

## **Avaliação estratégica de programas em educação no âmbito federal**

Detalhamento metodológico e caracterização dos dados para mapeamento de docentes e discentes das IES federais na área de saúde

# **Avaliação estratégica de programas em educação no âmbito federal**

Detalhamento metodológico e caracterização dos dados para mapeamento de docentes e discentes das IES federais na área de saúde



Brasília, DF  
Dezembro, 2021

# Centro de Gestão e Estudos Estratégicos (CGEE)

Organização social supervisionada pelo Ministério da Ciência, Tecnologia, Inovações (MCTI)

---

## Diretor-Presidente

*Marcio de Miranda Santos*

## Diretores

*Regina Maria Silverio*

*Luiz Arnaldo Pereira da Cunha Junior*

## Equipe de apoio

*Genilda Mota*

Detalhamento metodológico e caracterização dos dados para mapeamento de docentes e discentes das IES federais na área de saúde. Avaliação estratégica de programas em educação no âmbito federal. Brasília, DF: Centro de Gestão e Estudos Estratégicos, 2021.

19p.

1. Docentes e discentes. 2. Rede de competências. 3. Visualização de dados. I. Título. II. CGEE.

Centro de Gestão e Estudos Estratégicos - CGEE  
SCS Quadra 9 – Torre C – 4º andar – salas 401 a 405  
Edifício Parque Cidade Corporate  
70308-200 - Brasília, DF  
Telefone: (61) 3424.9600  
<http://www.cgee.org.br>

Todos os direitos reservados pelo Centro de Gestão e Estudos Estratégicos (CGEE). Os textos contidos nesta publicação poderão ser reproduzidos, armazenados ou transmitidos, desde que seja citada a fonte.

### Referência bibliográfica:

Centro de Gestão e Estudos Estratégicos - CGEE. **Mapeamento de competências das IFES públicas na área de saúde.** Avaliação estratégica de programas em educação no âmbito federal. Brasília, DF: 2021. 19p.

---

# **Avaliação estratégica de programas em educação no âmbito federal**

Detalhamento metodológico e caracterização dos dados para mapeamento de docentes e discentes das IES federais na área de saúde

## **Supervisão**

*Marcio de Miranda Santos*

## **Equipe técnica do CGEE**

*Sofia Daher (Coordenadora)*

*Adriana Badaró de Carvalho Villela*

*César Augusto Costa*

*Ivone Alves de Oliveira Lopes*

*João Vitor Rodrigues Martins*

*José Salomão*

*Marcelo Augusto Paiva dos Santos*

*Marcia Tupinambá*

*Matheus Figueiredo Pimenta*

## **Consultores**

*Antonio Pereira de Souza Júnior*

# SUMÁRIO

|                                                                                                                |           |
|----------------------------------------------------------------------------------------------------------------|-----------|
| <b>1. INTRODUÇÃO .....</b>                                                                                     | <b>6</b>  |
| OBJETIVO GERAL .....                                                                                           | 7         |
| OBJETIVOS ESPECÍFICOS .....                                                                                    | 7         |
| <b>2. ESTRATÉGIA DE CONDUÇÃO .....</b>                                                                         | <b>7</b>  |
| EQUIPE .....                                                                                                   | 7         |
| COORDENAÇÃO .....                                                                                              | 8         |
| <b>3. METODOLOGIA DE TRABALHO .....</b>                                                                        | <b>8</b>  |
| <b>4. DETALHAMENTO METODOLÓGICO .....</b>                                                                      | <b>11</b> |
| CARACTERIZAÇÃO DOS DADOS .....                                                                                 | 11        |
| REDES DE COMPETÊNCIAS HUMANAS NO PAÍS .....                                                                    | 12        |
| MODELAGEM DE TÓPICOS DAS INFORMAÇÕES DE PRODUÇÃO CIENTÍFICA E/OU TECNOLÓGICO NOS CURRÍCULOS SELECIONADOS ..... | 14        |
| VISUALIZAÇÃO DOS DADOS E OUTRAS EXPLORAÇÕES .....                                                              | 15        |
| <b>5. DETALHAMENTO DOS PRODUTOS .....</b>                                                                      | <b>17</b> |
| <b>6. NECESSIDADE DE DADOS E INFORMAÇÕES .....</b>                                                             | <b>19</b> |
| <b>7. CRONOGRAMA DE TRABALHO DAS PRÓXIMAS ETAPAS .....</b>                                                     | <b>19</b> |

## **1. INTRODUÇÃO**

O Ministério da Educação (MEC), por intermédio da Secretaria de Educação Superior (SESU) e sua Diretoria de Desenvolvimento da Educação em Saúde, demandou ao CGEE a realização de um estudo para mapear as áreas e os principais temas de atuação dos docentes e discentes pertencentes aos programas de graduação e pós-graduação das áreas da saúde das Instituições Federais de Ensino Superior (IFES).

Esse mapeamento temático, em perspectiva com outras informações, auxilia na identificação das principais competências temáticas locais e regionais, bem como compreender a contribuição efetiva e potencial dos programas das IFES para as prioridades e estratégias de pesquisa, desenvolvimento e inovação no âmbito nacional e à formulação de políticas públicas orientadas para o fortalecimento de áreas importantes para a saúde e bem estar social e formação de profissionais nas áreas de saúde.

Este documento descreve a abordagem metodológica dos serviços técnicos especializados a serem executados pelo CGEE em resposta à solicitação feita pela Diretoria da Educação em Saúde (SESU/MEC), bem como a estratégia de condução e o cronograma previsto. Quanto à caracterização dos dados inicialmente prevista para este produto, não foi efetuada devido à impossibilidade de acesso aos dados e informações necessárias para a realização do tratamento e da sistematização dos mesmos. Tão logo o CGEE receba da SESU/MEC os dados solicitados, serão executados e descritos os procedimentos referentes a esta etapa do projeto. Importante ressaltar, que o atraso da entrega dos dados deveu-se, principalmente, a questões envolvendo a adequação à Lei Geral de Proteção de Dados (Lei n.º13.709, de 14 de agosto de 2018) executados pelo MEC.

Em função da demanda apresentada pela Diretoria da Educação em Saúde (SESU/MEC) e expressa nos objetivos geral e específicos, o CGEE propõe a metodologia descrita a seguir, além da apresentação da estratégia geral de desenvolvimento do trabalho e o detalhamento por produto. Ao final, é apresentada uma proposta de cronograma de trabalho das próximas etapas.

## **Objetivo geral**

Identificar e analisar os principais temas e objetos de pesquisa dos programas de graduação e pós-graduação das áreas da saúde das Instituições Federais de Ensino Superior (IFES), por meio do mapeamento da produção técnica e científica dos docentes e discentes vinculados a esses programas.

## **Objetivos específicos**

- Identificar e analisar a produção científica e técnica dos docentes e discentes vinculados às IFES a partir dos currículos extraídos da Plataforma Lattes.
- Mapear as principais convergências e parcerias institucionais e regionais, bem como as redes temáticas com potencial para colaboração técnica e científica em temas selecionados.

## **2. ESTRATÉGIA DE CONDUÇÃO**

O trabalho será desenvolvido por especialistas e assessores técnicos do CGEE destacados pela Diretoria do Centro para este fim de acordo com suas capacidades e competências anteriormente expressas em condução ou elaboração de trabalhos correlatos, em estreita articulação com a Diretoria de Desenvolvimento da Educação em Saúde – DDES (SESU/MEC).

## **Equipe**

A equipe básica envolvida no contrato será composta pelos seguintes perfis profissionais:

- Supervisor: responsável pela supervisão geral da condução do projeto.
- Coordenador: responsável pela coordenação e acompanhamento técnico de todos os produtos e pela interlocução institucional com a SESU/MEC.
- Analistas de dados: responsáveis pela coleta, tratamento e sistematização dos dados, bem como na realização das análises relacionadas.
- Estatísticos: responsáveis pelo tratamento e análise estatística dos dados inseridos nos bancos de dados.
- Assistentes de projeto: responsáveis pelo apoio aos processos de desenvolvimento dos produtos, levantamento de necessidades e acompanhamento do cronograma.
- Assessores técnicos: assessores técnicos do Centro destinados aos desenvolvimentos dos produtos relacionados neste plano de trabalho.
- Assistentes administrativos: encarregados de providenciar trâmites administrativos de qualquer natureza relacionados com o objeto deste plano de trabalho.

## **Coordenação**

A estreita colaboração entre as partes desse projeto (SESU/MEC e CGEE) é um fator indispensável para o bom desenvolvimento dos trabalhos a serem executados, dado que a sua natureza impõe, entre outros aspectos, a troca de dados e informações para a obtenção dos resultados previstos. Além do acompanhamento de cada etapa do projeto, sugere-se a formação de um pequeno comitê de coordenação composto por membros do SESU/MEC e CGEE. Este comitê será responsável por discutir o conteúdo para os produtos e, em particular, os aspectos ligados a possíveis ajustes metodológicos no decorrer do trabalho.

## **3. METODOLOGIA DE TRABALHO**

Para atender ao objetivo do projeto, a abordagem metodológica prevê análises quantitativas associadas a análises qualitativas. Essa opção fundamenta-se no pressuposto de que os métodos quantitativos e qualitativos são complementares e permitem uma melhor aproximação ao objeto da avaliação quando realizados simultaneamente, pois possibilitam analisar os dados e informações referentes ao programa a partir de perspectivas distintas. Além disso, a metodologia de trabalho do



CGEE baseia-se na articulação entre conhecimento teórico, levantamento de dados primários e secundários e análises validadas em reuniões com especialistas.



**Figura 1** - Síntese da metodologia de trabalho CGEE  
Fonte: CGEE, 2016.

O levantamento de dados secundários é realizado por meio do acesso a bases de dados existentes (como a Plataforma Lattes, as bases *Web of Science* e Scopus, dentre outras a serem definidas pela necessidade de cada estudo) utilizando-se, sempre que possível, ferramentas já desenvolvidas pelo CGEE.

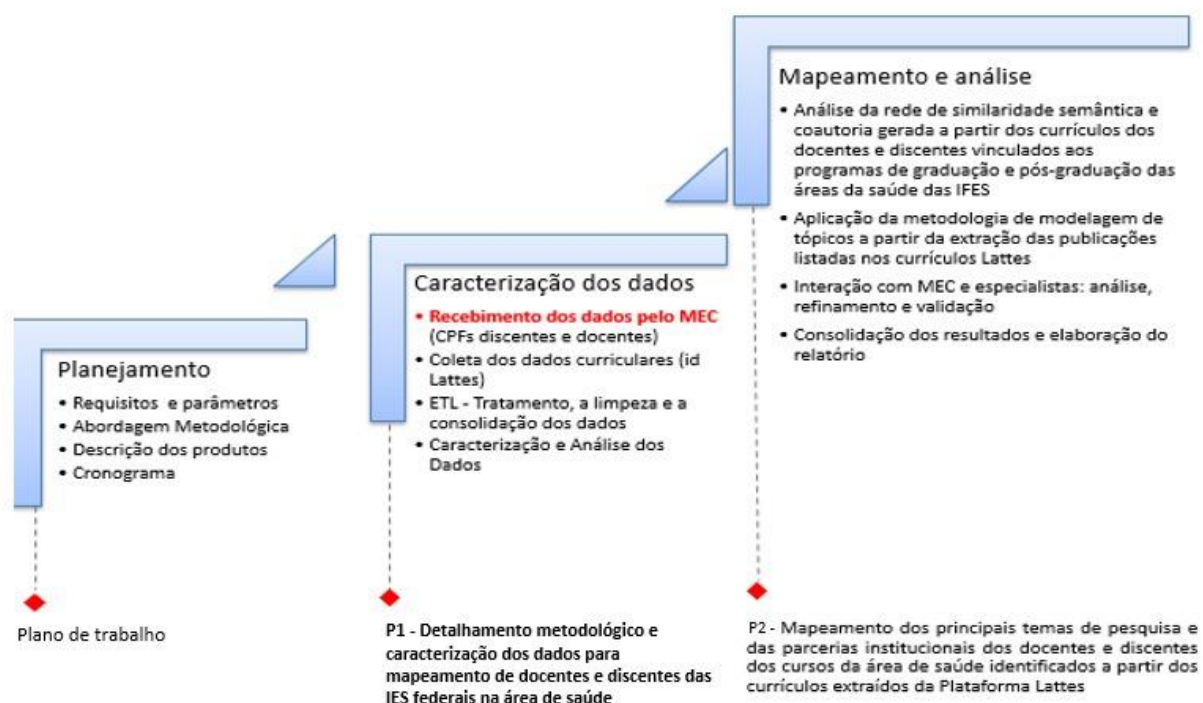
Além disso, por meio do Observatório em Ciência, Tecnologia e Inovação (OCTI), o CGEE tem desenvolvido metodologias e estudos que permitem o mapeamento e monitoramento sistemático da produção científica, no Brasil e no mundo, e das competências nacionais em Ciência, Tecnologia e Inovação. Essas abordagens permitem compreender melhor o estado da arte da produção do conhecimento, identificar tendências e temas emergentes, bem como mapear e compreender a dinâmica das redes de pesquisadores em temas estratégicos para o desenvolvimento sustentável no país, dando ênfase para sua distribuição regional.

Com base nessa *expertise*, a abordagem metodológica desse projeto consiste em gerar relatórios analíticos que permitam compreender como se organiza tematicamente as redes de recursos humanos em pesquisa, formada por docentes e discentes, nas IFES públicas da área de Saúde.

Este projeto utiliza, enquanto marco teórico-metodológico, a análise de redes científicas como instrumento para identificação e interpretação de competências instaladas no país. O levantamento de dados utilizados é realizado por meio do acesso a bases de dados existentes, mais especificamente a base Lattes, permitindo a mineração e exploração massiva de *big data*, objetivando elucidar sistematicamente relações entre suas variáveis, possíveis correlações e tendências (Lazaro & Giarolla, 2018).

Para execução dessas análises, ferramentas próprias do CGEE serão utilizadas neste estudo, dentre elas: i) o InsightNet, que permite a extração de currículos por diferentes entradas de buscas; ii) o InsightNet Browser, ferramenta que apoia a geração de análises e recursos de visualização dos dados apresentados; e iii) todos os recursos disponíveis no ambiente de análise da Plataforma do OCTI<sup>1</sup>.

A Figura 02 apresenta de forma esquemática as três etapas do estudo.



**Figura 2 – Etapas do projeto**

Fonte: Elaboração própria.

A Etapa 1 compreende o Planejamento, quando serão realizadas atividades preparatórias e de planejamento do estudo. A Etapa 2 envolve a coleta e tratamento dos dados para identificar as características populacionais e os metadados de identificação curricular do objeto de estudo, que possam permitir a extração dos currículos para a confecção do panorama temático, bem como estabelecer quais critérios e

<sup>1</sup> Disponível em: <<http://octi.cgee.org.br>>

características, a partir da análise de viabilidade, serão utilizadas para as análises temáticas do perfil estabelecido.

Por fim, a Etapa 3, destina-se ao mapeamento e análise, ou seja, nela serão aplicadas as metodologias previstas e realizadas as análises qualitativas. O objetivo é levantar informações e percepções cobrindo os objetivos geral e específicos do estudo. A partir dessas análises será elaborado o relatório final do projeto.

## **4. DETALHAMENTO METODOLÓGICO**

### **Caracterização dos dados**

Para a geração e análise do mapeamento destinado nessa ementa, a definição de seu escopo é passo exploratório inicial. Assim, devem ser acordados os recortes esperados para a identificação das competências de interesse do projeto. O levantamento de dados curriculares necessita da seleção prévia do conjunto de pesquisadores pertencentes ao escopo de área do mapeamento pretendido, a periodicidade desejada, bem como o nível de formação de maior interesse para os objetivos do estudo.

Recomenda-se a seleção de dez anos de informações curriculares para melhor aprofundamento em recortes cronológicos, com destaque para o levantamento de diferentes padrões temporais na rede de recursos humanos estudada.

O recorte também deverá levar em consideração qual período de vinculação dos pesquisadores – entre docentes e discentes – serão considerados de interesse do mapeamento com rastreabilidade nas IFES do estudo.

Recomenda-se selecionar: (i) ou o período exato de vinculação formal de cada pesquisador, o que demanda que esta informação esteja clara e objetiva na declaração curricular dos envolvidos; (ii) ou, que o período reflita apenas o início da vinculação formal de cada pesquisador, desconsiderando, portanto, sua eventual saída. Isto se deve ao fato de que seu conhecimento acumulado pode ainda ser identificado como um patrimônio associado às IFES, permitindo observar como a trajetória temática de seus pesquisadores se desdobrou no tempo, independente da finalização de seu vínculo formal.

Outro aspecto a ser considerado é o nível de formação pretendida. Geralmente, as informações em maior escala de disponibilidade e estruturação são aquelas associadas aos pesquisadores doutores, uma vez que possuem maior tempo de trajetória acadêmica. Cabe mencionar que, como a metodologia de redes, que será brevemente explanada no bloco seguinte, utiliza fonte de informações semânticas, a padronização geral dos currículos é uma condição desejada para um melhor resultado obtido.

De forma mais específica, currículos mais extensos, equivalentes e proporcionais entre si, produzem resultados mais robustos e de fácil identificação temática. Caso a opção de recorte seja também por mestres, recomenda-se que as redes semânticas – tanto de colaboração efetiva e potencial – sejam analisadas separadamente dos currículos associados aos doutores já formados.

Vale mencionar que uma vez definido o escopo de estudo, os dados curriculares passarão por avaliação da disponibilidade de informação, permitindo identificar o nível de preenchimento de seus currículos. Essa etapa também permite identificar possíveis lacunas esperadas em seus resultados. A possibilidade de recortar, cronologicamente, o início e fim de cada vínculo formal também será estudada nessa etapa.

Outras informações também serão levantadas, principalmente, com foco em dados institucionais, regionais e associados às produções técnicas e bibliográficas em cada currículo selecionado, bem como a atualização última de seus dados junto à Plataforma Lattes. Após tais explorações, medidas técnicas adicionais podem ser tomadas para o melhor tratamento dos dados necessários para as análises subsequentes.

### **Redes de competências humanas no país**

O trabalho técnico do presente estudo tem como abordagem principal a geração de redes complexas, com base em informações curriculares disponíveis na base nacional Lattes, de coordenação do CNPq. Estas redes se dividem em dois tipos. O primeiro tipo específico de monitoramento são as redes de colaboração científica, que utilizam da identificação de **coautorias internas** dos pesquisadores que compõem a população do levantamento e mede a intensidade de suas colaborações, indicando os principais domínios temáticos investigados, bem como a interação, em formato de rede científica, entre pesquisadores e suas instituições profissionais associadas. Este levantamento tem como resultado esperado a identificação de uma **rede de colaboração efetiva** entre os pesquisadores abordados pelo estudo.

Importante mencionar que a seleção da produção associada à coautoria pode ser selecionada, tanto de forma a incluir (i) trabalhos completos em anais, (ii) artigos completos publicados em periódicos, (iii) capítulos de livro e (iv) resumos publicados em anais, ou a selecionar foco em um subconjunto dessas opções. A utilização de informações oriundas da produção técnica, caso necessária, demandará tratamento e conversão adicionais e previamente realizadas dos dados.

Na sequência, também serão desenhadas as redes do tipo de **colaboração potencial**, que permite identificar convergências de atuação pela similaridade semântica das informações curriculares disponíveis. Este tipo de rede permite identificar domínios de pesquisa que mobilizam pesquisadores, sem necessariamente, indicar, uma colaboração efetiva entre eles. Desta forma, a rede produzida a partir de tal objetivo mede a intensidade de similaridade que seus currículos dividem, considerando cada dupla de currículos disponibilizados no recorte.

Ao fim de ambas gerações de redes, tanto as de colaboração efetiva quanto potencial, o estudo permitirá qualificar os principais agrupamentos temáticos que distribuem os recursos humanos em atuação nas áreas de Saúde, indicando, ainda, potenciais convergências temáticas entre instituições profissionais associadas e autodeclaradas, oportunidades de adensamento das redes científicas e a detecção de grupos potencialmente emergentes no recorte abordado.

Estes resultados irão compor relatório do Produto 2, que tem como objetivo principal qualificar a formação da rede de profissionais atuantes na área eleita para estudo. Além disso, o estudo permite gerar cruzamentos de informações curriculares, entre elas:

- Detecção de competências regionais em determinados agrupamentos temáticos;
- Identificação de temas associados aos pesquisadores mais recentemente formados nas áreas;
- Padrões de especialização temática por diferentes IFES identificados no recorte analisado.

Os resultados dependem da estruturação e disponibilização dos dados, que podem ser trabalhados de forma a gerar estatísticas descritivas do mapeamento, por meio do período selecionado para análise.

Um ponto adicional é a possibilidade de selecionar diferentes níveis de intensidade de similaridade curricular em cada rede, permitindo, assim, qualificar e aprofundar nas relações entre pesquisadores mais fortemente associados com temáticas estimadas.

### **Modelagem de tópicos das informações de produção científica e/ou tecnológico nos currículos selecionados**

A evolução e crescimento de aplicações e componentes digitais gerou consequentemente um volume de dados gigantesco. Diante a esse cenário surgiram diversas técnicas computacionais para sumarização e agrupamento de dados. A modelagem de tópicos, em especial, é uma das abordagens mais utilizadas para este fim. O princípio da modelagem de tópicos consiste em encontrar padrões em bases textuais, padrões estes que procuram descrever temas utilizando agrupamentos de palavras. A construção dos tópicos pode ser realizada por meio de algum dos diversos algoritmos presentes na literatura, os quais precisam de uma transformação nos dados para compreenderem uma linguagem natural. Seu funcionamento utiliza cálculos estatísticos ou fatoração de matrizes para compreender o contexto e temas de bases textuais.

Para isto, a primeira etapa na construção de tópicos é a transformação na representação dos dados. Algoritmos como o “*latent dirichlet allocation (LDA)*” utiliza a contagem de palavras nos documentos para estimar as probabilidades, enquanto outros algoritmos utilizam uma representação matricial, onde as linhas representam os documentos e as colunas as palavras. Para preencher essa matriz, uma técnica amplamente utilizada é o TF-IDF, esta metodologia calcula a importância de uma palavra em um documento em relação a uma coleção de documentos.

Em resumo, o LDA é um modelo probabilístico que descreve como os documentos são descritos. Nesse modelo as variáveis observáveis são os termos de cada documento e as não observáveis são as distribuições de tópicos. Os parâmetros das distribuições de tópicos, conhecidos como hiper-parâmetros, são dados *a priori*. Seguindo essa metodologia, utiliza-se a distribuição de Dirichlet na distribuição de tópicos, onde seu

resultado é utilizado para destinar as palavras aos documentos pertencentes a diferentes tópicos<sup>2</sup>.

Esta metodologia será utilizada para descrição mais detalhada dos agrupamentos temáticos observados na metodologia de redes semânticas, também através das informações curriculares, com foco nos títulos das publicações técnicas e/ou bibliográficas declaradas por cada pesquisador selecionado.

Serão produzidos dois inventários sobre os dados analisados, com foco em identificação de subtemas e também de possíveis temáticas mais multidisciplinares. Através da própria distribuição temática de tópicos estimados para cada agrupamento temático selecionado, serão disponibilizadas as possíveis linhas de atuação no interior de cada domínio de pesquisa previamente ilustrado na metodologia de redes. Desta forma, o relatório contará com um nível de especificação maior sobre quais são os temas que estruturam a rede de recursos humanos formados pelas IFEs pública em Saúde.

A metodologia também indicará os temas com maior distribuição em distintos agrupamentos temáticos, a partir de indicadores probabilísticos gerado com dados oriundos da modelagem, facilitando a observação de temas que ajudam a moldar mais de um domínio temático abordado pelos pesquisadores selecionados.

### **Visualização dos dados e outras explorações**

Para a análise dos dados, por meio de recursos instrumentais já citados, como o *InsightNet Browser*, o estudo produzirá redes navegáveis com as informações dos pesquisadores selecionados, permitindo compreender como estes recursos humanos interagem, tanto de forma efetiva quanto potencial.

Nesse ambiente interno de análise, as informações dos currículos analisados também conterão nuvens de palavras-chaves e quando possíveis, aplicações de outras técnicas de visualização de mapas científicos, produzidos pelo Observatório, como diagramas estratégicos, que revelam temas ascendentes, ou *timelines*, que permitem acompanhar no tempo, termos mais recentemente utilizados como descritores de pesquisas.

---

<sup>2</sup> Ver [https://www.researchgate.net/publication/221620547\\_Latent\\_Dirichlet\\_Allocation](https://www.researchgate.net/publication/221620547_Latent_Dirichlet_Allocation).

Abaixo, é possível verificar exemplo do instrumental utilizado para visualização dos dados:



**Figura 3** - Exemplo de rede navegável com informações curriculares e suas identificações temáticas, com dados oriundos da Plataforma Lattes

**Fonte:** elaboração própria.

A partir da figura 3, como exemplo, é possível identificar pesquisadores, representados pelos nós (ou círculos) distribuídos na imagem e também dos principais agrupamentos temáticos vinculados a formação das IFES em Saúde no país, representados pelas cores em cada região do mapeamento.

Estas redes poderão ser produzidas em diferentes temporalidades, a partir do marco temporal eleito para o estudo, facilitando, assim, a identificação de padrões na formação de recursos humanos na área no Brasil. Além desta funcionalidade, o Insightnet Browser, ferramenta utilizada para navegação das redes obtidas com os dados curriculares, também conterá informações associadas a cada pesquisador, quando requisitado e a depender do grau de sua disponibilidade, declarado na plataforma Lattes, entre elas:



- a. Origem institucional da última e/ou maior formação acadêmica
- b. UF de endereço profissional, quando disponível
- c. Instituição de endereço profissional, quando disponível
- d. Informações de projetos de pesquisa, entre elas, palavras-chave e seus resumos
- e. Lista de produções bibliográficas declaradas
- f. Lista de produções técnicas declaradas
- g. Lista de palavras-chaves indexadas nas produções citadas acima

## **5. DETALHAMENTO DOS PRODUTOS**

### **Produto – Caracterização dos dados para mapeamento de docentes e discentes das IES federais na área de saúde**

Este produto tem como objetivo definir a viabilidade técnica dos objetivos almejados sobre a produção do mapeamento temático, nos campos de interesse do escopo abordado. Para isso, é necessário construir um conjunto técnico de características populacionais com seus metadados de identificação curricular, que possam permitir a extração dos currículos para a confecção do panorama temático, bem como estabelecer quais critérios e características serão utilizadas para as análises temáticas do perfil estabelecido. As definições de escopo serão cotejadas com o levantamento da viabilidade de análises e recursos de dados.

Também devem ser acordados as faixas temporais que serão definidas como de interesse do mapeamento, o nível de formação acadêmica atribuído aos currículos analisados, os tipos de redes de recursos humanos pretendidas e outras características, tais como a especificação do tipo de produção que será analisada, com o objetivo de identificação dos contextos temáticos de pesquisa.

Essa etapa também consiste em promover o tratamento, a limpeza e a consolidação dos dados, a fim de evitar duplicações, sobreamentos ou ausência de informações. Os currículos selecionados deverão conter dados suficientes para a sua identificação junto a Plataforma Lattes, bem como a realização da tabulação de outras variáveis envolvidas nos currículos selecionados, com a finalidade de gerar um melhor manuseio e aprofundamento analítico no mapeamento temático analisado.

A realização do projeto bem como o prazo de entrega dos produtos depende do recebimento das informações necessárias, como por exemplo os CPFs da população a ser estudada.

**Produto - Mapeamento dos principais temas de pesquisa e das parcerias institucionais dos docentes e discentes dos cursos da área de saúde identificados a partir dos currículos extraídos da Plataforma Lattes**

Este produto tem como meta realizar um mapeamento temático geral da produção técnica-científica dos docentes e discentes, obtidos pelo desenho populacional aprovado previamente, a partir dos critérios de seleção curricular e de outras variáveis disponíveis e selecionadas, como faixa temporal, tipo de vínculo e/ou grau de formação, entre outras.

A partir desse mapeamento, serão geradas análises por agrupamentos temáticos identificados (conceitualizados como *clusters*) por meio das métricas de geração de rede (coautoria e similaridade semântica), identificando contextos temáticos envolvidos com o escopo determinado. Esses agrupamentos temáticos, bem como o panorama geral, serão analisados à luz das variáveis disponíveis pela plataforma Lattes, entre elas: região do endereço profissional, região da última titulação, instituições de obtenção do último título, ano de conclusão da titulação, entre outras.

Os agrupamentos temáticos permitirão identificar os assuntos de maior interesse pelos pesquisadores, por meio da indexação de palavras-chave autodeclaradas disponíveis, por currículo levantado. Algumas análises de tendências poderão ser aplicadas, quando disponíveis os metadados que são necessários para essa abordagem.

Além desses monitoramentos, os resultados permitirão manipular métricas de centralidade, para identificar os temas dos pesquisadores com características mais multidisciplinares e os assuntos mais recorrentes em perfis mais estrategicamente posicionados no conjunto temático analisado. A metodologia de modelagem de tópicos será utilizada para aprofundar nas subtemáticas de cada agrupamento previamente identificado e selecionado.

O produto também prevê análises por interação (semânticas e/ou coautorias) entre pesquisadores alocados profissionalmente em instituições diferentes, a fim de detectar possíveis padrões de parcerias institucionais. Este resultado dependerá da disponibilidade analisada no primeiro produto desse plano de trabalho. O panorama temático poderá também ser dividido entre as contribuições por discentes e as de

docentes, apontando para possíveis recorrências temáticas de interesse do escopo do estudo.

Por fim, o mapeamento poderá explicitar possíveis potenciais de parcerias, dada a afinidade temática identificada entre os currículos selecionados. O quadro geral dessas informações auxilia a compreender melhor os interesses acadêmico-científicos entre os pesquisadores analisados, bem como possíveis características institucionais por diferente segmento analisado. Para a realização das análises temáticas será necessária a interação com especialista (s) da área, bem como com representantes do MEC.

## 6. NECESSIDADE DE DADOS E INFORMAÇÕES

Para que este projeto possa ser realizado, o CGEE precisará ter acesso aos dados e às informações disponíveis na SESU/MEC sobre os docentes e discentes vinculados às IFES, em particular os CPF, para a extração dos currículos correspondentes na Plataforma Lattes. O recebimento desses dados é imprescindível para a realização do estudo. O tipo de informação utilizada para as análises variará de acordo com a qualidade dos dados disponíveis nos currículos levantados, bem como a disponibilidade e viabilidade de qualquer informação adicional a ser agregada.

O dimensionamento e análise dos trabalhos prevê a participação de colaboradores da SESU/MEC, com objetivo de viabilizar e acompanhar o andamento das atividades, definir medidas corretivas e garantir acesso aos dados e às informações, conforme mencionado anteriormente.

## 7. CRONOGRAMA DE TRABALHO DAS PRÓXIMAS ETAPAS

| Atividade                                                                                                                                                                                                                 | jan/22 | fev/22 | mar/22 | abr/22 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------|--------|--------|
| Recebimento dos dados                                                                                                                                                                                                     |        |        |        |        |
| Coleta, tratamento e carga dos dados                                                                                                                                                                                      |        |        |        |        |
| Identificação de especialistas                                                                                                                                                                                            |        |        |        |        |
| Rodadas de discussão com representante(s) do MEC                                                                                                                                                                          |        |        |        |        |
| Entrega do Produto - Caracterização dos dados                                                                                                                                                                             |        |        |        |        |
| Análise dos dados                                                                                                                                                                                                         |        |        |        |        |
| Rodadas de validação com especialista (s) e representante (s) do MEC                                                                                                                                                      |        |        |        |        |
| Entrega do Produto - Mapeamento dos principais temas de pesquisa e das parcerias institucionais dos docentes e discentes dos cursos da área de saúde identificados a partir dos currículos extraídos da Plataforma Lattes |        |        |        |        |