

ÊXODO CIENTÍFICO BRASILEIRO: UMA ANÁLISE DO PROCESSO DE MIGRAÇÃO PARA CAPACITAÇÃO NO BRASIL

Brazilian scientific exodus: an analysis of the migration process for training in Brazil

Higor Alexandre Duarte Mascarenhas (1); Patrícia Mascarenhas Dias (2); Thiago Magela Rodrigues Dias (3)

- (1) Engenheiro da Computação, Universidade do Estado de Minas Gerais, Divinópolis - MG, Brasil.
(2) Ms. Profa., Universidade do Estado de Minas Gerais, Divinópolis - MG, Brasil.
(3) Dr. Prof., Centro Federal de Educação Tecnológica de Minas Gerais, Divinópolis - MG, Brasil
Email para Correspondência: higoralexandre1996@gmail.com; (P) Apresentador

Resumo: O êxodo de indivíduos por diversos motivos ou circunstâncias vem ao longo dos últimos anos crescendo de forma significativa no cenário brasileiro e mundial. Um dos motivos identificados para esse fluxo migratório é a formação acadêmica desses indivíduos que buscam se capacitarem em instituições de ensino de melhor qualidade. Este trabalho, corresponde ao tipo de pesquisa quali-quantitativa, cujo objetivo será analisar como ocorre o êxodo científico do Brasil, em que indivíduos de diversas localidades do Brasil migram em busca de uma melhor formação acadêmica e científica. Logo, será caracterizada e analisada a ocorrência do êxodo científico brasileiro nas diversas áreas do conhecimento. Foi necessária a extração dos dados acadêmicos, a partir dos currículos de indivíduos cadastrados na Plataforma Lattes. Assim sendo, selecionou-se todos doutores, agregando cerca de 296.000 currículos cadastrados. A escolha deste grupo se justifica por se tratar do conjunto de mais alto nível de formação acadêmica, e que detém como característica, possuem dados atualizados em seus currículos. Logo, filtrou-se os dados, selecionando os atributos relevantes a pesquisa, e finalmente, efetuou-se o tratamento dos dados com o intuito de encontrar a localização geográfica das instituições em que se capacitaram. Como resultado será possível traçar um retrato sobre como ocorre o êxodo científico brasileiro com todo o conjunto de indivíduos analisados e, ainda, pelas suas áreas de atuação.

Palavras chaves: Plataforma Lattes; Êxodo científico brasileiro; Fluxo migratório; Análise de Dados.

Abstract: The exodus of individuals for various reasons or circumstances has grown significantly in recent years in the Brazilian and world scene. One of the reasons identified for this migratory flow is the academic formation of those individuals who seek to qualify in institutions of higher



education quality. This work corresponds to the type of qualitative-quantitative research, whose objective will be to analyze how Brazil's scientific exodus occurs, in which individuals from different Brazilian locations migrate in search of a better academic and scientific education. Therefore, the occurrence of the Brazilian scientific exodus in the various areas of knowledge will be characterized and analyzed. It was necessary to extract the academic data, from the curricula of individuals registered in the Platform Lattes. Thus, all individuals with a PhD degree were selected, adding about 296,000 registered curricula. The choice of this group of individuals is justified because it is the set of the highest the highest academic training level, and that they have as a characteristic, updated data in their curricula. Then, the data were filtered, selecting relevant attributes for the research, and finally, the data was processed in order to find the geographical location of the institutions in which they were trained. As a result, it will be possible to draw a view over how the Brazilian scientific exodus occurs with the whole set of the individuals analyzed, and also by their areas of activity.

Keywords: *Platform Lattes; Brazilian scientific exodus; Migratory flow; Data analysis.*

1 INTRODUÇÃO

A emigração de brasileiros para outros países e migração 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 (ALMEIDA; ZANLORENSSI, 2017). Em muitos casos, brasileiros saem em busca de emprego, ou estudos, visando sempre qualidade de vida.

Um dos principais motivos para migração é a necessidade de capacitação em alto nível de formação. Das principais causas por optar pela mobilidade de indivíduos nos no território brasileiro, refere-se à qualidade de ensino superior em outros estados, a busca de novas oportunidades e mais experiências em suas áreas (LOMBAS, 2017). Outro refúgio para tais estudantes condiz na ida para outros países, buscando assim, intercâmbio cultural e melhor investimento em bolsas de pesquisa. A saída do aluno para outros países não é interessante somente ao discente, mas também às instituições de origem, pois, o mesmo retorna na maioria das vezes mais produtivos, com rede de contato mais extensa, maior vivência, e podendo futuramente compartilhar suas experiências com outros estudantes da instituição de origem.

De acordo com Demartini (2017) 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 imigração de pesquisadores brasileiros para outros países pode 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 à um programa que ampara estudantes, oferecendo bolsas de estudos. Em 2015, o governo pretendia alcançar



101.000 bolsas de estudos para pesquisadores, graduandos, doutorandos, alunos ingressados no pós-doutorado, incentivando os discentes a se capacitarem em instituições de reconhecida relevância (AVEIRO, 2014). Hoje com apenas 5.000 bolsas disponíveis, o Programa perdeu bastante influência no ingresso de estudantes para outros países, por motivos de cortes de investimento.

Diante deste cenário, este trabalho apresentará um estudo quantiquantitativo sobre o êxodo de estudantes brasileiros que partiram do seu estado de nascimento para outros estados e/ou aqueles que foram para outros países em busca de capacitação. Para extração dos dados de estudantes brasileiros será utilizada a ferramenta LattesDataXplorer (DIAS, 2016), ferramenta responsável por extrair currículos de indivíduos cadastrados na Plataforma Lattes. Atualmente, o repositório de currículos da Plataforma Lattes é direcionado a área acadêmica/científica e profissional possuindo cerca de 6.000.000 de currículos cadastrados.

Logo, o objetivo geral é analisar o êxodo de estudantes brasileiros durante o seu processo de formação acadêmica, a partir de dados extraídos de seus de currículos cadastrados na Plataforma Lattes.

2 FUNDAMENTAÇÃO TEÓRICA

A escolha dos currículos da Plataforma Lattes, se justifica pelo fato destes serem direcionados principalmente a área acadêmica/científica e possuírem um vasto leque de informações dos indivíduos brasileiros. Ao se tratar de um trabalho, cujo objetivo é de analisar a mobilização de doutores brasileiros ao longo de sua formação acadêmica, foi definida a utilização destes currículos.

Medeiros e Mena-Chalco (2013) consideraram 633.508 currículos extraídos da Plataforma Lattes e 4.946.990 publicações com o propósito de explorar a rede social dos indivíduos de áreas distintas. Sendo assim calcularam as frequências dos termos dos títulos com o intuito de identificar quais são mais utilizadas em cada época. Logo após, foi utilizado mapas de termos e nuvem de palavras para os duzentos termos mais constantes de cada área por período determinado.

Para Ferraz, Quoniam e Maccari (2014), é inexistente um repositório curricular nacional único congênere à Plataforma Lattes, visto que somente repositórios de dados referenciais, de onde pode ser extraída referências bibliográficas e fontes de informação secundárias estão disponíveis para livre acesso. Sendo assim, a Plataforma Lattes pode convir como uma ferramenta de maior importância para o estudo da produção científica brasileira.

Já Mena-Chalco et al. (2014) utilizaram dados dos currículos da Plataforma Lattes a fim de identificar e representar a rede de colaboração de pesquisadores brasileiros. O trabalho tem por objetivo extrair os dados de currículos arquivados na Plataforma Lattes, apontar automaticamente a colaboração baseada em informações bibliométricas, gerando uma rede de colaboração, e aplicar métricas baseadas em análise topológica para entender como ocorre a interação entre os pesquisadores. Sendo assim, foi adotado o cálculo da distância de Levenshtein, para encontrar títulos iguais ou similares. Com as redes de



colaboração desenvolvidas, aspectos topológicos foram analisados, considerando apenas informações estruturais das redes, sem considerar pesos nos vínculos entre autores ou qualidade das publicações. Desse modo, identificou-se que em determinadas áreas, os pesquisadores pendem a realizar suas publicações com um número baixo de coautores em desvantagem a outras áreas, visto que a quantidade de coautores tende a ser elevada. Pôde ser observado que poucos pesquisadores possuem alto grau de conectividade. Além do mais, percebeu-se que a quantidade de colaboração entre os pesquisadores tem crescido significativamente, identificada pelo grau dos nós das redes analisadas em diferentes períodos de tempo.

Segundo Marques (2015) em artigo publicado na revista Fapesp, a Plataforma Lattes reunia mais de 4 milhões de currículos acadêmicos, e por isso tornou-se principal fonte de informações para estudos sobre a ciência brasileira, com o intuito de analisar fenômenos e tendências. Os currículos registram trajetória e a contribuição de cada estudante, pesquisador e técnico brasileiro. Em 2005 foi criado um extrator de dados chamado ScriptLattes, que tem por objetivo reunir e organizar, de forma automática, informações extraídas de grande aglomeração de dados. No artigo também é demonstrada uma vasta quantidade de trabalhos publicados com finalidade de análise de dados da retirados da Plataforma Lattes.

Oliveira, Dias e Moita (2018) realizaram um trabalho sobre análise de dados referentes às palavras-chave dos artigos publicados ao longo dos últimos 55 anos, proveniente dos doutores com currículos cadastrados na Plataforma Lattes. Foi observado que na medida que a quantidade de artigos eram publicadas, o número de palavras-chave também aumentavam, uma vez que cada artigo possui em média 6 palavras-chave. Os termos mais utilizadas como palavras-chave pelos doutores foram: educação, ensino e políticas públicas.

Aveiro (2014) realizou um estudo sobre os indivíduos que participam do programa Ciência Sem Fronteiras. Segundo o autor os estados que mais ingressam estudantes no programa Ciência Sem Fronteiras são as regiões do Sudeste, Sul e Nordeste respectivamente. Uma das razões pelas quais o Sudeste está em primeiro lugar é por ser o estado mais populoso. As faculdades que mais possuem bolsistas são USP, UFMG, UFRJ, UnB e UFSC. Quando se trata da distribuição de áreas existentes no cenário acadêmico brasileiro, as engenharias e demais áreas tecnológicas lideram este ranking com maior número de bolsistas registrados. No que se concerne país de destino foi observado que aqueles que mais receberam bolsistas são EUA, Reino Unido, Canadá, França, Austrália, Alemanha, Espanha, Portugal, Itália e Irlanda, respectivamente.

Araújo e Silva (2014) efetuaram um estudo em uma determinada instituição acadêmica portuguesa que abrangia os estudantes e pesquisadores brasileiros. Ao longo de seu estudo, os autores citam que foi realizada uma pesquisa por Franco (2002) e foi descoberto que Portugal era o sexto país com maior número de imigrantes brasileiros que possuem bolsa CAPES. Para analisar tais estudantes brasileiros os autores implementaram um questionário, com o objetivo de compreender o porquê da imigração. Pôde ser observado nas respostas dos questionários que a maioria dos estudantes vão para Portugal visando fazer novos contatos, aprimorar seus estudos e currículos, alguns por realização pessoal, tencionando expectativas de carreira. A maioria dos estudantes



brasileiros entrevistados pretendem retornar ao Brasil, mas muitos deles ainda querem conhecer outros lugares antes do regresso.

Chaves et al. (2016) realizou um estudo de dados da Plataforma Lattes de pesquisadores de pós-graduação, coletando dados da trajetória do pesquisador desde o nascimento até a sua última titulação. No trabalho realizado, o foco principal é voltado para doutores, e durante análises, chegou-se à conclusão que 95% são de origem dos estados 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. 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.

Pierro (2016) analisa a mobilidade de pesquisadores e estudantes brasileiros ao longo de sua formação acadêmica. Ressalta-se que 20% dos pesquisadores trabalham entretanto, a mais de 500 km de distância na instituição de onde ingressou na trajetória acadêmica, em contrapartida, a maioria trabalha cerca de 100 km de distância. Tal mobilidade fez com que pesquisadores entrevistados se envolvessem em várias linhas de pesquisa, tornando assim o trabalho do mesmo mais conhecido, nos locais da sua trajetória. O estudo indica que os estados da região sudeste, principalmente São Paulo, são aqueles que a maioria dos pesquisadores são oriundos da própria região; já os outros estados brasileiros possuem um padrão migratório temporário. A mobilidade de pesquisadores pode contribuir para a produção científica brasileira, mesclando conhecimento entre os pesquisadores dos diversos estados da federação. Os pesquisadores que encontram barreiras durante o limite da mobilidade optam por realizar redes de colaboração com cientistas de outras instituições.

Logo, é notório que uma larga escala de indivíduos que optam por obter uma capacitação em alto nível de formação partem da sua cidade de origem para outra, e com menor escala, partem do seu país de origem procurando obter capacitação. Vale ressaltar também que muitos dos trabalhos relacionados a este projeto, preferem extrair dados dos currículos cadastrados na Plataforma Lattes, por se tratar de uma ferramenta de muita importância para o estudo da produção científica brasileira.

3 METODOLOGIA

Como principal fonte de dados foram utilizados os currículos cadastrados na Plataforma Lattes. A justificativa da escolha da Plataforma se trata por: (1) registrar a trajetória e a contribuição de cada estudante, técnico e pesquisador brasileiros cadastrados (MARQUES, 2015); (2) os dados estarem disponíveis na internet (DIGIAMPIETRI, 2015); (3) representa a experiência do CNPq na integração de base de dados de currículos e de instituições da área de Ciência e tecnologia (SILVA; SMIT, 2009); (4) por se tratar de uma importante fonte de dados de alta qualidade com o intuito de medir e avaliar o desempenho acadêmico nacional (LANE, 2010).

Apesar de reconhecida relevância dos currículos cadastrados na Plataforma Lattes para análise e entendimento sobre a evolução da ciência brasileira conforme descrito

anteriormente, o acesso ao repositório de dados passa a ser um fator limitante para análises que considerem todos os indivíduos independentemente de suas áreas de atuação. Apesar da viabilidade de acesso individual a cada um dos currículos ser possível através de interface de consulta no portal da Plataforma Lattes, a análise de grandes grupos de indivíduos passa a ser um fator limitante para análises abrangentes. Logo, no contexto deste trabalho, para extração de todo o conjunto de currículos analisados, foi utilizado o LattesDataXplorer (DIAS, 2016) para extração e tratamento dos dados.

A extração dos dados foi realizada em maio de 2018 totalizando aproximadamente 296.000 currículos de indivíduos com doutorado concluído. O *framework* utilizado é responsável por conter uma coleção de componentes que visam realizar a coleta e tratamento dos dados, conforme ilustra a Figura 1.

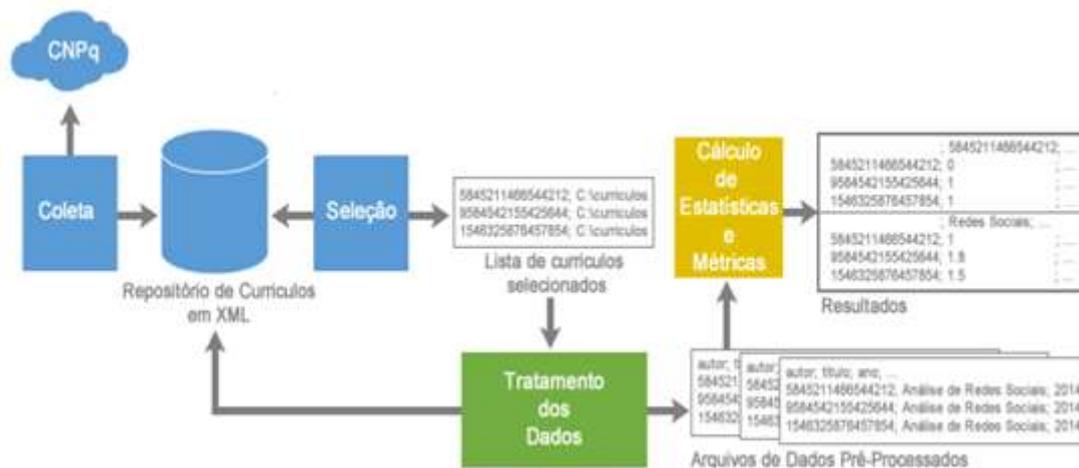


Figura 1. Visão geral do LattesDataXplorer

Fonte: (Dias, 2016)

Utilizando o *framework* é possível fazer um refinamento da busca de currículos específicos, baseado em parâmetros como nome, titulação, idioma, nacionalidade, grande área e área de atuação, dentre outros. Sendo gerado listas de currículos que atendem aos parâmetros informados.

Todo o procedimento de extração e tratamento de dados realizado pelo LattesDataXplorer se inicia a partir da aquisição dos códigos de currículos da Plataforma Lattes, baseado no refinamento executado na busca, com o propósito de no futuro estes códigos sejam armazenados localmente (Figura 2). A lista resultada da consulta refinada, possui todos os códigos de identificação de todos os currículos cadastrados, possibilitando ter o acesso individual em cada um destes existentes na Plataforma Lattes.

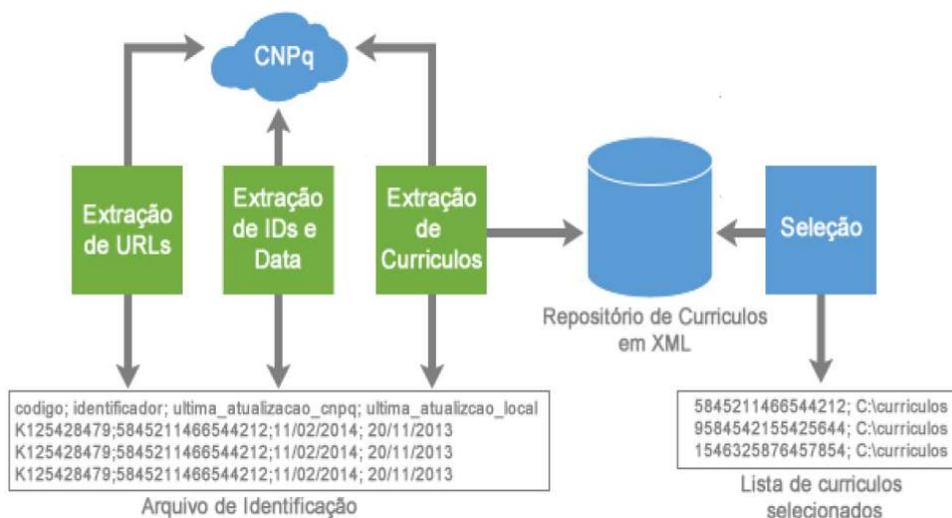


Figura 2. Componente de coletor do LattesDataXplorer.

Fonte: (Dias, 2016)

Todos currículos possuem uma URL (*Uniform Resource Locator*) a fim de permitir o acesso individual a cada um destes. A URL possui o endereço de busca da plataforma concatenada com o código do currículo do indivíduo. O componente de extração de IDs e Data tem a responsabilidade de acessar cada um dos currículos cujos códigos foram salvos, extraíndo seu identificador único e a data da última atualização do mesmo. Para esse fim, somente o cabeçalho de cada currículo em que estão presentes estas informações necessitam ser carregados. Tais dados foram armazenados em um arquivo de identificação contendo código, identificador, data da última atualização no CNPq e data da atualização do currículo armazenado localmente.

Consoante Dias (2016), afirma que para a atualização do repositório de currículos é necessário o arquivo de identificação, por se tratar da base para extração do mesmo. Durante a atualização o primeiro componente do processo de extração é executado, resultando na extração de todos os códigos cadastrados na plataforma. Os códigos já registrados no arquivo de identificação são ignorados, e no final do arquivo são adicionados aqueles novos códigos, representantes dos novos currículos ainda não extraídos.

Com o uso dos códigos, são acessados os cabeçalhos de cada um dos currículos assim sendo extraídos códigos identificadores e as datas de atualização junto à Plataforma Lattes, tanto para currículos já extraídos como para os novos currículos, ocorrendo a atualização do arquivo de identificação a cada nova extração. O acesso ao cabeçalho torna



mais ágil a extração de dados, por não haver a necessidade de esperar todo o currículo ser gerado.

Por fim, acontece a extração de currículos, sendo o extrator o responsável por verificar se possuem currículos cuja Data de Atualização Local é diferente da Data de Atualização junto ao CNPq; caso seja divergente, o currículo da Plataforma é extraído e substitui o currículo local, modificando a data de atualização. Quando não, o currículo permanece o mesmo sem alterações. Por fim são extraídos novos currículos cadastrados, por serem inseridos ao final do arquivo base. Inicialmente esses novos currículos não possuem data de atualização por terem sido extraídos pela primeira vez, sendo assim, é inserido sua data de atualização local.

Finalmente, todos os currículos são armazenados em XML, não necessitando de um novo repositório de dados feito por bancos relacionais, buscando um menor custo computacional.

Com todos currículos armazenados localmente, vale ressaltar que se torna mais fácil e flexível a manipulação de dados coletados a partir Plataforma Lattes. A fim de realizar uma análise de dados de grupos específicos, foi realizada a etapa de Seleção descrita abaixo.

Na etapa de Seleção é usada a linguagem de consulta XPath (*XML Path Language*) para pesquisa e posterior geração dos subgrupos. A linguagem XPath possibilita a construção de expressões que vão processar e percorrer um documento XML de forma similar ao uso de expressões regulares. Portanto, possibilita o agrupamento de um conjunto de currículos com parâmetros desejados. Assim sendo, em busca pelos parâmetros em cada um dos currículos, independentemente ou não de qual seção ele(s) seja(m) encontrado(s), tais currículos são selecionados e formam um grupo para análises. A partir de então os dados dos currículos são organizados em uma lista de currículos que foram selecionados.

A lista armazena os identificadores de cada currículo e o caminho que ele está armazenado localmente, sendo assim, foi possível analisar somente os currículos selecionados para um determinado tipo de grupo.

Diante do exposto foi coletado somente currículos de doutores, por se tratar do grupo de mais alto nível de formação; por se tratar de currículos que estão sempre atualizados e grande parte dos parâmetros necessários para o trabalho estão registrados. Quando se trata de atributos como cidade de formação da graduação, mestrado e doutorado, entre outros.

3.1 Filtragem dos dados

A fim de mapear o êxodo de indivíduos brasileiros cadastrados na Plataforma Lattes, foi efetuada a mineração de dados para filtrar dados relevantes, logo após serem filtrados ocorrerá um tratamento. A Figura 3 apresenta um aspecto geral do conjunto de componentes que foram desenvolvidos objetivando obter as análises desejadas.

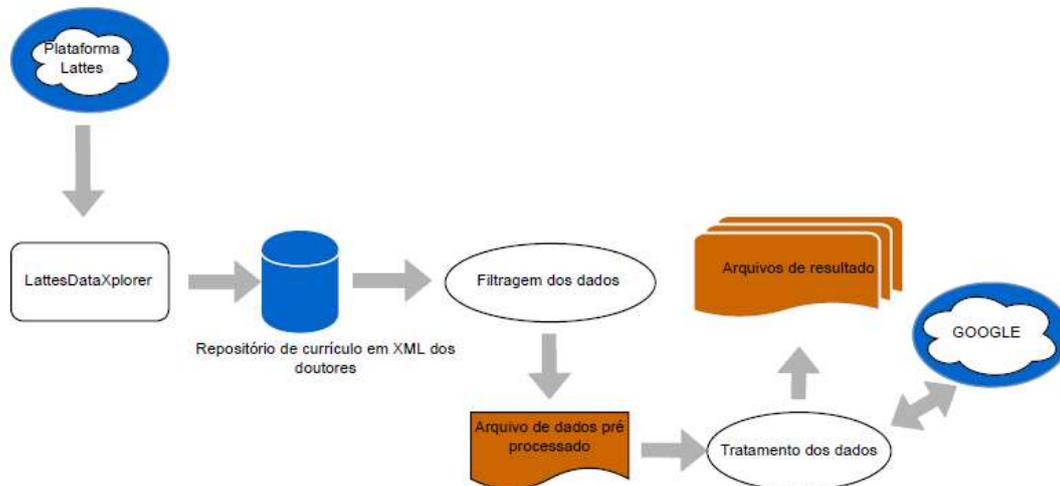


Figura 3. Aspecto geral do conjunto de componentes utilizados.

Fonte: Autores

Depois da extração de dados, ocorreu a “filtragem dos dados”. Esta fase será responsável por executar a mineração nos arquivos XML para extrair informações relevantes a pesquisa, armazenando-as em uma planilha (Arquivo de dados pré processado). Deste modo definiu-se o conjunto de dados centrais a serem processados. As informações dos currículos possuem: código do indivíduo; estado de nascimento; cidade de nascimento; grande área; área; código, nome e CEP da empresa que indivíduo trabalha atualmente, e possuem também o código de identificação, nome da instituição, início da formação e fim da formação das primeiras formações concluídas, iniciando na graduação até o doutorado.

O Módulo "Tratamento de dados" tem o intuito de processar os dados dos doutores coletados, tratá-los e caracterizá-los, resultando em outros arquivos, a fim de facilitar as análises dos dados. Esse processo realizou, basicamente três etapas: Busca pela localização geográfica; Limpeza e agrupamento de dados; Normalização dos dados.

- A etapa de "busca pela localização geográfica" é uma tarefa que foi realizada com a finalidade de geolocalizar uma instituição. Em linhas gerais, a etapa será efetuada da seguinte maneira: ao acessar o API (*Application Programming Interface*) de geolocalização do Google, foi informado o endereço da instituição, sendo assim, foi possível a extração da localização geográfica da instituição.
- Na etapa "limpeza e agrupamento de dados", ocorreu 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*; A normalização para extrair palavras acentuadas, e substituí-las pelo seu equivalente sem acentuação, entre outros.
- Já a etapa de "normalização dos dados" tem o intuito de reduzir a redundância de informações, descartando atributos de currículos com a ausência de dados.

Os "Arquivos de resultado" possui todos os dados específicos para a realização de algumas análises após as ações realizadas.

Logo, após toda a execução das etapas descritas anteriormente, e com os arquivos formatados de tal forma a facilitar as análises, algumas métricas foram aplicadas para compreensão sobre como tem ocorrido o êxodo científico brasileiro.

4 RESULTADOS

Como resultado inicial foi possível realizar uma caracterização do conjunto extraído. Por se tratar de um grupo de indivíduos que possuem alto nível de formação, sendo assim, levam no currículo o Status de "concluído" ao ser analisado a situação de formação da Graduação, Mestrado e Doutorado. A Figura 4 apresenta a quantidade de currículos que possuem atributos cadastrados, tais como: cidade de nascimento, instituição de atuação e instituições ao longo da formação.

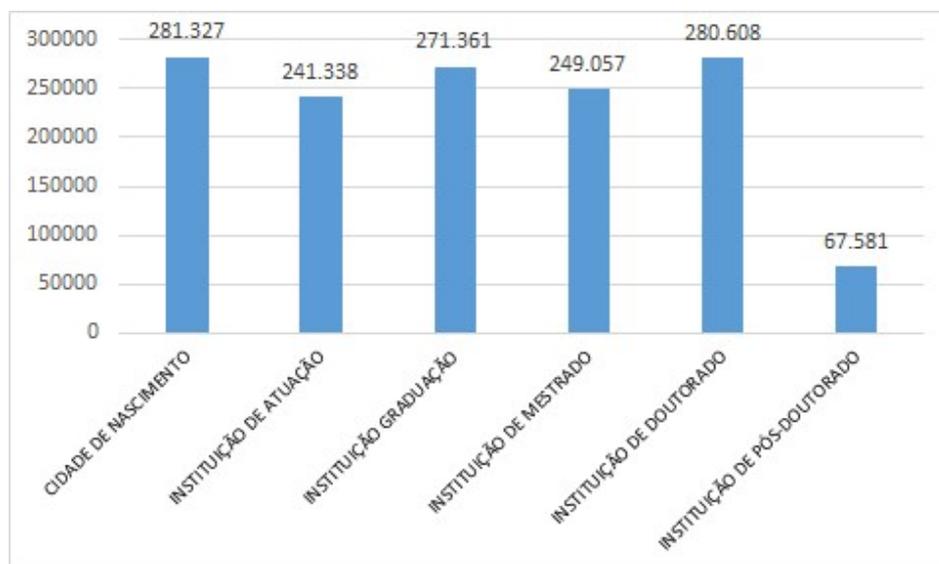


Figura 4. Quantitativo de currículos que possuem códigos de instituições cadastradas

Fonte: (Autores)

Como pode ser observado, a maioria dos indivíduos possuem a cidade de nascimento cadastrada com um total de 281.327 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 julga-se que são currículos mais antigos, em que não era obrigatório o cadastramento da cidade. Das demais instituições cadastradas demonstradas no gráfico, aquela que se identifica superior em dados quantitativos é a *instituição de doutorado* totalizando uma quantidade de 280.608 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 67.581 *instituições de pós-doutorado*, 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.

Sabe-se também que existem diversas Grandes Áreas para a destinação do estudo de indivíduos no ambiente acadêmico. Deste modo, por meio da Figura 5, é possível também ser observado a distribuição de doutores por todas as Grandes Áreas existentes na Plataforma Lattes, cadastradas pelos indivíduos ao longo do processo de registro do currículo.

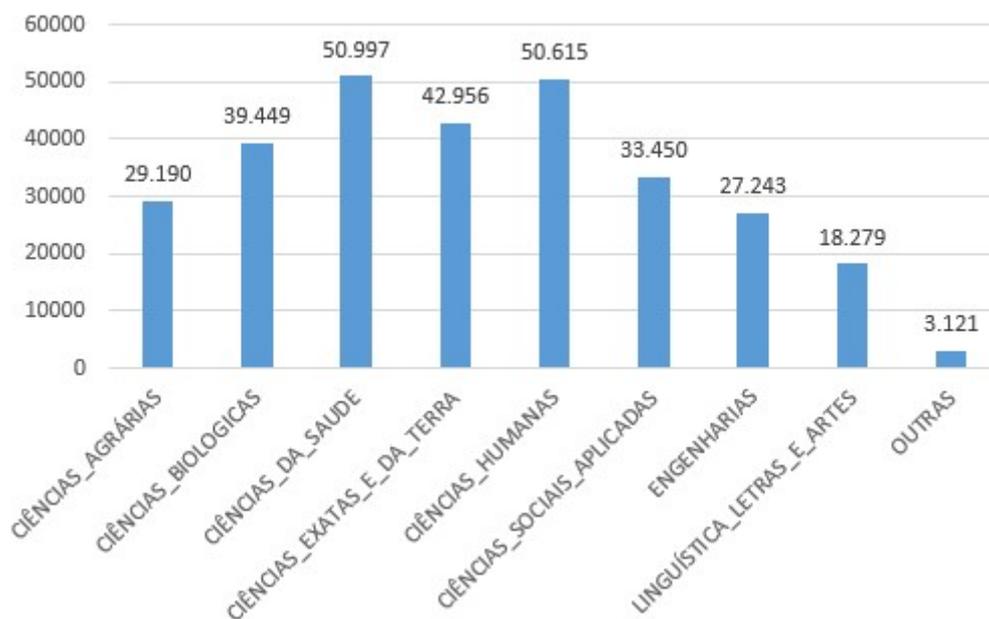


Figura 5. Distribuição de doutores por Grande Área

Fonte: (Autores)

É possível verificar que as Grandes Áreas que possuem mais doutores cadastrados são Grande Área Ciências da Saúde (50.997) e a Grande Área Ciências Humanas (50.615) respectivamente. A Grande Área Ciências Humanas se justifica pelo fato de ser mais destinada para indivíduos que pretendem trabalhar como docentes na sua carreira. Das Grandes Áreas classificadas na Plataforma Lattes, apesar de também possuir indivíduos voltados ao mercado acadêmico, aquela que possui menos cadastros é a Grande Área de Linguística, Letras e Artes totalizando 18.279 currículos de doutores que atuam nesta Grande Área.

Foi possível também identificar a distribuição dos doutores de acordo com seu estado de nascimento, e de sua instituição de atuação profissional. Conforme demonstra na Tabela 1.



XIII SIMMEC

Simpósio de Mecânica Computacional

29 de Outubro a 1º de Novembro de 2018
UFES - Campus Goiabeiras - Vitória, ES



Tablela 1. Distribuição de doutores por estado de nascimento e local onde trabalham

Estados	Nascimentos	Atuação profissional
AC	418	531
AL	1.910	1.862
AM	1.857	204
AP	283	381
BA	9.463	12.973
CE	8.123	6.277
DF	3.356	8.008
ES	3.721	252
GO	5.301	5.675
MA	2.477	2.217
MG	31.155	24.820
MS	2.298	3.101
MT	1.514	2.754
PA	5.442	4.573
PB	6.222	5.219
PE	9.389	15.105
PI	2.278	1.954
PR	16.481	16.513
RJ	32.334	29.784
RN	3.881	4.211
RO	400	294
RR	152	483
RS	28.536	21.704
SC	8.247	8.924
SE	1.559	1.940
SP	76.111	51.045
TO	333	1.170
NÃO INFORMADO	32.060	63.326



Quando se trata do estado de nascimento de doutores, observa-se que a grande parte dos indivíduos nasceram em estados da região Sudeste totalizando 143.321. Sendo que São Paulo possui a maioria de doutores nascidos, com um total de 76.111, Rio de Janeiro com 32.334, Minas Gerais com uma quantidade de 31.155 cadastros e Espírito Santo com apenas 3.721. Já aos demais estados de outras regiões, os que mais possuem doutores nascidos é o estado do Rio Grande do Sul, totalizando 28.536 doutores. Tal fato se justifica por se tratar de um estado que possui muitas universidades.

Apesar de São Paulo ser o estado com maior número de nascimento de doutores, quando se trata de atuação profissional, este estado se destaca pela ausência de doutores, em cerca de 25.000 doutores que migram para outros estados para atuar profissionalmente e/ou doutores que estão sem atuar. Porém ainda continua sendo o estado que possui mais doutores atuando, com cerca de 51.045 doutores trabalhando no estado. Isso ocorre por se tratar do estado mais populoso do Brasil e que possui a maior quantidade de universidades. Alguns estados como Pernambuco, Distrito Federal e Bahia, apesar de não possuir muitos doutores nascidos no estado, é o destino de muitos deles na atuação profissional. Deve-se destacar a quantidade de currículos cuja a atuação profissional não está cadastrada, totalizando 63.326, fato que pode gerar viés à pesquisa.

5 CONSIDERAÇÕES FINAIS

A partir dos resultados iniciais 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.

Pela caracterização geral realizada foi possível perceber como algumas grandes áreas como Ciências da Saúde e Ciências Humanas são bem mais representativas, em detrimento de outras quantidades bem inferiores como a grande área de Linguística, Letras e Artes. Além disso, também ficou nítido como região sudeste concentra a grande maioria dos doutores brasileiros, fato este influenciado diretamente pela concentração das principais universidades públicas do país.

REFERÊNCIAS

ALMEIDA, R.; ZANLORENSSI, G. *Fluxos migratrios: a distribuição da população de cada estado pelo país.* dec 2017. Disponível em: <<https://www.nexojornal.com.br/grafico/2017/12/01/Fluxos-migrat%C3%B3ria-distribui%C3%A7%C3%A3o-da-popula%C3%A7%C3%A3o-de-cada-estado-pelopa%C3%ADs>>.



ARAÚJO, E. R.; SILVA, S. M. F. Ecos do tempo. a mobilidade de investigadores e estudantes brasileiros em Portugal. *Sociologias*, Universidade Federal do Rio Grande do Sul. Curso de Pós-Graduação em Psicologia, v. 16, n. 37, p. 218-250, 2014.

AVEIRO, T. M. M. O programa ciência sem fronteiras como ferramenta de acesso à mobilidade internacional. *Tear: Revista de Educação Ciência e Tecnologia*, v. 3, n. 2, 2014.

CHAVES, L. C. R. et al. Analisando a mobilidade de pesquisadores através de registros curriculares na plataforma lattes. Universidade Federal da Paraíba, 2016.

DEMARTINI, M. *Falta de oportunidades mantém cientistas brasileiros no exterior*. Feb 2017. Disponível em: <<https://exame.abril.com.br/ciencia/falta-de-oportunidades-mantemcientistas-brasileiros-no-exterior/>>.

DIAS, T. M. R. *UM ESTUDO SOBRE A PRODUÇÃO CIENTÍFICA BRASILEIRA A PARTIR DE DADOS DA PLATAFORMA LATTES*. Tese (Doutorado) - Centro Federal de Educação Tecnológica de Minas Gerais, sep 2016.

DIGIAMPIETRI, L. A. Análise da rede social acadêmica brasileira. 2015.

FERRAZ, R. R. N.; QUONIAM, L.; MACCARI, E. A. The use of scriptlattes tool for extraction and on line availability of academic production from a departament of stricto sensu in management. *INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGY MANAGEMENT*, p. 663-679, 2014. 17

FRANCO, M. Globalização, internacionalização e cooperação interinstitucional. SOARES, MSA *Educação Superior no Brasil, Instituto Internacional para a Educação Superior na América Latina e no Caribe-IESALC-Unesco-Caracas*, p. 305-327, 2002.

LANE, J. Let's make science metrics more scientific. *Nature*, Nature Publishing Group, v. 464, n. 7288, p. 488, 2010.

LOMBAS, M. L. D. S. A mobilidade internacional acadêmica: características dos percursos de pesquisadores brasileiros. *Scielo*, v. 19, n. 44, p. 308-333, jan 2017.

MARQUES, F. Registros valiosos. *Pesquisa Fapesp*, p. 34-37, jul 2015.

MEDEIROS, C. B.; MENA-CHALCO, J. The dynamics of multidisciplinary research networks-mining a public repository of scientists cvs. In: *WORLD SOCIAL SCIENCE FORUM 2013. Social Transformations and the Digital Age*. Canada, Montreal, 2013. p. 1-17.

MENA-CHALCO, J. P. et al. Brazilian bibliometric coauthorship networks. *Journal of the Association for Information Science and Technology*, Wiley Online Library, v. 65, n. 7, p. 1424-1445, 2014.

OLIVEIRA, J. G.; DIAS, T. M. R.; MOITA, G. F. Uma análise dos principais tópicos de pesquisas investigados pelos pesquisadores doutores brasileiros. *Em Questão*, Universidade Federal do Rio Grande do Sul, Faculdade de Biblioteconomia e Comunicação, v. 24, n. 2, p.55-82, 2018.

PIERRO, B. de. Circulação limitada. *Pesquisa Fapesp*, p. 36-39, jan 2016.



SILVA, F.; SMIT, J. Organização da informação em sistemas eletrônicos abertos de informação científica & tecnológica: análise da plataforma lattes. *Perspectivas em ciência da informação*, v. 14, n. 1, p. 77-98, 2009.