

EIXO TEMÁTICO:

Organização e Representação da Informação e do Conhecimento

**CONTROLE DE VOCABULÁRIO EM PERIÓDICOS CIENTÍFICOS ELETRÔNICOS:
PROPOSTA DE IMPLANTAÇÃO DA VCPC TOOLS NO PERIÓDICO
“DISCURSOS FOTOGRÁFICOS”**

**VOCABULARY CONTROL IN ELECTRONIC SCIENTIFIC PERIODICALS:
DEPLOYMENT PROPOSAL OF VCPC TOOLS IN THE PERIODICALS
“DISCURSOS FOTOGRÁFICOS”**

Brígida Maria Nogueira Cervantes - brigidacervantes@gmail.com

Rodolfo Rorato Londero - rodolfolondero@hotmail.com

José Carlos Francisco dos Santos - jc.fa1982@gmail.com

Paula Regina Ventura Amorim Gonçalves - paulaventuramorim@gmail.com

Resumo: O controle de vocabulário é uma estrutura desenvolvida para ser aplicada nos processos de representação e acesso aos conteúdos temáticos informacionais. Este estudo tem como objetivo propor a implantação da VCPC *Tools* para o controle de vocabulário em periódicos científicos eletrônicos de áreas especializadas, disponíveis no Portal de Periódicos Científicos da Universidade Estadual de Londrina, visando aprimorar a representação de conteúdos por meio da atribuição de palavras-chave. Trata-se de uma pesquisa bibliográfica e documental que se concretizará por meio da pesquisa exploratória descritiva. Os procedimentos metodológicos envolvem a coleta de metadados e a validação das palavras-chave. Enunciam-se quatro requisitos funcionais mapeados conforme o reconhecimento do ambiente de periódicos científicos, considerando como se constituem, a ligação que promovem as palavras-chave ao estarem atribuídas do modo como são encontradas e que efeitos geram para a representação dos conteúdos temáticos. Dentre os resultados tem-se a obtenção de um índice de termos que apresenta também a correlação deste com os respectivos artigos científicos eletrônicos. Conclui-se que a aplicação do controle de vocabulário nas palavras-chave, atribuídas em periódicos científicos, possibilita maior consistência e relevância visando facilitar a recuperação e o acesso aos conteúdos temáticos informacionais de uma área especializada.

Palavras-chave: Controle de vocabulário. Periódicos Científicos Eletrônicos. VCPC *Tools*.

Abstract: Vocabulary control is a structure developed to apply in the representation processes and to access thematic information contents. The aim of this study is to propose the deployment of VCPC *Tools* for vocabulary control to improve the representation of contents by means of attribution of keywords in electronic scientific journals of specialized areas, available at Universidade Estadual de Londrina Journals Portal. This is a bibliographical and documental research, which will be achieved by means of descriptive exploratory research. The methodological procedures involve the metadata harvesting and the keyword validation. It is enunciated four functional requirements, which were mapped out according to the scientific journals environment recognition, considered as they are constituted, the connection that promote the keywords by the manner they are found and what effects are generated for the representation of thematic contents. Among the results, there is the acquisition of an index terms, which presents the correlation of this article with the respective electronic scientific articles. It concludes that the application of vocabulary

control in keywords assign to scientific journals, offers better consistence and relevance aiming to facilitate the retrieval and access to thematic informational contents of a specialized area.

Keywords: Vocabulary control. Electronic Scientific Periodicals. *VCPC Tools*.

1 INTRODUÇÃO

O interesse de investigação deste trabalho surge ao perceber a utilização da linguagem livre para a atribuição de palavras-chave a artigos científicos eletrônicos. Neste sentido, cabe ressaltar a ineficiência da linguagem livre como representação das palavras-chave diante de diversos fatores, como sintaxe dos termos (ora com plural, ora sem plural, termos compostos separados por um traço ou não), semântica dos termos e sua organização, que levam a uma estrutura de palavras-chave inconsistente na base de dados dos periódicos.

É notória a existência de dificuldades de visualizar as palavras-chave no momento em que vão sendo atribuídas, em meio à quantidade de expressões diretas como de termos derivados constantes nas bases de dados de periódicos científicos eletrônicos. Quando se pensa em realizar uma busca de todas as palavras-chave, é necessário formular uma expressão de busca, porém a falta de conhecimento das palavras-chave existentes na base de dados dificulta esta etapa da elaboração. Exatamente nesse contexto reside a relevância de buscar as palavras-chave sem deixar de conhecer as que estão registradas na base de dados do periódico.

Em razão disso, indaga-se: como realizar o controle do vocabulário das palavras-chave atribuídas aos artigos científicos disponíveis em periódicos científicos eletrônicos? A partir desta problemática, o que se propõe com este estudo é a aplicabilidade da ferramenta *VCPC Tools* para o controle de vocabulário das palavras-chave atribuídas aos artigos de periódicos científicos eletrônicos, por meio do processo de compatibilização destas palavras-chave com um vocabulário orientado a áreas especializadas. Para atender aos pressupostos teóricos e metodológicos, este estudo encontra-se fundamentado, principalmente, nos seguintes autores: Gil Leiva; Alonso Arroyo (2005) – na questão do controle de vocabulário e palavras-chave; ANSI/NISO. Z39.19 (2005) – norma para elaboração de vocabulário; Orengo; Huyck (2001); Orengo; Buriol; Coelho (2007) – na compatibilização de termos e análise léxica e semântica.

O estudo apresenta como objetivo geral: propor a implantação da VCPC *Tools* para o controle de vocabulário em periódicos científicos eletrônicos de áreas especializadas disponíveis no Portal de Periódicos Científicos da Universidade Estadual de Londrina. Para tanto, tem como objetivos específicos: desenvolver a ferramenta VCPC *Tools* para o periódico foco do estudo; coletar e armazenar os vocabulários controlados; compatibilizar as palavras-chave com os termos dos vocabulários controlados; e tratar de forma manual as palavras-chave não compatibilizadas automaticamente.

2 O PERIÓDICO CIENTÍFICO ELETRÔNICO “DISCURSOS FOTOGRÁFICOS”

2.1 PERIÓDICOS CIENTÍFICOS ELETRÔNICOS

Meadows (1999) caracteriza o surgimento dos periódicos no século XVII, apoiado em crenças como a perspectiva de lucratividade pelos editores e que, para construir novos descobrimentos, é necessário debate coletivo. Mueller (2000b, p. 73) confirma também que “os periódicos científicos apareceram no século XVII na Europa, em uma época marcada por mudanças em toda a sociedade, inclusive no campo científico” (MUELLER, 2000b, p. 73).

Anteriormente ao surgimento do periódico científico, a ciência era dominada somente por filósofos que faziam as comunicações e publicações de seus materiais à comunidade por meio de cartas, nas quais era possível demonstrar o estudo sobre determinado assunto. O periódico científico surge decorrente de uma necessidade de comunicação rápida e confiável de pesquisas, e este acontecimento se dá durante a ciência moderna (MEADOWS, 1999; MUELLER, 2000b).

Em meio às diversas evoluções, o periódico científico passa a ter um suporte computacional, o que origina o periódico científico eletrônico, cujas características fundamentais ultrapassam os limites do papel e integram-se com recursos como hipertexto, links e multimídias, ou seja, um formato de comunicação científica dentro de um modelo de processos eletrônicos de editoração.

Esta evolução das revistas científicas acontece a partir da década de 1960, com o tratamento bibliográfico realizado por computador. Além do armazenamento e edição de textos complexos e com grande quantidade de informações, surgem também as redes de telecomunicações e o “[...] mercado de informações on-line

[...]”, com a criação de revistas eletrônicas e bases de dados de textos completos (GONÇALVES; RAMOS; CASTRO, 2006, p. 168-169). Esta afirmação também é feita por Meadows (1999), que ainda adiciona que o computador visto como “devorador de números” (MEADOWS, 1999, p. 33) poderia ser utilizado no trabalho com informações alfabéticas.

O periódico científico eletrônico é “[...] um formato eletrônico de publicação [...]” que se utiliza dos benefícios de obter, em seu conteúdo, recursos como imagens, áudio, vídeo, links que direcionam para outros documentos. Usualmente é disponibilizado, por meio da Internet, com o suporte de hipertexto que pode conceder, ao leitor, o acesso às referências utilizadas na referida publicação. Uma exceção na disponibilidade: além da internet, o periódico científico eletrônico também pode estar no suporte de CD-ROM (LARA, 2006, p. 406). Mueller (2000b) complementa que o PCE é o termo utilizado para os periódicos que fazem uso de equipamentos eletrônicos para obtenção de acesso.

2.2 ÁREAS DO CONHECIMENTO: EM FOCO, A ÁREA DA COMUNICAÇÃO

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) define as Áreas do Conhecimento como uma classificação de “finalidade eminentemente prática” que permite “sistematizar e prestar informações concernentes a projetos de pesquisa e recursos humanos aos órgãos gestores da área de ciência e tecnologia” (CAPES, 2014). Essa classificação também é responsável por delimitar os campos de atuação dos periódicos científicos brasileiros.

Destacando o periódico em foco, *Discursos Fotográficos*, a Tabela de Áreas do Conhecimento da Capes classifica a subárea Comunicação Visual como um ramo da área Comunicação, sendo esta, por sua vez, um ramo da grande área Ciências Sociais Aplicadas I. Ao considerar apenas o nome, a subárea Comunicação Visual parece indicar um objeto evidente, circunscrevendo diversos meios (cinema, televisão, fotografia, cartaz, etc.) e atividades (publicidade, fotojornalismo, *design* gráfico, etc.) relacionadas. Entretanto, o problema não é tão simples, principalmente ao analisar as definições do termo “comunicação visual” apresentada pelos dicionários de comunicação.

Começando pelo mais antigo, o *Dicionário Básico de Comunicação*

estabelece uma definição restrita: “Parte do desenho industrial que trata da imagem visível de uma empresa ou entidade” (COSTA LIMA; DORIA; KATZ, 1975, p. 105). Já o *Dicionário de Comunicação* apresenta o termo apenas a partir da segunda edição, porém em uma definição mais ampla que a anterior: “Conjunto das técnicas, conhecimentos e procedimentos que buscam maior eficácia na transmissão visual de mensagens verbais e não-verbais através dos diversos meios de comunicação” (RABAÇA; BARBOSA, 1995, p. 177). Por fim, para citar um exemplo mais recente, o *Dicionário Houaiss de Comunicação e Multimídia* define “comunicação visual” de modo semelhante à anterior: “Aquela que se caracteriza pela ênfase nos aspectos gráfico-visuais da mensagem que está sendo transmitida” (NEIVA, 2013, p. 117).

Contudo, além das definições dos dicionários especializados, a subárea Comunicação Visual encontra alternativas relevantes em outras áreas. Veja, por exemplo, o caso da *cultura visual*, compreendida por Català Domènech (2011) como um novo campo de estudos interessado em problematizar as ideias de representação e visão. Ou mesmo o caso da *antropologia da comunicação visual* proposta por Canevacci (1990), interessada tanto nas técnicas audiovisuais como forma de documentar quanto na análise dos produtos da comunicação visual enquanto fenômeno global da cultura visual. Ou ainda o caso das *teorias da imagem*, pensadas como um campo de pesquisa em comunicação, mas também um “lugar de debates e apropriações teóricas de áreas afins, como Arte, Filosofia, Ciências da cultura, Sociologia, Antropologia e Semiótica” (BAITELLO JUNIOR; KLEIN, 2014, p. 487).

Esses poucos exemplos mostram como o desenvolvimento da área transcende o significado consensual de comunicação visual.

3 ESTRATÉGIAS METODOLÓGICAS

Caracteriza-se como uma pesquisa exploratória e descritiva com abordagem mista qualitativa e quantitativa. A coleta de dados será realizada pela ferramenta *VCPC Tools*, por meio do protocolo OAI-PMH, tendo como objeto de estudo as palavras-chave atribuídas aos artigos em periódicos científicos eletrônicos. A seguir, apresenta-se no Quadro 1 uma síntese das ações decorrentes dos objetivos do estudo.

QUADRO 1: Síntese das Ações a serem desenvolvidas

Objetivo geral: Propor a implantação da VCPC <i>Tools</i> para o controle de vocabulário em periódicos científicos eletrônicos de áreas especializadas disponíveis no Portal de Periódicos Científicos da Universidade Estadual de Londrina, para aprimorar o acesso aos conteúdos temáticos.	
Objetivos específicos	Ações
Desenvolver a ferramenta VCPC <i>tools</i> para o periódico foco do estudo – Etapa 1.	<ul style="list-style-type: none"> • Identificar o endereço do repositório digital do periódico (3.1.1); • Desenvolver o projeto da ferramenta VCPC <i>tools</i> (3.1.2), • Desenvolver o projeto físico do banco de dados (3.1.3); • Criar a base de dados no sistema gerenciador de banco de dados (3.1.4).
Elaborar a estrutura conceitual do domínio e subdomínio do periódico – Etapa 2.	<ul style="list-style-type: none"> • Identificar o escopo do periódico (3.2.1); • Levantar os instrumentos terminológicos do domínio e subdomínio (3.2.2); • Analisar as Tabelas de Áreas do Conhecimento (3.2.3).
Analisar as palavras-chave com os termos dos instrumentos terminológicos – Etapa 3.	<ul style="list-style-type: none"> • Coletar e armazenar os metadados (3.3.1); • Verificar as palavras-chave que são idênticas entre si (3.3.2); • Verificar as palavras-chave que tenham termos correspondentes nos instrumentos terminológicos (3.3.3); • classificar tematicamente as palavras-chave conforme a estrutura conceitual do periódico (3.3.4); • Inserir a palavra-chave como candidato a termo no vocabulário controlado da revista (3.3.5).

Fonte: Elaborado pelos autores.

3.1 ETAPA 1 - DESENVOLVER A FERRAMENTA VCPC TOOLS

3.1.1 Identificar o Endereço do Repositório Digital do Periódico

A ação de identificação e localização do endereço do repositório digital da revista é pré-requisito para estruturar as demais fases do processo de coleta de dados. Conseguida essa identificação e validação do endereço do repositório, serão efetuadas as diversas requisições para alcançar a resposta por meio da estrutura de um arquivo XML. A partir desta ação será possível obter informações para o desenvolvimento da ferramenta VCPC *Tools*. A escolha para a aplicação do protótipo da ferramenta VCPC *Tools* foi o periódico “Discursos Fotográficos”, da área da Comunicação, disponível no Sistema Eletrônico de Editoração de Revistas (SEER). Esse sistema customizado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), a partir do programa Open Journal Systems (OJS), contribui para a inserção dos periódicos brasileiros no cenário de pesquisa

internacional por meio do acesso aberto (SOUZA, 2010, p. 2). Deste modo, “o SEER/OJS é um software livre, de código aberto, que permite a criação de uma revista ou um portal de revistas. Fornece facilidades para criação de revistas eletrônicas disponibilizadas na Web” (BRITO; GUEDES; SHINTAKU, 2013, p. 25). A automatização do processo de editoração eletrônica de revistas é uma funcionalidade de grande relevância. Portanto, o resultado principal desta ação é a identificação do endereço do protocolo OAI-PMH do repositório digital da revista.

3.1.2 Desenvolver o Projeto da Ferramenta VCPC Tools

O projeto da ferramenta *VCPC Tools* foi desenvolvido na linguagem de programação PHP, por se tratar de uma linguagem livre para web e permitir uma perfeita integração com o sistema gerenciador de banco de dados MySQL. A ferramenta executa os seguintes procedimentos, aqui tratados de requisitos funcionais, enumerados com a sigla RF: RF01 Coleta de metadados; RF02 Análise de Metadados e Vocabulário controlado; RF03 Tratamento Palavras-chave; RF04 Resultados. Os requisitos funcionais serão descritos nos próximos itens.

- **RF01 Coleta de metadados:** Neste requisito funcional, o principal objetivo é coletar os metadados do periódico e armazená-los na base de dados. Desse modo, o RF01 executará uma requisição ao protocolo OAI-PMH: `?verb=ListIdentifiers&metadataPrefix=oai_dc`, o qual retorna os identificadores de todos os documentos deste repositório. Cabe frisar que, conforme a quantidade de registros nos arquivos XML de retorno, passa a constar a tag `resumptionToken`. O RF01 faz o registro de todos esses identificadores em uma variável denominada de array (vetor), exceto os registros de identificadores com o atributo `status="deleted"` da tag header. A cada registro de identificador, é realizada uma nova requisição ao protocolo OAI-PMH, abaixo descrita. A requisição seguinte ao protocolo OAI-PMH: `?verb=GetRecord&identifier= identificador&metadataPrefix=oai_dc`, é realizada para todos os identificadores dos documentos, “identificador” significa o identificador de cada documento acima descrito. Esta requisição retorna os registros dos metadados de cada documento, seguindo o padrão de metadados *Dublin Core*. De posse desses registros, é chegada a fase de inserir, no banco de dados, os descritivos de cada metadado, ou seja, o conteúdo de cada tag do arquivo XML, em resposta à requisição enviada ao protocolo OAI-PMH. As palavras-chave serão coletadas nas

seguintes seções da revista: artigos, pontos de vista, relatos de experiência, comunicações, resenhas de livros/mídia e teses e dissertações, excluindo-se editoriais e expedientes. No processo de inserção é realizada a verificação da existência da palavra-chave na base de dados; caso já exista, estabelece os vínculos com a palavra-chave.

- **RF02 Análise de Metadados e Vocabulário controlado:** O objetivo deste requisito funcional é verificar as palavras-chave que remetem à igualdade dos termos com os vocabulários controlados. Neste processo, o foco é a sintaxe das palavras-chave comparadas aos vocabulários controlados. Para a compatibilização, utilizam-se os seguintes mecanismos de comparação: 1) idênticas - comparação de cada palavra-chave com a busca do respectivo termo do vocabulário controlado, caracterizando, aqui, a compatibilização de palavras-chave exatamente idênticas com um termo do vocabulário controlado; 2) idênticas retirando as palavras vazias - exclusão das palavras-chave, palavras vazias. O mesmo procedimento é realizado com os termos do vocabulário controlado. Este procedimento fundamenta-se em Gil Leiva (1999). A seguir, realizam-se os testes comparativos, desconsiderando os caracteres maiúsculos, minúsculos e acentuação; 3) idênticas retirando os plurais - exclusão dos plurais das palavras-chave. Nesta fase, utilizam-se os resultados da exclusão das palavras vazias executadas na segunda compatibilização. São tratados os plurais, utilizando o conjunto de regras e exceções de Orengo; Buriol e Coelho (2007) que propõem melhorias para o algoritmo de stemming em língua portuguesa RSPL, proposto por Orengo e Huyck (2001) - Removedor de Sufixos. Retira-se o plural das palavras-chave e gravam-se os resultados em um novo campo no banco de dados, para preservar o registro original. O mesmo procedimento é aplicado aos termos do vocabulário controlado; 4) idênticas retirando os sufixos - redução de sufixos empregando regras e exceções de Orengo, Buriol e Coelho (2007), utilizadas na compatibilização anterior. Inicia-se com a redução do plural, executado na terceira compatibilização; 5) idênticas por meio do índice contido - execução a partir do índice contido que será calculado e armazenado no banco de dados entre cada palavra-chave e cada termo do vocabulário controlado. A partir do propósito de Gil Leiva (1999) de medir o índice de consistência dos termos em sua proposta de automatização da indexação, surgiu a iniciativa de desenvolver o índice contido, porém com o propósito de compatibilizar, caso alcance os parâmetros determinantes de equivalência de termos. As compatibilizações são realizadas

considerando os mecanismos de comparação com os vocabulários controlados. É estabelecido o vínculo das palavras-chave com o vocabulário controlado. Este processo é recursivo a todos os mecanismos de comparações já descritos. Na possibilidade de existir mais de um vocabulário controlado para a realização das compatibilizações, a comparação ocorre com ambos os vocabulários controlados, entretanto, será considerado o primeiro como o principal.

- **RF03 Tratamento Palavras-chave:** O objetivo do RF03 é realizar o tratamento das palavras-chave não compatibilizadas com os vocabulários controlados, com a intervenção do humano. Neste processo, o foco é na semântica das palavras-chave para identificar os possíveis relacionamentos entre palavras-chave e termos dos vocabulários controlados. Está centrado em três operações: 1) compatibilização com termo correspondente no primeiro vocabulário controlado; 2) compatibilização com o segundo vocabulário controlado; 3) inserção em vocabulário controlado da própria revista.

- **RF04 Resultados:** O objetivo do RF04 é mostrar os resultados quantitativos dos processamentos referentes à compatibilização de palavras-chave por processos de comparação.

3.1.3 Desenvolver o projeto físico do banco de dados

Neste planejamento está contido o desenvolvimento do projeto físico do banco de dados – base estruturada relacional. Segundo Heuser (2004), o projeto do banco de dados passa por dois estágios: modelo conceitual – modelagem das necessidades de armazenamento por meio de um diagrama de entidade relacionamento – e o modelo lógico – modelagem do diagrama entidade relacionamento obtido para um SGBD específico.

3.1.4 Criar a Base de Dados no Sistema Gerenciador de Banco de Dados

A construção do banco de dados no sistema de gerenciamento de banco de dados é realizada por meio da linguagem SQL, utilizando o comando CREATE TABLE. A base de dados foi criada no SGBD MySQL, escolhido por se tratar do mesmo sistema gerenciado de banco de dados utilizado no sistema SEER.

3.2 ETAPA 2 - ELABORAR A ESTRUTURA CONCEITUAL DO DOMÍNIO E SUBDOMÍNIO DO PERIÓDICO

3.2.1 Identificar o escopo do periódico

A revista “Discursos Fotográficos” é uma publicação semestral e está vinculada ao Programa de Pós-Graduação em Comunicação da Universidade Estadual de Londrina. Lançada em 2005, publica trabalhos inéditos (artigos científicos, entrevistas e resenhas) voltados à Comunicação Visual (Fotografia, Televisão, Cinema, Semiótica, *Design*, Produção gráfica, Antropologia visual e outros).

3.2.2 Levantar os instrumentos terminológicos do domínio e subdomínio

A partir da realização de levantamento bibliográfico, constatou-se que a área de Comunicação, até o momento, carece de instrumentos terminológico-documentários, como tesouros e vocabulários controlados. Em razão disso, a pesquisa foi realizada, especialmente, em dicionários e glossários na área de Comunicação.

3.2.3 Analisar Tabelas de Áreas do Conhecimento

A Tabela de Áreas do Conhecimento do CNPq classifica a subárea Comunicação Visual como um ramo da área Comunicação, sendo esta, por sua vez, um ramo da grande área Ciências Sociais Aplicadas I. Cabe esclarecer que esta etapa encontra-se em andamento.

3.3 ETAPA 3 - ANALISAR AS PALAVRAS-CHAVE COM OS TERMOS DOS INSTRUMENTOS TERMINOLÓGICOS

3.3.1 Coletar e armazenar os metadados

A importação dos metadados do periódico para a base de dados é o processo de realizar as requisições ao protocolo OAI-PMH e tratar o retorno da requisição, por meio do arquivo XML. Este processo faz parte da modelagem do RF01 da

ferramenta proposta. Portanto, nesta etapa, se dá a execução dessa funcionalidade.

3.3.2 Verificar as palavras-chave que são idênticas entre si

O levantamento das palavras-chave do periódico “Discursos Fotográficos” com a ferramenta *VCPC Tools* demonstrou que o universo da pesquisa era formado por 978 palavras-chave. Este levantamento possibilitou que fosse realizada a verificação das palavras-chave idênticas e o seu tratamento.

A partir desse levantamento foi realizada a primeira higienização dos dados, isto é, foram retiradas as palavras-chave idênticas entre si, porém, nesse momento não foram retirados os sinônimos, nem as palavras no plural. Após a conclusão desta etapa, obteve-se como resultado 639 palavras-chave. Destaca-se que o levantamento foi executado do ano de 2005, quando foi lançado o periódico, até o ano de 2015, portanto, desde a criação da Revista até sua última publicação. Assim, as palavras-chave foram coletadas dos 19 números já publicados.

3.3.3 Verificar as palavras-chave que tenham termos correspondentes nos instrumentos terminológicos

Para a verificação das palavras-chave que tenham termos correspondentes no vocabulário controlado consultou-se a documentação terminológica da área de Comunicação. Além disso, verificou-se no artigo se a mesma palavra representava sentido idêntico em sua aplicação.

3.3.4 Classificar tematicamente as palavras-chave conforme a estrutura conceitual do periódico

Ao pensar em palavras-chave e termos recorreu-se à definição de Dahlberg, (1978, p. 102) que argumenta da seguinte forma:

Pode-se definir a formação dos conceitos como a reunião e compilação de enunciados verdadeiros a respeito de determinado objeto. Para fixar o resultado dessa compilação necessitamos de um instrumento. Este é constituído pela palavra ou por qualquer signo que possa traduzir ou fixar essa compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico.

Para a organização das palavras-chave coletadas, considerou-se como

divisão de assuntos os termos descritos no periódico, assim, foram levantados nove (9) termos que compõem as classes de assuntos: Fotografia, Televisão, Cinema, Semiótica, Design, Produção Gráfica, Antropologia Visual, Jornalismo e Outros Assuntos. Buscou-se por meio de pesquisa nos artigos, isto é, leitura do resumo e quando necessário do corpo do artigo, verificar em qual divisão de assunto o termo melhor representava o texto. Também foi verificado em glossários e em dicionários qual o significado e a representação que tais termos tinham do texto. Assim, tornou-se possível o conhecimento e a familiarização com termos pertinentes ao domínio Comunicação e ao subdomínio Comunicação Visual.

Segundo Fujita (1998, p. 33), quando há escassez de materiais, pode-se utilizar “três modelos: o currículo, a estrutura departamental e o conhecimento do especialista”. Nesse contexto, um especialista, professor adjunto do Departamento de Comunicação, integrou-se ao Projeto de Pesquisa e tem colaborado nos estudos terminológicos da área da Comunicação, o que possibilita a ligação dos termos, ou não, com as bases conceituais do periódico “Discursos Fotográficos”. Além disso, o aprofundamento do trabalho conjunto com o pesquisador da área de Comunicação levou a ampliação das classes e subclasses de assuntos do Domínio e do Subdomínio em foco. A seguir, o Quadro 2 apresenta um exemplo desta expansão.

QUADRO 2: Parte da Estrutura conceitual do periódico “Discursos Fotográficos”

COD. ASS.	DIVISÃO DE ASSUNTOS	SUBDIVISÃO DE ASSUNTOS	CLASSES E SUBCLASSES DE ASSUNTOS
1			FOTOGRAFIA
	1.1		FOTOJORNALISMO
	1.2		FOTOGRAFIA DOCUMENTAL
	1.3		FOTOGRAFIA PUBLICITÁRIA
	1.4		ACERVOS FOTOGRÁFICOS
	1.5		FOTO CLUBISMO
	1.6		ANÁLISE EM FOTOGRAFIA
	1.7		HISTÓRIA DA FOTOGRAFIA
	1.8		TECNICAS DE FOTOGRAFIA
	1.9		TEORIAS DA FOTOGRAFIA
	1.10		FOTOGRAFIA E EDUCAÇÃO
	1.11		FOTOGRAFIA E HISTÓRIA
	1.12		FOTOGRAFIA E LEGISLAÇÃO
	1.13		FOTOGRAFIA E MEMÓRIA
	1.14		FOTOGRAFIA E OUTRAS ARTES

Fonte: Elaborada pelos autores.

Conforme o quadro acima, pode-se perceber a ampliação da Classe de Assunto Fotografia em dezessete subclasses, possibilitando desta maneira que os

assuntos referentes à Fotografia fossem contemplados.

3.3.5 Inserir a palavra-chave como candidato a termo no vocabulário controlado da revista

Esta ação é realizada nos procedimentos do RF03 Tratamento Palavras-chave, a partir das atividades das ações 3.3.3 e 3.3.4, sendo possível inserir a palavra-chave, no vocabulário controlado da revista, após a análise terminológica das características para se considerar um termo.

4 RESULTADOS PRELIMINARES

Com os resultados obtidos a partir de um índice de termos que apresente correlação com os respectivos artigos publicados no periódico científico eletrônico, buscou-se proporcionar o acesso aos conteúdos temáticos e contribuir para a apropriação de novos conhecimentos. Considerou-se que a aplicação do controle de vocabulário nas palavras-chave, atribuídas em artigos de periódicos científicos, permitiu maior consistência e relevância, visando facilitar o acesso a conteúdos temáticos de áreas especializadas, disponíveis no Portal de Periódicos da UEL. Os resultados alcançados serão apresentados na mesma estrutura das estratégias metodológicas, portanto, para melhor compreensão, será utilizada a mesma numeração.

A ETAPA 1 - DESENVOLVER A FERRAMENTA VCPC *TOOLS* ocorreu conforme delineada nos procedimentos metodológicos, que compreendem os itens 3.1.1 Identificar o Endereço do Repositório Digital do Periódico; o 3.1.2 Desenvolver o Projeto da Ferramenta VCPC *Tools*; o 3.1.3 Desenvolver o projeto físico do banco de dados; e o 3.1.4 Criar a Base de Dados no Sistema Gerenciador de Banco de Dados. Resultante desta etapa, na Figura 1 apresenta-se a interface inicial do ambiente do administrador.

Já na ETAPA 2 - ELABORAR A ESTRUTURA CONCEITUAL DO DOMÍNIO E SUBDOMÍNIO DO PERIÓDICO foram realizados os itens: 2.1 Identificar o escopo do periódico, 2.2 Levantar os instrumentos terminológicos do domínio e subdomínio, 2.3 Analisar Tabelas de Áreas do Conhecimento.

FIGURA: Interface inicial do VCPC Tools

Ambiente Administrador

Importar MD/XML | Análise MD/CV | Resultados | Tratamento PC | Gerenciar PC

Gerenciar VC

Nesta etapa os metadados serão importados da base de dados conforme o endereço digitado no campo abaixo.

- protocolo OAI-PMH, e a primeira etapa de análise: Eliminação de redundância (palavras repetidas).

Endereço OAI-PMH:

Limpar | Importar | Finalizar

Fonte: Elaboração própria.

Na ETAPA 3 - ANALISAR AS PALAVRAS-CHAVE COM OS TERMOS DOS INSTRUMENTOS TERMINOLÓGICOS, do item 3.1 Coletar e armazenar os metadados, obteve-se um total de 341 documentos e 1020 palavras-chave inseridas na base de dados da ferramenta. Apresenta-se no Quadro 3 a distribuição do total de documentos por seção do periódico científico.

QUADRO 3: Coleta de metadados

Tipo de documento	Descrição	Quantidade de documentos	Quantidade de palavras-chave
discursosfotograficos:ART	ARTIGOS	166	627
discursosfotograficos:Resenhas	RESENHAS	59	122
discursosfotograficos:ED	EDITORIAIS	19	-
discursosfotograficos:DIS	DISSERTAÇÕES	51	229
discursosfotograficos:EN	ENTREVISTAS	23	9
discursosfotograficos:EXP	EXPEDIENTE	14	-
discursosfotograficos:DOS	DOSSIÊ	9	33
TOTAL		341	1020

Fonte: Elaborado pelos autores.

Observa-se a predominância da quantidade de documentos denominados artigos, em seguida as resenhas e dissertações. Em relação à quantidade de palavras-chave, o maior número está relacionado com os artigos, posteriormente com dissertações e resenhas, ocorrendo uma inversão em relação à quantidade de documentos.

3.2 Verificar as palavras-chave que são idênticas entre si. Neste item o VCPC Tools realiza os procedimentos comparativos e relaciona as palavras-chave idênticas, e reduz a redundância de palavras-chave iguais. O VCPC Tools fez esta

contagem dos registros idênticos, totalizando 362. Portanto, das 1020 palavras-chave coletadas, subtraem as iguais e totalizam 658, sendo este o *corpus* de estudo das próximas etapas.

QUADRO 4: Coleta de metadados.

Qtde de palavras-chave	Quantidade de documentos				
	discursosfotograficos: ART	discursosfotograficos: Resenhas	discursosfotograficos: DIS	discursosfotograficos: EM	discursosfotograficos: DOS
0	7	16	1	14	1
1	0	11	0	9	0
2	0	7	1	0	0
3	20	7	5	0	1
4	128	17	17	0	5
5	11	0	20	0	2
6	0	0	5	0	0
7	0	0	2	0	0
8	0	1	0	0	0

FONTE: Elaborado pelos autores.

3.3 Verificar as palavras-chave que tenham termos correspondentes nos instrumentos terminológicos

Esta verificação encontra-se em andamento.

3.4 Classificar tematicamente as palavras-chave conforme a estrutura conceitual do periódico

Esta classificação encontra-se em andamento.

3.5 Inserir a palavra-chave como candidata a termo no vocabulário controlado da revista. Esta ação é realizada nos procedimentos do RF03 Tratamento Palavras-chave, a partir das atividades das ações 3.3 e 3.4. Optou-se por inserir, no vocabulário controlado da revista, a palavra-chave, caso tenha características para ser considerada um termo.

5 CONSIDERAÇÕES PARCIAIS

O controle de vocabulário das palavras-chave de artigos científicos eletrônicos e a compatibilização entre as palavras-chave e os termos presentes nos instrumentos terminológicos de áreas especializadas foram realizados conforme os parâmetros estabelecidos na proposta metodológica, e dentro das expectativas de validação. Para a execução de cada etapa do processo utilizou-se a *VCPC Tools*, o que tornou possível observar a diversidade de sintaxe das palavras-chave atribuídas aos artigos científicos eletrônicos.

O índice de termos proporcionado pela ferramenta torna-se um meio para conhecer a revista com uma visão holística. O perfil da revista é mais uma opção de

visualização do conteúdo por ela tratado e seus vieses de produção científica. Além disso, este modo de visualização dos conteúdos temáticos disponíveis na base do periódico “Discursos Fotográficos” apresentou-se como alternativa eficaz para auxiliar pesquisadores e os autores, por meio do índice de termos, na atribuição das palavras-chave às suas produções científicas.

Cabe salientar que o projeto de pesquisa supracitado se encontra em andamento e espera-se, com o aprimoramento da *VCPC Tools*, possibilitar a aplicação da proposta em periódicos de áreas especializadas, gerenciados pelo sistema SEER.

REFERÊNCIAS

STANDARDS ORGANIZATION - **NISO. Z39:19-2005**: guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2005.

BAITELLO JUNIOR, N.; KLEIN, A. Teorias da Imagem. In: CITELLI, A. et al (orgs.). **Dicionário de Comunicação**: escolas, teorias e autores. São Paulo: Contexto, 2014.

BRITO, R. F.; GUEDES, M. G.; SHINTAKU, M. **Atribuição de identificadores digitais para publicações científicas**: DOI para o SEER/OJS /. Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2013.

CANEVACCI, M. **Antropologia da Comunicação Visual**. São Paulo: Brasiliense, 1990.

CAPES. Tabela de Áreas de Conhecimento/Avaliação (2014). Disponível em: <<http://www.capes.gov.br/avaliacao/instrumentos-de-apoio/tabela-de-areas-do-conhecimento-avaliacao>>. Acesso em: 28 maio 2016.

CATALÀ DOMÈNECH, J. M. **A Forma do Real**: introdução aos estudos visuais. São Paulo: Summus, 2011.

COSTA LIMA, L.; DORIA, F. A.; KATZ, C. S. **Dicionário Básico de Comunicação**. Rio de Janeiro: Paz e Terra, 1975.

DAHLBERG, I. Teoria do conceito. *Ciência da Informação*, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978c. Disponível em: <<http://revista.ibict.br/ciinf/issue/view/98>>. Acesso em: 22 dez. 2015.

FUJITA, M. S. L. Modelos de categorização para a construção de tesouros: metodologia de ensino. In: BOCCATO, Vera Regina Casari; GRACIOSO, Luciana de Souza (Org.). **Estudos de linguagem em Ciência da Informação**. Campinas: Alínea, 2011. Cap. 2.

FUJITA, M. S. L. A elaboração de categorias do tesouro: modelos de elaboração. **Cadernos da Faculdade de Filosofia e Ciências**, Marília, v. 7, n. 1/2, p. 107-120, 1998.

GIL LEIVA, I.; ALONSO ARROYO, A. La relación entre las palabras clave aportadas por los autores de artículos de revista y su indización en las bases de datos ISOC, IME e ICYT. **Revista Española de Documentación Científica**, Madrid, v. 28, n. 1, 2005. Disponível em: <<http://redc.revistas.csic.es/index.php/redc/article/view/165>>. Acesso em: 9 set. 2015.

GIL LEIVA, I. La automatización de la indización de documentos. [S. l.]: Ediciones Trea, 1999.

GONÇALVES, A. L. Uso de resumos e palavras-chave em ciências sociais: uma avaliação. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 13, n. 26, p. 1-15, 2008. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2008v13n26p78/6636>>. Acesso em: 9 set. 2015.

GONÇALVES, A.; RAMOS, L. M. S. V. C.; CASTRO, R. C. F. Revistas científicas: características, funções e critérios de qualidade. In: POBLACION, D. A.; WITTER, G. P.; SILVA, J. F. M. (Org.). **Comunicação e produção científica: contexto, indicadores, avaliação**. São Paulo: Angellar, 2006. p. 163-190.

HEUSER, C. A. Projeto de banco de dados. 5. ed. Porto Alegre: Sagra Luzzatto, 2004.

MEADOWS, A. J. **A comunicação científica**. Brasília: Brique de Lemos, 1999.

LARA, M. L. G. Glossário: termos e conceitos da área de comunicação e produção científica. In: POBLACION, D. A.; WITTER, G. P.; SILVA, J. F. M. **Comunicação e produção científica: contexto, indicadores e avaliação**. São Paulo: Algelara, 2006. p. 387-414.

MUELLER, S. P. M. O periódico científico. In: CAMPELLO, B. S.; CEDÓN, B. V.; KREMER, J. M. (Org.). **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: UFMG, 2000b.

NEIVA, E. **Dicionário Houaiss de Comunicação e Multimídia**. São Paulo: Publifolha, 2013.

ORENGO, V. M.; HUYCK, C. A Stemming Algorithm for the Portuguese Language. In: INTERNATIONAL SYMPOSIUM ON STRING PROCESSING AND INFORMATION RETRIEVAL (spire), Laguna de San Raphael, Chile, 2001, p.183-193.

ORENGO, V. M.; BURIOL, L. S.; COELHO, A. R. A Study on the Use of Stemming for Monolingual Ad-Hoc Portuguese Information Retrieval. In: CLEF 2006, Alicante, Berlin Heidelberg: Springer-Verlag. v. 4730, 2007, p. 91-98.

RABAÇA, C. A.; BARBOSA, G. G. **Dicionário de Comunicação**. São Paulo: Ática, 1995.

SOUZA, J. L. A. Revistas eletrônicas com uso de software livre. **Datagrama zero**: Revista de Ciência da Informação, Brasília, v. 11, n. 4, ago. 2010. Disponível em: <http://www.dgz.org.br/ago10/Art_06.htm>. Acesso em: 10 abr. 2014.