

Adoção de Técnicas de Análises de Redes para Compreensão do Processo de Capacitação no Brasil

Adoption of Network Analysis Techniques to Understand the Training Process in Brazil

Higor Alexandre Duarte Mascarenhas ¹, Thiago Magela Rodrigues Dias ², Patrícia Mascarenhas Dias ³

Como citar (APA): Mascarenhas, H.A.D.; Dias, T.M.R.; Dias, P.M. (2020). Adoção de Técnicas de Análises de Redes para Compreensão do Processo de Capacitação no Brasil. AWARI; 1(1). <https://doi.org/10.47909/awari.63>

Recebido: 16-04-2020

Acceptado: 20-06-2020

Editor: Adilson Luiz Pinto

Copyright : © 2020 Mascarenhas. This is an open access article distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming and building upon the material as long as the license terms are followed.

RESUMEN

A migração de brasileiros tem se tornado cada vez mais frequente nos dias atuais, com a principal finalidade obter melhores condições de vida. Estudos apontam que um dos principais motivos para a migração é a busca por capacitação em alto nível de formação. Logo, neste cenário, esta pesquisa tem como principal objetivo analisar o êxodo de estudantes brasileiros durante o seu processo de formação acadêmica, a partir de dados extraídos de seus currículos cadastrados na Plataforma Lattes com a adoção de técnicas de análises de redes. Utilizou-se da Plataforma Lattes por se referir a um dos principais repositórios acadêmicos brasileiros, e por possuir informações relevantes para esta pesquisa. Sendo assim, foi utilizado o framework LattesDataXplorer para a extração e tratamento dos dados, posteriormente, selecionou-se o conjunto de dados de indivíduos com doutorado concluído por se tratar de indivíduos com mais alto nível de formação e que mantem uma atualização constante de seus currículos. Feito isso, foi efetuado um enriquecimento de dados com geolocalização e informações das instituições onde se capacitaram, a fim de obter resultados de distancias percorridas pelos doutores. Como forma de visualização de dados, foi utilizado análise de redes, e utilização de métricas para obter uma visão geral de como ocorre o êxodo científico brasileiro. Percebe-se uma alta concentração de doutores em cidades com maior concentração de universidades que possuem programas de pós-graduação em nível de mestrado e doutorado, bem como, se caracterizam por possuir

1 CEFET-MG, Brasil. Email: higoralexandre1996@gmail.com, ORCID: 0000-0001-6057-3888.

2 CEFET-MG, Brasil. Email: thiagomagela@cefetmg.br, ORCID: 0000-0001-5057-9936.

3 CEFET-MG, Brasil. Email: patricia.dias@uemg.br, ORCID: 0000-0002-8448-6874.

maiores rendas per capita.

Palavras-chave: Análise de Redes; Plataforma Lattes; Êxodo Científico Brasileiro

ABSTRACT

The migration of Brazilians has become more and more frequent nowadays, with the main purpose of obtaining better living conditions. Studies indicate that one of the main reasons for migration is the search for training at a high level of training. Therefore, in this scenario, this research has as main objective to analyze the exodus of Brazilian students during their academic formation process, from data extracted from their curricula registered in the Lattes Platform with the adoption of network analysis techniques. The Lattes Platform was used for referring to one of the main Brazilian academic repositories, and for having relevant information for this research. Therefore, the LattesDataXplorer framework was used for the extraction and treatment of the data. Subsequently, the data set of individuals with a doctorate completed were selected because they are individuals with a higher level of education and who maintain a constant update of their curricula. . Once this was done, data was enriched with geolocation and information from the institutions where they trained, in order to obtain results from distances covered by doctors. As a way of visualizing data, network analysis was used, and metrics were used to obtain an overview of how the Brazilian scientific exodus occurs. A high concentration of doctors is perceived in cities with a higher concentration of universities that have postgraduate programs at the master's and doctoral level, as well as being characterized by having higher incomes per capita.

Keywords: Network Analysis; Lattes Platform; Brazilian Scientific Exodus

1 INTRODUÇÃO

A emigração de brasileiros para outros países e para outros estados tem aumentado de forma significativa, de modo que, no Brasil, estudos demonstram que em algumas cidades possuem taxa de 10 a 30% de migrantes que não vivem no seu estado de origem , em muitos casos, brasileiros saem em busca de emprego, ou estudos, visando sempre qualidade de vida.

Dentre os principais motivos para migração, um que se destaca é a necessidade de capacitação em alto nível de formação, e uma causas por optar pela mobilidade de indivíduos no território brasileiro refere-se à qualidade de ensino superior em outros estados, a busca de novas oportunidades e mais experiências em suas áreas . A migração de estudantes para outros países também tem se tornado frequente, buscando assim, intercâmbio cultural e melhor investimento em bolsas de pesquisa. A emigração do estudante para outros países é de interesse do discente e das instituições de origem, pois, o mesmo retorna na maioria das vezes mais produtivo, com rede de contato mais extensa, maior vivência, e podendo futuramente compartilhar suas experiências com outros estudantes da instituição de origem.

Pesquisas apontam que a cada dia tem se tornado mais difícil produzir pesquisa científica no Brasil, devido a cortes de investimentos destinados a bolsas. Um dos principais motivos para a emigração de pesquisadores brasileiros para outros países podem ser apontado pela falta de apoio do governo. Logo, com esse cenário pesquisadores brasileiros saem do país, dificultando assim o retorno pela falta de oportunidades. Grande parte dos cientistas brasileiros que voltam para o Brasil não conseguem emprego na sua área de formação,

fazendo assim que não progridam nas suas carreiras. Um programa que facilitou e auxiliou bastante o ingresso de estudante às instituições com sedes em outros países é o Ciência Sem Fronteiras, por se referir a um programa que ampara estudantes, oferecendo bolsas de estudos.

Diante deste cenário, este trabalho apresentará um estudo sobre o êxodo de estudantes brasileiros que partiram do seu estado/cidade de nascimento para outros estados/cidades e/ou aqueles que foram para outros países em busca de capacitação. Como fonte de dados foi utilizado o repositório de currículos da Plataforma Lattes que atualmente registra informações acadêmicas/científicas e profissionais possuindo cerca de 6.750.000 currículos cadastrados. Logo, foi possível obter um retrato de como ocorre o êxodo científico brasileiro através de análise baseadas em métricas de redes sociais, resultando em uma visão ampla, e inédita da mobilidade dos doutores brasileiros ao longo da formação acadêmica.

2 Trabalhos relacionados

em seu estudo efetuaram uma análise do fluxo migratório de indivíduos nascidos em 196 países de todos os continentes, sua pesquisa realizou a análise em meados dos anos 1990 até o ano de 2010, com a finalidade de entender padrões e tendências dos fluxos de migrações dos países e continentes selecionados. O estudo possibilitou identificar os fluxos de migrações dos indivíduos obtidos para estudo de acordo com nível de desenvolvimento dos países, sendo assim efetuou análises de 1990 a 1995, 1995 a 2000, 2000 a 2005 e 2005 a 2010, deste modo, o autor destaca que a maior porcentagem dos indivíduos de países menos desenvolvidos dos anos 1990 a 2000 optam por imigrar para países mais desenvolvidos, enquanto indivíduos de países desenvolvidos imigram para outros países desenvolvidos. Já dos anos 2000 a 2010 os indivíduos de países desenvolvidos optam por imigrar para outros países menos desenvolvidos. Os autores também fazem uma análise com relação a imigrações de indivíduos nos continentes e perceberam que os maiores movimentos ocorrem entre o sul e oeste da Ásia, da América Latina para a América do Norte e entre os países África.

No trabalho de os autores realizaram um estudo baseado em dados da Plataforma Lattes considerando indivíduos com doutorado concluído, coletando dados da trajetória do pesquisador desde o nascimento, formações acadêmicas e atuação profissional. Para efetuar análises utilizou-se como forma de visualização de dados Redes denominado Gráfico de Mobilidade. Os nós são compostos pelas cidades que possuem Registro de Nascimento (RN), Formação (RF) e Trabalho (RT). Já as arestas são compostas pelo agrupamento de Fluxos de Nascimento para a primeira Formação (FNF), de Formação para outra Formação (FFF) e da última Formação para o local de Trabalho (FFT) de cada pesquisador entre duas cidades. A partir de então devido à dificuldade para interpretar os dados os autores resolveram utilizar métricas para viabilizar a análise de mobilidade. Durante análises, chegou-se à conclusão que 95% dos indivíduos são de origem das regiões Sul, Sudeste e Nordeste. Foi mencionado que 40% das primeiras formações dos doutores foram realizadas em suas cidades de origem, e que 87% daqueles que deslocam para outras cidades não ultrapassam o limite de 1.000 km, também foi citado que que 61% das pessoas saem da cidade de última formação para atuar profissionalmente. A cidade com maior formação de doutores é São Paulo e os continentes com maior número de formação é América do Sul, América do Norte e Europa.

Já em foi efetuado um estudo com o objetivo de entender como ocorre a colaboração científica entre os principais pesquisadores do Brasil, analisando também as diversas modalidades de bolsas, levando em conta apenas os pesquisadores com bolsas de

produtividade vinculadas ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Sendo assim utilizaram como base de dados o repositório de currículos constituído na Plataforma Lattes. A partir de então, foi efetuada análise de redes para alcançar o objetivo do estudo. Obtiveram como resultados que os indivíduos pertencentes a principal modalidade de bolsa (1A) tendem a publicar de forma igualitária em anais de congresso e periódicos, diferente das outras modalidades que em maior quantidade publicam em anais de congresso; outra observação que os autores destacam é que nas modalidades mais altas de bolsas, a colaboração ocorre de maneira mais intensa, já nos níveis inferiores, os bolsistas tendem a colaborar em menor escala entre os indivíduos com o mesmo nível de bolsa, em detrimento da colaboração com os bolsistas de modalidade de bolsas mais altas.

Em , os autores realizaram um estudo com o intuito de descrever a produção científica dos membros dos grupos de pesquisa científica das áreas de informação no Brasil, utilizaram como fonte de informação o Diretório de Grupos de Pesquisa e a Plataforma Lattes, com as informações cadastradas no período de 1992 a 2016. Para extração dos dados utilizaram a ferramenta *ScriptLattes*, sendo assim efetuaram análises estatísticas para apresentarem os resultados. Os autores concluíram que 50% da produção dos membros dos grupos de pesquisas das áreas de informação foram publicados entre 2008 e 2015, e que um terço dos grupos foram formados nos últimos 5 anos. Destacaram também que a região Sudeste se destaca das demais com maior produção e logo atrás dela, vem a região Nordeste. Com relação aos canais de publicação se destacam anais de congresso e artigos em revistas referendadas.

Com a revisão realizada, foi possível observar a viabilidade em adotar a Plataforma Lattes como fonte de dados, pela vasta quantidade de dados nela cadastrados e ainda aplicar técnicas de análises de redes como método para a visualização dos resultados.

3 METODOLOGIA

No presente trabalho a principal fonte de dados utilizada foi o repositório curricular disponível na Plataforma Lattes. Inicialmente foi necessária a utilização do LattesDataXplorer para extração dos dados, dada a dificuldade de obtê-los, uma vez que a interface de consulta aos currículos da Plataforma Lattes possibilita somente o acesso de apenas um currículo por vez, sendo assim a análise de grandes grupos de indivíduos passa a ser um fator limitante. A extração dos dados foi realizada em maio de 2019 totalizando 308.317 currículos de indivíduos com doutorado concluído, considerando todos os doutores independentemente da data de conclusão de seus doutorados.

Logo após ter sido realizada a extração dos dados, foram realizados tratamentos com o objetivo de obter extratos de dados formatados a fim de facilitar futuras análises. Assim sendo, etapas como “Seleção dos Dados” e “Tratamento dos Dados” foram efetuadas conforme esquema apresentado na Figura 1.

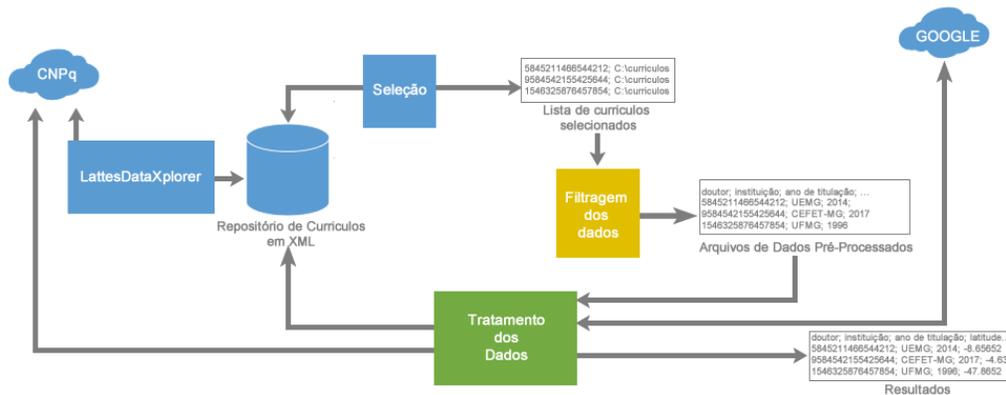


Figura 1. Aspecto geral do conjunto de componentes utilizados.

Na Etapa de “Seleção” é usada a linguagem de consulta XPath (*XML Path Language*) para pesquisa e posterior geração dos subgrupos a serem analisados. A linguagem XPath possibilita a construção de expressões que vão percorrer um documento XML de forma similar ao uso de expressões regulares. Portanto, possibilita o agrupamento de um conjunto de currículos com parâmetros desejados, como formação acadêmica ou áreas de atuação.

A lista armazena os identificadores de cada currículo e o caminho em que ele está armazenado localmente, sendo assim, será possível analisar somente os currículos selecionados. Diante do exposto foi coletado somente currículos de indivíduos com doutorado concluído, por se tratar do grupo com o maior nível de formação acadêmica; por se tratar de currículos que são frequentemente atualizados e grande parte dos parâmetros necessários para o presente trabalho estarem registrados em seus currículos.

Após a seleção do conjunto a ser analisado, o módulo de “Filtragem dos dados”, que é responsável por analisar os currículos em arquivos XML objetivando obter informações relevantes a pesquisa, caracteriza um extrato de dados formatados (Arquivos de dados pré-processados). As informações dos currículos registradas no arquivo possuem: identificador do currículo; estado e cidade de nascimento do indivíduo; código da instituição, nome e CEP do vínculo atual de atuação do indivíduo, além do código de identificação e nome da instituição para cada nível de formação concluído, considerando desde a graduação até o doutorado.

Posteriormente, é executado o módulo de “Tratamento dos dados” (Figura 2), em que são realizadas quatro etapas: Obtenção de CEP da instituição; Busca pela localização geográfica; Limpeza e agrupamento de dados e Normalização dos dados.

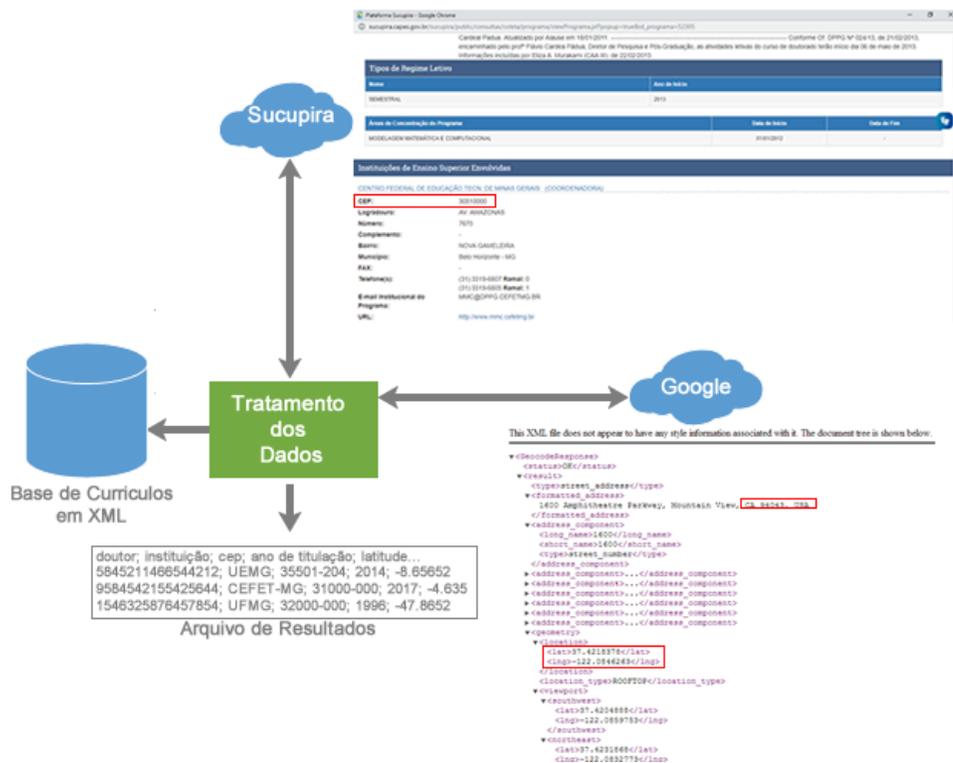


Figura 2. Tratamento dos dados.

A primeira etapa efetuada é a de "**Obtenção de CEP da instituição**" em que a partir do código da instituição recuperado no currículo, o mesmo é consultado no diretório de instituições da Plataforma Lattes no intuito de obter os dados da instituição, e dessa forma, recuperar da seção de endereço os dados de localização da instituição, as características da instituição, a partir de então, o site retornará as informações das instituições e assim ocorrerá a obtenção do CEP da instituição.

A etapa de "**busca pela localização geográfica**" é uma tarefa realizada com a finalidade de geolocalizar uma instituição. Em que ao acessar a API (*Application Programming Interface*) de geolocalização do Google, será enviado o endereço da instituição para posteriormente ter como retorno a localização geográfica (latitude e longitude) da instituição.

Na etapa de "**Limpeza e agrupamento de dados**", ocorre a realização de exclusão de possíveis termos irrelevantes para a pesquisa, com o intuito de diminuir o volume de dados a serem processados e analisados. Como exemplo: a remoção de stopWords nos nomes das cidades; a normalização para extrair palavras acentuadas, e substituí-las pelo seu equivalente sem acentuação.

Já a etapa de "**Normalização dos dados**" tem o intuito de reduzir a redundância de informações, descartando atributos com a ausência de dados, como por exemplo CEP com ausência de algarismos. Posteriormente, é gerado o "Arquivo de Resultados", representando uma sumarização de todos os dados obtidos nos currículos de doutores brasileiros, não necessitando de consultas dos arquivos XML dos currículos extraídos. Logo, após toda a execução das etapas descritas anteriormente, diversas métricas são aplicadas para compreensão sobre como tem ocorrido a mobilidade dos doutores brasileiros ao longo do seu processo de formação acadêmica.

4 RESULTADOS

4.1 Caracterização Geral

Inicialmente foi possível realizar uma caracterização do conjunto analisado. Por se tratar de um grupo de indivíduos que possuem o mais alto nível de formação, para serem considerados somente foi incluído na análise as formações com o Status de "concluído", resultando em um total de 308.317 indivíduos. No intuito de avaliar o potencial de análise dos dados extraídos, a Figura 3 apresenta a quantidade de currículos que possuem atributos cadastrados, tais como: cidade de nascimento, instituição de atuação e instituições de formação.

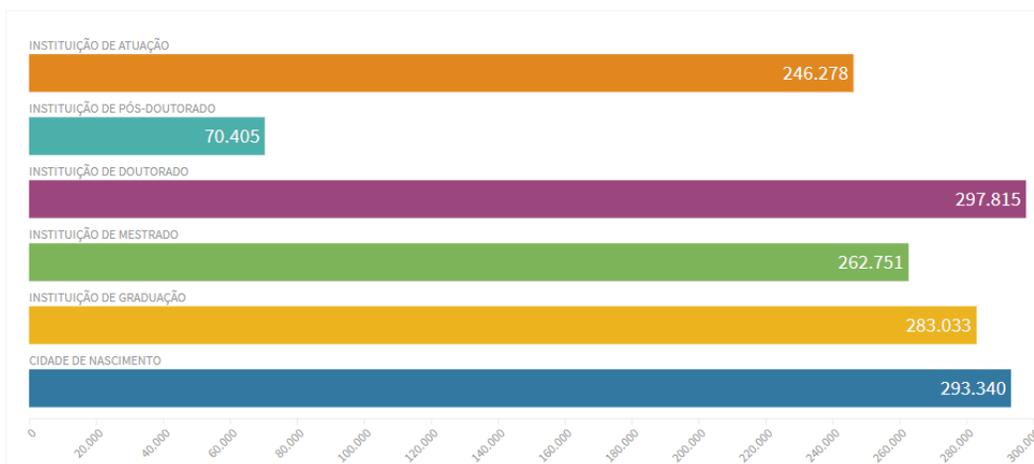


Figura 3. Quantitativo de currículos que possuem informação que são objetos de análise.

Como pode ser observado, a maioria dos indivíduos possuem a cidade de nascimento cadastrada com um total de 293.340 (95%) cadastros, por se tratar de um campo obrigatório ao ser realizado o cadastro na Plataforma Lattes. Aqueles indivíduos que não possuem cadastro de cidade de nascimento julgam-se que são currículos mais antigos, em que não era obrigatório o cadastramento da cidade. Das demais instituições demonstradas no gráfico, aquela que se identifica superior em dados quantitativos é a *instituição de conclusão do doutorado* totalizando uma quantidade de 297.815 (96%) cadastrados, por se tratar do grupo selecionado para a realização do estudo. A *instituição de pós-doutorado* é aquela que possui uma menor quantidade de cadastros, com um montante de 70.405 (22%), tal fato se justifica, pois, a seleção dos grupos foi direcionada aos indivíduos com doutorado concluído; por este motivo, nem sempre um indivíduo que possui o título de doutorado possui o pós-doutorado.

Uma informação importante a ser analisada nos currículos diz respeito as áreas de atuação dos indivíduos. Deste modo, por meio desta informação (Figura 4), é possível observar a distribuição do composto em análise por todas as Grandes Áreas de atuação informadas.

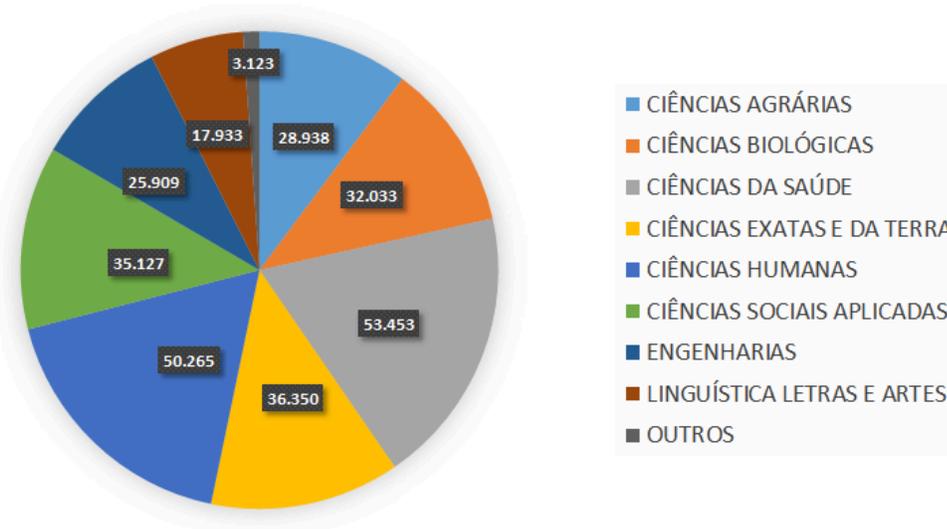


Figura 4. Distribuição de doutores por Grande Área.

É possível verificar que as Grandes Áreas que possuem mais doutores atuando são as Grandes Áreas de Ciências da Saúde (53.493) e a Grande Área de Ciências Humanas (50.265) respectivamente. A Grande Área de Ciências Humanas se justifica pelo fato de ser mais destinada para indivíduos que pretendem trabalhar na academia. A Grande Área de Linguística, Letras e Artes se caracteriza por ser a menos representativa, totalizando 17.933 currículos de doutores que atuam nesta Grande Área.

Uma vez caracterizado todo o conjunto de dados, foi possível obter um mapeamento das distâncias percorridas pelos indivíduos ao longo de seu processo de capacitação, a cada nível de formação. Na Tabela 1 é apresentado o quantitativo de registros obtidos, para ser calculada a distância média percorrida pelos indivíduos de um nível de formação acadêmica até outra.

Tabela 1. Quantidade de registros obtidos (Fonte: autores).

Quantidade	Graduação	Mestrado	Doutorado
Nascimento	239.948 (78%)	229.874 (74%)	247.866 (80%)
Graduação	-	217.875 (70%)	233.362 (75%)
Mestrado	-	-	230.271 (74%)
Total	308.314		

Quando se trata de quantidade de registros, deve ser destacada a perda quantitativa daqueles que não possibilitam a medição de um determinado nível de capacitação para seu posterior. A maior perda sofrida é dos registros de graduação-mestrado, fato que pode ser justificado pela inserção incorreta de dados de localização geográfica das instituições. Diferentemente da menor perda sofrida que são dos registros nascimento-doutorado, tal fato se explica, pois, o nascimento é obrigatório no preenchimento quando ocorre cadastro na Plataforma Lattes e também porque o doutorado foi o grupo extraído da Plataforma Lattes

para a análise da pesquisa. Tendo em vista a quantidade de registros e a distância média entre cidades que foram realizados um nível de formação até outro, foi possível apresentar na Tabela 2, o cálculo das distâncias médias em quilômetros entre um nível de formação até outro.

Tabela 2. Distância média em Km entre os níveis de formação (Fonte: autores).

Distância (Km)	Graduação	Mestrado	Doutorado
Nascimento	291,58	548,59	1.000,75
Graduação	-	432,38	901,26
Mestrado	-	-	619,00

É possível observar o resultado da distância média de todas as etapas de capacitação dos doutores brasileiros, durante sua formação acadêmica. Pode ser observado que a distância média entre as etapas tem uma variação considerada. Inicialmente, analisando a distância média do local de nascimento para graduação percebe-se que está é a menor distância calculada. Um dos fatores que influenciam tal fenômeno é que grande parte das cidades brasileiras possuem instituições que proporcionam ao estudante cursos de graduação, e aquelas que não possuem na maioria das vezes, ficam próximas à outras cidades que detêm os cursos neste nível de capacitação de interesse dos estudantes. Já as maiores distâncias estão entre o local de nascimento e de capacitação a nível de doutorado, seguindo da graduação/doutorado em que o deslocamento é maior que os outros níveis de formação. Na Figura 5, é possível observar a distribuição entre os níveis considerando a mediana das distâncias percorridas no processo de formação acadêmico dos indivíduos analisados.

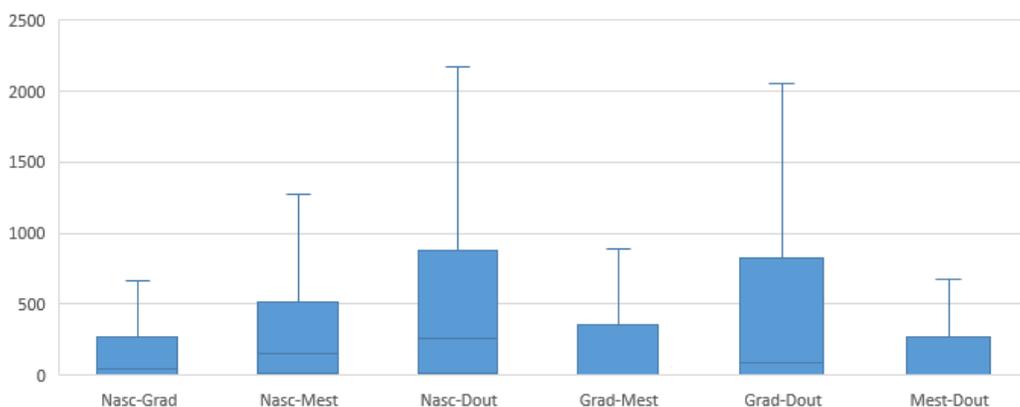


Figura 5. Distribuição das distâncias entre os níveis de formação.

Percebe-se pela distribuição que os menores valores de mediana são entre nascimento-graduação e mestrado-doutorado, pois grande parte dos indivíduos optam por obter o título de “Doutor” na mesma instituição em que foi realizado o mestrado, e aqueles que não escolhem obter os dois títulos na mesma instituição realizam em instituições mais distantes, o que contribui para que eleve a média entre estes níveis de formação. As medianas das distâncias percorridas pelos indivíduos brasileiros ao longo de sua formação acadêmica, estão separadas entre nascimento-graduação, nascimento-mestrado, nascimento-doutorado, graduação-mestrado, graduação-doutorado e mestrado-doutorado.

A maior mediana observada é a de nascimento-doutorado, por se tratar de o local de nascimento do indivíduo até o mais alto nível de formação resultando em uma representação de toda a sua vida acadêmica. Vale ressaltar também a distância mínima das medianas entre graduação-mestrado e mestrado-doutorado, isso se justifica, pois, os estudantes na maioria das vezes decidem seguir em maior parte dos casos na mesma instituição durante essas transições em seus processos formativos.

4.2 Redes de migração

Inicialmente foi possível caracterizar um retrato das migrações dos indivíduos por meio da obtenção da localidade com que os indivíduos se capacitaram a cada nível de formação. Logo, inicialmente obteve-se uma visão baseada em redes da mobilidade dos doutores visando vínculos em um âmbito nacional, durante os níveis de formação: nascimento-graduação; graduação-mestrado; mestrado-doutorado e doutorado-atuação profissional.

Na Figura 6 os nós são representados por diversas cidades do território brasileiro, levando em conta a cidade de nascimento do indivíduo e a sua cidade de graduação.

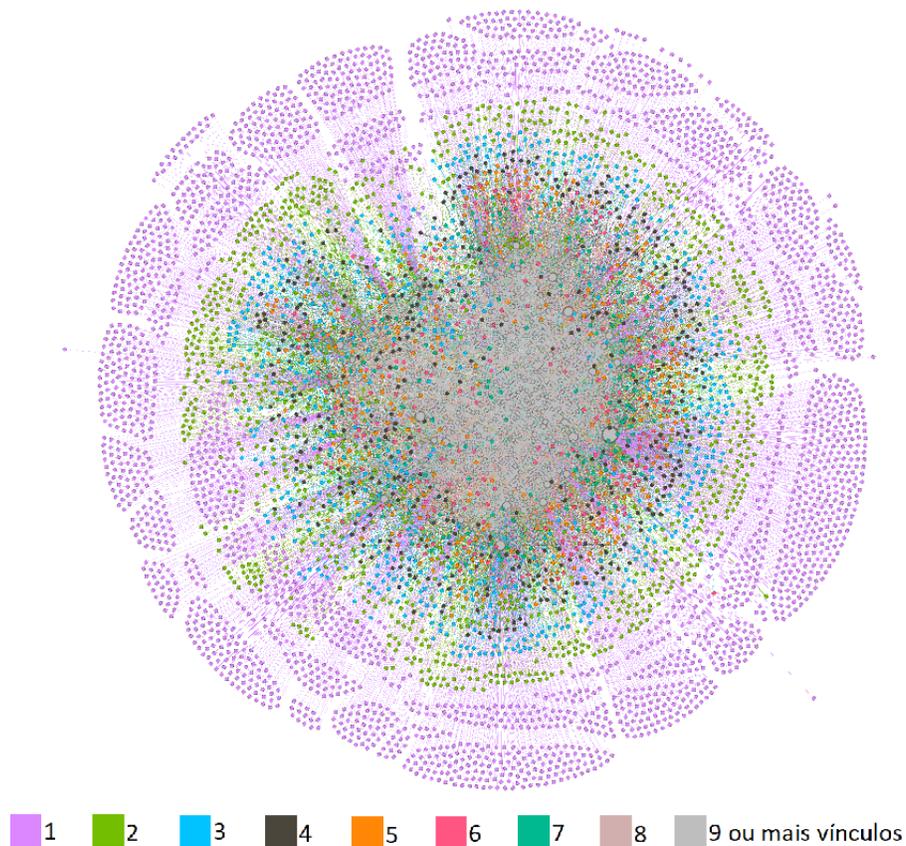


Figura 6. Rede de vínculos do nascimento para graduação.

Foi possível observar que a densidade da rede é baixa (0,001), uma vez que os nós da rede estão pouco conectados, verifica-se também que a quantidade de conexões entre os nós é baixa, dada a quantidade de nós presentes na rede. Destaca-se na rede a quantidade de nós com grau equivalente a 1, o que influencia na baixa densidade. Identificou-se que a

maioria dos vínculos entre os nós são de nós com baixo grau (que representam cidades com menor quantidade de universidades), com aquele nó com mais alto grau (cidades com maior quantidade de universidades). Vale ressaltar a formação de clusters, em que ao analisá-los foi possível perceber que representa o grupo de cidades pertencentes a uma mesma região ou cidades bem próximas umas das outras que tendem a atrair estudantes de cidades próximas.

Quando se trata dos vínculos nos níveis graduação-mestrado (Figura 7) e mestrado-doutorado (Figura 8), percebe-se que as redes são bem similares, e há uma redução da quantidade de nós, devido a menor migração já que os doutores tendem a optarem por permanecer na mesma instituição nestas etapas de capacitação.

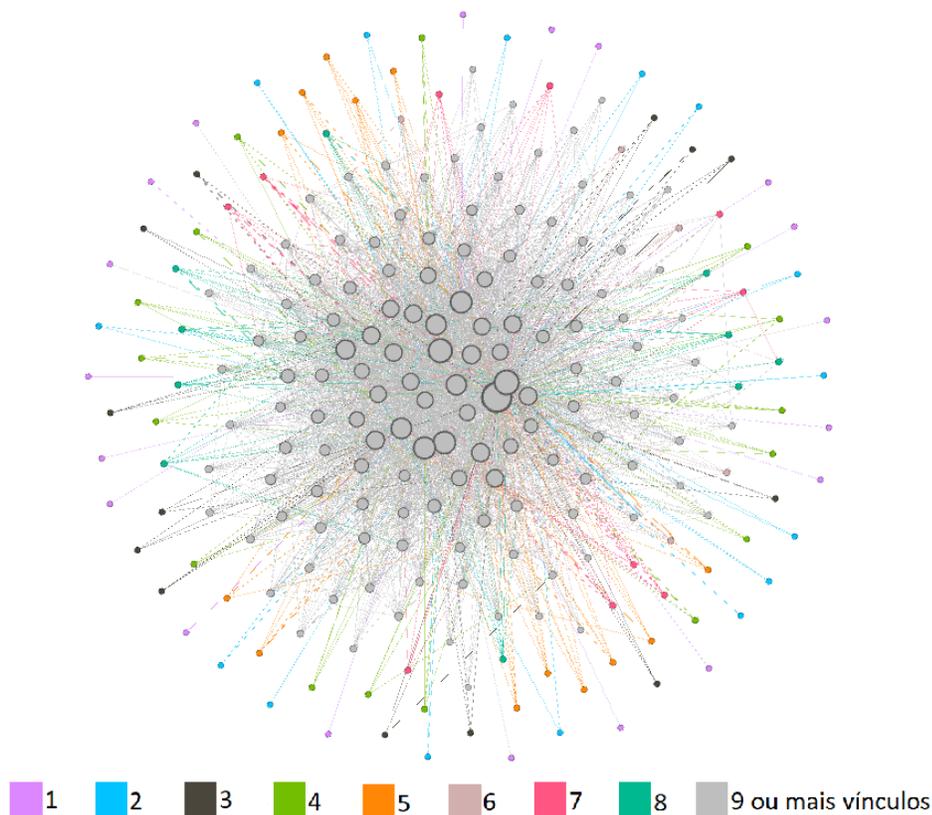


Figura 7. Rede de vínculos da graduação para mestrado.

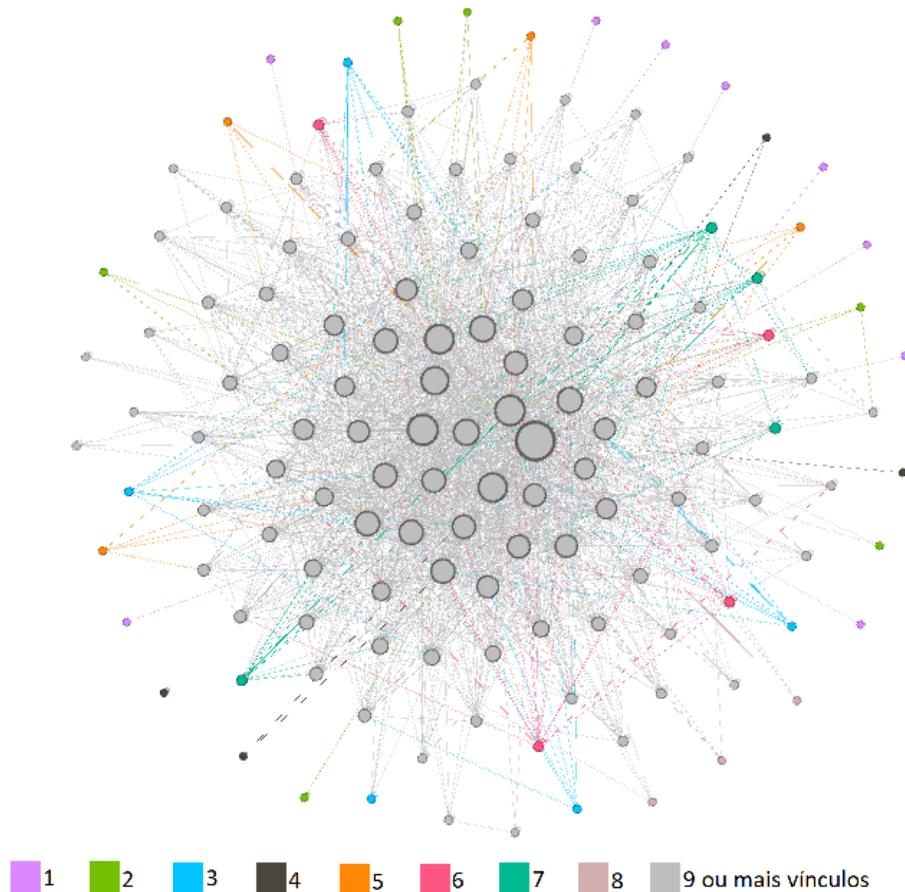


Figura 8. Rede de vínculos de mestrado para doutorado.

É possível perceber que aqueles nós com maiores graus tanto de entrada como também de saída são representados por aquelas cidades que mais ofertam cursos em nível de pós-graduação, como por exemplo São Paulo, Rio de Janeiro, Porto Alegre e Belo Horizonte.

Já na Figura 9 é possível observar maior distribuição entre os nós com base em seus respectivos graus. Destaca-se a maior quantidade de nós na rede, tendo em vista a atuação dos doutores em diversas cidades. O diâmetro da rede é igual a 5 o que nos confirma que os nós estão relativamente próximos uns dos outros, em que as cidades com menores números de vínculos, estão diretamente conectadas as cidades com maiores quantidades de vínculos.

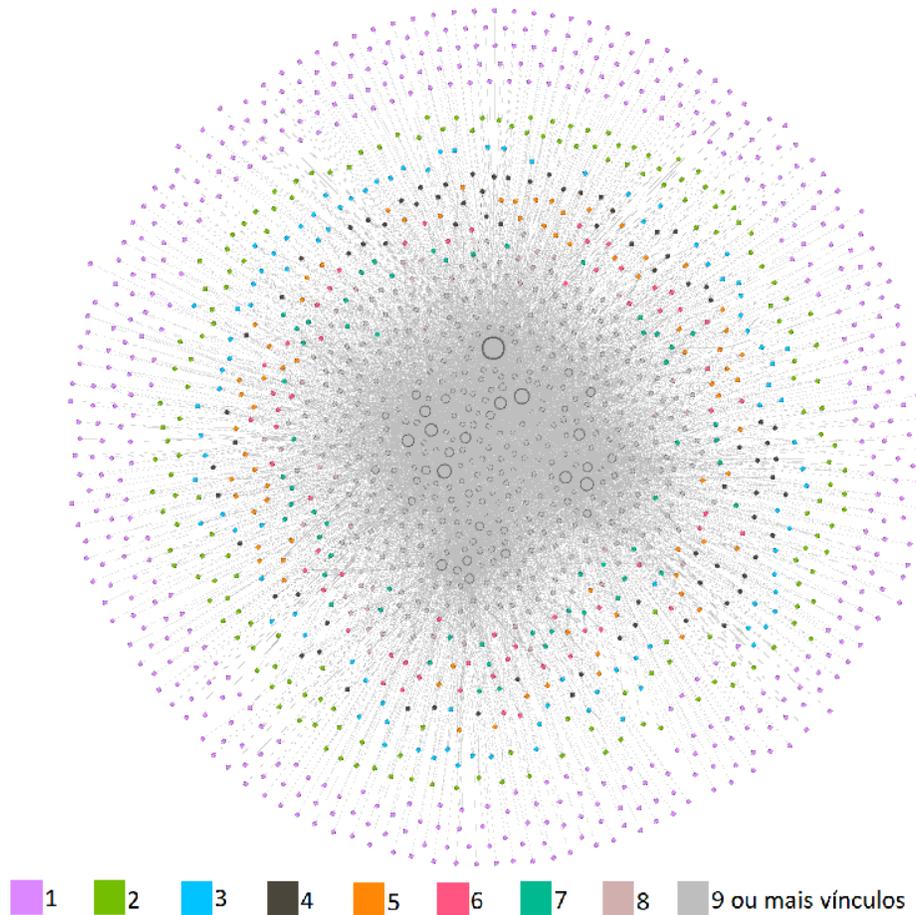


Figura 9. Rede de vínculos de doutorado para atuação profissional.

Com densidade baixa (0,005), a análise desta revela rede que diversos indivíduos concluem seus doutorados em uma cidade, e acabam por ter que atuar profissionalmente em uma outra cidade. No centro da rede um nó se destaca dos demais, este nó é representado pela cidade de São Paulo, o que novamente confirma a maior opção de ocupação dos doutores nesta cidade.

Com o intuito de melhor explorar a trajetória dos doutores, e obter dessa forma uma visualização quando se trata de origem (cidade de nascimento) e destino (cidade de atuação), foi caracterizada a rede da Figura 10.

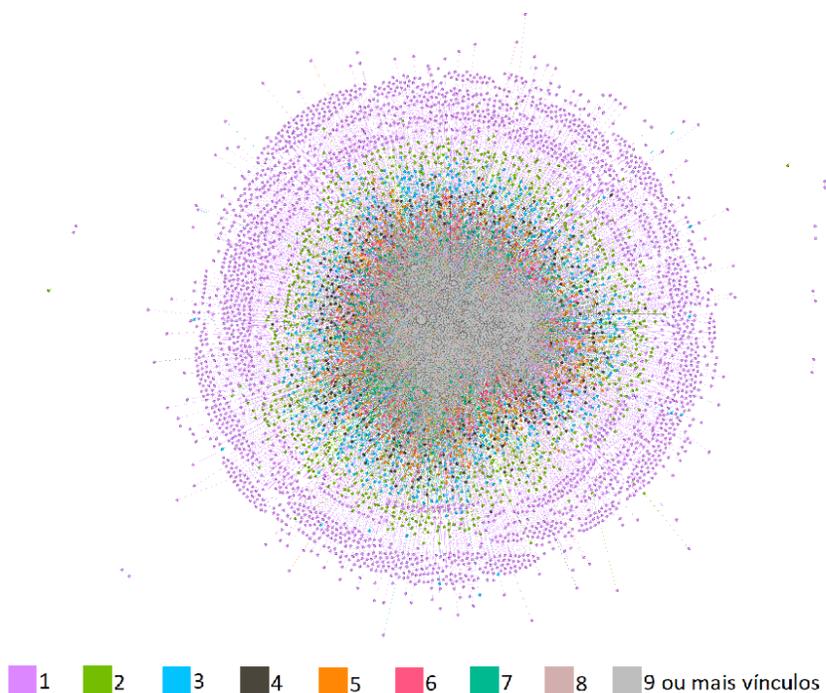


Figura 10. Rede de vínculos do nascimento para atuação profissional.

Percebe-se o quão disperso estão alguns nós, sendo que em sua maioria são díades ou até mesmo nós isolados (caracterizando o vínculo com ele mesmo). Foi possível observar que, os nós isolados são em sua grande maioria representados por cidades cadastradas com erro de digitação.

A fim de explorar o conhecimento dos resultados a partir das redes de vínculos caracterizadas, foram analisadas as principais métricas utilizadas na literatura (Tabela 3).

Tabela 3. Métricas extraídas das redes nacionais caracterizadas (Fonte: autores). **Nota:** C.G.: Componente Gigante.

Métricas	Nasc-Grad	Grad-Mest	Mest-Dout	Dout-Atuação	Nasc-Atuação
Total de Nós	6.080	237	138	1.497	6.241
Total de Arestas	27.219	4.583	3.009	10.602	53.640
Grau Médio dos Nós	4,477	19,338	21,804	7,082	8,585
Total de Nós no C.G.	6.078	237	137	1.497	6.217
% de Nós no C.G.	99,99	100	99,97	100	99,65

Métricas	Nasc-Grad	Grad-Mest	Mest-Dout	Dout-Atuação	Nasc-Atuação
Total de Arestas no C.G.	27.218	4.583	3.008	10.602	53.625
% de Arestas no C.G.	99,99	100	99,99	100	99,97
Densidade da Rede	0,001	0,082	0,159	0,005	0,001
Diâmetro da Rede	6	4	4	5	7
Caminho Mínimo médio	2,614	2,057	1,905	2,347	2,87

Representado aqueles vínculos que possuem maiores quantidade de nós pertencentes, e conseqüentemente uma maior quantidade de arestas, os vínculos de nascimento-graduação e nascimento-atuação se destacam, devido se tratar da localidade de nascimento do indivíduo, em que estão espalhados em diversas cidades, e também a atuação devido possuir ofertas de empregos em cidades diversas. Com menores quantidade de nós, porém superiores aos relacionamentos de graduação-mestrado e mestrado-doutorado.

Ao ser analisado o grau médio dos nós, destaca-se os vínculos entre graduação-mestrado e mestrado-doutorado, uma vez que possuem menores quantidades de nós, o que possivelmente é causado pela menor oferta de cursos de pós-graduação nas cidades que não estão em grandes centros, o que ocasiona no aumento da taxa de densidade destas duas redes comparadas com as demais. As redes caracterizadas pelos vínculos de graduação-mestrado e mestrado-doutorado, também se destacam por possuírem um menor diâmetro de rede (4).

Ao se analisar as Componentes Gigantes, com as redes globais, observa-se que todas as redes possuem taxas bastante semelhantes, e que a quantidade de nós e arestas na componente gigante também são próximos a da rede global.

Ao verificar o caminho mínimo médio, é possível observar que os valores dos níveis graduação-mestrado e mestrado-doutorado são inferiores aos outros, resultado de uma menor quantidade de vínculos e conseqüentemente interligação entre as cidades.

5 CONSIDERAÇÕES

A partir dos resultados obtidos foi possível verificar a viabilidade em adotar os currículos cadastrados na Plataforma Lattes como fonte de dados para análises sobre como ocorre o Êxodo Científico brasileiro. A escolha do grupo de doutores se caracteriza como uma parcela significativa de todo o conjunto de dados cadastrados na Plataforma Lattes, tendo em vista que são os indivíduos com maior nível de formação acadêmica concluída. Percebeu-se também que em geral seus currículos são recentemente atualizados e a maioria possui endereço profissional cadastrado.

Após a realização da caracterização geral do conjunto de indivíduos a serem analisados, foi possível observar como algumas grandes áreas como as de Ciências da Saúde e de

Ciências Humanas são bem mais representativas, em detrimento de outras como a grande área de Linguística, Letras e Artes.

A próxima etapa do trabalho foi a obtenção dos dados de geolocalização de todas as instituições em que os doutores se capacitaram em algum nível de formação acadêmica. Verificou-se a distância média percorrida pelos indivíduos ao longo de sua formação acadêmica, além de identificar a mediana das distâncias entre os níveis de formação, sendo observado que as distâncias percorridas foram pequenas, e que muitas das vezes os indivíduos optaram em se capacitar na mesma instituição em que iniciam seu processo de formação.

Posteriormente, através de análise baseadas em redes obteve-se um retrato de como são distribuídos os doutores em âmbito nacional, em que se pode observar como ocorrem relacionamentos entre as cidades/estados, quando se trata de interações entre as localidades de formação/atuação dos doutores ao longo da formação acadêmica.

Percebe-se como se destacam as cidades que possuem maiores quantidades de universidades, como por exemplo capitais de estados brasileiros, por se tratar da maior oferta de cursos de pós-graduação, e maior quantidade de ofertas de emprego.

Como era de se esperar, a cidade de São Paulo possui um maior valor de entrada por se tratar de uma cidade que oferece diversas oportunidades de emprego e possui um dos maiores centros acadêmicos no Brasil, e conseqüentemente maior oferta de cursos de pós-graduação, o que acarreta em ser um atrativo para os doutores de outras cidades.

REFERÊNCIAS

Abel, G. J., & Sander, N. (2014). Quantifying global international migration flows. *Science*, 343(6178), 1520–1522. <https://doi.org/10.1126/science.1248676>

Chaves, L. C. R., & Duarte, A. (2015, agosto 1). Analisando a mobilidade de pesquisadores através de registros curriculares na Plataforma Lattes. *Anais do Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*. IV Brazilian Workshop on Social Network Analysis and Mining. <https://doi.org/10.5753/brasnam.2015.6774>

Dias, T. M., Moreira, T. H., & Dias, P. M. (2020). Fomento à pesquisa científica no Brasil: A colaboração dos pesquisadores de excelência. *Parcerias Estratégicas*, 24(48), 151–166.

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

IBGE, C. D. (2010). *Instituto Brasileiro de Geografia e Estatística, 2010*. <http://www.ibge.gov.br/home/>

Lombas, M. L. de S. (2017). A mobilidade internacional acadêmica: Características dos percursos de pesquisadores brasileiros. *Sociologias*, 19(44), 308–333.

Moreira, J. R., Mueller, S. P. M., & Vilan Filho, J. L. (2020). Produção científica dos membros dos grupos de pesquisa das áreas de Informação no Brasil. *Informação & Informação*, 25(1), 1–20. <http://dx.doi.org/10.5433/1981-8920.2020v25n1p1>