

ESTRATÉGIA DE PESQUISA BIBLIOGRÁFICA UTILIZANDO O MÉTODO MULTICRITÉRIO DE BORDA NA PRIORIZAÇÃO DE ARTIGOS

Fernando Andrade (UFF)

andradefernando@yahoo.com.br

Jose Rodrigues de Farias Filho (UFF)

fariasfilho@gmail.com

EIXO TEMÁTICO: Métodos, Técnicas e Ferramentas Para Estudos
Bibliométricos e Cientométricos

MODALIDADE: Apresentação pôster

1 INTRODUÇÃO

Na década de 70, Umberto Eco escreveu o clássico “Como se faz uma tese”. Esta obra propõe a construção de um trabalho científico em 5 etapas: A escolha do tema; A procura do material; O plano de trabalho e a elaboração de fichas; A redação; e por fim A redação definitiva. O presente trabalho tem como objetivo apresentar uma estratégia contemporânea para etapa “A procura do material”.

Com o desenvolvimento da internet, o crescimento significativo do conteúdo acadêmico e o desenvolvimento dos motores de busca, o problema de acesso à informação científica passou da escassez para o excesso. Assim, frente a uma vasta disponibilidade de artigos em amplas bases de dados bibliográficas, um dos atuais desafios dos pesquisadores é a escolha dos artigos mais adequados para construção da argumentação teórica de seus estudos. Neste contexto, o presente trabalho tem como objetivo apresentar uma estratégia simplificada de pesquisa bibliográfica utilizando ferramentas de gestão bibliográfica para seleção de artigos e o método multicritério de Borda para priorização dos mesmos.

O trabalho está organizado em quatro tópicos, o primeiro trata da exploração bibliográfica cujo objetivo é a construção de uma base ou portfólio de artigos. No segundo é apresentado um método de análise de alinhamento para selecionar os artigos com maior aderência ao estudo. O terceiro traz a aplicação do Método de Borda como ferramenta de apoio na priorização dos artigos selecionados. Por fim, um exemplo de aplicação do método proposto é apresentado.

2 EXPLORAÇÃO BIBLIOGRÁFICA: CONSTRUÇÃO DA BASE OU PORTFÓLIO DE ARTIGOS

O mapeamento bibliométrico ou mapeamento científico é uma campo da bibliometria que objetiva, através de representações gráficas, representar como disciplinas, campos de estudos, documentos ou autores estão relacionados entre si . Existem uma série de ferramentas que auxiliam na execução deste mapeamento . Para uma etapa inicial de pesquisa, a partir dos estudos da metodologia elaborada por , o presente trabalho traz uma proposta simplificada de exploração bibliográfica em que a coleta de material bibliográfico pode ser estruturada em 4 etapas (Figura 1).

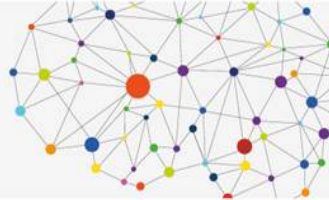


Figura 1 - Modelo simplificado para exploração bibliográfica.

Fonte: Produção do próprio autor

Tendo o problema de pesquisa definido, deve-se construir uma árvore de palavras-chave que serão as bases das pesquisas avançadas nos motores de busca. Pode-se ter mais de uma árvore por problema de pesquisa. Para uma melhor visualização, sugere-se a utilização de ferramentas para construção de mapas mentais, como por exemplo, Xmind™ e FreeMind™. Os operadores booleanos são utilizados para aumentar a abrangência da pesquisa (conexão de palavras ou expressões sinônimas com “OU”) e para restringir os resultados de buscas avançadas (conexão de expressões com “E”). Os trabalhos de e abordam com detalhes a construção destas árvores de palavras-chave.

O próximo passo é realizar a pesquisa avançada em motores de busca (ScienceDirect, SCOPUS, Wiley Online Library, Google Acadêmico, SciELO, etc). A especificidade sintática de cada motor é encontrada em seus campos de ajuda . Com a árvore de palavras-chave e conhecendo a sintaxe de pesquisa avançada do motor de busca escolhido pode-se elaborar os algoritmos de busca avançada. Esta é uma etapa dinâmica, ou seja, se o resultado da pesquisa for muito baixo (p.ex., menor que 50 artigos), deve-se aumentar a abrangência do algoritmo de pesquisa reduzindo a quantidade de ligações “E”. Nos casos de resultados muito altos (p.ex., maior que 1.000 artigos), deve-se aumentar o número de palavras-chave com conexão “E” ou utilizar operadores de restrição, como “AND NOT” e “NOT”, por exemplo.



Normalmente, os resultados das buscas são conjuntos de centenas à milhares de artigos, o que torna adequada a utilização de um software de gestão bibliográfica (p.ex., Zotero™ e EndNotes™). A maior parte dos motores de busca possuem ferramentas para exportar os resultados de pesquisa em arquivos compatíveis com os softwares de gestão bibliográfica. Vale ressaltar que a quantidade de referências que podem ser exportadas de cada vez varia consideravelmente entre os diferentes motores, por exemplo, a base Wiley Online Library permite a exportação em blocos de 20 (vinte) artigos, já com ScienceDirect pode-se exportar blocos de até 1.000 (mil) artigos.

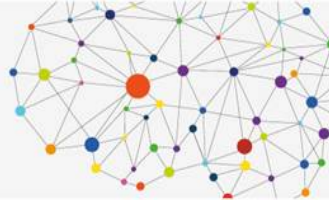
Com o resultado das pesquisas avançadas, já importadas em uma ferramenta de gestão bibliográfica, o próximo passo é extrair deste software os seguintes dados da base de artigos: (i) ano da publicação, (ii) periódico, (iii) título e (iv) resumo. Sugere-se carregar estas informações em uma planilha eletrônica. Esta planilha, intitulada “base_artigos”, é o produto desta primeira etapa do processo de pesquisa bibliográfica.

3 SELEÇÃO DE ARTIGOS: FILTROS COM BASE NA ANÁLISE DE RELEVÂNCIA

Com a base de artigos organizada em formato de planilha eletrônica com quatro colunas (ano da publicação, periódico, título e resumo), pode-se iniciar a análise de relevância. Para isso devem ser criadas mais três colunas na planilha “base_artigo”, são elas: alinhamento do periódico, alinhamento do título e alinhamento do resumo.

O processo de análise de relevância inicia-se pelo periódico. Como o mesmo periódico se repete em várias linhas da “base_artigos”, sugere-se trabalhar com tabelas dinâmicas permitindo listar os periódicos sem repetição. Com isso, inicia-se a análise de alinhamento do periódico atribuindo “sim” aos periódicos que possuem aderência com o tema da pesquisa e “não” para os demais. No término desta análise, com utilização da função PROCV do Excel, ou equivalente se outra ferramenta de planilha eletrônica estiver sendo utilizada, pode-se transferir a classificação dos periódicos feitos na tabela dinâmica para a coluna “alinhamento do periódico”.

O próximo passo é preencher a coluna “alinhamento do título”. Para cada linha com “não” na coluna “alinhamento do periódico” propõem-se preencher com “não” a coluna “alinhamento do título”. Após esta tratativa, inicia-se a análise dos títulos restantes atribuindo “sim”, caso tenham aderência com o tema a ser estudado, e “não” caso contrário. A mesma tratativa deve ser executada com “alinhamento do resumo”.



Esta proposta de análise de alinhamento, que pode-se denominar de binária, é utilizada para definição de materiais contundentes com o estudo, ou seja, aqueles que vão contribuir para base teórica do mesmo. Para uma análise exploratória do conhecimento, a análise de relevância pode ser trabalhada na forma de escala, por exemplo fornecendo notas de 1 a 5 (quanto maior a nota mais alinhado com o tema), sendo esta a proposta apresentada por .

Adotando o modelo binário descrito acima, encerra-se a etapa de filtro com base na relevância. Os artigos que receberam “sim” nas três colunas terão valor 1 na coluna “alinhamento com o tema” e estarão aptos para entrar na etapa de priorização, os demais ficarão com 0 e como consequência serão desconsiderados da etapa de priorização.

4 PRIORIZAÇÃO DE ARTIGOS: APLICAÇÃO DO MÉTODO MULTICRITÉRIO DE BORDA

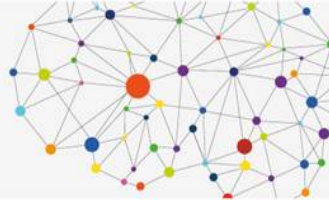
O objetivo desta etapa é ranquear os artigos selecionados. Para isso propõe-se utilizar quatro indicadores para priorização: (i) o ano de publicação, (ii) JCR (Fator de Impacto), (iii) JCR (Fator de Impacto 5 anos) e (iv) Meia Vida. Os últimos três indicadores são obtidos na base JCR Journal Citation Reports (Thomson Scientific / ISI Web Services) que pode ser acessada através do portal Periódicos Capes. Ressalta-se que os 4 indicadores propostos são do tipo “quanto maior melhor”.

Com a definição dos indicadores, torna-se necessário definir o peso de cada um na priorização dos artigos. Para tanto, propõem-se a utilização do Método de Borda aplicado a um grupo de alunos e/ou profissionais que possuem contato com o tema. Para aplicação do Método de Borda, sugere-se utilização de ferramentas de pesquisa *online*, como por exemplo o Qualtrics™ .

Com o preenchimento das variáveis e a definição dos pesos através do Método de Borda, pode-se atribuir pontos para cada artigo utilizando a Equação (1).

$$Pontuação_{i_1}^n = Relevância_i * \sum_{j=1}^m peso_{variável_j} * \frac{variável_j}{máximo(vi)} \quad (1)$$

Onde, i representa cada um dos n artigos a serem priorizados, n a quantidade de artigos a serem priorizados, j representa cada uma das m variáveis ou atributos dos artigos e m a quantidade de variáveis ou atributos considerados



A aplicação da Equação (1) para o caso dos 4 indicadores citados resulta ao disposto na Equação (2).

$$Pontuação_i = Relevância_i * \left\{ \left(peso_{ano} * \frac{ano_i}{máximo(ano)} \right) + \left(peso_{JCR} * \frac{JCR_i}{máximo(JCR)} \right) + \left(peso_{JCR5anos} * \frac{JCR5anos_i}{máximo(JCR5anos)} \right) + \left(peso_{msiavida} * \frac{msiavida_i}{máximo(msiavida)} \right) \right\}$$

(2)

A partir da atribuição de pontos os artigos são ranqueado. Dependendo da quantidade de material nesta etapa, pode-se utilizar o primeiro quartil ou os primeiro 30 artigos, deixando os demais para consultas ou aprofundamentos.

5 EXEMPLO DE APLICAÇÃO DA METODOLOGIA PROPOSTA

A proposta deste tópico é apresentar a aplicação da metodologia descrita nos itens anteriores no seguinte problema de pesquisa estruturado de acordo com o modelo de definição de escopo proposta por .: “(i) Estamos estudando regulação do setor elétrico na América Latina (ii) porque queremos descobrir como esta variável influencia na atratividade do setor para investimentos estrangeiros, (iii) para entender como aspectos regionais influenciam na internacionalização de multinacionais do setor elétrico”.

A árvore de palavras-chave inicial para este problema de pesquisa com lógica booleana pode ser: (“Electric Power Distribution”) AND (Regulation) AND (“Return on Capital” OR Remuneration) AND (“Latin America” OR “South America”). Adaptando estas palavras-chave na sintaxe de diferentes motores de busca obteve-se os seguintes resultados: na base Wiley Online Library mais de 90% dos periódicos não foram aderentes ao tema de pesquisa. Na base IEEE Xplore, a pesquisa retornou uma pequena quantidade de artigos, somente 17. O SciELO e Google Acadêmico retornaram uma pequena quantidade de artigos com baixo alinhamento ao tema de pesquisa. Finalmente, a pesquisa com ScienceDirect teve retorno significativo, a primeira pesquisa retornou 948 artigos, restringindo a pesquisa para os últimos cinco anos o número reduziu para 257. Com a leitura transversal, identificou-se um considerável volume de artigos relacionados com energia eólica, tema fora do escopo da pesquisa. Assim, com a utilização do operador booleano “AND NOT” restringiu-se a palavra “Wind” resultando em 131 artigos. Estes foram importados para a ferramenta de gestão bibliográfica EndNote™ e em seguida exportado em formato texto (.txt) com os campos básicos: (i) ano, (ii) periódico, (iii) título e (vi) resumo.

Dos 131 artigos, 26 foram descartados por não possuírem um dos 4 campos básicos supracitados. A seguir, a análise de relevância do periódico eliminou 7 artigos, a análise do título eliminou 50 e do resumo eliminou 22. Assim, foram selecionados 26 artigos.

Para priorização dos artigos, utilizou-se o Método de Borda para definir por votação o peso dos indicadores que caracterizam cada artigo e/ou periódico. Os eleitores foram os alunos da disciplina Administração Estratégica do Programa de Pós Graduação em Engenharia de Produção da Universidade Federal Fluminense ministrada pelo professor José Rodrigues de Farias Filho no segundo semestre de 2013. A ferramenta de pesquisa *online* utilizada foi o Qualtrics™. O resultado foi o seguinte: peso 23 para ano da publicação, 31 para fator de impacto, 22 para fator de impacto 5 anos e 24 para meia vida. Pela quantidade de artigos resultantes (menor que 30), todos foram considerados no estudo em questão.

5 CONCLUSÃO

Com aumento considerável no número de publicações científicas e a disponibilidade deste vasto material em grandes bases de pesquisa, os métodos para busca de referencial teórico foram sensivelmente alterados. Uma série de trabalhos apresentam propostas para seleção e priorização de artigos neste novo cenário, como por exemplo o trabalho de e .

O presente trabalho propõe um modelo geralista simplificado para seleção e priorização de artigos, utilizando para isso o motor de busca ScienceDirect, ferramenta de gestão bibliográfica EndNote™, gestor de planilha eletrônica Excel™, avaliação de periódicos JCR, modelo de votação baseado no Método de Borda e ferramenta de enquete Qualtrics™.

REFERÊNCIAS