<p>1 - ACESSIBILIDADE 2 - BIBLIOTECAS 3 - CATEGORIAS 4 - USUARIOS 5 - UNIDADES DE INFORMAÇÃO</p>
<p>Artigo 52</p> <p>FREIRE, Isa Maria et al. Estudos de usuários: o padrão que une três abordagens. Ciência da Informação, Brasília, v. 31, n. 3, set. dez., 2002.</p>	<p>1 - ESTUDOS DE USUÁRIOS; 2 - EDUCAÇÃO AMBIENTAL; 3 - INTERNET; 4 - HIPERTEXTO; 5 - PESQUISA PARTICIPANTE.</p>	<p>1 - USUARIOS 2 - CIENCIA DA INFORMAÇÃO 3 - EDUCAÇÃO 4 - INTERNET 5 - PESQUISA 6 - PAPEL 7 - CIENCIAS SOCIAIS APLICADAS 8 - AREAS DO CONHECIMENTO 9 - CIENCIAS HUMANAS 10 - REDES DE TELECOMUNICAÇÕES 11 - REDES DE COMUNICAÇÃO 12 - SUPORTES DE INFORMAÇÃO</p>
<p>Artigo 53</p> <p>AIRES, R. V. N. X.; ALUÍSIO, S. M. Como incrementar a qualidade dos resultados das máquinas de busca: da análise de logs à interação em português. Ciência da Informação, v. 32, n. 1, p. 5-16, 2003.</p>	<p>1 - ANÁLISE DE LOGS; 2 - MÁQUINAS DE BUSCA; 3 - RECUPERAÇÃO DE INFORMAÇÃO; 4 - COMPORTAMENTO DE USUÁRIOS; 5 - ESTRATÉGIAS DE BUSCA.</p>	<p>1 - PROFESSORES 2 - RECUPERAÇÃO DA INFORMAÇÃO 3 - USUARIOS 4 - COMUNIDADES CIENTIFICAS</p>
<p>Artigo 54</p> <p>DUDZIAK, Elisabeth Adriana. Information literacy:: princípios,</p>	<p>1 - INFORMATION LITERACY; 2 - COMPETÊNCIA EM INFORMAÇÃO; 3 - ALFABETIZAÇÃO INFORMACIONAL; 4 - BIBLIOTECA APRENDENTE;</p>	<p>1 - FILOSOFIA 2 - ACESSO 3 - AVALIAÇÃO 4 - BIBLIOTECAS</p>

<p>filosofia e prática. Ciência da Informação, Brasília, v. 32, n. 1, jan. 2003.</p>	<p>5 - BIBLIOTECÁRIO EDUCADOR; 6 - SOCIEDADE DE APRENDIZAGEM; 7 - HABILIDADES INFORMACIONAIS.</p>	<p>5 - BIBLIOTECARIOS 6 - COMPETENCIA EM INFORMAÇÃO 7 - EDUCAÇÃO 8 - PAPEL 9 - CIENCIAS HUMANAS 10 - AREAS DO CONHECIMENTO 11 - PROCESSOS DE GESTÃO 12 - UNIDADES DE INFORMAÇÃO 13 - PESSOAL DA BIBLIOTECA 14 - SUPORTES DE INFORMAÇÃO</p>
<p>Artigo 55 FERREIRA, Danielle Thiago. Profissional da informação:: perfil de habilidades demandadas pelo mercado de trabalho. Ciência da Informação, Brasília, v. 32, n. 1, jan. abr., 2003.</p>	<p>1 - PROFISSIONAL DA INFORMAÇÃO; 2 - PROFISSIONAL DA INFORMAÇÃO – HABILIDADES; 3 - PERFIL E ATUAÇÃO PROFISSIONAL; 4 - MERCADO DE TRABALHO.</p>	<p>1 - MERCADO DE TRABALHO 2 - FORMAÇÃO PROFISSIONAL 3 - CIENCIA DA INFORMAÇÃO 4 - EMPRESAS 5 - PROFISSIONAIS DE INFORMAÇÃO 6 - CIENCIAS SOCIAIS APLICADAS 7 - AREAS DO CONHECIMENTO</p>
<p>Artigo 56 FREIRE, Isa Maria. O olhar da consCiência possível sobre o campo científco. Ciência da Informação, Brasília, v. 32, n. 1, jan. abr., 2003.</p>	<p>1 - TEORIA DA CIÊNCIA DA INFORMAÇÃO; 2 - SOCIOLOGIA DA INFORMAÇÃO; 3 - HISTÓRIA DA CIÊNCIA DA INFORMAÇÃO; 4 - COMUNICAÇÃO CIENTÍFICA; 5 - RESPONSABILIDADE SOCIAL.</p>	<p>1 - CIENCIA DA INFORMAÇÃO 2 - CIENTISTAS 3 - CIENTISTAS DA INFORMAÇÃO 4 - INFORMAÇÃO CIENTIFICA E TECNOLÓGICA 5 - PESQUISA 6 - RELEVANCIA 7 - CIENCIAS SOCIAIS APLICADAS 8 - AREAS DO CONHECIMENTO 9 - COMUNIDADES CIENTIFICAS 10 - PROFISSIONAIS DE INFORMAÇÃO 11 - CONTEUDOS DA INFORMAÇÃO 12 - ATRIBUTOS DA INFORMAÇÃO</p>
<p>Artigo 57 GÓMEZ, Maria Néida González de. As relações entre Ciência, Estado e sociedade:: um domínio de visibilidade para as questões da informação. Ciência da Informação, Brasília, v. 32, n. 1, jan. abr.,2003.</p>	<p>1 - RECUPERAÇÃO DA INFORMAÇÃO; 2 - INTELIGÊNCIA CIENTÍFICA; 3 - INTEGRAÇÃO DOS CONHECIMENTOS; 4 - ESTADO; 5 - CIÊNCIA; 6 - SOCIEDADE; 7 - INFORMAÇÃO.</p>	<p>1 - AVALIAÇÃO 2 - CIENCIA DA INFORMAÇÃO 3 - PESQUISA 4 - PROCESSOS DE GESTÃO 5 - CIENCIAS SOCIAIS APLICADAS 6 - AREAS DO CONHECIMENTO</p>
<p>Artigo 58 LIMA, G. N. B. M. O. Interfaces entre a Ciência da informação e a Ciência cognitiva. Ciência da Informação, v. 32, n. 1, jan. abr., 2003.</p>	<p>1 - CIÊNCIA DA INFORMAÇÃO; 2 - CIÊNCIA COGNITIVA; 3 - PROCESSAMENTO DA INFORMAÇÃO; 4 - CATEGORIZAÇÃO; 5 - INDEXAÇÃO; 6 - RECUPERAÇÃO DA INFORMAÇÃO; 7 - INTERAÇÃO HOMEM-COMPUTADOR.</p>	<p>1 - CIENCIA DA INFORMAÇÃO 2 - CIENCIA COGNITIVA 3 - INDEXAÇÃO 4 - INTERAÇÃO HOMEM-COMPUTADOR 5 - RECUPERAÇÃO DA INFORMAÇÃO 6 - CIENCIAS SOCIAIS APLICADAS 7 - AREAS DO CONHECIMENTO 8 - PSICOLOGIA 9 - CIENCIAS HUMANAS 10 - ORGANIZAÇÃO DO CONHECIMENTO</p>
<p>Artigo 59 MOSTAFA, Solange Puntel; MÁXIMO, Luis Fernando. A produção científica da Anped e da Intercom no GT da Educação e Comunicação.</p>	<p>1 - COMUNICAÇÃO CIENTÍFICA; 2 - BIBLIOMETRIA; 3 - COMUNICAÇÃO E EDUCAÇÃO; 4 - ESTUDO DE CITAÇÕES; 5 - CIENTOMETRIA.</p>	<p>1 - EDUCAÇÃO 2 - ESTUDOS DE CASO 3 - FRENTE DE PESQUISA 4 - PESQUISA 5 - CIENCIAS HUMANAS 6 - AREAS DO CONHECIMENTO</p>

<p>Ciência da Informação, Brasília, v. 32, n. 1, jan. abr., 2003.</p>		<p>7 - METODOS DE PESQUISA 8 - METODOS DE PESQUISA E ANALISE</p>
<p>Artigo 60</p> <p>FREIRE, A Maria. SILVA, Helena Pereira da. Inteligência competitiva na Internet:: um processo otimizado por agentes inteligentes. Ciência da Informação, Brasília, v. 32, n. 1, jan. abr., 2003.</p>	<p>1 - INTELIGÊNCIA COMPETITIVA; 2 - INTERNET; 3 - <i>MONITORAMENTO DE FONTES DE INFORMAÇÃO</i>; 4 - AGENTES INTELIGENTES.</p>	<p>1 - INTELIGENCIA COMPETITIVA 2 - INTERNET 3 - AGENTES INTELIGENTES 4 - ESTUDOS DE CASO 5 - GESTÃO 6 - GESTÃO DA INFORMAÇÃO 7 - INOVAÇÃO 8 - <i>MONITORAMENTO</i> 9 - REDES DE TELECOMUNICAÇÕES 10 - REDES DE COMUNICAÇÃO 11 - PROGRAMAS DE RECUPERAÇÃO DA INFORMAÇÃO 12 - APLICATIVOS 13 - PROGRAMAS DE COMPUTADOR QUANTO AO USO 14 - PROGRAMAS DE COMPUTADOR 15 - METODOS DE PESQUISA 16 - METODOS DE PESQUISA E ANALISE 17 - PROCESSOS DE GESTÃO</p>

APÊNDICE H– Modelo de um artigo científico formatado segundo critérios do SISA ⁽¹⁾

#CTI#

Auto-arquivamento: uma opção inovadora para a produção científica

#FTI#

#CR#

RESUMO: Trata do conceito inovador de auto-arquivamento e suas implicações no sistema de publicações científicas. Esta nova filosofia procura minimizar as conseqüências provocadas pelo controle editorial, pela revisão severa entre os pares e pela reserva dos direitos autorais. A experiência da Budapest Open Access Initiative (BOAI) é relatada com o objetivo de mostrar uma ação efetiva que viabiliza o auto-arquivamento. Fundamentada no acesso livre (open access), a BOAI pretende uma reorganização dos mecanismos de produção do meio científico, baseada em conceitos mais democráticos de acesso ao conteúdo.

Palavras chave: Arquivos-abertos, Sistema de Publicação, Budapest Open Access Initiative, Acesso Livre, Auto-arquivamento

#FR#

#CTE#

INTRODUÇÃO

A nova era da Internet tem provocado grandes mudanças no seio dos sistemas de publicações. Ao estabelecer um paralelo entre publicação eletrônica e publicação impressa, observa-se que a primeira apresenta grandes vantagens em relação a segunda, tornando-se uma das principais fontes de informação e de meio de disseminação de pesquisas científicas, além de proporcionar uma maior visibilidade e acessibilidade ao conteúdo.

Taubes (1996) aponta algumas vantagens da publicação eletrônica, em relação à tradicional.

[...]

No que diz respeito ao valor quantitativo da produção científica do pesquisador, as publicações eletrônicas dão um passo a frente do modelo tradicional. Além de contabilizarem a produção individual (quantos artigos o autor publicou) e o número de citações (índice de impacto), as novas tecnologias de informação e comunicação possibilitam obter o somatório de vezes que o artigo foi acessado.

[...]

As editoras comerciais criam obstáculos que dificultam o acesso às informações, e conseqüentemente à sua disseminação. Os custos elevados das publicações científicas, tanto na forma impressa quanto eletrônica, são barreiras para muitos pesquisadores, bibliotecas, instituições de ensino e de pesquisa. Mesmo as instituições e universidades com grandes recursos não têm acesso a todas as publicações científicas.

[...]

6.CONCLUSÃO

No decorrer dos últimos anos, tem-se observado importantes mudanças no sistema de publicação científica. Se por um lado a Internet e as novas tecnologias da informação e da comunicação têm oferecido os meios tecnológicos para viabilizar estas transformações, por outro, o mundo da ciência tem buscado novas estratégias de divulgação e acesso dos resultados de pesquisas. O auto-arquivamento tem sido uma opção de grande aceitação no mundo científico. Este meio tem encontrado espaço em iniciativas semelhantes ao BOAI, citado no presente trabalho. Fundamentado na filosofia do acesso livre, ela pretende uma reorganização do sistema de publicações, baseada em conceitos mais democráticos de acesso ao conteúdo. Procuramos neste artigo mostrar esta nova concepção de publicação científica, esclarecendo o conceito inovador de auto-arquivamento e sua aplicação na BOAI.

#FTE#

(1)

Fonte: CAFÉ, Lígia; LAGE, Marcia Basílio. Auto-arquivamento: uma opção inovadora para a produção científica. **Datagramazero: Revista de Ciência da Informação**, Brasília, v. 3, n. 3, jun. 2002.

ANEXO A – Lista de Descritores do TBCI

Lista de Descritores - Tesauro Brasileiro de Ciência da Informação (TBCI)

1. AACR1 USE CÓDIGO DE CATALOGAÇÃO ANGLO-AMERICANO
2. AACR2 USE CÓDIGO DE CATALOGAÇÃO ANGLO-AMERICANO
3. ABNT USE ASSOCIAÇÕES DE ELABORAÇÃO DE NORMAS
4. ABOUTNESS USE TEMATICIDADE
5. ABSTRACTS USE RESUMOS
6. ACERVOS (BIBLIOTECA)
7. ACERVOS BIBLIOGRÁFICOS USE ACERVOS (BIBLIOTECA)
8. ACERVOS PESSOAIS
9. ACESSIBILIDADE
10. ACESSO
11. ACESSO À BIBLIOTECA
12. ACESSO A DISTÂNCIA USE ACESSO REMOTO
13. ACESSO À INFORMAÇÃO
14. ACESSO A RECURSOS USE ACESSO
15. ACESSO ABERTO USE ACESSO LIVRE
16. ACESSO AO DOCUMENTO
17. ACESSO ÀS PUBLICAÇÕES USE ACESSO AO DOCUMENTO
18. ACESSO BIBLIOGRÁFICO USE ACESSO AO DOCUMENTO
19. ACESSO LIVRE
20. ACESSO REMOTO
21. ACESSO SEM FIO
22. ACESSO UNIVERSAL
23. ACOPLAMENTO BIBLIOGRÁFICO
24. ACOPLAMENTO DE REFERÊNCIAS BIBLIOGRÁFICAS USE ACOPLAMENTO BIBLIOGRÁFICO
25. ACORDOS DE UTILIZAÇÃO
26. AD USE AUDIODESCRIÇÃO
27. ADMINISTRADORES USE GESTORES
28. ADMINISTRADORES DE BIBLIOTECAS USE GESTORES DE BIBLIOTECAS
29. ADMINISTRAÇÃO (ÁREA DO CONHECIMENTO)
30. ADMINISTRAÇÃO BIBLIOTECÁRIA USE GESTÃO DE BIBLIOTECAS
31. ADMINISTRAÇÃO DE BIBLIOTECAS USE GESTÃO DE BIBLIOTECAS
32. AGÊNCIAS DE FOMENTO
33. AGÊNCIAS DE FOMENTO À PESQUISA USE AGÊNCIAS DE FOMENTO
34. AGÊNCIAS DE NOTÍCIAS
35. AGÊNCIAS GOVERNAMENTAIS
36. AGENTES DE BUSCA USE AGENTES INTELIGENTES
37. AGENTES DE INFORMAÇÃO USE CONSULTORES DE INFORMAÇÃO
38. AGENTES DE RECUPERAÇÃO DA INFORMAÇÃO USE AGENTES INTELIGENTES
39. AGENTES INTELIGENTES
40. AGLUTINADORES DE INFORMAÇÃO USE GATEKEEPERS
41. AGRONOMIA
42. ALDEIA GLOBAL USE SOCIEDADE DA INFORMAÇÃO
43. ALFABETAÇÃO USE ARRANJO ALFABÉTICO
44. ALFABETAÇÃO LETRA POR LETRA
45. ALFABETAÇÃO PALAVRA POR PALAVRA
46. ALFABETIZAÇÃO DIGITAL USE COMPETÊNCIA NO USO DE COMPUTADORES
47. ALFABETIZAÇÃO ELETRÔNICA USE COMPETÊNCIA NO USO DE COMPUTADORES
48. ALFABETIZAÇÃO EM INFORMAÇÃO USE COMPETÊNCIA EM INFORMAÇÃO

.....

- 1718. TERMOS
- 1719. TERMOS CANDIDATOS A DESCRITOR
- 1720. TERMOS DE BUSCA
- 1721. TERMOS DE INDEXAÇÃO
- 1722. TERMOS DE USO USE ACORDOS DE UTILIZAÇÃO
- 1723. TERMOS GENÉRICOS MAIORES
- 1724. TERMOS NÃO-AUTORIZADOS USE TERMOS NÃO-PREFERIDOS
- 1725. TERMOS NÃO-PREFERIDOS
- 1726. TERMOS PREFERIDOS USE DESCRITORES
- 1727. TERMOS PROIBIDOS USE TERMOS NÃO-PREFERIDOS
- 1728. TESAUROS
- 1729. TESAUROS BILÍNGUES USE TESAUROS MULTILÍNGUES
- 1730. TESAUROS FACETADOS
- 1731. TESAUROS MULTILÍNGUES
- 1732. TESES USE DISSERTAÇÕES E TESES
- 1733. TESTE
- 1734. TICS USE TECNOLOGIAS DA INFORMAÇÃO E COMUNICAÇÃO
- 1735. TIPOS DE DOCUMENTO
- 1736. TÍTULOS USE TÍTULOS DE DOCUMENTOS
- 1737. TÍTULOS DE DOCUMENTOS
- 1738. TOLERÂNCIA A FALHAS
- 1739. TOMADA DE DECISÕES
- 1740. TRABALHO COOPERATIVO USE COOPERAÇÃO
- 1741. TRADUTORES
- 1742. TRADUÇÃO AUTOMÁTICA
- 1743. TRÁFEGO NOS SÍTIOS WEB
- 1744. TRANSDISCIPLINARIDADE
- 1745. TRANSFERÊNCIA DA INFORMAÇÃO
- 1746. TRANSFERÊNCIA DE ARQUIVOS
- 1747. TRANSMISSÃO DE DADOS
- 1748. TRANSMISSÃO DE VOZ
- 1749. TRANSMISSÃO FAC-SIMILAR
- 1750. TRATAMENTO DE MATERIAIS BIBLIOGRÁFICOS USE PROCESSAMENTO TÉCNICO
- 1751. TRATAMENTO TÉCNICO DO ACERVO USE PROCESSAMENTO TÉCNICO
- 1752. TREINAMENTO
- 1753. TREINAMENTO DE USUÁRIOS
- 1754. TREINAMENTO EM SERVIÇO USE TREINAMENTO
- 1755. TREINAMENTO PROFISSIONAL USE TREINAMENTO
- 1756. TRUNCAGEM USE TRUNCAMENTO
- 1757. TRUNCAMENTO
- 1758. TURISMO
- 1759. TUTORIAIS
- 1760. UNIDADES DE INFORMAÇÃO
- 1761. UNIFORM RESOURCE IDENTIFIER USE URI
- 1762. UNIFORM RESOURCE LOCATORS USE URL
- 1763. UNIFORM RESOURCE NAMES USE URN
- 1764. UNIFORMIDADE NA INDEXAÇÃO USE CONSISTÊNCIA NA INDEXAÇÃO
- 1765. UNIVERSIDADES
- 1766. UPLOAD USE TRANSFERÊNCIA DE ARQUIVOS
- 1767. URBANISMO USE ARQUITETURA E URBANISMO
- 1768. URI
- 1769. URL
- 1770. URN
- 1771. USABILIDADE (PROGRAMAS DE COMPUTADOR)
- 1772. USO LEGÍTIMO

- 1773. USOS DA BIBLIOTECA USE USOS DA INFORMAÇÃO
- 1774. USOS DA INFORMAÇÃO
- 1775. USUÁRIOS
- 1776. USUÁRIOS DA BIBLIOTECA USE USUÁRIOS
- 1777. USUÁRIOS DE INFORMAÇÃO USE USUÁRIOS
- 1778. USUÁRIOS EXPERIENTES
- 1779. USUÁRIOS FINAIS
- 1780. USUÁRIOS NOVATOS
- 1781. USUÁRIOS NOVOS USE USUÁRIOS NOVATOS
- 1782. UTILIZADORES DA INFORMAÇÃO USE USUÁRIOS
- 1783. VALIDADE
- 1784. VALIDAÇÃO DE DADOS
- 1785. VERIFICAÇÃO DE DADOS
- 1786. VIDA MÉDIA
- 1787. VIDA ÚTIL DE ARMAZENAMENTO USE VIDA MÉDIA
- 1788. VIDA ÚTIL DO DOCUMENTO USE VIDA MÉDIA
- 1789. VIDEOCLIPES
- 1790. VIDEOCONFERÊNCIAS
- 1791. VIDEODISCOS
- 1792. VIDEOTECAS
- 1793. VIDEOTEIPES USE FITAS DE VÍDEO
- 1794. VIOLAÇÃO DE DIREITOS AUTORAIS
- 1795. VIOLAÇÃO DE MARCAS REGISTRADAS
- 1796. VOCABULÁRIOS CONTROLADOS USE LINGUAGENS DOCUMENTÁRIAS
- 1797. VULGARIZAÇÃO DA CIÊNCIA USE DIVULGAÇÃO CIENTÍFICA
- 1798. W3C USE ASSOCIAÇÕES DE ELABORAÇÃO DE NORMAS
- 1799. WAN USE REDES DE LONGA DISTÂNCIA
- 1800. WEB USE WORLD WIDE WEB
- 1801. WEB DESIGNERS
- 1802. WEB INTELIGENTE USE WEB SEMÂNTICA
- 1803. WEB INVISÍVEL
- 1804. WEB OCULTA USE WEB INVISÍVEL
- 1805. WEB PROFUNDA USE WEB INVISÍVEL
- 1806. WEB SEMÂNTICA
- 1807. WEBLOGS USE BLOGS
- 1808. WEBMETRIA
- 1809. WORLD WIDE WEB
- 1810. WORLD WIDE WEB CONSORTIUM USE ASSOCIAÇÕES DE ELABORAÇÃO DE NORMAS
- 1811. WWW USE WORLD WIDE WEB
- 1812. XML
- 1813. ZOOLOGIA
- 1814. ZOOTECNIA

ANEXOB - Lista de palavras vaziasLista de Palavras Vazias

A	AOS	AVALIAR
À	APENAS	AVALIARAM
ABAIXO	APESAR	BAIXA
ABORDA	APLICA	BAIXAS
ABORDADOS	APLICADO	BAIXO
ABORDAMOS	APLICADOS	BAIXOS
ABORDANDO	APLICAM	BASTANTE
ABORDARAM	APLICAMOS	BEM
ACASO	APLICANDO	BILHÃO
ACERCA	APLICAR	BOM
ACIMA	APLICOU	CÁ
AFINAL	APONTA	CADA
AFIRMA	APONTAM	CAMINHO
AFIRMAR	APONTAR	CASO
AFIRMARAM	APONTARAM	CATORZE
AFIRMOU	APÓS	CEDO
AFORA	APRESENTA	CEM
AGORA	APRESENTAÇÃO	CENTENA
AI	APRESENTADA	CENTÉSIMO
AINDA	APRESENTADAS	CERTO
ALÉM	APRESENTADO	CIMA
ALGO	APRESENTADOS	CINCO
ALGUÉM	APRESENTAM	CINQUENTA
ALGUM	APRESENTAMOS	COM
ALGUMA	APRESENTANDO	COMIGO
ALGUMAS	APRESENTAR	COMO
ALGUNS	APRESENTARAM	COMPRIDO
ALI	APRESENTARÃO	CONFORME
ALIÁS	APRESENTAREM	CONHECEM
ALTA	APRESENTAREMOS	CONHECER
ALTO	APRESENTAVAM	CONHECIDA
AMANHA	APRESENTOU	CONHECIDAS
AMANHÃ	AQUELA	CONHECIDO
AMBAS	AQUELAS	CONHECIMENTOS
AMBOS	AQUELE	CONOSCO
AMPLA	AQUELES	CONSIGO
AMPLAMENTE	AQUÉM	CONTANTO
AMPLO	AQUI	CONTIGO
ANALISA	AQUILO	CONTRA
ANALISADA	AS	CONTUDO
ANALISADAS	ÀS	CONVOSCO
ANALISADO	ASSIM	CUIDADO
ANALISADOS	ATÉ	CUJA
ANALISANDO	ATRÁS	CUJAS
ANALISAR	ATRAVÉS	CUJO
ANALISAREMOS	ATUAL	DA
ANTE	ATUALMENTE	DADA
ANTES	AUMENTAR	DADAS
AO	AUMENTO	DAQUELA
DAQUELAS	DO	ESTIVÉRAMOS
DAQUELE	DOBRO	ESTIVEREM

DAQUELES	DOS	ESTIVERMOS
DAQUILO	DOZE	ESTIVESSE
DAS	DUAS	ESTIVESSEM
DE	DUPLO	ESTIVÉSSEMOS
DEBAIXO	DURANTE	ESTOU
DÉCIMO	DUZENTAS	ESTUDA
DELA	DUZENTOS	ESTUDADA
DELAS	DÚZIA	ESTUDADAS
DELE	E	ESTUDADO
DELES	É	ESTUDADOS
DEMAIS	EIS	ESTUDAMOS
DENTRO	ELA	ESTUDAR
DEPOIS	ELAS	ESTUDIOSOS
DESDE	ELE	ESTUDOU
DESENVOLVEM	ELES	EU
DESENVOLVEMOS	EM	EXCETO
DESENVOLVER	EMBORA	EXCLUSIVA
DESENVOLVEU	ENFIM	EXCLUSIVO
DESENVOLVIDA	ENQUANTO	EXEMPLO
DESENVOLVIDAS	ENTÃO	EXEMPLOS
DESENVOLVIDO	ENTRE	FARÁ
DESENVOLVIDOS	ENTRETANTO	FAZ
DESENVOLVO	ERAM	FAZER
DESLIGADO	ÉRAMOS	FAZIA
DESSA	ESSA	FEZ
DESSAS	ESSAS	FIM
DESSE	ESSE	FOI
DESSES	ESSES	FOMOS
DESTA	ESTA	FOR
DESTAS	ESTÁ	FORA
DESTE	ESTAMOS	FORAM
DESTES	ESTANDO	FÔRAMOS
DESTOA	ESTÃO	FOREM
DETERMINADO	ESTAR	FORMAÇÃO
DEVE	ESTARÁ	FORMADA
DEVEM	ESTARÃO	FORMADO
DEVERÁ	ESTARIA	FORMADORAS
DEVIDO	ESTARIAM	FORMAL
DEZ	ESTAS	FORMANDO
DEZENA	ESTAVA	FORMAR
DEZENOVE	ESTAVAM	FORMARAM
DEZESSEIS	ESTÁVAMOS	FORMARIA
DEZESSETE	ESTE	FORMOS
DEZOITO	ESTEJA	FORMOU
DIFERENTE	ESTEJAM	FORTE
DIFERENTES	ESTEJAMOS	FORTES
DIMINUIR	ESTES	FOSSE
DIMINUIÇÃO	ESTEVE	FOSSEM
DIREITA	ESTIVE	FÔSSEMOS
DISSO	ESTIVEMOS	FRACA
DIVERSOS	ESTIVER	FRACAS
DIZ	ESTIVERA	FRACO
DIZER	ESTIVERAM	FRACOS

FUI	ME	NOVAS
HÁ	MEDE	NOVE
HAJA	MEDIDO	NOVECENTAS
HAJAM	MEDINDO	NOVECENTOS
HAJAMOS	MEDIR	NOVENTA
HÃO	MEIOS	NOVO
HAVEMOS	MELHOR	NUM
HAVIA	MENOR	NUMA
HEI	MENOS	NUMAS
HOJE	MESMA	NUNCA
HORAS	MESMAS	NUNS
HOUVE	MESMO	O
HOUVEMOS	MESMOS	OBA
HOUVER	METADE	OBJETIVA
HOUVERA	MEU	OBJETIVAS
HOUVERÁ	MEUS	OBJETIVOS
HOUVERAM	MIL	OBTENDO
HOUVÉRAMOS	MILÉSIMO	OBTER
HOUVERÃO	MILÉSIMOS	OBTIDA
HOUVEREI	MILHÃO	OBTIDAS
HOUVEREM	MILHAR	OBTIDO
HOUVEREMOS	MILHÕES	OBTIDOS
HOUVERIA	MIM	OH
HOUVERIAM	MINHA	OITAVO
HOUVERÍAMOS	MINHAS	OITENTA
HOUVERMOS	MODO	OITO
HOUVESSE	MODOS	OITOCENTAS
HOUVESSEM	MUITO	OITOCENTOS
HOUVÉSSEMOS	MUITOS	OLHA
HUM	NA	ONDE
INCLUSIVE	NADA	ONTEM
INFORMADOS	NÃO	ONZE
INFORMAR	NAS	OPA
INFORMARAM	NEM	ORA
INFORMATIVOS	NENHUM	OS
INICIAR	NESTA	OU
INICIO	NESTAS	OUTRA
IR	NESTE	OUTRAS
IRÁ	NESTES	OUTRO
ISSO	NESSE	OUTRORA
ISTO	NESSES	OUTROS
JÁ	NESSA	PARA
JAMAIS	NESSAS	PARTE
LHE	NINGUÉM	PARTES
LHES	NO	PARTIDA
LIGADO	NOME	PARTINDO
LOGO	NONO	PARTIR
LONGE	NOS	PEGAR
MAIOR	NÓS	PELA
MAIORIA	NOSSA	PELAS
MAIORIAS	NOSSAS	PELO
MAIS	NOSSO	PELOS
MANEIRA	NOSSOS	PEQUENO
MANEIRAS	NOVA	PERANTE
MAS	NOVAMENTE	PERTO
PESQUISADAS	QUATROCENTOS	SEREMOS
PESQUISADO	QUE	SERIA
PESQUISADORAS	QUÊ	SERIAM

PESQUISADORES	QUEM	SERÍAMOS
PESQUISADOS	QUER	SESSENTA
PIOR	QUICA	SETE
PODE	QUIETO	SETECENTAS
PÔDE	QUINHENTAS	SETECENTOS
PODEM	QUINHENTOS	SETENTA
PODEMOS	QUINTA	SÉTIMO
PODENDO	QUINTO	SEU
PODERÁ	QUINZE	SEUS
PODERÃO	REALIZAÇÃO	SEXTO
PODEREMOS	REALIZADA	SI
PODERIA	REALIZADAS	SIDO
PODERIAM	REALIZADO	SIM
PODEROSAS	REALIZADOS	SÓ
PODIA	REALIZAMOS	SOB
POIS	REALIZAR	SOBRETUDO
POR	REALIZARAM	SOMENTE
PÔR	REALIZAREM	SOMOS
POREM	REALIZAVAM	SOU
PORÉM	RELATIVA	SUA
PORQUANTO	RELATIVAS	SUAS
PORQUE	RELATIVO	TAIS
PORTANTO	RELATIVOS	TAL
PORVENTURA	RESULTA	TALVEZ
POSSÍVEL	RESULTADO	TAMBÉM
POUCO	RESULTADOS	TAMPOUCO
PRIMEIRA	RESULTAM	TANTO
PRIMEIRAS	RESULTANTE	TÃO
PRIMEIRO	RESULTANTES	TARDE
PRIMEIROS	RESULTARAM	TE
PRINCIPAIS	SABER	TEM
PRINCIPALMENTE	SALVE	TÊM
PROJETADOS	SALVO	TEMOS
PROJETISTA	SÃO	TENDO
PROJETISTAS	SE	TENHA
PRÓPRIA	SEGUNDA	TENHAM
PRÓPRIAS	SEGUNDO	TENHAMOS
PRÓPRIO	SEIS	TENHO
PUXA	SEISCENTAS	TENTAR
QUAIS	SEISCENTOS	TENTARAM
QUAL	SEJA	TENTE
QUALQUER	SEJAM	TENTEI
QUANDO	SEJAMOS	TER
QUANTA	SEMELHANTE	TERÁ
QUANTAS	SEMPRE	TERÃO
QUANTOS	SENÃO	TERÇA
QUÃO	SENDO	TERCEIRA
QUARENTA	SEQUER	TERCEIRO
QUARTO	SER	TERÇO
QUATORZE	SERÁ	TEREI
QUATRO	SERÃO	TEREMOS
QUATROCENTAS	SEREI	TERIA
TERIAM	UMAS	
TERÍAMOS	ÚNICAS	
TEU	ÚNICO	
TEUS	ÚNICOS	
TEVE	UNS	
TI	USA	

TINHA	USAR	
TINHAM	USARAM	
TÍNHAMOS	USE	
TIVE	USOU	
TIVEMOS	UTILIZA	
TIVER	UTILIZAÇÃO	
TIVERA	UTILIZADA	
TIVERAM	UTILIZADAS	
TIVÉRAMOS	UTILIZADO	
TIVEREM	UTILIZADOS	
TIVERMOS	UTILIZAM	
TIVESSE	UTILIZAMOS	
TIVESSEM	UTILIZANDO	
TIVÉSSEMOS	UTILIZAR	
TODAVIA	UTILIZARAM	
TODOS	UTILIZEI	
TRABALHA	UTILIZOU	
TRABALHAM	VÁRIAS	
TRABALHANDO	VÁRIOS	
TRABALHAR	VEJA	
TRABALHOSAS	VEM	
TRÁS	VER	
TRATA	VEZ	
TRATADA	VEZES	
TRATADO	VIGÉSIMO	
TRATADOS	VINTE	
TRATAM	VIVA	
TRATAMENTO	VOCÊ	
TRATAMOS	VOCÊS	
TRATANDO	VOS	
TRATAR	VÓS	
TRÊS	VOSSA	
TREZE	VOSSAS	
TREZENTAS	VOSSO	
TREZENTOS	VOSSOS	
TRIGÉSIMO		
TRILHÃO		
TRINTA		
TRÍPLICE		
TRIPLO		
TU		
TUA		
TUAS		
TUDO		
UI		
ÚLTIMO		
UM		
UMA		
UMA		

ANEXOC - Arquivo de Termos Gerais

Arquivo de Termos Gerais – TCITG

1. ACERVOS (BIBLIOTECA) TG COLEÇÕES
2. ACERVOS PESSOAIS TG COLEÇÕES
3. ACESSO À BIBLIOTECA TG ACESSO
4. ACESSO À INFORMAÇÃO TG ACESSO
5. ACESSO AO DOCUMENTO TG ACESSO
6. ACESSO LIVRE TG ACESSO
7. ACESSO UNIVERSAL TG ACESSO
8. ACOPLAMENTO BIBLIOGRÁFICO TG ANÁLISE DE CITAÇÃO
9. ADMINISTRAÇÃO (ÁREA DO CONHECIMENTO) TG CIÊNCIAS SOCIAIS APLICADAS
10. AGÊNCIAS DE NOTÍCIAS TG MÍDIAS DE NOTÍCIAS
11. AGENTES INTELIGENTES TG PROGRAMAS DE RECUPERAÇÃO DA INFORMAÇÃO
12. AGRONOMIA TG CIÊNCIAS AGRÁRIAS
13. ALFABETAÇÃO LETRA POR LETRA TG ARRANJO ALFABÉTICO
14. ALFABETAÇÃO PALAVRA POR PALAVRA TG ARRANJO ALFABÉTICO
15. ALGORITMOS TG CONTEÚDOS DA INFORMAÇÃO
16. ALMANAQUES TG OBRAS DE REFERÊNCIA
17. AMBIGUIDADE TG ATRIBUTOS DA INFORMAÇÃO
18. ANAIS DE CONGRESSOS TG PUBLICAÇÕES
19. ANAIS DE CONGRESSOS TG PUBLICAÇÕES SERIADAS
20. ANÁLISE COMPARATIVA TG MÉTODOS DE PESQUISA E ANÁLISE
21. ANÁLISE DE AGRUPAMENTOS TG ANÁLISE DE DADOS
22. ANÁLISE DE CITAÇÃO TG BIBLIOMETRIA
23. ANÁLISE DE COCORRÊNCIA TG ANÁLISE DE DADOS
24. ANÁLISE DE CUSTOS TG ANÁLISE COMPARATIVA
25. ANÁLISE DE CUSTOS TG MÉTODOS DE ANÁLISE NA INTELIGÊNCIA COMPETITIVA
26. ANÁLISE DE DADOS TG ANÁLISE COMPARATIVA
27. ANÁLISE DE MACROAMBIENTE TG MÉTODOS DE ANÁLISE NA INTELIGÊNCIA COMPETITIVA
28. ANÁLISE DE TAREFAS TG MÉTODOS DE PESQUISA E ANÁLISE
29. ANÁLISE DO FLUXO DE TRABALHO TG ANÁLISE DE TAREFAS
30. ANÁLISE LINGUÍSTICA TG PROCESSAMENTO DE TEXTOS
31. ANÁLISE QUALITATIVA TG MÉTODOS DE PESQUISA E ANÁLISE
32. ANÁLISE QUANTITATIVA TG MÉTODOS DE PESQUISA E ANÁLISE
33. ANÁLISE SEMÂNTICA TG ANÁLISE LINGUÍSTICA
34. ANIMAÇÃO ASSISTIDA POR COMPUTADOR TG COMPUTAÇÃO GRÁFICA
35. ANTONÍMIA TG RELAÇÕES ENTRE CONCEITOS
36. ANTROPOLOGIA TG CIÊNCIAS HUMANAS
37. ANUÁRIOS TG OBRAS DE REFERÊNCIA
38. ANUÁRIOS TG PUBLICAÇÕES SERIADAS
39. APLICATIVOS TG PROGRAMAS DE COMPUTADOR QUANTO AO USO
40. APRENDIZADO DE MÁQUINA TG INTELIGÊNCIA ARTIFICIAL
41. AQUISIÇÃO DE CONHECIMENTO TG ENGENHARIA DO CONHECIMENTO
42. AQUISIÇÃO DE DOCUMENTOS TG DESENVOLVIMENTO DE COLEÇÕES
43. ARMAZENAMENTO COMPACTO TG ARMAZENAMENTO DE DOCUMENTOS
44. ARQUEOLOGIA TG CIÊNCIAS HUMANAS
45. ARQUITETOS DE INFORMAÇÃO TG PROFISSIONAIS DE INFORMAÇÃO
46. ARQUITETURA E URBANISMO TG CIÊNCIAS SOCIAIS APLICADAS
47. ARQUIVISTAS TG PROFISSIONAIS DE INFORMAÇÃO
- [...]
879. TRANSMISSÃO FAC-SIMILAR TG TELECOMUNICAÇÕES
880. TREINAMENTO DE USUÁRIOS TG SERVIÇOS DE BIBLIOTECA
881. TRUNCAMENTO TG ESTRATÉGIAS DE BUSCA

- 882. TURISMO TG CIÊNCIAS SOCIAIS APLICADAS
- 883. TUTORIAIS TG OBRAS DE REFERÊNCIA
- 884. UNIVERSIDADES TG INSTITUIÇÕES DE ENSINO E PESQUISA
- 885. URI TG IDENTIFICADORES PERSISTENTES
- 886. URL TG URI
- 887. URN TG URI
- 888. USABILIDADE (PROGRAMAS DE COMPUTADOR) TG ATRIBUTOS DE PROGRAMAS DE COMPUTADOR
- 889. USO LEGÍTIMO TG PROPRIEDADE INTELECTUAL
- 890. USUÁRIOS EXPERIENTES TG USUÁRIOS
- 891. USUÁRIOS FINAIS TG USUÁRIOS
- 892. USUÁRIOS NOVATOS TG USUÁRIOS
- 893. VALIDADE TG ATRIBUTOS DA INFORMAÇÃO
- 894. VALIDAÇÃO DE DADOS TG PROCESSAMENTO DE DADOS
- 895. VERIFICAÇÃO DE DADOS TG PROCESSAMENTO DE DADOS
- 896. VIDA MÉDIA TG ANÁLISE DE CITAÇÃO
- 897. VIDEOCLIPES TG GRAVAÇÕES DE VÍDEO
- 898. VIDEOCONFERÊNCIAS TG TELECONFERÊNCIAS
- 899. VIDEODISCOS TG DISCOS ÓTICOS
- 900. VIDEOTECAS TG BIBLIOTECAS
- 901. VIOLAÇÃO DE DIREITOS AUTORAIS TG CRIME
- 902. VIOLAÇÃO DE MARCAS REGISTRADAS TG CRIME
- 903. WEB DESIGNERS TG PROFISSIONAIS DE INFORMAÇÃO
- 904. WEB SEMÂNTICA TG WORLD WIDE WEB
- 905. WEBMETRIA TG METRIAS DA INFORMAÇÃO E COMUNICAÇÃO
- 906. WORLD WIDE WEB TG INTERNET
- 907. WORLD WIDE WEB TG MEIOS DE COMUNICAÇÃO DE MASSA
- 908. XML TG LINGUAGENS DE MARCAÇÃO
- 909. ZOOLOGIA TG CIÊNCIAS BIOLÓGICAS
- 910. ZOOTECNIA TG CIÊNCIAS AGRÁRIAS