Artigo Original

Encontros Bibli

MÉTODO PARA AVALIAÇÃO DIRETA DA INDEXAÇÃO AUTOMÁTICA VIA JULGAMENTO POR INDEXADORES

Method for direct evaluation of automatic indexing through assessment by indexers

Renato Fernandes Correa
Universidade Federal de Pernambuco, Departamento de Ciência da Informação
Recife, PE, Brasil
renato.correa@ufpe.br
https://orcid.org/0000-0002-9880-8678

Mariângela Spotti Lopes Fujita
Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP)
Programa de Pós-Graduação em Ciência da Informação
Marília, SP, Brasil
mariangela.fujita@unesp.br
https://orcid.org/0000-0002-8239-7114

A lista completa com informações dos autores está no final do artigo

RESUMO

Objetivo: Com a finalidade de avaliar um sistema de indexação automática aplicado a *e-books*, este trabalho propõe e aplica um método para avaliação direta da indexação automática via julgamento por indexadores, quanto à qualidade dos termos de indexação atribuídos automaticamente aos documentos.
Método: Realizou-se uma pesquisa metodológica de natureza descritiva e aplicada, utilizando-se procedimentos técnicos da pesquisa bibliográfica e pesquisa empírica. Inicialmente, por meio de revisão de literatura, delimitaram-se as etapas do método proposto de avaliação direta da indexação automática via julgamento por indexadores, para depois proceder à construção de instrumento para coleta de dados e aplicação do método proposto na avaliação da indexação automática do sistema SISTRA, na indexação de *e-books* técnicos-científicos.
Resultado: O método proposto de avaliação é apresentado através de diagrama e descrição de etapas. O método consiste, primeiramente, no julgamento por indexadores da qualidade dos termos atribuídos pelo sistema de indexação automática a uma amostra de documentos digitais e, posteriormente, na análise de valores calculados para métricas de qualidade da indexação automática. A aplicação do método proposto se mostrou útil numa primeira avaliação de um sistema de indexação automática.
Conclusões: Conclui-se que o método proposto de avaliação direta da indexação automática via julgamento por indexadores propicia a padronização da avaliação e sua prática por profissionais da informação, e que a avaliação direta é uma atividade necessária para a aplicação e a adoção da indexação automática na indexação por assunto de documentos digitais, no âmbito das unidades de informação.
PALAVRAS-CHAVE: Indexação automática. Sistemas de indexação automática. Avaliação da indexação automática. Avaliação direta da indexação automática. Trajetória metodológica.

ABSTRACT

Objective: In order to evaluate an automatic indexing system applied to e-books, this paper proposes and applies a method for direct evaluation of automatic indexing through assessment by indexers of the quality of the index terms automatically assigned to the documents.
Method: A methodological research of descriptive and applied nature was carried out, using technical procedures of bibliographic research and empirical research. Initially, by means of a literature review, the steps of the proposed method of direct evaluation of automatic indexing through assessment by indexers were delimited, to then proceed with the construction of an instrument for data collection and application of the proposed method in the evaluation of automatic indexing of the SISTRA system in the indexing of technical-scientific e-books.
Result: The proposed evaluation method is described by means of a diagram and systematic description. The method consists first in the judgment by indexers of the quality of terms assigned by the automatic indexing system to a sample of digital documents, and then in the analysis of values calculated for quality metrics of automatic indexing. The application of the proposed method proved useful in a first evaluation of an automatic indexing system.
Conclusions: We conclude that the proposed method of direct evaluation of automatic indexing through assessment by indexers provides the standardization of evaluation and its practice by information professionals, and that direct evaluation is a necessary activity for the application and adoption of automatic indexing in indexing by subject of digital documents in the scope of information units.

KEYWORDS: Automatic indexing. Automatic indexing system. Evaluation of automatic indexing. Direct evaluation of automatic indexing. Methodological pathway.

1 INTRODUÇÃO

A indexação por assunto de documentos digitais tem importante papel no enriquecimento de metadados, na estruturação de coleções por temática, na busca e recuperação pelos usuários e no estabelecimento de conexões semânticas entre recursos.

Contudo, definir quais termos representam os assuntos principais de documentos não é uma tarefa simples, exigindo do indexador o conhecimento e a prática de indexação. Adicionalmente, o processo de indexar lotes de documentos pode depender de um longo tempo de atividade intelectual, além de sofrer influência da dispersão terminológica característica do uso de termos em linguagem natural e da subjetividade envolvida na atribuição de tais termos.

Alternativamente, a indexação por assunto pode ser realizada pela indexação automática. A indexação automática é assim definida por Corrêa e Lapa (2013, p. 258):

> Um conjunto de operações realizadas pelo computador, de natureza estatística, linguística, ou de programação, destinado a selecionar termos como elementos descritivos de um documento pelo processamento automático de seu conteúdo.

A indexação automática visa a tornar o processo de indexação mais rápido e menos custoso, agilizando o processo de atribuição de termos relevantes aos documentos digitais, por meio do processamento computacional dos termos presentes no conteúdo textual de resumos e textos completos.

A indexação automática pode ser por atribuição ou por extração (LANCASTER, 2004). Quando um vocabulário controlado é utilizado no processo, ocorre a indexação automática por atribuição; caso contrário, tem-se a indexação automática por extração.

Por instância, alguns trabalhos recentes têm investigado a indexação automática por atribuição de artigos de periódicos da área de Ciência da Informação escritos no idioma português do Brasil (BANDIM; CORREA, 2018, 2019; SILVA; CORREA, 2020; SILVA; CORREA; GIL-LEIVA, 2020) e em espanhol (GIL-LEIVA; ORTUÑO; CORRÊA, 2022).

Na última década, observa-se um aumento de relatos de aplicação da indexação automática por atribuição na indexação por assunto de documentos técnico-científicos em formato digital. Como exemplos, podem ser citadas as iniciativas nas bibliotecas nacionais da Finlândia (SUOMINEN, 2019; SUOMINEN; INKINEN; LEHTINEN, 2022), Alemanha (JUNGER, 2018; JUNGER; SCHOLZE, 2021) e Estônia (ASULA *et al*., 2021).

Entretanto, a avaliação da qualidade da indexação automática se faz tão necessária quanto sua aplicação em unidades de informação, pois é por intermédio dela que se decide utilizar uma ferramenta existente ou melhorar seu funcionamento (GOLUB *et al*., 2016).

A avaliação da indexação automática não se encontra normatizada ou padronizada, porém, procedimentos para avaliação direta têm-se consolidado na literatura científica (KIM *et al*., 2010; HASAN; NG, 2014; GOLUB *et al*., 2016). A finalidade de tal avaliação é aferir a qualidade dos termos de indexação associados automaticamente aos documentos, levando em conta o conhecimento especializado do indexador.

Golub *et al*. (2016) propuseram um *framework*, baseado em revisão da literatura e na experiência coletiva dos autores, consistindo em abordagens e um conjunto de recomendações associadas para a avaliação da indexação automática. Entre as abordagens, destaca-se a avaliação direta da indexação automática, a qual pode ser dividida em duas: por meio do julgamento por avaliador ou da comparação com um padrão-ouro. No presente artigo, adotamos a primeira, atribuindo o papel de avaliador ao indexador, devido ao conhecimento especializado na atividade profissional da indexação.

Diante da problemática da avaliação direta da indexação automática e de sua realização, no âmbito das bibliotecas universitárias, na incorporação em lote de *e-books* ao catálogo, este artigo parte do seguinte problema de pesquisa: como efetuar a avaliação direta da indexação automática via julgamento por indexadores? A fim de responder a esse questionamento, o objetivo deste trabalho é propor e aplicar um método para avaliação direta da indexação automática via julgamento por indexadores.

A justificativa deste estudo se baseia na concepção de instrumento, na forma de método, que guie os profissionais da informação na realização da atividade de avaliação direta da indexação automática via julgamento por indexadores, a qual se faz necessária para a aplicação ou o aperfeiçoamento de sistemas de indexação automática.

2 AVALIAÇÃO DA INDEXAÇÃO AUTOMÁTICA

Na realização do processo de indexação automática, os termos de indexação são definidos via análise automática do texto do documento digital realizada por *software*, sendo esse configurado previamente pelo profissional da informação, a fim de efetuar a seleção de termos sem interferência humana.

As pesquisas sobre indexação automática de documentos envolvem diferentes áreas de conhecimento, com maiores contribuições das áreas de Ciência da Informação e

Ciência da Computação. Recomendam-se os artigos de Gil-Leiva, Ortuño e Corrêa (2022) e Golub (2021), para a leitura de uma revisão bibliográfica internacional sobre a indexação automática.

O emprego da indexação automática nas unidades de informação visa principalmente a garantir a inclusão em lote de um grande número de documentos digitais no catálogo *on-line* ou em repositório digital, poupar tempo e trabalho do profissional indexador na indexação e possibilitar uma melhor uniformidade e homogeneidade com relação aos termos de indexação.

Porém, uma questão importante, a qual antecede a aplicação da indexação automática nas unidades de informação, é a avaliação da indexação automática quanto à qualidade dos termos de indexação, objetivando a tomada de decisão quanto ao uso operacional ou a necessidade de aprimoramento de um sistema de indexação automática.

A avaliação da indexação automática é compreendida como um processo empírico ou experimental, que busca medir a qualidade da indexação pela análise do resultado da mesma, isto é, dos termos de indexação.

De acordo com Gil-Leiva (2008), a avaliação da indexação automática pode ser feita de forma extrínseca quantitativa, utilizando duas abordagens: comparando os termos das indexações automática e intelectual e aferindo a interconsistência; ou mensurando a eficácia na recuperação da informação, empregando-se índices resultantes de ambas as indexações. A última abordagem é categorizada por Golub *et al*. (2016) como uma abordagem de avaliação indireta da indexação automática, já que avalia a qualidade da indexação indiretamente, por meio da análise da performance na recuperação de informação. Por sua vez, a primeira abordagem é categorizada como avaliação direta da indexação automática (GOLUB *et al*., 2016), pois avalia a indexação automática quanto ao alcance dos termos da indexação intelectual.

Adicionalmente, a avaliação extrínseca da indexação automática pode ser feita por meio de índices de revocação, precisão e medida F, tendo a indexação intelectual como padrão de referência ou padrão-ouro (SILVA; CORREA; GIL-LEIVA, 2020). Nesse contexto, a revocação corresponde ao percentual de termos da indexação intelectual propostos pela indexação automática, a precisão corresponde ao percentual de termos da indexação automática que correspondem a termos da indexação intelectual, e a medida F corresponde à média harmônica entre as métricas de revocação e precisão. Uma discussão mais detalhada sobre a formulação e aplicação de tais índices, pode ser consultada em (SILVA; CORREA; GIL-LEIVA, 2020).

Golub *et al.* (2016) discutem questões relacionadas com as abordagens de avaliação da indexação automática existentes e propõem um *framework* de avaliação, com base numa revisão sistemática da literatura e na experiência coletiva dos autores. O *framework* consiste em um conjunto de recomendações e distingue as seguintes três abordagens: avaliação da qualidade da indexação diretamente por intermédio do julgamento por avaliador ou da comparação com um padrão-ouro; avaliação da qualidade da indexação assistida por computador diretamente no contexto de um fluxo de trabalho de indexação; e avaliação da qualidade da indexação indiretamente através da análise de desempenho na recuperação da informação.

Quanto à abordagem da avaliação direta da indexação automática, como proposta por Golub *et al*. (2016), percebe-se que essa pode ser dividida em duas abordagens específicas: via julgamento por avaliador ou por comparação com um padrão-ouro.

A avaliação direta da indexação automática por comparação com um padrão-ouro (GOLUB *et al.*, 2016) toma os termos da indexação intelectual como padrão de referência de qualidade ou padrão-ouro (do inglês *gold standard*) e assume a completude e corretude dos termos da indexação intelectual. Essa abordagem de avaliação direta da indexação automática ocorre por meio da comparação dos termos das indexações intelectual e automática, e cálculo de valores para métricas, como precisão, revocação e medida F, que refletem diretamente a qualidade da indexação automática, em função da proposição de termos da indexação intelectual. Tal abordagem é a mais reportada na literatura científica (HASAN; NG, 2014; GOLUB *et al.*, 2016), devido principalmente à sua alta reprodutibilidade, porém, exige mais recursos humanos na construção de uma coleção de documentos indexados como padrão-ouro.

Alternativamente, a avaliação direta da indexação automática via julgamento por avaliador (GOLUB *et al*., 2016) se efetiva pelo julgamento por avaliador da qualidade dos termos da indexação automática como termos de indexação. Os bons termos de indexação são aqueles considerados como relevantes, corretos e apropriados aos documentos aos quais foram atribuídos. Essa abordagem é menos reportada na literatura científica, por ter uma reprodutibilidade limitada, contudo, tem a vantagem de exigir menos recursos humanos, numa primeira avaliação.

Golub *et. al*. (2016) citam apenas um trabalho que realizou uma avaliação direta da indexação via julgamento por avaliador. Trata-se da investigação de Rosenberg (1971 *apud* GOLUB *et al*., 2016), que comparou dois métodos de indexação de assuntos, utilizando um painel de juízes para classificar os termos de indexação atribuídos a um determinado

documento, numa escala de cinco pontos, baseada tanto na adequação ao documento quanto na utilidade como ponto de acesso. Os demais trabalhos citados implementaram uma avaliação no contexto da classificação de páginas da *web* e da indexação automática de imagens, bem como adotaram como avaliadores alguns usuários finais, sendo de contextos muito diferentes da avaliação da qualidade da indexação automática, no âmbito das unidades de informação.

Ademais, a abordagem recomendada para avaliar diretamente a qualidade da indexação proposta por Golub *et al*. (2016) é constituída por alguns passos, almejando a construção de um padrão-ouro como produto da avaliação. Tais passos envolvem: a seleção de documentos via amostragem estratificada por área de assunto dos documentos; a realização da indexação intelectual dos documentos por indexadores, especialistas no assunto, usuários finais e sistemas de indexação automática; a integração da lista de termos de indexação atribuídos aos documentos; e a avaliação da lista de termos atribuídos a cada documento por indexadores experientes e usuários finais, pela seleção e adição de termos relevantes, e remoção de termos erroneamente atribuídos.

Assim, com base nos trabalhos discutidos nesta seção e no melhor conhecimento dos autores, não se encontra especificado na literatura científica um método para a avaliação direta da indexação automática via julgamento por indexadores. Portanto, visando preencher essa lacuna, a proposta do presente artigo é a criação de tal método de avaliação.

Visando a contribuir com a padronização da avaliação direta da indexação automática amparada na avaliação dos termos de indexação por indexadores, a próxima seção detalha a metodologia usada na construção e aplicação do método proposto.

3 METODOLOGIA

Quanto aos objetivos ou fins, realizou-se uma pesquisa metodológica de natureza descritiva. Quanto aos meios, adotaram-se procedimentos técnicos da pesquisa bibliográfica pautada em revisão de literatura e da pesquisa empírica.

Baseando-se na categorização proposta por Vergara (2009), no presente artigo, procedeu-se a uma pesquisa metodológica, a qual compreende estudo relacionado ao desenvolvimento de instrumentos de captação ou manipulação da realidade, sendo associado, portanto, a caminhos, formas, maneiras e procedimentos para atingir

determinado fim. Nesse contexto, este trabalho propõe método para avaliação direta da indexação automática via julgamento por indexadores.

Adicionalmente, a pesquisa tem natureza descritiva, pois expõe e descreve os procedimentos do método proposto para avaliação direta da indexação automática via julgamento por indexadores, sistematizando-os na forma de método, além de descrever a análise dos resultados da aplicação do método proposto.

Quanto aos meios, constitui-se em pesquisa bibliográfica, por se fundamentar nos relatos de avaliação da indexação automática presentes na literatura científica, os quais foram discutidos na seção 2 deste trabalho. Os procedimentos de revisão de literatura foram empregados objetivando tornar explícitos os procedimentos de avaliação direta da indexação automática via julgamento por avaliador, sendo complementada com conhecimento e experiência dos autores, na criação do método de avaliação proposto. A escolha dos avaliadores como sendo indexadores se justifica pelo contexto de aplicação e avaliação da indexação automática nas unidades de informação ou bibliotecas, visando ao uso de sistemas de indexação automática como ferramenta de auxílio ao bibliotecário na realização mais célere da indexação para lotes de documentos digitais.

Além disso, o *Business Process Management* (BPM) (ABPMP,, 2013) foi utilizado para representar, na forma de diagrama, o fluxo e processos do método proposto de avaliação direta da indexação automática via julgamento por indexadores.

O método proposto de avaliação direta da indexação automática via julgamento por indexadores foi projetado para ter caráter genérico e poder ser particularizado e aplicado com uso de qualquer sistema de indexação automática por atribuição, desde que seja possível o acesso aos documentos de entrada e a saída do sistema para tal conjunto de documentos, na forma de lista de termos da indexação.

Após a proposição do método, foi efetivada uma pesquisa experimental ou empírica envolvendo a aplicação do método proposto na avaliação direta da indexação automática de *e-books* técnico-científicos, por intermédio do sistema SISTRA, no contexto da rede de bibliotecas da UNESP, onde foi possível validar a aplicabilidade e analisar os resultados obtidos com o método proposto.

O *software* SISTRA (SIStema para o TRatamento Automático de *e-book* multilíngue para textos técnico-científicos) foi desenvolvido por pesquisadores da UNESP, em parceria com pesquisadores da Universidade de Murcia e UFPA (GIL-LEIVA *et al.*, 2022), sendo um produto de projeto de pesquisa com fomento FAPESP (FUJITA, 2020).

O SISTRA foi escolhido pela inovação de ser um sistema de indexação automática que processa *e-books* no formato PDF com conteúdo no idioma português do Brasil, com a detecção de campos semânticos do documento (como título, sumário, capítulos, legendas e referências) e aplicação de regras heurísticas de indexação automática por atribuição, isto é, com a utilização de vocabulário controlado como fonte de descritores. Além da indexação automática, o sistema também promove a classificação automática, via atribuição de códigos de classificação, automaticamente, aos documentos. Porém, essa última funcionalidade do sistema não é explorada no presente trabalho, cujo foco é a avaliação da indexação automática.

O SISTRA emprega um conjunto de regras heurísticas baseadas na posição dos termos ou na frequência dos termos, para obter a indexação de assuntos, usando vocabulários controlados. Os vocabulários controlados são fornecidos pelo utilizador como parâmetros de entrada na configuração do sistema, estando disponível para comparação de termos durante o processo de indexação automática, e para consulta na checagem dos resultados obtidos.

As regras de posicionamento definem que, a depender da ocorrência ou posicionamento dos termos, em determinadas partes nos documentos, eles são atribuídos como termos de indexação. As regras estatísticas, amparadas na frequência de aparecimento dos termos em um documento e no conjunto de documentos, determinam que termos com valores de frequência acima de um limiar são atribuídos ao documento como termos de indexação. É possível também especificar regras mistas, compostas por condições de posição e regras estatísticas, ao mesmo tempo.

O SISTRA já se encontra configurado previamente com um conjunto de nove regras heurísticas, as quais podem ser visualizadas nas linhas da Figura 1.

Figura 1 – Regras heurísticas do SISTRA

Regra	Título	Sumário	Capítulo	Legenda	Referências Bibliográficas	DF	TF-IDF	IDF
R9	✓	✓	✗	✗	✗	0,0	0,0	1,5
R8	✓	✓	✗	✗	✗	0,0	0,02	0,0
R7	✗	✗	✗	✗	✗	200,0	0,0	0,0
R6	✗	✓	✗	✓	✓	0,0	0,0	0,0
R5	✓	✗	✗	✗	✓	0,0	0,0	0,0
R4	✓	✗	✗	✓	✗	0,0	0,0	0,0
R3	✗	✗	✗	✓	✓	0,0	0,0	0,0
R2	✗	✓	✗	✓	✗	0,0	0,0	0,0
R1	✓	✓	✗	✗	✗	0,0	0,0	0,0

Fonte: Dados da pesquisa (2023)

As regras de posicionamento no SISTRA correspondem às condições de um termo de indexação ocorrer em diferentes partes do documento para ser automaticamente atribuído a ele, nomeadamente: títulos, sumários, capítulos, legendas (figuras, tabelas, quadros etc.) e referências bibliográficas. Uma regra pode ser formada por um único elemento de posição ou por vários desses elementos, ao mesmo tempo. Um exemplo de regra desse tipo é a regra R3 da Figura 1, a qual pode ser assim expressa: se um termo do vocabulário controlado ocorre em legenda e também em referência bibliográfica, então, ele é atribuído ao documento como termo de indexação.

As regras estatísticas são fundamentadas na frequência de ocorrência dos termos e correspondem às condições de frequência a que os termos do vocabulário controlado devem atender, para serem automaticamente atribuídos ao documento. Utiliza-se DF (da sigla em inglês para *Document Frequency*) para representar a frequência absoluta de um termo, em um determinado documento, TF (da sigla em inglês para *Term Frequency*) para representar a frequência normalizada de um termo, em um determinado documento, IDF (da sigla em inglês para *Inverse Document Frequency*), para representar a frequência inversa nos documentos da coleção, e TF-IDF, para representar a multiplicação da frequência do termo normalizada em um documento pela frequência inversa nos

documentos. Uma regra pode ser formada, estabelecendo-se um valor limiar para um único elemento estatístico ou para cada um de vários elementos, ao mesmo tempo. Um exemplo de regra desse tipo é a regra R7 da Figura 1, a qual pode ser expressa como: se um termo do vocabulário controlado tiver DF acima de 200 ocorrências, então, ele é atribuído ao documento como termo de indexação.

Embora o SISTRA seja um protótipo de sistema de indexação automática com acesso restrito, a avaliação do mesmo se faz pertinente, e pode ser utilizada para fins de validação do método proposto, sem trazer prejuízo aos procedimentos metodológicos deste trabalho.

Na seção seguinte, são descritos a elaboração do método de avaliação direta da indexação automática, concebido em etapas, bem como os resultados de sua aplicação.

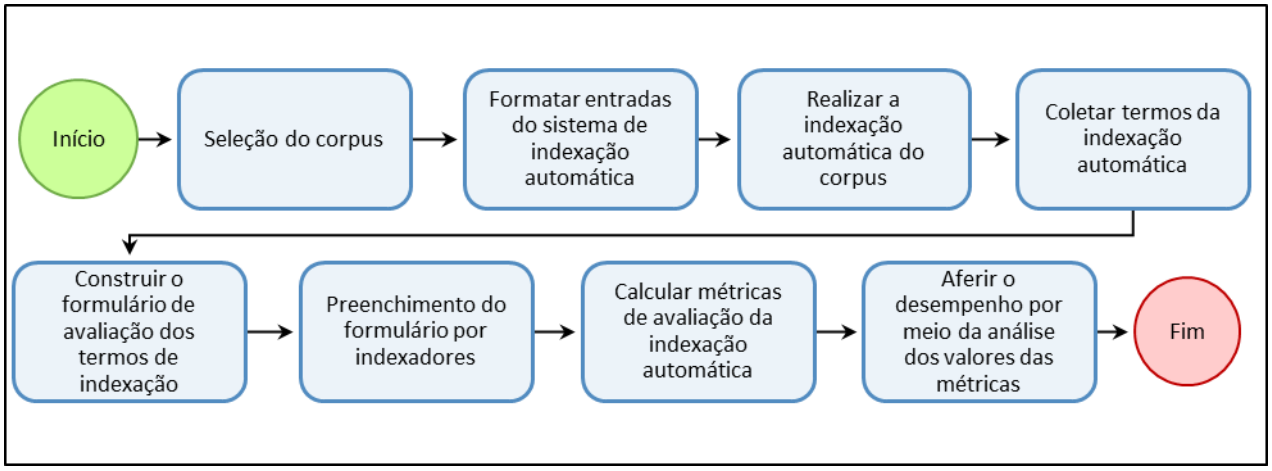
4 ANÁLISE DE RESULTADOS

Nesta seção, são apresentados os resultados do presente artigo, a saber: o método proposto de avaliação direta da indexação automática via julgamento por indexadores e os resultados da aplicação desse método.

4.1 Método para avaliação direta da indexação automática via julgamento por indexadores

De sorte a garantir uma melhor descrição do método proposto, a Figura 2 apresenta os passos ou processos pertinentes ao fluxo da avaliação direta da indexação automática via julgamento por indexadores.

Figura 2 – Diagrama da avaliação direta da indexação automática via julgamento por indexadores



Fonte: Dados da pesquisa (2023)

Cada etapa do método de avaliação direta da indexação automática via julgamento por indexadores é descrita a seguir:

1. Seleção do *corpus* – essa etapa consiste numa seleção via amostragem de um conjunto de documentos digitais, os quais serão indexados automaticamente por um sistema de indexação automática, definindo um *corpus* de documentos. Golub *et al*. (2016) recomendam que essa amostragem seja aleatória e contenha, no mínimo, 20 documentos por área de assunto, buscando-se evitar que os resultados na avaliação sejam suscetíveis a variações aleatórias. A granularidade da área de assunto não é pré-definida, dependendo da coleção de documentos e podendo assumir de uma Ciência a uma grande área de assunto, tais como “Ciências Humanas e Sociais”, “Ciências Exatas”, “Ciências Biológicas e da Saúde” etc. Porém, intuitivamente, o número total de documentos a ser analisado paralelamente por dois indexadores deve ser limitado a 30 documentos, a fim de se evitar a influência do cansaço intelectual e do tempo excessivo despendido na tarefa de avaliação. O planejamento para divisão do *corpus* para análise por duplas de membros de uma equipe de indexadores se faz necessário, pois, para a definição de *corpus* maiores;
2. Formatar entradas do sistema de indexação automática – as entradas do sistema de indexação automática envolvem o *corpus*, os vocabulários controlados e a lista de palavras-vazias, as quais precisam estar em formato processável pelo sistema de indexação automática. Geralmente, cada documento digital do *corpus* deve estar no formato PDF ou no formato de arquivo de texto, enquanto os vocabulários controlados devem estar no formato SKOS (que é um formato de texto baseado em RDF\XML), e as listas de palavras-vazias, no formato de arquivo de texto sem formatação. Um aspecto importante é o emprego de uma mesma tabela de caracteres na codificação dos arquivos de texto, que geralmente deve ser “UTF-8 sem BOM”;
3. Realizar a indexação automática do *corpus* – os parâmetros de configuração e as entradas formatadas devem ser fornecidos ao sistema de indexação automática, para a atribuição de termos de indexação aos documentos do *corpus*. Para sistemas que utilizam aprendizagem de máquina supervisionado, deverá ser fornecido um modelo de indexação automática, previamente treinado em um subconjunto diferente de documentos com descritores associados e originados de uma indexação intelectual, com o uso dos mesmos vocabulários controlados (SILVA; CORREA; GIL-LEIVA, 2020);

4. Coletar termos da indexação automática – os descritores atribuídos automaticamente a cada documento do *corpus* devem ser capturados de relatórios do sistema de indexação automática por atribuição, a fim de serem inseridos em formulário de avaliação dos termos da indexação automática;
5. Construir o formulário de avaliação dos termos de indexação – o formulário deve permitir acesso ao texto completo de cada documento do *corpus*, depois à lista de termos atribuídos automaticamente a cada documento e, para cada termo, prover campos para entrada do julgamento do indexador quanto à qualidade do termo de indexação atribuído automaticamente ao documento. O julgamento da qualidade de cada termo de indexação para cada documento deve ser capturado por meio da inserção de notas para determinados aspectos, como a qualidade do termo de indexação como ponto de acesso ao documento numa escala de três pontos (1-Não relevante, 2-Neutro, 3-Relevante) e uma nota em escala de cinco pontos, quanto à adequação do termo em ser correto e apropriado ao documento (1-Ruim, 2-Pouco aceitável, 3-Aceitável, 4-Bom, 5-Muito bom). Um exemplo de formulário será apresentado na seção 4.2;
6. Preenchimento do formulário por indexadores – o formulário de avaliação qualitativa dos termos de indexação deve ser preenchido paralelamente por dois indexadores e conter descritores para, no máximo, 30 documentos. Para *corpus* maiores, torna-se necessária a divisão dos documentos em múltiplos formulários a serem preenchidos por pares de indexadores, na avaliação dos termos de indexação. Visando a manter controle sobre a subjetividade na avaliação, é importante que o formulário seja apresentado aos indexadores juntamente com um documento-guia, o qual explique o contexto da avaliação e como proceder à avaliação, através do formulário. O formulário deve ser preenchido pelos indexadores com a seleção de nota para a qualidade de cada termo de indexação como ponto de acesso a cada documento e nota para a adequação de cada termo em ser correto e apropriado a cada documento, buscando, com isso, representar a qualidade de cada termo na representação temática do conteúdo de cada documento;
7. Calcular as métricas de avaliação da indexação automática – o cálculo de valores para métricas de qualidade da indexação automática se faz necessário, de sorte a se proceder com a avaliação. Tais métricas mensuram a semelhança dos descritores escolhidos pela indexação automática e pela indexação intelectual, sendo esta última feita a partir da avaliação dos termos. Para cálculo de métricas, é imperioso determinar

os bons termos de indexação apontados pelos indexadores. A determinação dos bons descritores ocorre a partir dos julgamentos quanto à qualidade dos termos, via uma expressão booleana envolvendo os valores numéricos médios das opções selecionadas no julgamento da qualidade e nota de cada termo. Sugere-se a expressão: QUALIDADE_MÉDIA > 2 AND NOTA_MÉDIA ≥ 3 (a qual corresponde à condição de a qualidade média ser maior que dois e a nota média ser maior ou igual a três). Se verdadeira a expressão para um termo atribuído automaticamente a um documento, ele é considerado um bom termo de indexação; caso contrário, ele é tomado como um termo ruim de indexação. A circunstância de a qualidade média ser maior que dois corresponde à escolha pelos indexadores da opção “Neutro” ou “Relevante”, para a qualidade do termo, sendo a opção “Relevante” escolhida por pelo menos um deles. O fato de a nota média ser maior ou igual a três corresponde à escolha pelos indexadores principalmente das opções “Aceitável”, “Bom”, ou “Muito bom” para a nota do termo, mas pode também ser originada das combinações das notas “Ruim” e “Muito bom”; “Pouco aceitável” e “Bom”, ou “Pouco aceitável” e “Muito bom”. A expressão composta com o e-lógico busca indicar, por conseguinte, como bons termos de indexação, aqueles termos apontados como tais por pelo menos um dos indexadores, via escolha de opções mais altas nas escalas para qualidade do termo e nota do termo. Com a determinação dos bons descritores, é possível então mensurar as seguintes métricas de avaliação da indexação automática: o número de termos comuns entre as indexações automática e intelectual e a precisão. A métrica de precisão mede o percentual de termos considerados como bons termos de indexação pelos indexadores do total de termos propostos pelo sistema de indexação automática;

8. Aferir o desempenho, por meio da análise dos valores das métricas – analisar os valores médios das métricas de avaliação da indexação automática, procurando determinar o desempenho do sistema de indexação automática quanto à qualidade dos termos de indexação atribuídos aos documentos do *corpus*. No presente método, tem-se a possibilidade do cálculo da precisão média do sistema de indexação automática.

Segundo Narukawa, Gil-Leiva e Fujita (2009), a avaliação da qualidade da indexação pode ser feita via análise da consistência na indexação, a qual reflete o grau de concordância na representação da informação de um documento por diferentes indexadores. Porém, como os bons termos de indexação destacados pelos indexadores pertencem a um subconjunto dos termos da indexação automática que equivale ao conjunto dos termos em comum entre as duas indexações, logo o número de termos da indexação

intelectual será igual ao número de termos em comum entre as duas indexações. Então, com base na fórmula de cálculo da consistência empregada por Narukawa, Gil-Leiva e Fujita (2009), se deduz que a consistência entre a indexação automática e indexação intelectual tem o mesmo valor do índice de precisão, como pode ser observado na equação a seguir:

$$C = \frac{T_{co}}{A + B - T_{co}} = \frac{T_{co}}{A + T_{co} - T_{co}} = \frac{T_{co}}{A} = P$$

Onde C é a consistência, T_{co} é o número de termos em comum entre as duas indexações, A é o número de termos da indexação automática, B é o número de termos da indexação intelectual, e sendo P a precisão, que corresponde a razão entre o número de termos em comum entre as duas indexações e o número de termos da indexação automática.

Consequentemente, nesse contexto de igualdade dos índices de consistência e precisão, a precisão média do sistema de indexação automática tem o mesmo valor da consistência média entre a indexação automática e a indexação intelectual. De posse do valor da precisão média, é possível ter uma primeira estimação quantitativa do desempenho do sistema de indexação automática.

Não se faz possível o cálculo das métricas de revocação e medida F, pois, para calculá-las, seria necessária a possibilidade de adição no formulário, pelos indexadores, de todos os termos omitidos pelo sistema de indexação automática a cada documento do *corpus*, o que tornaria a avaliação mais complexa e mais trabalhosa. Optou-se, assim, pela simplicidade e viabilidade do método proposto como uma primeira avaliação da indexação automática, ao não trazer a possibilidade de inclusão de termos de indexação pelos indexadores.

Portanto, o resultado quantitativo do processo de avaliação direta da indexação automática via julgamento por indexadores consiste dos valores médios das métricas de avaliação da indexação automática para o *corpus,* os quais devem ser dispostos em uma tabela, contemplando valores médios para os seguintes campos: número de termos atribuídos pela indexação automática, número de bons termos de indexação – que equivale ao número de termos comuns com a indexação intelectual – e a precisão. Uma análise qualitativa da qualidade da indexação automática pode ser realizada, pois, através dos valores médios apresentados na tabela e dos dados que deram origem aos valores médios. Um exemplo desse tipo de tabela é a Tabela 1, exposta na próxima subseção.

Na subseção a seguir, o método é aplicado na avaliação direta da indexação automática via julgamento por indexadores, obtida pelo sistema SISTRA aplicado a *e-books* técnico-científicos, no âmbito da rede de bibliotecas da UNESP.

4.2 Experimento de avaliação direta da indexação automática via julgamento por indexadores

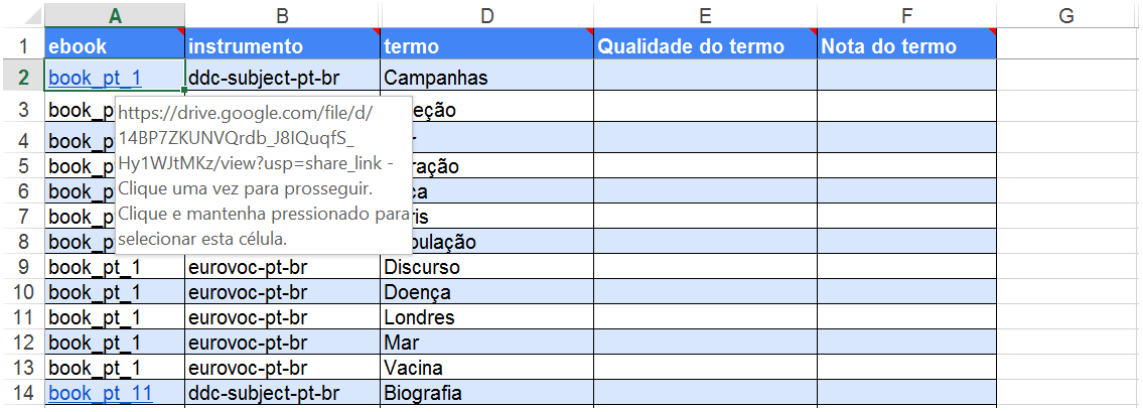
Para efetuar a indexação automática por atribuição de um conjunto de *e-books*, foi aplicado o *software* SISTRA (SIStema para o TRatamento Automático de *e-book* multilíngue para textos técnico-científicos), com a utilização de dois vocabulários controlados: a lista alfabética de assuntos relacionados com códigos de classificação, derivada de dados exportados do catálogo da rede de bibliotecas da UNESP, e o tesauro EuroVoc.

As etapas do método proposto de avaliação direta da indexação automática via julgamento por indexadores, descritas na subseção anterior, foram instanciadas para o contexto da avaliação do SISTRA, na indexação automática por atribuição de *e-books* técnico-científicos, no âmbito da rede de bibliotecas da UNESP. A particularização ou desdobramento de cada etapa do método proposto é descrito a seguir:

1. Seleção do *corpus* – foi criada uma amostra de 25 *e-books* da área de Ciências Humanas, todos em formato PDF, com conteúdo textual no idioma português do Brasil, processados em sua totalidade pelo SISTRA, com geração de termos de indexação;
2. Formatar entradas do sistema de indexação automática – os vocabulários controlados usados como instrumentos de indexação pelo SISTRA, na atribuição de termos, consistiram de uma lista alfabética de assuntos e o tesauro EuroVoc. A lista alfabética de assuntos foi inserida em arquivo de texto sem formatação, o qual contempla termos preferidos ou autorizados dispostos em uma linha do arquivo para cada termo. A lista alfabética de assuntos foi gerada a partir do processamento de dados do catálogo, referentes à atividade de catalogação de assunto de itens do acervo. No formato original, os dados correspondiam a assuntos associados com códigos de classificação em registros bibliográficos do catálogo, sendo fornecidos pela rede de bibliotecas da UNESP. O tesauro EuroVoc, no idioma português, foi baixado em formato SKOS da URL: https://eur-lex.europa.eu/browse/eurovoc.html?locale=pt. Uma lista de palavras-vazias (*stopwords*) para o idioma português do Brasil foi também inserida em arquivo de texto sem formatação;

3. Realizar a indexação automática do *corpus* – Os parâmetros de configuração e as entradas formatadas foram fornecidos como entrada ao sistema SISTRA, para a atribuição de termos de indexação dos vocabulários controlados aos documentos do *corpus*. O SISTRA é um sistema especialista, baseado em regras heurísticas, que não utiliza algoritmo de aprendizagem de máquina supervisionado, não necessitando de treinamento para gerar um modelo de indexação. Nesse processo, foram carregados no sistema os arquivos dos vocabulários controlados e lista de *stopwords*. Já se encontravam previamente especificadas, na configuração do SISTRA, um conjunto de nove regras heurísticas de atribuição de termos de indexação, conforme consta na Figura 1. Depois, foram carregados no sistema os arquivos no formato PDF dos *e-books* e, ao fazer isso, o SISTRA deu início ao processamento da indexação automática dos *e-books*, resultando na atribuição automática de termos de indexação a cada documento;
4. Coletar termos da indexação automática – o resultado da indexação automática, que consistiu dos termos de indexação atribuídos pelo SISTRA aos *e-books*, foi exportado para um arquivo no formato de planilha eletrônica, a fim de serem inserido em um formulário de avaliação dos termos da indexação automática;
5. Construir o formulário de avaliação dos termos de indexação – o formulário foi construído no formato de planilha eletrônica, segundo pode ser observado na Figura 3. O formulário contém os seguintes campos especificados nas colunas da planilha: *e-book* – evidencia o identificador de um e-book, sendo a primeira ocorrência um *link* para o texto completo do *e-book*, a fim de possibilitar a leitura técnica pelos indexadores; *instrumento* – inclui o identificador de um vocabulário controlado fonte do termo, isto é, aponta para a origem do termo extraído, o qual, na presente avaliação, pode assumir dois possíveis valores (ddc-subject-pt-br – representando a lista alfabética de assuntos relacionados com códigos de classificação, ou eurovoc-pt-br – representando o tesauro EuroVoc); *termo* – contém a cadeia de caracteres do termo da indexação automática atribuído; *qualidade do termo* – deve ser preenchido com uma opção referente a qualidade do termo como ponto de acesso ao *e-book*; *nota do termo* – deve ser preenchido com uma opção referente a nota do termo como representativo e apropriado ao e-*book*.

Figura 3 – Formulário de avaliação no formato de planilha eletrônica.



	A	B	D	E	F	G
1	ebook	instrumento	termo	Qualidade do termo	Nota do termo	
2	book_pt_1	ddc-subject-pt-br	Campanhas			
3	book_p	https://drive.google.com/file/d/	eção			
4	book_p	14BP7ZKUNVQrdb_J8IQuqfS_				
5	book_p	Hy1WJtMKz/view?usp=share_link -	ração			
6	book_p	Clique uma vez para prosseguir.	ca			
7	book_p	Clique e mantenha pressionado para	is			
8	book_p	selecionar esta célula.	pulação			
9	book_pt_1	eurovoc-pt-br	Discurso			
10	book_pt_1	eurovoc-pt-br	Doença			
11	book_pt_1	eurovoc-pt-br	Londres			
12	book_pt_1	eurovoc-pt-br	Mar			
13	book_pt_1	eurovoc-pt-br	Vacina			
14	book_pt_11	ddc-subject-pt-br	Biografia			

Fonte: Dados da pesquisa (2023)

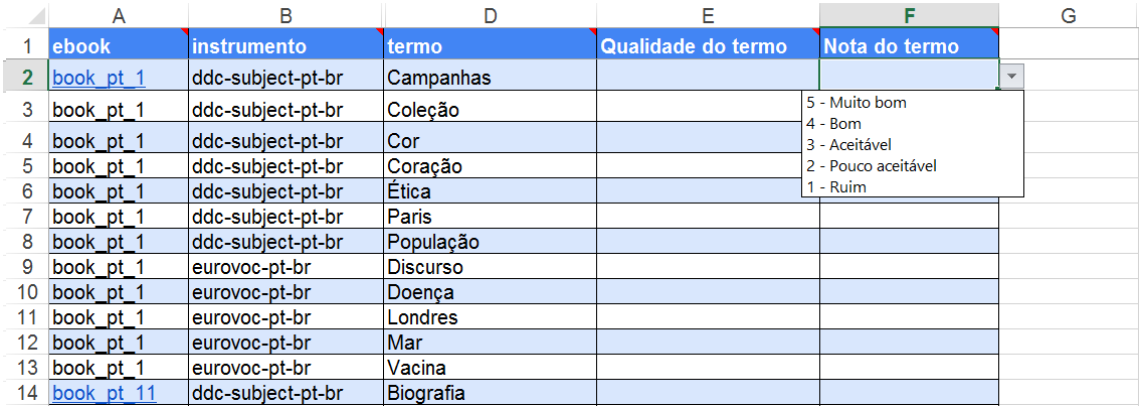
Em cada linha, tem-se o registro da atribuição a um identificador de *e-book* de um termo da indexação automática. A fim de permitir o registro do julgamento da qualidade de cada termo de indexação para cada *e-book*, foi adicionado um menu suspenso na célula nas colunas: *Qualidade do termo* – referente à qualidade do termo de indexação como ponto de acesso ao documento numa escala de três pontos (1-Não relevante, 2-Neutro, 3-Relevante), e *Nota do termo* – relativo a uma nota em escala de cinco pontos, quanto à adequação do termo em ser correto e apropriado ao documento (1-Ruim, 2-Pouco aceitável, 3-Aceitável, 4-Bom, 5-Muito bom). A Figura 4 e a Figura 5 mostram, respectivamente, os menus suspensos para seleção de valores para a qualidade e nota de um termo;

Figura 4 - Menu de opções para registro da qualidade do termo no formulário.

	A	B	D	E	F	G
1	ebook	instrumento	termo	Qualidade do termo	Nota do termo	
2	book_pt_1	ddc-subject-pt-br	Campanhas			
3	book_pt_1	ddc-subject-pt-br	Coleção	3 - RELEVANTE		
4	book_pt_1	ddc-subject-pt-br	Cor	2 - NEUTRO		
5	book_pt_1	ddc-subject-pt-br	Coração	1 - NÃO RELEVANTE		
6	book_pt_1	ddc-subject-pt-br	Ética			
7	book_pt_1	ddc-subject-pt-br	Paris			
8	book_pt_1	ddc-subject-pt-br	População			
9	book_pt_1	eurovoc-pt-br	Discurso			
10	book_pt_1	eurovoc-pt-br	Doença			
11	book_pt_1	eurovoc-pt-br	Londres			
12	book_pt_1	eurovoc-pt-br	Mar			
13	book_pt_1	eurovoc-pt-br	Vacina			
14	book_pt_11	ddc-subject-pt-br	Biografia			

Fonte: Dados da pesquisa (2023)

Figura 5 - Menu de opções para registro da nota do termo no formulário.



	A	B	D	E	F	G
1	ebook	instrumento	termo	Qualidade do termo	Nota do termo	
2	book_pt_1	ddc-subject-pt-br	Campanhas			
3	book_pt_1	ddc-subject-pt-br	Coleção			
4	book_pt_1	ddc-subject-pt-br	Cor			
5	book_pt_1	ddc-subject-pt-br	Coração			
6	book_pt_1	ddc-subject-pt-br	Ética			
7	book_pt_1	ddc-subject-pt-br	Paris			
8	book_pt_1	ddc-subject-pt-br	População			
9	book_pt_1	eurovoc-pt-br	Discurso			
10	book_pt_1	eurovoc-pt-br	Doença			
11	book_pt_1	eurovoc-pt-br	Londres			
12	book_pt_1	eurovoc-pt-br	Mar			
13	book_pt_1	eurovoc-pt-br	Vacina			
14	book_pt_11	ddc-subject-pt-br	Biografia			

Fonte: Dados da pesquisa (2023)

6. Preenchimento do formulário por indexadores – o formulário de avaliação qualitativa dos termos de indexação automática atribuídos aos 25 *e-books* foi preenchido paralelamente e de forma independente por dois indexadores, ambos bibliotecários experientes da UNESP. Almejando manter controle sobre a subjetividade na avaliação, o formulário foi apresentado aos indexadores individualmente, juntamente com um documento-guia, explicando o contexto da avaliação e como proceder à avaliação, por meio do formulário. Este, no formato de planilha eletrônica, foi preenchido pelos indexadores com a seleção de nota para a qualidade do termo de indexação como ponto de acesso e nota para a adequação do termo referente a ser correto e apropriado a cada *e-book*, com o objetivo de representar a qualidade de cada termo, na representação temática do conteúdo de cada *e-book*;
7. Calcular as métricas de avaliação da indexação automática – para determinação dos bons termos de indexação apontados pelos indexadores, os formulários preenchidos foram mesclados em uma única planilha, como ilustra a Figura 6. Os campos iniciando com A1 foram preenchidos com os valores numéricos das opções selecionadas pelo indexador 1, e os campos iniciando com A2 foram preenchidos com valores numéricos das opções selecionadas pelo indexador 2. Isso possibilitou o cálculo dos valores numéricos médios para os campos “Qualidade do termo” e “Nota do termo”, para cada termo de indexação automaticamente atribuído a cada *e-book*, sendo os valores registrados respectivamente nas colunas “Qualidade_Média” e “Nota_Média” da planilha. Em seguida, na coluna “Bom termo” da planilha, foi aplicada a expressão booleana envolvendo os valores numéricos médios para qualidade e nota dos termos (Qualidade_Média > 2 AND Nota_Média ≥ 3), visando à determinação dos bons termos

de indexação destacados pelos indexadores. Com a determinação dos bons termos de indexação indicados pelos indexadores, foi possível calcular o valor do índice de precisão que mede o percentual de termos considerados como bons termos de indexação pelos indexadores do total de termos atribuídos pelo sistema de indexação automática;

Figura 6 – Trecho da planilha para determinação dos bons termos de indexação.

ebook	instrumento	termo	A1-Qualidade	A1-Nota	A2-Qualidade	A2-Nota	Qualidade_Média	Nota_Média	Bom termo
book_pt_1	ddc-subject-pt-br	Campanhas	2	2	1	1	1,5	1,5	0
book_pt_1	ddc-subject-pt-br	Coleção	1	1	1	1	1	1	0
book_pt_1	ddc-subject-pt-br	Cor	1	1	1	1	1	1	0
book_pt_1	ddc-subject-pt-br	Coração	1	1	1	1	1	1	0
book_pt_1	ddc-subject-pt-br	Ética	3	4	2	2	2,5	3	1
book_pt_1	ddc-subject-pt-br	Paris	1	1	1	1	1	1	0
book_pt_1	ddc-subject-pt-br	População	2	3	2	2	2	2,5	0
book_pt_1	eurovoc-pt-br	Discurso	3	5	2	2	2,5	3,5	1
book_pt_1	eurovoc-pt-br	Doença	1	1	1	1	1	1	0
book_pt_1	eurovoc-pt-br	Londres	1	1	1	1	1	1	0
book_pt_1	eurovoc-pt-br	Mar	1	1	1	1	1	1	0
book_pt_1	eurovoc-pt-br	Vacina	1	1	2	3	1,5	2	0
book_pt_11	ddc-subject-pt-br	Biografia	2	3	1	1	1,5	2	0

Fonte: Dados da pesquisa (2023)

8. Aferir o desempenho através da análise dos valores das métricas – o desempenho do sistema de indexação automática. Quanto à qualidade dos termos de indexação atribuídos aos documentos do *corpus*, esta foi aferida por meio do cálculo da precisão média do sistema de indexação automática, que tem o mesmo valor da consistência média entre a indexação automática e a indexação intelectual. Os valores médios das métricas de avaliação da indexação automática para o *corpus* foram dispostos em uma tabela, contemplando valores médios para os seguintes indicadores: número de termos da indexação automática, o qual corresponde ao número de termos atribuídos pela indexação automática, ao número de termos comuns com a indexação intelectual, que equivale ao número de bons termos de indexação, assim como à precisão. Uma análise qualitativa da qualidade da indexação automática foi então realizada, por intermédio dos valores médios e dos dados que deram origem aos mesmos.

Os parâmetros descritivos dos valores das métricas de avaliação da indexação automática para o *corpus* estão sintetizados na Tabela 1.

Tabela 1: Valores das métricas obtidas pelo SISTRA no *corpus*

	Número de termos da indexação automática	Número de termos comuns com a indexação intelectual	Precisão
Mínimo	4	0	0%
Máximo	19	8	77,8%
Média	9,2	4,1	44,6%
Desvio-padrão	3,9	2,2	19,1%

Fonte: Dados da pesquisa (2023)

A partir dos resultados dispostos na Tabela 1, observou-se um mínimo de quatro termos e um máximo de 19 termos atribuídos pela indexação automática, enquanto os indexadores apontaram um mínimo de zero termo e um máximo de oito termos, como bons termos de indexação por *e-book*. Esses valores resultam em uma média de aproximadamente nove descritores atribuídos por *e-book* pela indexação automática, de um total de 230 atribuições de descritores, e uma média de aproximadamente quatro bons descritores por *e-book*, de um total de 103 atribuições de bons descritores destacados pelos indexadores.

Outro resultado relevante diz respeito à precisão: percebe-se uma média de 44,6% nos índices de precisão da indexação automática, variando entre 0% e 77,8%. Nesse contexto, a precisão retrata o percentual de termos propostos pela indexação automática que foram considerados como bons descritores pelos indexadores. Como, no presente método de avaliação, a consistência média tem o mesmo valor da precisão média obtida, a partir de tal consistência média, pode-se categorizar a indexação automática com um nível de desempenho de bom a ótimo, segundo a categorização proposta por Bandim e Correa (2018). Entretanto, verifica-se que esse valor de consistência média é um valor superestimado para o valor real de consistência média entre a indexação automática e a indexação intelectual, dado que, à medida que seja possibilitado aos indexadores acrescentar termos omitidos pelo sistema de indexação automática, consequentemente, a consistência média diminuirá.

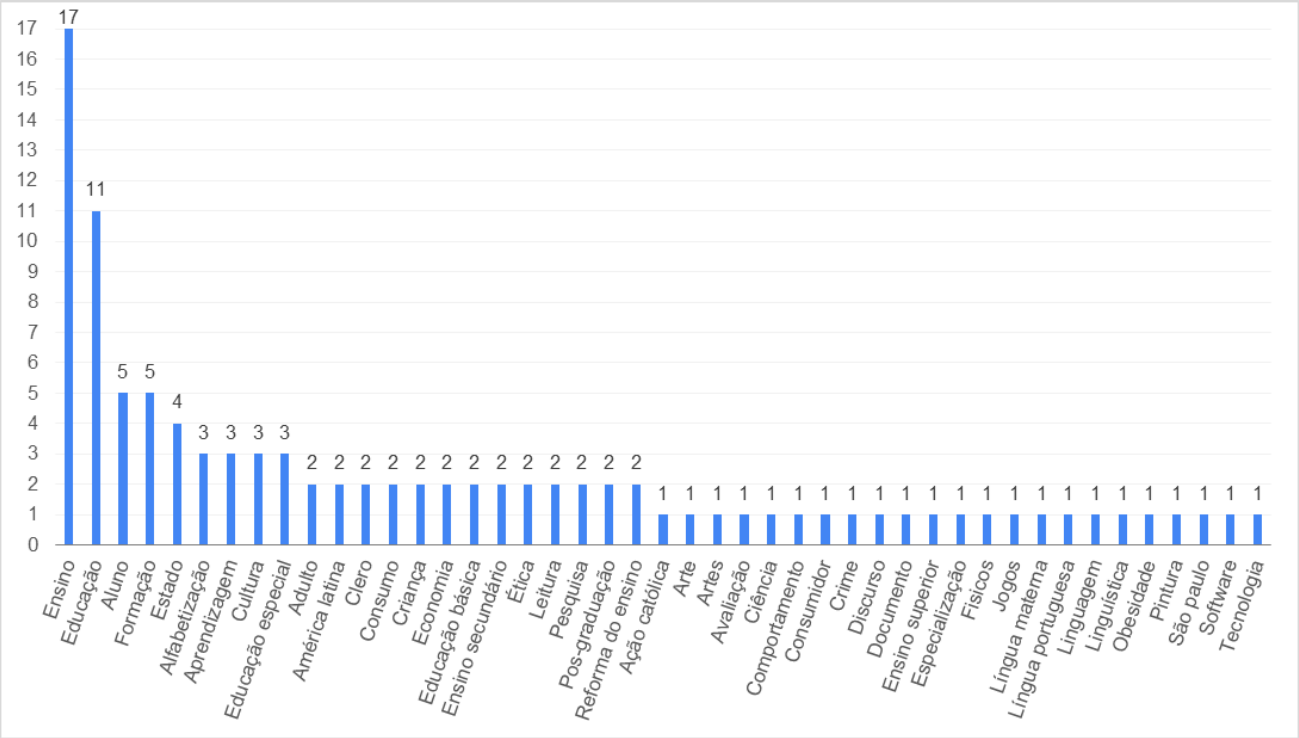
Com relação à precisão, esta identifica o percentual médio de termos relevantes recuperados pelo sistema de indexação automática, isto é, o percentual médio de bons termos de indexação do total de termos atribuídos pela indexação automática. Nesse caso, se nenhum termo relevante fosse atribuído, seria considerado o valor de zero ou 0%; se todos os termos atribuídos fossem relevantes, o valor assumido seria de um ou 100%. Dada a média de 44,6%, pode-se afirmar que quase metade dos termos atribuídos pela

indexação automática são relevantes, isto é, são bons termos para a indexação intelectual. Isso implica que o SISTRA alcança uma boa precisão média, podendo ser incorporado num fluxo de trabalho de indexação semiautomática para *e-books*, no contexto da rede de bibliotecas da Unesp.

O Gráfico 1 apresenta a distribuição de frequência absoluta na atribuição dos bons termos de indexação no *corpus*. São focalizadas as frequências de atribuição para os 45 termos considerados como bons descritores, pelos indexadores, totalizando 103 atribuições de bons termos de indexação.

A distribuição de frequência do Gráfico 1 é compatível com a Lei de Zipf, ao identificar um pequeno conjunto de termos com alta ocorrência no *corpus* e um grande conjunto de termos com baixa frequência. Constatou-se que os termos “Ensino” e “Educação” são os termos dominantes, evidenciando o principal núcleo temático do *corpus*. Em acréscimo, os termos “Aluno”, “Formação” e “Estado” se encontram entre os cinco mais frequentes no *corpus*, em destaque com frequências no topo do Gráfico 1.

Gráfico 1 – Frequência de atribuição dos bons termos de indexação



Fonte: Dados da pesquisa (2023)

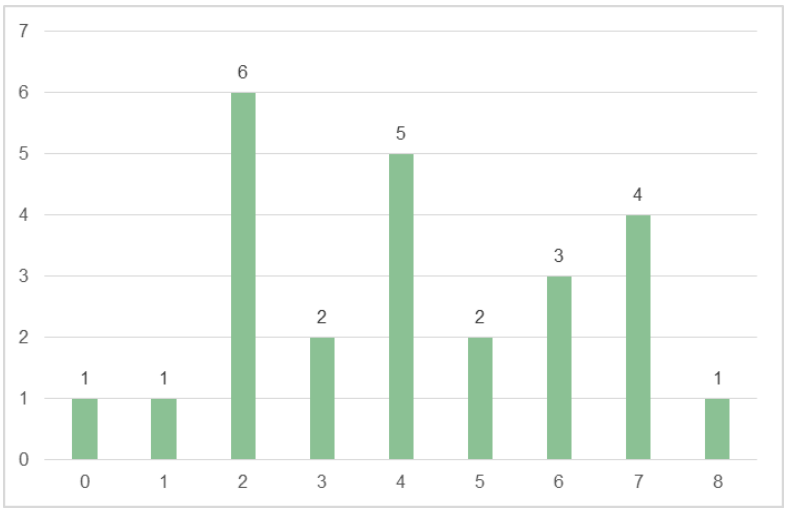
Percebe-se, na distribuição do Gráfico 1, a predominância de termos simples, formados por uma única palavra, sendo um total de 35 termos simples, os quais correspondem a 87 ocorrências ou atribuições, no *corpus*. No gráfico, podem ser

observados dez termos compostos, constituídos por duas ou três palavras, que correspondem a 16 atribuições. Entre os termos compostos mais frequentes, destacam-se “Educação especial”, “Educação básica”, “Ensino secundário” e “Reforma do ensino”, por pertencerem ao núcleo temático principal do *corpus*.

Adicionalmente, apurou-se que, das 230 atribuições automáticas de termos pelo SISTRA, 207 foram atribuições de termos simples, das quais 87 atribuições foram apontadas como bons descritores e 120 atribuições foram rejeitadas pelos indexadores, por não se consistirem em bons descritores. Quantos aos termos compostos, o SISTRA atribuiu automaticamente 23 termos compostos, dos quais 16 atribuições foram consideradas de bons descritores e sete atribuições foram rejeitadas, com as seguintes quantidades: “São Paulo”, por três vezes; “Ensino primário”, por uma vez; “Século XIX”, por uma vez; “Rio de Janeiro”, por uma vez; e “Contrato de trabalho”, por uma vez. Assim, percebe-se uma tendência de rejeição da maioria dos termos simples, bem como de termos compostos que representam local e tempo, embora alguns termos compostos que denotam local tenham sido apontados como bons descritores, tais como “América Latina” e “São Paulo”.

O Gráfico 2 mostra a distribuição do número de *e-books* com determinado número de bons descritores atribuídos. Nesse gráfico, no eixo horizontal, tem-se o quantitativo de bons termos de indexação atribuídos, o qual varia de zero a oito termos, enquanto, no eixo vertical, está representado o número absoluto de *e-books* com determinado quantitativo de bons termos atribuídos.

Gráfico 2 – Frequência de atribuição de quantitativo de bons termos de indexação



Fonte: Dados da pesquisa (2023)

Pelo Gráfico 2, pode-se notar que foram atribuídos dois bons descritores para seis *e-books*, sendo o valor de frequência de atribuição mais alto registrado no gráfico. Esses seis *e-books* correspondem a 24% do *corpus*, e, para cada um deles, o sistema de indexação automática obteve um desempenho satisfatório, ao atribuir dois bons descritores. Adicionalmente, para 17 *e-books*, que significam 68% do *corpus*, o sistema teve um desempenho bom, pois foram atribuídos três ou mais bons descritores, número que se enquadra dentro da faixa prevista para o número total de descritores a serem atribuídos, na catalogação de assunto dos itens do acervo, constante na política de indexação da rede de bibliotecas da Unesp, a qual varia de três a 12 descritores. Entretanto, na região esquerda do Gráfico 2, observa-se que, para dois *e-books*, que correspondem a 8% do *corpus*, obteve-se a atribuição de nenhum (zero) e um bom descritor, respectivamente, representando os piores casos, para os quais o sistema de indexação automática teve um desempenho insatisfatório.

Portanto, o processo de avaliação da indexação automática via julgamento por indexadores aponta para um bom desempenho do sistema SISTRA, na representação temática dos *e-books* do *corpus*. Caso os resultados não se enquadrassem no padrão estabelecido como bom, o processo de indexação automática precisaria ser revisado, com o propósito de atingir bons índices de precisão e consistência, envolvendo a codificação da tabela de caracteres dos arquivos de entrada, o uso dos vocabulários controlados nas indexações e as ferramentas de processamento de linguagem natural do sistema de indexação automática.

Como atividade complementar à avaliação da indexação automática realizada, visando à melhoria ou ao aperfeiçoamento do sistema, pode ser implementada uma análise de termos omitidos na indexação automática. Por instância, apresenta-se uma análise de termos omitidos na indexação automática para os dois piores casos. A análise de termos omitidos engloba a concretização da etapa de análise de assunto, na indexação intelectual de cada *e-book*, com foco na atribuição de termos em evidência no título, ficha catalográfica e sumário do *e-book*, seguida da etapa de inspeção dos relatórios gerados pelo SISTRA, do processo de indexação automática para os *e-books*.

Por exemplo, houve um *e-book* com nenhum bom termo de indexação atribuído. A esse poderiam ser atribuídos os assuntos “Economia” e “Política econômica”, porém, esses termos não foram atribuídos pelo SISTRA e nem aparecem no relatório de termos propostos por regras heurísticas disparadas para o *e-book*. Além disso, esse *e-book* contém três capítulos escritos no idioma espanhol, o que pode ter influenciado negativamente a

indexação automática. Para um *e-book*, foi atribuído apenas o termo “Ética” como um bom termo de indexação. A esse *e-book* poderiam ser atribuídos os assuntos “Psicologia” e “Psicólogos”, mas esses termos não foram atribuídos pelo SISTRA, embora variações sintáticas de tais termos apareçam relacionadas na lista de termos propostos por algumas regras heurísticas do SISTRA, disparadas na indexação automática do *e-book*. A análise de termos omitidos pela indexação automática pode ser realizada para todos os *e-books do corpus* ou por meio de amostragem, de sorte a trazer elementos para aperfeiçoamento do sistema de indexação automática.

Outra atividade complementar à avaliação da indexação automática e que reúne elementos para a melhoria do sistema de indexação automática consiste no levantamento dos fatores intervenientes na indexação automática, que busca caracterizar as diferenças morfológicas, sintáticas e semânticas dos termos da indexação intelectual e da indexação automática (BANDIM; CORREA, 2019). Tal análise de fatores intervenientes na indexação automática extrapola e enriquece os dados obtidos na avaliação, incorporando a análise dos procedimentos da realização da indexação intelectual e da indexação automática para cada documento do *corpus*.

A aplicação do método proposto de avaliação da indexação automática via julgamento por indexadores se mostrou pertinente, numa primeira avaliação, pois permitiu estimar o desempenho do sistema e contribuiu efetivamente para o posterior levantamento de termos omitidos e fatores intervenientes, na indexação automática do sistema avaliado, almejando a sua melhoria.

Os resultados desse tipo de avaliação também podem ser utilizados na posterior construção de um padrão-ouro, objetivando uma avaliação direta da indexação automática por comparação com um padrão-ouro, na qual é possível o cálculo de métricas adicionais, como revocação e medida F, não alcançáveis pelo método proposto.

5 CONSIDERAÇÕES FINAIS

Os resultados apresentados apontam para a pertinência do método proposto, na avaliação da indexação automática via julgamento por indexadores. A aplicação desse método, na avaliação da indexação automática por atribuição do sistema SISTRA aplicado a *e-books* técnicos científicos, no contexto da rede de bibliotecas da Unesp, possibilitou validar seus processos e resultados alcançáveis.

Ademais, foi possível constatar que o método proposto contribui para a sistematização e a padronização do processo de avaliação da indexação automática via julgamento por indexadores, no contexto das unidades de informação. Ao aplicar a indexação automática em um *corpus* de documentos e possibilitar aos indexadores o julgamento dos termos atribuídos automaticamente, quanto à qualidade na representação temática, o método minimiza a sobrecarga intelectual do indexador na função de avaliador e permite a determinação de métricas de qualidade da indexação automática, importantes para estimar o desempenho do sistema de indexação automática.

O método proposto favorece o cálculo do número médio de termos comuns com a indexação intelectual e da precisão média, tornando possível caracterizar o desempenho do sistema de indexação automática, além de ensejar a análise da frequência de atribuição de bons termos de indexação e do quantitativo de bons termos de indexação atribuídos aos documentos do *corpus*. Com isso, permite visualizar os principais termos relevantes atribuídos automaticamente e validados pelos indexadores, o que auxilia na compreensão e delimitação dos domínios temáticos do *corpus* analisado, facilitando a determinação dos melhores e piores casos na indexação automática.

Com o método proposto, pode-se acelerar o processo de avaliação da indexação automática, por trazer maior agilidade na identificação de bons descritores alcançados pela indexação automática, sendo útil como uma primeira avaliação da qualidade alcançada por um sistema.

Como limitações do método proposto, é possível apontar: a dependência à avaliação pelos indexadores, o que limita a reprodutibilidade e a neutralidade na avaliação de outros sistemas; a impossibilidade de obtenção dos índices de revocação e medida F para o sistema avaliado, sem que se faça a indexação intelectual do *corpus*; e a limitação na identificação de bons descritores como pertencentes a um subconjunto dos termos atribuídos pela indexação automática para o *corpus*.

Estudos futuros propõem ampliar as análises obtidas, por meio da aplicação do método proposto na avaliação para o mesmo *corpus*, de outros sistemas de indexação automática por atribuição. Além disso, pretende-se aplicar o método proposto, utilizando-se diferentes *corpora* de publicações técnico-científicas, com documentos de diversas áreas do conhecimento. Adicionalmente, almeja-se pesquisar a incorporação da análise de agrupamento e de coocorrência de bons descritores, como parte das análises realizadas com os dados obtidos na avaliação. Outra linha de investigação diz respeito ao emprego

dos dados da presente avaliação na posterior construção de um padrão-ouro, visando a uma avaliação da indexação automática por comparação com um padrão-ouro.

REFERÊNCIAS

ABPMP (Brasil). **BPM CBOK** – Guia para o gerenciamento de processos de negócio: corpo comum de conhecimento. ABPMP BPM CBOK, v. 3.0, 1. ed. ABPMP, 2013. Disponível em: https://www.abpmp-br.org/educacao/bpm-cbok/. Acesso em: 24 ago. 2023.

ASULA, M.; MAKKE, J.; FREIENTHAL, L.; KUULMETS, H. A.; SIREL, R. Kratt: Developing an Automatic Subject Indexing Tool for the National Library of Estonia. **Cataloging & Classification Quarterly**, v. 59, n. 8, p. 775-793, 2021. DOI: 10.1080/01639374.2021.1998283.

BANDIM, M. A. S.; CORREA, R. F. A consistência na indexação automática por atribuição de artigos científicos na área de Ciência da Informação. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, *[S. l.]*, v. 23, n. 53, p. 64-77, 2018. DOI: 10.5007/1518-2924.2018v23n53p64.

BANDIM, M. A. S.; CORREA, R. F. Indexação automática por atribuição de artigos científicos em português da área de Ciência da Informação. **Transinformação**, Campinas, v. 31, p. e180004, 2019. DOI: 10.1590/2318-0889201931e180004.

CORRÊA, R. F.; LAPA, R. C. Panorama de estudos sobre indexação automática no âmbito da ciência da informação no Brasil (1973-2012). **Ciência da Informação**, *[S. l.]*, v. 42, n. 2, p. 255-273, 2013. Disponível em: https://revista.ibict.br/ciinf/article/view/1385. Acesso em: 14 set. 2023.

FUJITA, M. S. L. **Representação Documental Automática e Multilíngue de Textos Científicos (SISTRA)**. 2020. (Projeto de pesquisa FAPESP processo 2019/25470-6). Disponível em: https://bv.fapesp.br/pt/auxilios/107480/representacao-documental-automatica-e-multilingue-de-textos-tecnico-cientificos-sistra/. Acesso em: 24 ago. 2023.

GIL-LEIVA, I.; FUJITA, M. S. L.; REDIGOLO, F. M.; SARAN, J. F. Extracción de información de documentos pdf para su uso en la indización automática de e-books. **Transinformação**, Campinas, v. 34, p. 1-11, 2022. DOI: 10.1590/2318-0889202234e210069.

GIL-LEIVA, I.; ORTUÑO, P. D.; CORRÊA, R. F. Indización automática de artículos científicos sobre Biblioteconomía y Documentación con SISA, KEA y MAUI. **Revista Española de Documentación Científica**, *[S. l.]*, v. 45, n. 4, p. e338, 2022. DOI: 10.3989/redc.2022.4.1917.

GOLUB, K. Automated Subject Indexing: An Overview. **Cataloging & Classification Quarterly**, v. 59, n. 8, p. 702-719, 2021. DOI: 10.1080/01639374.2021.2012311.

GOLUB, K.; SOERGEL, D.; BUCHANAN, G.; TUDHOPE, D.; LYKKE, M.; HIOM, D. A framework for evaluating automatic indexing or classification in the context of retrieval.

Journal of the Association for Information Science and Technology (JASIST), v. 67, n.1, p. 3-16, 2016. DOI: 10.1002/asi.23600.

HASAN, K. S.; NG, V. Automatic keyphrase extraction: a survey of the state of the art. *In*: ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 52., 2014, Baltimore. **Proceedings** […] Association for Computational Linguistics, 2014. (Volume 1: Long Papers), p. 1262-1273.. DOI: 10.3115/v1/P14-1119.

JUNGER, U. Automation first – the subject cataloguing policy of the Deutsche Nationalbibliothek. *In*: **IFLA WLIC 2018** – Transform Libraries, Transform Societies, 2018, Kuala Lumpur, Malaysia. IFLA, 2018. Disponível em: https://library.ifla.org/id/eprint/2213. Acesso em: 29 maio 2023.

JUNGER, U.; SCHOLZE, F. Neue Wege und Qualitäten – Die Inhaltserschließungspolitik der Deutschen Nationalbibliothek. *In*: FRANKE-MAIER, M.; KASPRZIK, A.; LEDL, A.; SCHÜRMANN, H. (ed.). **Qualität in der Inhaltserschließung**. Berlin, Boston: De Gruyter Saur, 2021. p. 55-70. DOI: 10.1515/9783110691597-004. Disponível em: https://www.degruyter.com/document/doi/10.1515/9783110691597-004/html. Acesso em: 29 maio 2023.

KIM, S.N.; MEDELYAN, O.; KAN, M.Y.; BALDWIN, T. SemEval-2010 task 5: automatic keyphrase extraction from scientific articles. *In*: INTERNATIONAL WORKSHOP ON SEMANTIC EVALUATION, 5., 2010, Uppsala, Sweden. **Proceedings** […] Association for Computational Linguistics, 2010. p. 21-26. Disponível em: https://aclanthology.org/S10-1004. Acesso em: 29 maio 2023.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. Tradução de Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos, 2004. [Tradução de: **Indexing and abstracting in theory and practice**].

LAPA, R. C.; CORRÊA, R. F. Indexação automática no âmbito da ciência da informação no brasil. **Informação & Tecnologia**, *[S. l.]*, v. 1, n. 2, p. 59-76, 2014. Disponível em: https://periodicos.ufpb.br/index.php/itec/article/view/21408. Acesso em: 14 set. 2023.

NARUKAWA, C. M.; GIL-LEIVA, I.; 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**: Estudos, João Pessoa, v. 19, n. 2, p. 99-118, 2009. Disponível em: https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/2925. Acesso em: 14 set. 2023.

ROSENBERG, V. Comparative evaluation of two indexing methods using judges. **Journal of the American Society for Information Science**, v. 22, n. 4, p. 251-259, 1971. DOI: 10.1002/asi.4630220404.

SILVA, S. R. B.; CORREA, R. F. Sistemas de Indexação automática por atribuição: uma análise comparativa. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, *[S. l.]*, v. 25, p. 1-25, 2020. DOI: 10.5007/1518-2924.2020.e70740.

SILVA, S. R. B.; CORREA, R. F.; GIL-LEIVA, I. Avaliação direta e conjunta de Sistemas de Indexação Automática por Atribuição. **Informação & Sociedade:** Estudos, João Pessoa, v. 30, n. 4, p. 1-27, 2020. DOI: 10.22478/ufpb.1809-4783.2020v30n4.57259.

SUOMINEN, O. Annif: DIY automated subject indexing using multiple algorithms. **LIBER Quarterly**: The Journal of the Association of European Research Libraries, v. 29, n. 1, 2019. DOI:10.18352/lq.10285.

SUOMINEN, O.; INKINEN, J.; LEHTINEN, M. Annif and Finto AI: Developing and Implementing Automated Subject Indexing. **JLIS.It**, v. 13, n. 1, p. 265-282, 2022. DOI: 10.4403/jlis.it-12740.

VERGARA, S. C. **Projetos e relatórios de pesquisa em administração**. 11. ed. São Paulo: Atlas, 2009.

NOTAS

AGRADECIMENTOS
À Unesp por aprovar e receber o primeiro autor como participante em programa de pós-doutoramento. À UFPE por conceder afastamento para realização de pós-doutorado na Unesp do primeiro autor. À FAPESP pelo fomento ao Projeto de pesquisa FAPESP processo 2019/25470-6, que resultou no desenvolvimento do sistema SISTRA. Aos pesquisadores do grupo de pesquisa "Representação Temática da Informação", por compartilhar seus conhecimentos e possibilitar acesso ao sistema SISTRA, bem como aos dados e relatórios de pesquisa. Aos dois bibliotecários da Unesp que de forma anônima e voluntária contribuíram como avaliadores dos termos de indexação.

CONTRIBUIÇÃO DE AUTORIA
Os papéis descrevem a contribuição específica de cada colaborador para a produção acadêmica inserir os dados dos autores conforme exemplo, excluindo o que não for aplicável. Iniciais dos primeiros nomes acrescidas com o último Sobrenome, conforme exemplo.

Concepção e elaboração do manuscrito: R. F. Correa, M. S. L. Fujita
Coleta de dados: R. F. Correa, M. S. L. Fujita
Análise de dados: R. F. Correa
Discussão dos resultados: R. F. Correa
Revisão e aprovação: R. F. Correa, M. S. L. Fujita

Caso necessário veja outros papéis em: https://credit.niso.org

CONJUNTO DE DADOS DE PESQUISA

Todo o conjunto de dados que dá suporte aos resultados deste estudo foi publicado no próprio artigo.

FINANCIAMENTO
Não se aplica.

CONSENTIMENTO DE USO DE IMAGEM
Não se aplica

APROVAÇÃO DE COMITÊ DE ÉTICA EM PESQUISA
não se aplica.

CONFLITO DE INTERESSES
Não se aplica

PUBLISHER
Universidade Federal de Santa Catarina. Programa de Pós-graduação em Ciência da Informação. Publicação no Portal de Periódicos UFSC. As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

EDITORES
Edgar Bisset Alvarez, Ana Clara Cândido, Patrícia Neubert, Genilson Geraldo, Jônatas Edison da Silva, Mayara Madeira Trevisol.

HISTÓRICO
Recebido em: 26-09-2023 – Aprovado em: 19-01-2024 - Publicado em: 23-02-2024