REVISTA INTERDISCIPLINAR DE PESQUISA EM ENGENHARIA

CILAMCE 2016
XXXVII IBERIAN LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGINEERING
BRASÍLIA - DF - BRAZIL

ANÁLISE DOS TÓPICOS DE PESQUISA DOS PESQUISADORES BRASILEIROS QUE ATUAM NAS ÁREAS DE ENGENHARIAS

Jether Gomes

jethergomes@yahoo.com.br

Thiago M. R. Dias

thiago@div.cefetmg.br

Gray F. Moita

gray@dppg.cefetmg.br

Programa de Pós-graduação em Modelagem Matemática e Computacional

Centro Federal de Educação Tecnológica de Minas Gerais

Av. Amazonas, 7576, Nova Gameleira, 30510-000, Belo Horizonte, MG

Abstract. *A produção e publicação de trabalhos científicos têm aumentado significativamente nos últimos anos, sendo à internet o principal fator de acesso e difusão destes. Diante disto, vários estudos sobre dados de produções científicas têm sido realizados por pesquisadores de diversas áreas do conhecimento a fim de analisar fenômenos e tendências acerca da ciência. Sendo que, o entendimento de como as pesquisas têm evoluído, pode por exemplo, servir de base para construção de políticas científicas visando novos avanços na ciência e ou impulsionar grupos de pesquisas a se tornarem mais produtivos. Neste contexto, o objetivo deste trabalho é analisar os principais tópicos de pesquisa investigados ao longo da trajetória da ciência brasileira de pesquisadores que atuam nas áreas de engenharias, com intuito de mapear o conhecimento científico e identificar temas em destaques. Para isso, são realizados estudos sobre a frequência e relacionamento das palavras-chave do conjunto de artigos científicos cadastrados nos currículos existentes na Plataforma Lattes de cada um dos pesquisadores selecionados, contando para tanto com o auxílio de características de análise bibliométrica.*

Palavras-chave: *Tópicos de Pesquisa; Bibliometria; Plataforma Lattes;*

1 INTRODUÇÃO

O grande número de informações disponibilizadas pela internet e a sociabilização da atividade científica por parte de redes de pesquisadores são fatores essenciais para o atual desenvolvimento da ciência (Castells, 2010). Serviços como bibliotecas digitais, redes de relacionamentos, repositórios bibliográficos e sítios para registro individual de produção científica, são alguns exemplos de como a internet tem contribuído consideravelmente na quantidade de trabalhos publicados, permitindo que usuários não apenas acessem conteúdo disponível, mas também possam registrar a sua produção técnica e científica a partir de sua interação com esse meio. Desta forma, trabalhos publicados e disponibilizados podem ser acessados instantaneamente contribuindo para a expansão do conhecimento (Dias, 2016).

O conhecimento é o elemento fundamental para a geração do desenvolvimento. Além da divulgação científica contribuir para a democratização do conhecimento, aproxima o cidadão comum dos benefícios que ele tem direito de reivindicar para a melhoria do bem-estar social, dando-lhe uma visão mais clara sobre as verdadeiras causas e efeitos dos problemas que enfrenta no dia a dia (Carneiro, 2003).

Numa sociedade competitiva em escala mundial, implementar o conhecimento técnico e científico é tarefa indispensável para o desenvolvimento econômico e social, principalmente para os países em desenvolvimento que são consumidores deste conhecimento de forma a identificar suas necessidades e peculiaridades (Saes, 2005). Porém, muitas das vezes, seu plano de implementação é a existência de recursos limitados e uma exigência cada vez maior de racionalidade e objetividade na aplicação dos poucos recursos disponíveis.

Tão importante quanto ter investimentos é ter habilidade para controlar, entender e medir o patamar científico das nações e/ou grupos individualizados, negócios e fundações que devem decidir suas prioridades científicas. Portanto, torna-se indispensável estudar, para conhecer e medir, a produção científica para a implementação desse conhecimento, superação das dificuldades e alcance dos pressupostos de racionalidade e objetividade (Saes, 2005)

Segundo Yi e Choi (2012), o entendimento da produção científica pode promover novos avanços na ciência. Em Brito et. al. (2016), os autores destacam que trabalhos desta natureza são considerados urgentes no Brasil e podem retratar o que é desenvolvido e publicado em ciência, tecnologia e inovação, possibilitando gerar parâmetros para nortear esforços e investimentos com intuito de impulsionar resultados de pesquisa.

Um crescente interesse por partes de pesquisadores ao redor do mundo das mais diversas áreas quanto à extração de conhecimentos em base de dados científicos tem sido revelado (Digiampietri 2015; Mena-Chalco, et. al. 2014; Dias et. al., 2014; Zhu et. al, 2013; Yi e Choi, 2012). Entretanto, os trabalhos utilizam convencionalmente, bases de dados científicas internacionais. Logo, por serem internacionais, podem não representar o que é produzido no Brasil (Brito et. al, 2016). Com isso, analisar fonte de dados que englobe diversos tipos de

publicação, principalmente em veículos nacionais e de diversas áreas, passa a ser uma tarefa relevante para compreensão da ciência brasileira.

Para Pritchard (1969), a bibliometria se destaca como uma das principais ciências métricas de análise de conteúdo, podendo ser aplicada a fontes de dados científicos com o intuito de se obter informações quantitativas sobre publicações. Dias (2016), destaca que com a utilização da bibliometria é possível identificar as tendências e o crescimento do conhecimento científico em diversas áreas, observar a dispersão do conhecimento científico, auxiliar políticas para investimentos e entender como acontece a evolução científica.

Assim sendo, este trabalho tem como objetivo realizar uma análise sobre palavras-chave de publicações científicas extraídas dos currículos dos indivíduos que atuam nas áreas de engenharias. O estudo contempla análises bibliométricas, afim de mapear os tópicos de investigação dos pesquisadores brasileiros, evidenciando-os que possuem destaque. Com isso, é apresentado uma visão detalhada sobre os principais temas de investigação que os pesquisadores que atuam nas áreas de engenharia vêm produzindo.

2 TRABALHOS RELACIONADOS

Esforços para identificar tópicos de pesquisa constitui uma forma de melhorar a compreensão do que têm-se produzido acerca da ciência. Estes estudos podem ser baseados tanto em contagem de palavras extraídas dos títulos das publicações ou das palavras-chaves de produções bibliográficas (Choi et. al., 2011). Por exemplo, Trucolo e Digiampietri (2014) desenvolveram e aplicaram regressões lineares e não lineares dos índices de importância baseado em frequência dos termos extraídos de títulos de publicações científicas de uma base histórica, com intuito de identificar tendências de assuntos e ramos de pesquisa em ciência da informação, para curto, médio e longo prazo.

Já Medeiros e Mena-Chalco (2013) analisaram mais de 650 mil currículos da Plataforma Lattes a fim de estudar a rede social composta por todas as pessoas que declararam atuar em ao menos uma das seguintes grandes-áreas: Ciências Humanas, Ciências Sociais Aplicadas ou Linguística, Letras e Artes. Adicionalmente, analisaram frequência das palavras dos títulos das publicações para identificar quais estão sendo mais utilizadas por período de tempo e que, de certa forma, são as mais importantes para estas áreas. Eles ressaltaram que a utilização de rede para a visualização das interações de pesquisa para grandes grupos é inviável devido ao grande número de nós e conexões que aparecem. Portanto, como apoio a análise, utilizaram os conceitos de mapas de palavras para as 200 palavras mais frequentes de cada área, por períodos de tempo estudados.

No trabalho de Cataldi et al. (2010), os autores reconhecem o importante papel do Twitter na disseminação de informações e propõe uma técnica de detecção em tempo real dos tópicos emergentes expressos pela comunidade sob restrições de tempo especificado pelo usuário. Primeiro, extraíram o conteúdo dos *tweets* e modelaram o conjunto de termos de acordo com uma métrica de envelhecimento para destacar os termos emergentes. Além disso,

foram analisadas as relações sociais na rede com o algoritmo Page Rank, a fim de determinar a autoridade dos usuários. Por fim, foi gerado um grafo direcionado que liga os termos emergentes com outras palavras relacionadas semanticamente.

Souza et al. (2014) analisaram redes semânticas construídas a partir de termos extraídos dos títulos de chamados de um sistema de atendimento eletrônico com uso de técnicas de análise de redes socais. Os autores destacaram três tipos de centralidade para análise da rede construída. A centralidade de grau, que trata a importância de um vértice nas conexões que estabelece com vértices vizinhos. A centralidade de proximidade, que mostra a importância de uma palavra em relação aos vizinhos mais próximos e também a sua importância em relação a toda a rede de palavras. E por fim, a centralidade de intermediação, que quantifica o número de vezes que uma palavra age como ponte ao longo do caminho mais curto entre outras duas palavras. Os resultados apresentados permitiram visualizar quais os problemas mais recorrentes a fim de facilitar a tomada de decisão e a definição de tendências nas solicitações de determinados usuários e/ou setores.

Em Zhu et al. (2013), com base na rede composta de 111.444 palavras-chave de publicações da área de ciência da informação, métricas de análises de redes sociais foram aplicadas e identificado o efeito mundo pequeno, mostrando que as palavras estão próximas umas das outras. Além disto, também foi identificado com o cálculo do grau de centralidade, que algumas palavras possuem um número alto de vínculos com outras e que isto demonstra a sua importância na rede. Diante de tal constatação, os autores realizaram um estudo preliminar sobre como detectar os termos mais relevantes de uma linha de pesquisa. Este método também foi comparado com a estratégia de identificação por frequência das palavras, justificando que a análise baseada em grau de centralidade tende a ser mais eficiente.

Khan e Wood (2015) analisaram as redes de palavras-chave e palavras extraídas dos títulos de artigos científicos referentes a área de Gestão da Tecnologia da Informação (GIT) através de técnicas de análise de redes sociais e do algoritmo *Burst Detection* com a proposta de mapear o conhecimento científico e destacar os tópicos emergentes que se têm produzido acerca do assunto, dado a importância do tema para indústria e sua constante evolução (Figura 1). As informações para estudo foram extraídas a partir da interface de consulta da biblioteca Web of Science. As redes foram construídas com base em aproximadamente 2000 palavras dos 893 artigos publicados entre 1995 e 2014 referentes a 40 revistas, 64 países, 914 instituições e 1914 pesquisadores. As métricas utilizadas para análise foram: número de componentes, diâmetro, densidade, coeficiente de clusterização, grau médio, intermediação e medidas de centralidade. Contudo, apesar da validade dos resultados, os autores sugerem limitações no estudo devido os dados analisados não serem suficientes para generalizar as conclusões acerca da área da GIT.

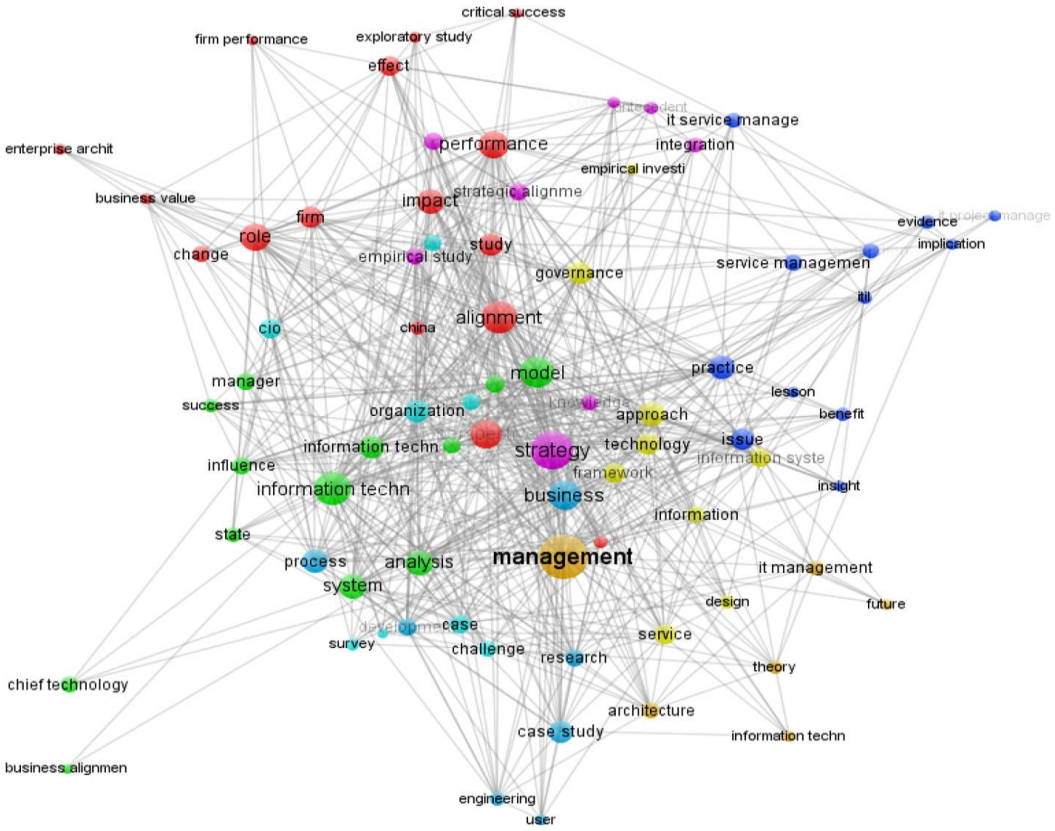

Figura 1. Redes de palavras extraídas dos títulos de produções científicas (Khan e Wood, 2015)

3 DESENVOLVIMENTO

A fonte de dados explorada para realização deste trabalho foi a Plataforma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Tal plataforma tem como objetivo integrar os sistemas de informação dos órgãos federais brasileiros, otimizando o processo de gestão de Ciência e Tecnologia (C&T) (Dias et al., 2014).

A escolha da plataforma está relacionada ao fato de tratar da integração de dados científicos de currículos e de instituições de todas as áreas de C&T existentes ao longo de toda trajetória da ciência brasileira. Atualmente, a plataforma conta com mais de 4,6 milhões de currículos (Dias, 2016). Esses dados disponíveis livremente na internet, ainda não foram amplamente analisados, apesar de, tipicamente, serem utilizados para avaliar ou verificar dados de pesquisadores e/ou grupos destes (Digiampietri, 2015). Também é utilizada como fonte de informação para órgãos que avaliam o Sistema Nacional de Pós-graduação do Brasil, e para as agências de fomento e financiamento de pesquisas e/ou ofertas de bolsas.

Apesar dos dados disponibilizados, estes são apenas visualizados através de uma interface de consulta que apresenta cada currículo individualmente, e com isso, não permite uma análise mais abrangente. Diante disto, para uma análise mais detalhada de grupos, instituições ou até de toda a nação de cientistas brasileiros, técnicas e ferramentas para a extração e análises desses dados científicos se fazem necessárias.

Para a aquisição dos currículos a serem analisados, foi utilizado o framework de extração de dados científicos desenvolvido por Dias et al. (2014). Figura 2.

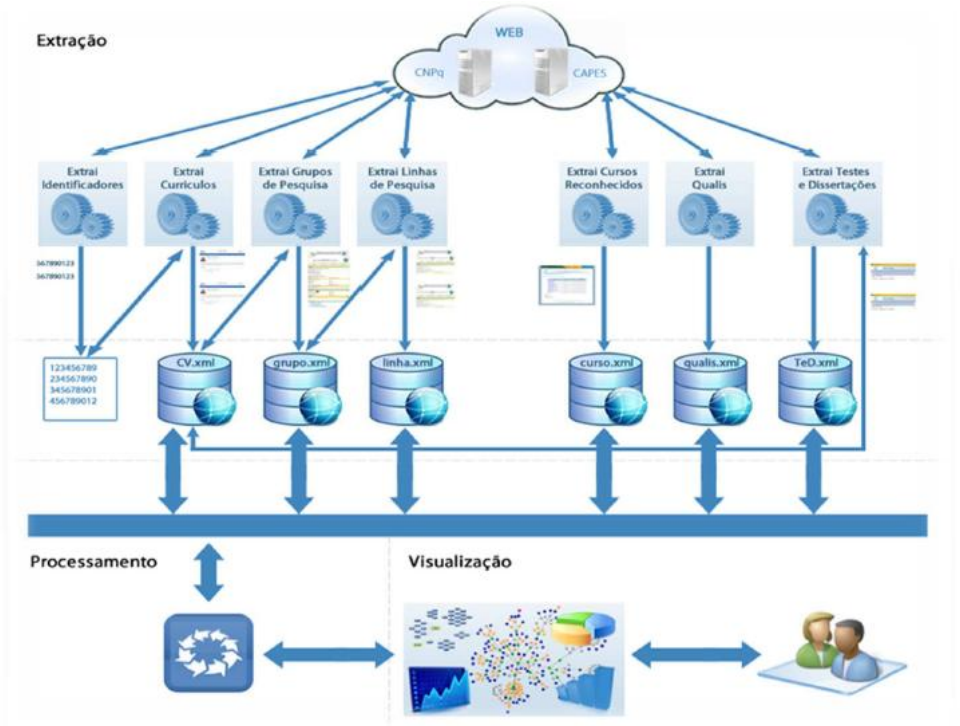

Figura 2. Framework para extração de dados da Plataforma Lattes (Dias et al., 2014)

O processo de extração dos dados se inicia pela aquisição de uma lista dos códigos dos currículos obtidos através da requisição na interface de consulta da Plataforma Lattes, para que, em seguida, os identificadores possam ser armazenados localmente. De posse dos identificadores, o framework realiza o download dos arquivos, armazenando-os em disco no formato XML (eXtensible Markup Language). Apesar de ser possível realizar a extração de currículos em HTML (HyperText Markup Language), a versão XML é a mais adequada para o processamento automático, pois possui todas as seções e campos bem delimitados, além de conter as palavras-chave dos artigos científicos, principal objeto de estudo deste trabalho.

A justificativa para o estudo de palavras-chave em detrimento dos títulos das publicações, estes amplamente utilizado em pesquisas na literatura, se dá pelo fato que as palavras-chave de um determinado trabalho têm como objetivo principal descrever os temas de pesquisa que norteiam o conteúdo descrito sem se preocupar com sua semântica.

A coleta dos dados ocorreu em março de 2016, totalizando 208.079 currículos de indivíduos que declararam ter como sua principal área de atuação a Engenharia. Em seguida, foi realizado a mineração dos arquivos XML com intuito de extrair informações dos indivíduos e de seus artigos científicos. Essa etapa é importante, pois extrai as informações que realmente precisam ser processadas e analisadas, e com isso, diminui o tempo de processamento computacional.

Ferraz et al. (2014) destaca que, analisar palavras-chave de artigos cadastrados na Plataforma Lattes não é uma tarefa trivial, tendo em vista que a escolha das palavras não segue um padrão pré-definido. Como resultado disto, geralmente tem-se uma coleção muito

grande e sem nenhum padrão. Na tentativa de contornar este problema, foi desenvolvido um método que realiza o processamento das palavras-chave com intuito de excluir possíveis palavras com ruídos ou que não representam um tópico de pesquisa.

O método inicia-se obtendo a quantidade de palavras-chave extraídas de cada artigo e suas referências para este. Diante disso, cada palavra passa por um processo de detecção de idioma, informação utilizada na etapa de *stemming*. Em continuação, na etapa de *lowercase*, as palavras são convertidas para minúsculo com a proposta de padronizar o conjunto.

Já na etapa de *StopWords*, são removidas as palavras que não possui valor semântico. Em seguida, é realizado o processo de normalização para extrair as letras acentuadas e substituí-las pelo seu equivalente sem acentuação. Na etapa de *stemming*, cada palavra é reduzida a seu radical, e com isso, evita a inclusão de palavras com o mesmo significado de formas distintas. No caso de palavras compostas, o processo é executado em cada palavra individualmente, e em seguida, são concatenadas formando uma única palavra.

Cada palavra resultante de todo o processo é inserida no dicionário mantendo seu formato original e a referência através do código do artigo. Caso a palavra já esteja presente no dicionário, um contador correspondente é incrementado. A Tabela 1 apresenta um exemplo de transformação de uma palavra. Enquanto a Tabela 2 ilustra um exemplo do dicionário.

Tabela 1: Exemplo de transformação de uma palavra extraída de um artigo científico

Etapa Algoritmo	Palavra Original
1	Gerência de Dados, 1000
2	Gerência de Dados, 1000, Português
3	gerência de dados, 1000, Português
4	gerência dados, 1000, Português
5	gerencia dados, 1000, Português
6	gerenc dad, 1000, Português
7	gerenc-dad, 1000

Tabela 2: Exemplo de dicionário onde a palavra transformada é inserida

Frequência	Radical	Cód. Artigo	Palavra 1	Cód. Artigo	Palavra 2	...
2	gerenc-dad	1000	Gerência de dados	500	Gerência de dados	**...**
25	mineraca-dad	1000	Mineração de Dados	100	Mineração de Dados	...
5	banc-dad	30	Banco de Dados	1000	Banco de Dados	...

O dicionário é composto pela quantidade das palavras que se reduziram a um determinado radical, pelo radical propriamente dito, os códigos referentes aos artigos que elas apareceram e as palavras-chave originais. Essa estratégia foi adotada para que todas as palavras-chave radicalizadas fossem possíveis de serem rastreáveis perante sua respectiva

publicação. Por exemplo, neste caso, pode-se notar que, no artigo de código 1000 apareceram as palavras-chave: Gerência de Dados, Mineração de Dados e Banco de Dados.

4 RESULTADOS E DISCUSSÕES

De todos os currículos cadastrados na Plataforma Lattes em março de 2016 (4.457.260) um total de 208.079 informaram como sua principal grande área, as Engenharias. Tal grande área é a quinta com a maior quantidade de currículos. Além da grande área de atuação de um determinado indivíduo, os currículos também possuem informações sobre as áreas de atuação, logo é possível visualizar quais as áreas com a maior representatividade (Figura 4).

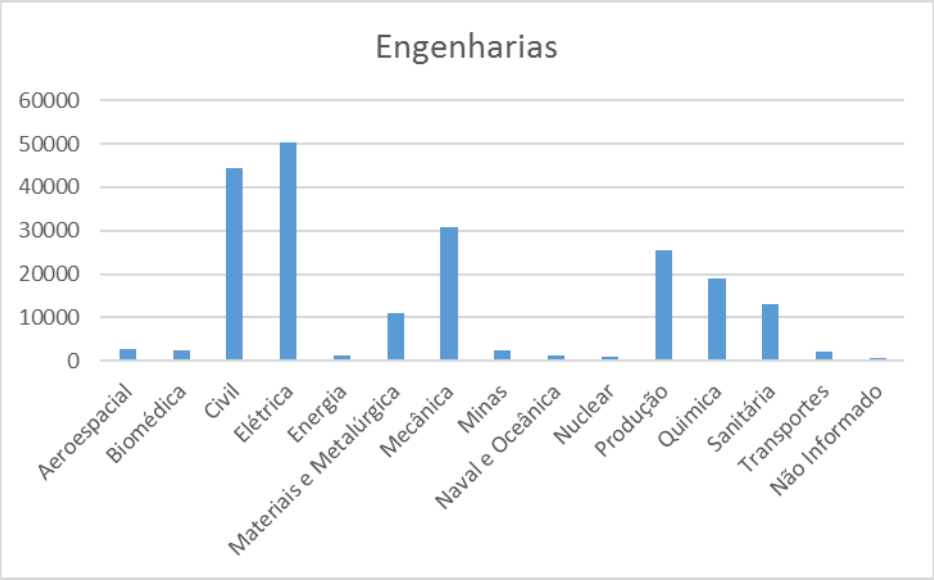

Figura 3. Distribuição dos currículos por áreas na grande área de Engenharias

Como pode ser observado, a área de Engenharia Elétrica é detentora de aproximadamente de 24% dos currículos, seguida pela Engenharia Civil e Engenharia Mecânica. Consequentemente é de se esperar que tais áreas também sejam responsáveis pela maior quantidade de artigos publicados e dessa forma, tenham influência direta nas principais palavras-chave. Ressalta-se ainda que apenas 0,36% dos currículos vinculados a grande área de Engenharias não especificaram área de atuação. Apesar da maioria dos currículos terem sido atualizados a pelo menos dois anos (76%), percebe-se uma grande quantidade não possuem publicações científicas cadastradas (Tabela 3.)

Tabela 3: Currículos sem produção científica informada

Grande Área	Total de currículos	Artigos em Anais de Congresso		Artigos em Periódicos	
		Sem Produção	%	Sem Produção	%
Engenharias	208.079	153.833	73,93	177.538	85,32

Ao verificar os currículos com produção científica cadastrada, é possível identificar quais as principais palavras-chave utilizadas e que representam os principais tópicos de estudos (Tabela 4). O resultado apresentado teve como critério de ranqueamento a frequência das principais palavras.

Tabela 4: Principais tópicos de estudos utilizados pelos indivíduos que atuam nas Engenharias

Palavra-chave	Frequência
Simulação	1.713
Ergonomia	1.712
Otimização	1.670
Meio Ambiente	1.618
Qualidade	1.596
Sustentabilidade	1.479
Biodiesel	1.440
Modelagem	1.331
Geoprocessamento	1.310
Redes Neurais	1.195
Corrosão	1.158
Elementos Finitos	1.109
Arquitetura	1.091
Ensino	1.068
Sensoriamento Remoto	1.032
Reciclagem	1.031
Design	1.006
Inteligência Artificial	993
Concreto	960
Irrigação	938

Percebe-se que pelo conjunto das 20 palavras-chave mais frequentes, várias estão relacionadas as áreas de atuação com a maior quantidade de indivíduos com currículos Lattes cadastrados, como por exemplo: Meio Ambiente, preocupação frequente da Engenharia Civil; Ergonomia muito referenciado por quem atua na Engenharia de Produção; como também Arquitetura frequentemente investigada por pesquisadores da Engenharia Civil. No entanto, destaca-se palavras-chave que são temas de investigação genéricos que frequentemente são aplicados em trabalhos de áreas distintas como é o caso de Simulação, Otimização, Qualidade, Modelagem e Redes Neurais.

Diante disso, observa-se que dentre os tópicos mais frequentes ao considerar as palavras-chave utilizadas nas publicações registradas nos currículos Lattes de indivíduos que têm atuado nas áreas de engenharias, estes sofrem influência direta de temas vinculados as áreas com a maior quantidade de currículos, bem como tópicos adotados por pesquisadores de

diversas áreas. Logo, análises que considerem outros critérios além da frequência, como as baseadas em métricas de análises de redes sociais podem proporcionar maior precisão na identificação dos tópicos mais relevantes ou impactantes no conjunto analisados.

5 CONCLUSÃO

Tendo em vista o estudo aqui realizado, ressalta-se que este tipo de análise se caracteriza como importante mecanismo, pois possibilita identificar quais os tópicos de investigação mais impactantes dentre uma comunidade de pesquisa. Ao analisar as palavras-chave dos currículos Lattes é possível considerar publicações realizadas em anais de congressos, o que não seria factível verificar em outras fontes de dados internacionais. Com isso, é possível obter uma visão precisa dos tópicos mais investigados.

Percebe-se que dentre a análise de frequência das palavras, esta sofre influência das áreas mais representativas, bem como, dos tópicos adotados em diversas áreas. Com isso, espera-se que como trabalhos futuros, sejam incorporadas análises que considerem fatores temporais para determinar a relevância de um tópico, e ainda, análises baseadas em métricas de redes sociais para determinar aquelas palavras mais centrais e consequentemente, com maior grau de importância.

AGRADECIMENTOS

Os autores agradecem ao CEFET-MG e a FAPEMIG pelo auxílio na pesquisa.

REFERÊNCIAS

Brito, A.G.C., Quoniam, L. & Mena-Chalco, J. P., 2016. Exploração da Plataforma Lattes por assunto: proposta de metodologia. *Transinformação*, vol.28, n.1, pp. 77-86. Disponível em: http://dx.doi.org/10.1590/2318-08892016002800006

Carneiro, D.M., 2003. *C&T em prol da cidadania*. Fapemig. Disponível em: http://memoria.ebc.com.br/agenciabrasil/noticia/2003-10-03/ct-em-prol-da-cidadania

Castells, M., 2010. *A sociedade em rede.* Paz e Terra.

Cataldi, M., Caro, L. D. & Schifanella, C., 2010. Emerging topic detection on Twitter based on temporal and social terms evaluation. *In Proceedings of the Tenth International Workshop on Multimedia Data Mining*, pp. 1-10.

Choi, J., Yi, S. & Lee, K.C., 2011. Analysis of keyword networks in MIS research and implications for predicting knowledge evolution. *Information & Management*, vol. 48, n.8, pp. 371-381. Disponível em: http://www.sciencedirect.com/science/article/pii/S0378720611000784

Dias, T. M. R., 2016. *Um estudo da produção científica brasileira a partir de dados da Plataforma Lattes*. Tese de Doutorado, Pós-Graduação em Modelagem Matemática e Computacional, Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG).

Dias, T. M. R.; Moita, G. F.; Dias, P. M; Moreira, T. H., 2014. Identificação e caracterização de redes científicas de dados currículares. *Revista Brasileira de Sistemas de Informação*, vol.7, n.3, pp. 5-18.

Digiampietri, L.A., 2015. *Análise da rede social brasileira*. Tese Livre Docência, Escola de Artes, Ciências e Humanidade, Universidade de São Paulo (USP).

Ferraz, R.R.N., Quoniam, L. & Maccari, E.A., 2014. The use of ScriptLattes tool for extraction and on line availability of academic production from a departament of stricto sensu in management. *In 11th International Conference on Information Systems and Technology Management* (CONTECSI), pp. 663-679.

Khan, G.F. & Wood, J., 2015. Information technology management domain: emerging themes and keyword analysis. *Scientometrics*, vol. 105, n. 2, pp. 959-972. Disponível em: http://link.springer.com/article/10.1007/s11192-015-1712-5

Medeiros, C. B. & Mena-Chalco, J. P., 2013. The dynamics of multidisciplinary research networks - mining a public repository of scientists CVs. *In: Council, I. S. S. (Ed.). World Social Science Forum*, pp. 1–17.

Mena-Chalco, J. P., Digiampietri, L.A., Lopes, F.M. & Junior, R.M.C., 2014. Brazilian bibliometric coauthorship networks. *Journal of the Association for Information Science and Technology*, vol. 65, n. 7, pp. 1424-1445. Disponível em: http://dx.doi.org/10.1002/asi.23010

Pritchard, A., 1969. Statistical bibliography or bibliometrics? *Journal of Documentation*, vol. 25, n. 4, pp. 348-349.

Saes, S.G., 2005. *Aplicação de métodos bibliométricos e da co-word analysis na avaliação da literatura científica brasileira em ciências da saúde de 1990 a 2002*. Tese de Doutorado, Pós-Graduação em Saúde Pública, Universidade de São Paulo (USP).

Souza, J. Júnior, M.L.M., Brito, A.V. & Duarte, A.N., 2014. Análise de redes de palavras baseada em títulos extraídos de um sistema de atendimento. *In III Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*, pp. 87-96

Trucolo, C.C. & Digiampietri, L.A., 2014. Análise de tendências da produção científica nacional da área de Ciência da Computação. *Revista de Sistemas de Informação da FSMA*, vol. 14, pp. 2–10.

Yi, S. & Choi, J., 2012. The organization of scientific knowledge: the structural characteristics of keyword networks. *Scientometrics*, vol. 90, n. 3, pp. 1016-1026. Disponível em: http://dx.doi.org/ 10.1007/s11192-011-0560-1

Zhu, D., Wang, D., Hassan, S., Haddawy, P., 2013. Small-world phenomenon of keywords network based on complex network. *Scientometrics*, vol. 97, n. 2, pp. 435-442. Disponível em: http://dx.doi.org/10.1007/s11192-013-1019-3

.