



UMA VISÃO GERAL DO PROCESSO DE ORIENTAÇÃO E DA PRODUÇÃO CIENTÍFICA DA RELAÇÃO ORIENTADOR- ORIENTADO NAS ÁREAS DE ENGENHARIAS

Moreira, T. H. J.

tales.info@gmail.com

Dias, T. M. R.

thiago@div.cefetmg.br

Moita, G. F.

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 relação orientador-orientado, além da evidente propagação de conhecimento, pode proporcionar um aumento na produção científica dos orientadores. Especificamente em programas de pós-graduação, onde orientados submetem diversos trabalhos em diferentes meios de publicação, estes acabam por impulsionar a produção de seu orientador, já que em geral os orientadores surgem como coautores, responsáveis por instruir e auxiliar no desenvolvimento do trabalho. Portanto, visualizar o processo de orientação e a produção científica resultado desta relação, é mais um importante meio de análise da colaboração científica nas diversas áreas do conhecimento. Neste trabalho, são utilizados os dados de orientações e supervisões de pós-graduação provenientes dos currículos Lattes, dos principais orientadores que atuam na grande área Engenharias, com o objetivo de se obter uma visão geral do processo de orientação deste conjunto, e ainda, produzir árvores genealógicas acadêmicas, onde é possível verificar como o conhecimento tem se propagado nas diversas áreas da engenharia.*

Palavras-chave: *Orientação, Genealogia Acadêmica, Plataforma Lattes*

1 INTRODUÇÃO

A evolução das pesquisas científicas tem forte influência do processo de formação, onde pesquisadores orientadores inserem novos pesquisadores que contribuem para que novos estudos sejam realizados em diversas áreas do conhecimento. Grande parte dos trabalhos realizados com orientação no Brasil são decorrentes de Programas de Pós-Graduação (PPGs), impulsionados pela necessidade de capacitação e titulação de docentes e pesquisadores.

Para Ferreira *et al.* (2009), o binômio ou díade orientador-orientado é indubitavelmente a base dos PPGs, o que determina o crescimento e a expansão dos cursos de Pós-Graduação (PG) e a demanda de orientação. Além disso, os autores ressaltam que o aluno de PG é um pesquisador em potencial, em estágio avançado de desenvolvimento, ou seja, a caminho da autonomia científica, mas ainda dependente de um professor, o que justifica as atividades de orientação como efetivamente necessárias.

Adicionalmente aos dados básicos que caracterizam os vínculos sobre o processo de orientação, informações como publicações, área de atuação e projetos de pesquisa podem ser extraídos de todos os elementos que compõem a rede a ser analisada com o intuito de se obter dados quantitativos que representam a difusão do conhecimento a partir de um orientador em particular. Tendo em vista as possibilidades de visualização e entendimento do histórico de orientação e conseqüentemente a difusão do conhecimento de um determinado orientador, realizar a modelagem e caracterização de árvores genealógicas acadêmicas surge como uma alternativa interessante para a análise de como a ciência se propaga nas várias áreas de estudo.

Para isso, dados sobre orientações bem como de características individuais de cada um dos indivíduos que compõem a rede são analisados. Neste trabalho, os dados utilizados são os currículos que compõem a Plataforma Lattes sob a coordenação do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

Neste caso, os currículos Lattes dos principais orientadores que atuam nas áreas de Engenharias são processados com o objetivo de se obter uma visão geral do processo de orientação deste conjunto, e ainda, produzir árvores genealógicas acadêmicas, integradas com informações de produção bibliográfica, onde é possível verificar como o conhecimento tem se propagado nas diversas áreas da engenharia.

O objetivo deste trabalho é aplicar a partir de toda a base de currículos Lattes previamente extraída, uma busca, a fim de encontrar os pesquisadores mais produtivos das áreas de Engenharias com orientações em pós-graduação e, conseqüentemente, todos os seus descendentes. Dessa forma, é possível obter toda a genealogia acadêmica a partir dos pesquisadores em questão.

Posteriormente, as árvores geradas podem ser manipuladas com o intuito de melhorar a visualização dos elementos que a compõem e servirem de subsídios para análises diversas que visam compreender como a relação orientador-orientado tem acontecido na grande área de Engenharias. Diante de tais estudos, e ainda, tendo em vista a quantidade de dados sobre o processo de orientação registrados nos currículos Lattes, este trabalho tem como principal fonte de dados, os registros de orientação e formação acadêmica dos currículos analisados.

2 TRABALHOS RELACIONADOS

Para Sugimoto (2014), genealogia acadêmica é o estudo quantitativo da herança intelectual adquirido através do relacionamento orientador-orientado. Logo, compreender tal relação pode proporcionar mecanismos para analisar como as orientações estão sendo conduzidas. Diante disso, este trabalho propõe uma abordagem baseada em árvores para a modelagem da genealogia acadêmica de pesquisadores orientadores, possibilitando visualizar todo o histórico de orientação de forma hierárquica, e conseqüentemente como o seu conhecimento tem se propagado. Com isso, diversos outros trabalhos podem ser propostos objetivando analisar as árvores genealógicas caracterizadas.

Árvores genealógicas acadêmicas são caracterizadas como árvores que representam hierarquicamente o histórico de um orientador e todos os seus orientados. Logo, caracterizando uma rede genealógica acadêmica é possível observar todo o histórico de um determinado orientador e como o seu conhecimento foi repassado ao longo do tempo (Moreira *et al.*, 2014).

Tuesta *et al.* (2012) apresentam uma análise temporal da relação orientador-orientado, com um estudo de caso sobre a produtividade dos pesquisadores doutores da área de Ciência da Computação extraindo os dados de análise dos currículos Lattes. No trabalho, são analisadas as principais características do grupo e as relações de coautoria.

Em Leite Filho e Martins (2006), são verificadas as influências da relação orientador-orientado no processo de produção de teses e dissertações dos PPGs em Contabilidade da cidade de São Paulo. Os autores citam como justificativas para o estudo a importância de se analisar aspectos que teriam ligação com a construção do conhecimento, especificamente em se tratando da área de Contabilidade e, a tentativa de sinalizar a importância da temática orientação.

Diversos outros trabalhos têm explorado a Plataforma Lattes como principal fonte de informações para caracterização e análise sobre orientações acadêmicas (Miyahara *et al.*, 2011; Dias *et al.*, 2014; Moreira *et al.*, 2014; Delgado *et al.*, 2015; Rossi e Mena-Chalco, 2015). No entanto, a grande maioria dos trabalhos citados anteriormente atuam em pequenos conjuntos de dados previamente definidos, obtidos manualmente ou de forma limitada.

3 METODOLOGIA

Este trabalho tem como principal fonte de dados os currículos da Plataforma Lattes. A Plataforma Lattes representa a experiência do CNPq na integração de bases de dados de Currículos, de Grupos de pesquisa e de Instituições em um único Sistema de Informações, se tornando um padrão nacional no registro da vida pregressa e atual dos estudantes e pesquisadores do país (Cnpq, 2016).

O Currículo Lattes se tornou um padrão nacional no registro da vida pregressa e atual dos estudantes e pesquisadores do país, e é hoje adotado pela maioria das instituições de fomento, universidades e institutos de pesquisa do País. Por sua riqueza de informações e sua crescente confiabilidade e abrangência, se tornou elemento indispensável e compulsório à análise de mérito e competência dos pleitos de financiamentos na área de ciência e tecnologia (CNPQ, 2015).

Diversos outros trabalhos para análise de dados científicos têm explorado a Plataforma Lattes como principal fonte de informações (Alves *et al.*, 2011b; c; a; Digiampietri *et al.*, 2012; Mena-Chalco *et al.*, 2012; Dias e Moita, 2013; Dias *et al.*, 2013; Mena-Chalco e Cesar-Jr, 2013).

Apesar de os dados dos currículos da Plataforma Lattes serem disponibilizados livremente, esses são visualizados através de interface de consulta que apresenta os currículos individualmente. Diante disso, técnicas e ferramentas para a extração e integração dos dados com outras bases de dados científicos para complementação das informações se fazem necessárias (Dias *et al.*, 2013). Para a extração dos currículos, foi utilizado o framework de extração e integração de dados científicos, desenvolvido por Dias *et al.* (2013) para obtenção dos dados a serem analisados, que os armazena em formato XML (Extensible Markup Language) conforme figura 1.

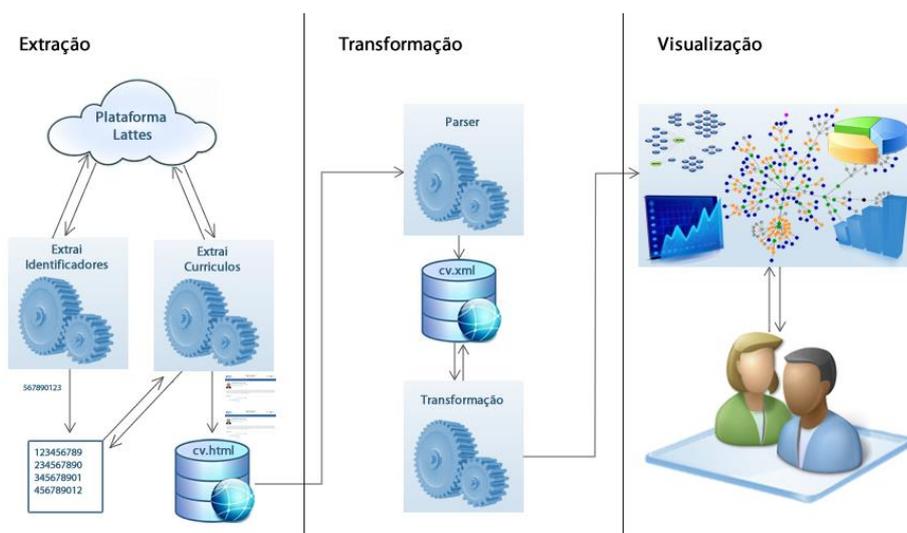


Figure 1: Arquitetura do framework de Extração e Integração de Dados Científicos.

De posse dos dados, o módulo de seleção é aplicado afim de se obter uma lista com todos os pesquisadores da grande área de Engenharias. Tais pesquisadores são ordenados de acordo com a quantidade de orientações de forma decrescente, retornando assim os pesquisadores deste conjunto que mais orientaram.

Para a construção das árvores, uma estratégia recursiva é utilizada, na qual uma raiz é previamente selecionada, fazendo a identificação de todas as orientações e supervisões concluídas, obtendo uma lista de orientados que estão diretamente vinculados ao nó raiz. Este processo é repetido para cada item da nova lista, identificando assim os vários níveis de descendentes do nó raiz. Consequentemente, currículos que não possuem orientações concluídas vão se tornando nós folhas da árvore. Esse processo acontece até que todos os descendentes sejam processados, encontrando por fim apenas nós folhas.

Portanto, cada currículo encontrado pelo módulo de seleção passa a ser o nó raiz a cada iteração, gerando uma nova árvore caso o mesmo não tenha sido inserido na árvore de outro pesquisador, neste caso, não seja um descendente de outro pesquisador da área.

De posse desses dados, também foram obtidos os trabalhos realizados em conjunto, a partir da relação orientador-orientado. Para a identificação da produção do orientador proveniente de colaboração com o aluno orientado, foi utilizado o módulo de Identificação e Caracterização proposto por Dias e Moita (2015). Utilizando o referido módulo, os dados

foram processados, produzindo como resultado a rede de colaboração científica do orientador. Tais resultados consistem em nós e arestas, que possibilitam a análise da rede caracterizada com técnicas de análise de redes sociais, possibilitando obter conhecimento sobre como acontece a colaboração científica dos pesquisadores analisados.

4 RESULTADOS

A plataforma Lattes conta hoje com aproximadamente 4.700.000 currículos, sendo que após aplicar o módulo de seleção, foram selecionadas 146.840 orientações com a grande área de Engenharias em programas de pós-graduação. Apesar de não ser a grande área com a maior quantidade de indivíduos (257.144), esta possui uma grande quantidade de produção científica e de orientações, o que a torna um interessante objeto de estudo.

Se tratando de orientações, a grande área de Ciências Sociais Aplicadas é a grande área que mais com maior volume (Figura 2). No entanto, grande parte deste se deve a orientações em trabalhos de conclusão de curso, diferentemente das Engenharias, em que muito se orienta alunos de pós-graduação, estando a frente até mesmo da grande área de Ciências Sociais Aplicadas dada como exemplo.

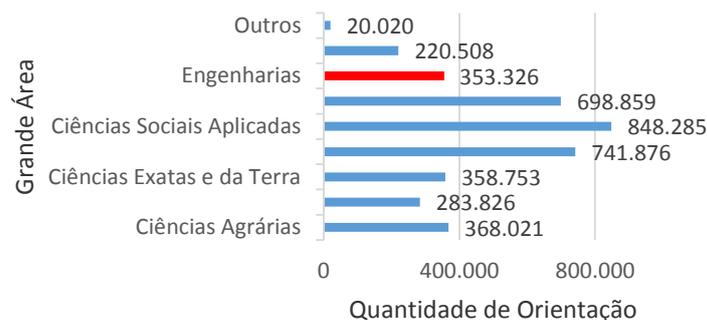


Figure 2: Orientações distribuídas por grandes áreas com destaque em Engenharias

A figura 3 demonstra a grande quantidade de orientações em cursos de Pós-graduação, que somadas ultrapassam as orientações de Graduação, reforçando a importância da área em cursos de Pós-graduação.

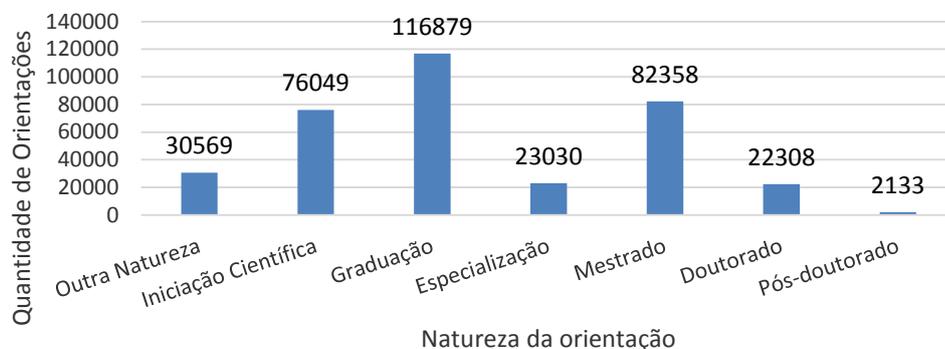


Figure 3: Natureza das orientações da grande área de Engenharias

Após a seleção dos indivíduos que atuam na grande área Engenharias, os mesmos foram ranqueados de duas formas: indivíduos que orientam em maior quantidade e as maiores árvores de orientação da grande área Engenharias (Tabelas 1, 2 respectivamente). Neste caso, a quantidade de refere-se apenas a orientações de pós-graduação.

Tabela 1: Maior quantidade de orientações

Nº	Instituição*	Quantidade
1	COPPE - UFRJ	259
2	UFRJ	246
3	COPPE - UFRJ	218
3	UFSC	198
5	UNICAMP	195
6	UFF	188
7	UFSC	181
8	UFSC	181
9	UFSC	175
10	UFRJ	171

* Instituição com maior quantidade de orientações para cada indivíduo

As maiores quantidades de orientações indicam a grande capacidade do indivíduo no processo de orientação. Porém, nem sempre isto significa que seu conhecimento está sendo repassado para outras gerações. Em casos como o do indivíduo nº 1 da tabela 1, observa-se que possui um grande volume de orientações. Inicialmente, pode-se presumir que seja uma das maiores redes encontradas. Porém, sua árvore genealógica possui cinco gerações de descendentes, onde a maior neste quesito conta com 9 gerações. Este ainda está em 7º se tratando do tamanho total da rede.

Diferentemente, o indivíduo 1 da tabela 2, apesar da pequena quantidade de orientações diretas (112) se comparado ao indivíduo anterior, teve seu conhecimento repassado ao longo do tempo e isto possibilitou que sua rede seja a maior entre todos os orientadores da grande área de engenharias, apesar da pequena quantidade de gerações (apenas quatro).

Apesar das diferenças apresentadas, pode-se notar que existe um equilíbrio entre os indivíduos, principalmente no tamanho das maiores redes de orientação. Isto se deve ao fato de que todos eles orientam em cursos de pós-graduação. Alunos de cursos de pós-graduação tendem a permanecer no meio acadêmico, resultando em novos orientadores, repassando seu conhecimento para outros alunos, e conseqüentemente, incrementando sua rede de orientações.

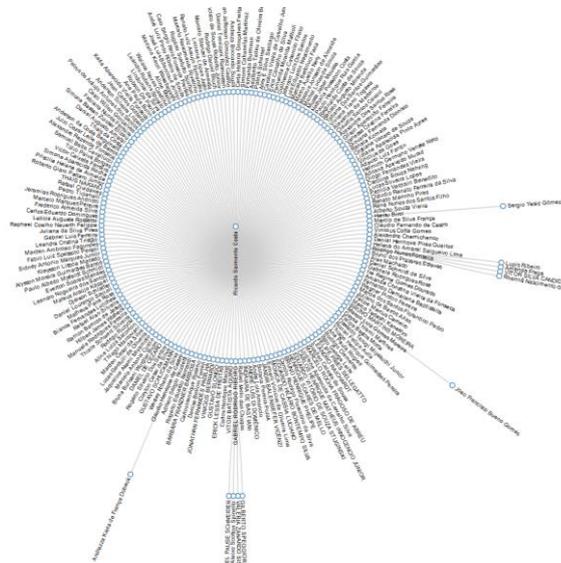
Tabela 2: Maior quantidade de descendentes (maiores redes)

Nº	Instituição*	Gerações	Tamanho
1	UFRJ	4	1846
2	ITA	6	1714
3	UNICAMP	6	1324
3	UFSC	4	1275
5	UFSC	4	1268
6	UFRJ	6	1210
7	COPPE - UFRJ	5	1208
8	UFRJ	5	1146
9	USP	6	1081
10	UFSC	5	1040

* Instituição com maior quantidade de orientações para cada indivíduo

Analisando os resultados na grande área de engenharias independentemente da natureza da orientação, foi identificado um indivíduo com orientações na instituição Fundação Getúlio Vargas.

A figura 4 representa a árvore genealógica do indivíduo citado acima (maior quantidade de orientações diretas), exibindo este como o nó central e todos os seus descendentes nos diferentes níveis da árvore. A partir deste método, diversas árvores podem ser produzidas com o intuito de representar orientadores e como seu conhecimento foi repassado ao longo do tempo.

**Figura 4: Indivíduo com maior quantidade de orientações diretas.**

Este indivíduo conta com um total de 736 orientações diretas, sendo uma quantidade mais que o dobro do 1º indivíduo da tabela 1, com maior quantidade em programas de pós-graduação. Porém, mesmo com grande quantidade de orientações, seu conhecimento pouco foi repassado para outras gerações (conta com apenas duas). Isto se deve ao fato de que suas orientações não são de alunos de pós-graduação, onde nem todos permanecem no meio acadêmico.

Diante disto, também foi produzida a árvore com a maior rede de orientações diretas da grande área Engenharias (figura 5). Esta possui 5 gerações de descendentes e 1208 nós, sendo o nó central o indivíduo n° 1 da tabela 1, também presente na tabela 2.

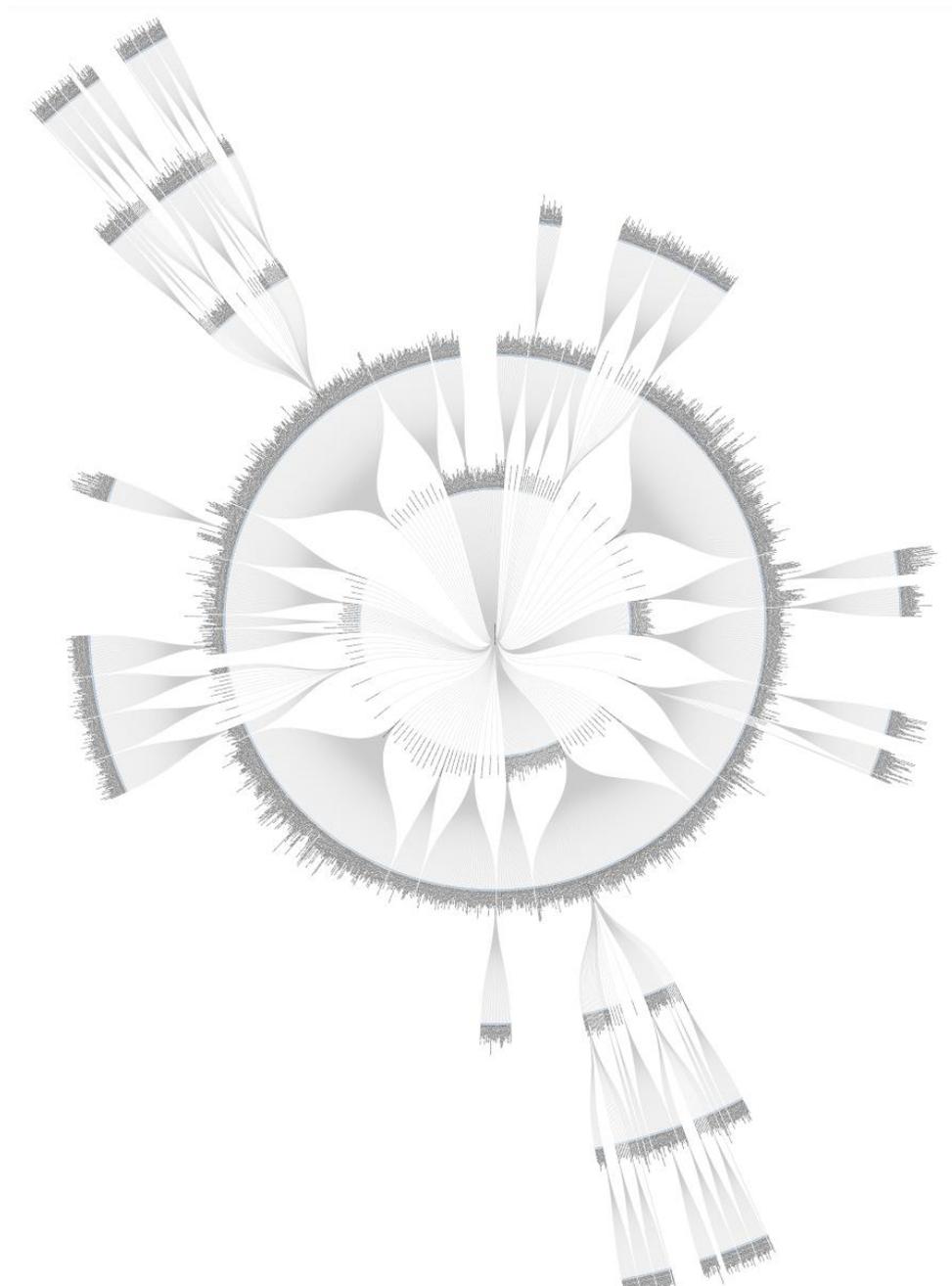


Figura 5: Rede do indivíduo com maior quantidade de descendentes diretos

Diante disto, observamos que um grande volume de orientações não resulta necessariamente em grandes redes, mas sim a qualidade e natureza das orientações.

5 CONCLUSÕES

Neste trabalho foram processados todos os currículos da plataforma, afim de se obter uma visão geral do processo de orientações acadêmicas na grande área de Engenharias.

A grande maioria destes não possuem orientações identificadas pela plataforma, sendo estes transformados em nós folhas das árvores geradas. Ou seja, não se sabe exatamente qual indivíduo da plataforma foi orientado em grande parte dos casos. Isto se deve à falta de vínculos definidos pelos usuários dentro da plataforma Lattes.

Existe uma grande dificuldade de identificar tais indivíduos, já que não existe vínculo entre este e seu orientador. Quanto maior o número de orientações, maior é a rede de colaboração e sua produção passa a ser realizada de forma mais acentuada com os orientados.

Foi possível encontrar características de determinados orientadores de acordo com as classificações aplicadas: orientadores com uma quantidade de orientações acima do normal, sendo estas em grande parte dos casos de menor importância; e orientadores com menor quantidade de orientações diretas, porém, com maiores redes e conseqüentemente, maiores quantidades de descendentes. Isto se deve principalmente a orientações em cursos de pós-graduação, onde muitos destes orientados possuem a intenção de ingressar no meio acadêmico, incrementando a rede de seus orientadores.

Foi ainda possível gerar árvores de importantes orientadores de acordo com a classificação realizada, permitindo constatar as características citadas através de um meio visual.

AGRADECIMENTOS

Os autores agradecem o apoio da FAPEMIG, bem como ao CEFET-MG, para a realização desse trabalho.

REFERÊNCIAS

ALVES, A. D.; YANASSE, H. H.; SOMA, N. Y. LattesMiner: uma linguagem de domínio específico para extração automática de informações da Plataforma Lattes. XII Workshop de Computação Aplicada - WORCAP, 2011a. p.6.

_____. Perfil dos Bolsistas PQ das Áreas de Engenharia de Produção e de Transportes do CNPq: Enfoque na Subárea de Pesquisa Operacional. XLIII Simpósio Brasileiro de Pesquisa Operacional, 2011b. p.12.

_____. SUCUPIRA: um Sistema de Extração de Informações da Plataforma Lattes para Identificação de Redes Sociais Acadêmicas. Actas da 6ª Conferência Ibérica de Sistemas e Tecnologias de Informação - CISTI, 2011c. p.371- 376.

CNPQ. Sobre a plataforma Lattes. **Conselho Nacional de Desenvolvimento Científico e Tecnológico**, 2016. Disponível em: < <http://lattes.cnpq.br/> >. Acesso em: 2016-07-18.

- DELGADO, K. V. et al. Análise Comparativa da Produtividade dos Pares Orientador-Orientado em Ciência da Computação. **Revista Eletrônica de Sistemas de Informação ISSN 1677-3071 doi: 10.5329/RESI**, v. 14, 2015.
- DIAS, T. M. R.; MOITA, G. F. Extração e Modelagem de Redes de Colaboração Científica. Conferência IADIS Ibero-Americana WWW/Internet, 2013.
- DIAS, T. M. R. et al. Identificação e Caracterização de Redes Científicas de Dados Curriculares. **iSys-Revista Brasileira de Sistemas de Informação**, v. 7, n. 3, p. 5-18, 2014. ISSN 1984-2902.
- DIAS, T. M. R. et al. Modelagem e Caracterização de Redes Científicas: Um Estudo Sobre a Plataforma Lattes. BRASNAM - II Brazilian Workshop on Social Network Analysis and Mining, 2013. p.10-20.
- DIGIAMPIETRI, L. A. et al. Minerando e Caracterizando Dados de Currículos Lattes. I Brazilian Workshop on Social Network Analysis and Mining (BrasNAM), 2012. p.12.
- FERREIRA, L. M.; FURTADO, F.; SILVEIRA, T. S. Relação Orientador-Orientando. O Conhecimento Multiplicador. **Acta Cirúrgica Brasileira**, v. 24, p. 170-172, 2009.
- LEITE FILHO, G. A.; MARTINS, G. D. A. Relação orientador-orientando e suas influências na elaboração de teses e dissertações. **Revista de Administração de Empresas**, v. 46, p. 99-109, 2006.
- MENA-CHALCO, J. P.; CESAR-JR, R. M. Prospecção de dados acadêmicos de currículos Lattes através de Scriptlattes. **Hayashi, MCPI; Mugnaini, RM; Hayashi, CR M.(Orgs.). Bibliometria e Cientometria: reflexões teóricas e interfaces. São Carlos: Pedro e João**, p. 109-128, 2013.
- MENA-CHALCO, J. P.; DIGIAMPIETRI, L. A.; CESAR-JR, R. M. Caracterizando as redes de coautoria de currículos Lattes. BraSNAM - Brazilian Workshop on Social Network Analysis and Mining, 2012. p.12.
- MIYAHARA, E. K.; MENA-CHALCO, J. P.; CESAR-JR, R. M. **Genealogia Acadêmica Lattes**. Instituto de Matemática e Estatística. São Paulo. 2011
- MOREIRA, T. H. J. et al. Genealogia Científica: Uma Análise Hierárquica De Pesquisadores Orientadores. Proceedings of the XI SIMMEC (Simpósio de Mecânica Computacional), Juiz de Fora, MG, Brazil, 2014. p.9.
- ROSSI, L.; MENA-CHALCO, J. P. Índice-h genealógico expandido: Uma medida de impacto em grafos de orientação acadêmica. IV Brazilian Workshop on Social Network Analysis and Mining (BraSNAM 2015), 2015. p.12.
- SUGIMOTO, C. R. **Academic genealogy**. MIT Press, 2014. 365-382 ISBN 9780262323277.
- TUESTA, E. F. et al. Análise temporal da relação orientador-orientado: um estudo de caso sobre a produtividade dos pesquisadores doutores da área de Ciência da Computação. Proceedings of the Brazilian Workshop on Social Network Analysis and Mining (BraSNAM), 2012. p.11.