

**REDES DE PRODUÇÃO/COLABORAÇÃO
CIENTÍFICA NO ESTADO DO TOCANTINS**

SCIENTIFIC PRODUCTION/COLLABORATION
NETWORKS IN THE STATE OF TOCANTINS

*REDES DE PRODUCCIÓN/COLABORACIÓN CIENTÍFICA
EN EL ESTADO DE TOCANTINS*

João Pedro Noronha de Moraes Souza

Graduando em Ciência da Computação. Universidade Federal do Tocantins (UFT). E-mail: pedro.noronha@uft.edu.br | [Orcid.org/0009-0001-2618-3066](https://orcid.org/0009-0001-2618-3066)

Tanilson Dias Santos

Doutor em Engenharia de Sistemas e Computação, Professor do curso de Ciência da Computação. Universidade Federal do Tocantins (UFT). E-mail: tanilson.dias@uft.edu.br | [Orcid.org/0000-0002-0636-5751](https://orcid.org/0000-0002-0636-5751)

RESUMO:

Este trabalho apresenta os resultados da aplicação da Teoria dos Grafos na modelagem de redes de produção científica no estado do Tocantins, analisando colaborações entre pesquisadores e instituições. A modelagem considera os pesquisadores como vértices e as colaborações entre eles como arestas, permitindo visualizar a estrutura das interações acadêmicas. O estudo analisou dados coletados da plataforma OpenAlex entre janeiro e março de 2024, abrangendo publicações realizadas no período de 1998 a 2024. A análise focou em três instituições públicas — UFT, Unitins e IFTO — e utilizou as ferramentas NetworkX e Gephi para modelagem e visualização das redes. Foram aplicadas métricas como centralidade e detecção de comunidades para avaliar a conectividade e a integração das colaborações. Os resultados revelaram um crescimento expressivo das coautorias ao longo dos anos, especialmente na UFT, onde a rede se expandiu significativamente entre 2014 e 2024. A maior componente conexa da rede também cresceu, indicando maior integração entre os pesquisadores. Entretanto, desafios como a coleta incompleta de dados em algumas instituições foram identificados, reforçando a necessidade de uma infraestrutura mais acessível para o armazenamento e compartilhamento de informações científicas. As conclusões deste estudo podem auxiliar na formulação de políticas de incentivo à pesquisa, promovendo maior colaboração e integração no cenário acadêmico tocantinense.

PALAVRAS-CHAVE: Teoria dos Grafos; Redes de Coautoria; Produção Científica.

ABSTRACT:

This study presents the results of applying Graph Theory to model scientific production networks in the state of Tocantins, analyzing collaborations between researchers and institutions. The modeling considers researchers as vertices and collaborations between them as edges, allowing visualization of the structure of academic interactions. The study analyzed data collected from the OpenAlex platform between January and March 2024, covering publications from 1998 to 2024. The analysis focused on three public institutions — UFT, Unitins, and IFTO — using the NetworkX and Gephi tools for network modeling and visualization. Metrics such as centrality and community detection were applied to assess connectivity and collaboration integration. The results revealed a significant increase in co-authorships over the years, particularly at UFT, where the network expanded considerably between 2014 and 2024. The largest connected component of the network also grew, indicating greater integration among researchers. However, challenges such as incomplete data collection from some institutions were identified, emphasizing the need for more accessible infrastructure for storing and sharing scientific information. The

conclusions of this study can assist in the formulation of research support policies, promoting greater collaboration and integration within the academic landscape of Tocantins.

KEYWORDS: *Graph Theory; Co-authorship Networks; Scientific Production.*

RESUMEN:

Este trabajo presenta los resultados de la aplicación de la Teoría de Grafos en la modelización de redes de producción científica en el estado de Tocantins, analizando las colaboraciones entre investigadores e instituciones. La modelización considera los investigadores como vértices y las colaboraciones entre ellos como aristas, permitiendo visualizar la estructura de las interacciones académicas. El estudio analizó datos recopilados de la plataforma OpenAlex entre enero y marzo de 2024, abarcando publicaciones del período 1998 a 2024. El análisis se centró en tres instituciones públicas — UFT, Unitins e IFTO — y utilizó las herramientas NetworkX y Gephi para modelar y visualizar las redes. Se aplicaron métricas como centralidad y detección de comunidades para evaluar la conectividad y la integración de las colaboraciones. Los resultados revelaron un crecimiento significativo de las coautorías a lo largo de los años, particularmente en la UFT, donde la red se expandió considerablemente entre 2014 y 2024. El mayor componente conexo de la red también creció, indicando una mayor integración entre los investigadores. Sin embargo, se identificaron desafíos como la recopilación incompleta de datos en algunas instituciones, destacando la necesidad de una infraestructura más accesible para el almacenamiento y el intercambio de información científica. Las conclusiones de este estudio pueden ayudar en la formulación de políticas de apoyo a la investigación, promoviendo una mayor colaboración e integración en el ámbito académico de Tocantins.

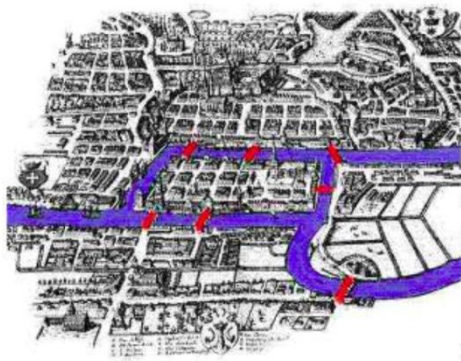
Palabras clave *Teoría de Grafos; Redes de Coautoría; Producción Científica.*

INTRODUÇÃO

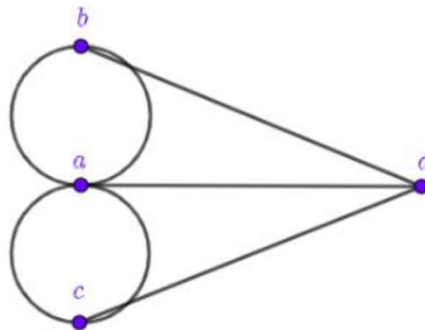
A Teoria dos Grafos é um ramo da Matemática e da Ciência da Computação que estuda as relações entre objetos de um determinado conjunto. Com essas relações é possível modelar diversos problemas reais utilizando estruturas de dados chamadas de grafos, (SZWARCFITER, 2018). Com os grafos, somos capazes de modelar problemas do cotidiano, tais como: problemas que envolvem fluxos em redes, utilizados em algoritmos de roteamento de veículos; problemas de conjunto dominante, investigação de clusters em redes sociais.

O relato histórico mais remoto registrado da Teoria dos Grafos aponta para a cidade de Königsberg em 1736 e é atribuído ao matemático suíço Leonhard Euler (1707-1783). Ele apresentou o, agora famoso, problema das 7 pontes de Königsberg. Euler associou cada ilha da cidade a um ponto, que hoje chamamos de vértice, e a cada ponte que liga esses pontos o que hoje chamamos de aresta. Com isso surgiu o que a literatura reconhece como sendo o primeiro grafo da história (BIGGS; LLOYD; WILSON, 1986). A Figura 1(a) ilustra um mapa simplificado da cidade de Königsberg, enquanto a Figura 1(b) apresenta o grafo relacionado ao problema.

Figura 1 – Abstração do problema das 7 pontes de Königsberg



(a) Mapa simplificado da cidade de Königsberg



(b) Grafo das 7 pontes de Königsberg

A Figura 1, é útil, em particular, porque representa o cerne do que foi a proposta desta pesquisa. Para entender o contexto, perceba que na época em que Euler viveu existia um enigma em sua cidade. A pergunta para a qual não havia resposta era a seguinte: Seria possível sair de uma determinada ilha da cidade de Königsberg, passar por todas as pontes uma única vez e retornar à ilha inicial?

As ferramentas matemáticas disponíveis à época não permitiam responder a esta pergunta com exatidão, logo foi necessário o desenvolvimento de uma técnica de análise adequada. Essa técnica deveria abstrair detalhes irrelevantes na modelagem do problema. Dessa forma, de posse do mapa da Figura 1 (a), Euler propôs o grafo apresentado na Figura 1(b) que representa a modelagem exata da

cidade de Konisberg com a seguinte semântica linguística: cada ilha da cidade recebeu um label (a, b, c, d) que foi modelado como um vértice do grafo. E quando existia uma ponte entre duas ilhas, era inserida uma aresta entre os vértices que representavam essas ilhas, dessa forma a Figura 1(b) representa efetivamente o modelo simplificado da cidade de Konisberg. Esses conceitos importantes da Teoria dos Grafos serão detalhados mais adiante neste trabalho.

Nos dias atuais, com frequência buscamos soluções muito complexas para problemas que talvez pudessem ser solucionados de forma mais prática. Imaginemos por um segundo que se ao invés de modelar as ilhas e pontes da cidade de Konisberg, nós fôssemos capazes de modelar as cidades e rodovias do estado do Tocantins. Quais informações o grafo resultante desta modelagem poderia nos trazer? E se ao invés de modelar ruas e cidades, fôssemos capazes de modelar linhas de transmissão de água, de energia elétrica e esgoto? Talvez possamos encontrar um conjunto de características que seja importante manter ou que seja importante que não exista no grafo obtido. Essas questões podem ser modeladas como problemas de conectividade em grafos.

Diagnosticar quais são as cidades ou rodovias mais importantes do estado seria equivalente ao problema de identificar medidas de centralidade em redes sociais. Em tempo, se conseguíssemos um recorte do subgrafos da internet com os principais portais de notícias do estado, seríamos capazes de identificar clusters, agrupamentos e comunidades de websites do estado. Outro tópico interessante de investigação seria a capilaridade, profundidade e alcance do conhecimento científico produzido no estado do Tocantins, talvez pudéssemos ser capazes de modelar essa rede de produção científica com base nas colaborações entre pesquisadores, instituições ou ainda temas de interseção de pesquisa entre diferentes grupos locais. Esse tipo de rede de pesquisa permite identificar potenciais de estudos e mapear a produtividade em diversas instâncias (tanto interno à universidade, quanto em todo o estado). Esse mapeamento pode ajudar a nortear o desenvolvimento de editais de fomento e produtividade, por exemplo. Todas essas informações são úteis do ponto de vista prático para, por exemplo, auxiliar na elaboração de planos de contingência, ou ainda diagnosticar necessidades de implantação de redundância em redes.

Muitos trabalhos na área da Ciência da Computação propõem modelagem de problemas diversos utilizando a Teoria dos Grafos, e.g. o trabalho de Santos et al (2020) estuda a propriedade Helly em Grafos de Interseção, e a principal motivação da pesquisa é a modelagem do problema de Layout de Circuitos digitais; há ainda trabalhos que estudam as propriedades estruturais de elementos químicos como hidrocarbonetos, através de análise espectral de um grafo, como Huckel (1931); outras aplicações e problemas modelados com grafos podem ser encontrados em Abreu et al (2007) e Szwarcfiter (2018); trabalhos que abordam

as redes sociais na produção científica dos programas de pós-graduação brasileiros, como o de Nascimento e Beuren (2011).

A análise das redes de produção científica no estado do Tocantins foi escolhida como foco desta pesquisa devido à escassez de estudos que explorem a produção e colaboração científica e institucional na região. Essa lacuna torna o tema não apenas relevante do ponto de vista acadêmico, mas também de interesse para a comunidade em geral.

Os resultados desta pesquisa destacam a importância de incentivar a produção científica no Tocantins, evidenciando a necessidade de políticas e estratégias eficazes para seu fortalecimento. Além disso, o estudo contribui para a avaliação das iniciativas já adotadas no fomento à pesquisa, oferecendo subsídios para seu aprimoramento.

A modelagem proposta fundamenta-se tanto em precedentes históricos significativos quanto em uma base teórica e bibliográfica sólida, garantindo o suporte necessário ao desenvolvimento da investigação.

ARCABOUÇO TEÓRICO

Nesta seção serão apresentados os fundamentos sobre a Teoria dos Grafos e os devidos métodos para análise das redes relacionadas à produção científica que podem ser modeladas no estado do Tocantins. A Teoria dos Grafos é um ramo da Matemática e da Ciência da Computação que estuda as relações entre objetos de um determinado conjunto.

Um *grafo* é uma estrutura composta por *vértices*, que representam os elementos individuais, e *arestas*, que correspondem às conexões ou relações entre esses elementos. Esses grafos podem ser de natureza diversa, dependendo das características das relações que representam. Denotamos um *grafo não-direcionado* por $G = (V, E)$, onde G representa o grafo, $V(G)$ faz referência ao conjunto de vértices do grafo G , que deve ser finito e não vazio. E , por último, $E(G)$ representa o conjunto de arestas do grafo G , que é composto por pares de elementos tomados de $V(G)$. Grafos não-direcionados, ou seja, não há uma distinção de direção na conexão entre dois vértices. Por outro lado, os grafos direcionados, grafos dirigidos ou digrafos, contém arestas com direção específica e são apropriados para modelar relações assimétricas, onde a direção da interação é significativa. Um grafo direcionado é definido por $G = (V, E)$ onde V é o conjunto de vértices e $E \subseteq V \times V$ é o conjunto de arestas, representadas como pares ordenados (u, v) , com $u, v \in V$. Em um grafo direcionado, cada aresta (u, v) indica uma relação dirigida do vértice u para o vértice v , ou seja, u é a origem e v é o destino da aresta. Além disso, os grafos podem ser ponderados, atribuindo-se valores às arestas, o que permite modelar a intensidade ou a frequência das interações. Grafos ponderados são definidos

como $G = (V, E, w)$, onde V é o conjunto de vértices, $E \subseteq V \times V$ é o conjunto de arestas, e $w: E \rightarrow \mathbb{R}$ é uma função de peso que atribui um valor real a cada aresta $e \in E$. Este é um aspecto interessante para análise de redes científicas, onde o peso de uma aresta pode refletir, por exemplo, a quantidade de publicações conjuntas entre dois pesquisadores ou a frequência de colaborações entre instituições. Em grafos, um caminho é uma sequência de vértices e arestas onde cada aresta conecta dois vértices consecutivos na sequência. É definido como (v_1, v_2, \dots, v_n) tal que cada par de vértices consecutivos (v_i, v_{i+1}) está conectado por uma aresta no grafo. Um *caminho* é considerado simples se todos os vértices e arestas na sequência são distintos, ou seja, nenhum vértice ou aresta é repetido ao longo do caminho. Um *caminho simples* não tem ciclos e não passa pelo mesmo vértice mais de uma vez. Um *ciclo* é um caminho que começa e termina no mesmo vértice, formando um laço. Caminho mínimo é o caminho de menor comprimento (número de arestas) entre dois vértices. Em grafos ponderados, onde as arestas têm pesos associados, o caminho mínimo é aquele cuja soma dos pesos das arestas é a menor possível.

O grau de um vértice é definido como o número de arestas que estão conectadas a ele. Em termos simples, o grau de um vértice corresponde à quantidade de outros vértices com os quais ele está diretamente ligado. Em grafos não direcionados o grau de um vértice v , denotado por $\deg(v)$, é o número total de arestas que incidem sobre v . Já em grafos direcionados o grau de um vértice é dividido em duas partes: grau de entrada, o número de arestas que chegam ao vértice v ; grau de saída, o número de arestas que saem do vértice v . O grau de um vértice é uma propriedade fundamental que influencia a conectividade e a importância de um vértice dentro da rede. A conectividade em grafos refere-se à capacidade de os vértices de um grafo estarem ligados uns aos outros por caminhos. Ela descreve a robustez da estrutura do grafo e a facilidade com que as informações ou interações podem se propagar entre seus vértices. Um grafo não direcionado é considerado conexo se, para qualquer par de vértices u e v , existe pelo menos um caminho que os conecta. Isso significa que é possível percorrer o grafo inteiro sem desconexões. Caso contrário, o grafo é dito desconexo e pode ser dividido em componentes conexos, que são subgrafos onde todos os vértices de cada componente estão conectados entre si, mas não com vértices de outros componentes. Em um grafo direcionado, a conectividade é definida em termos de caminhos direcionados. Um grafo direcionado é considerado fortemente conexo se existir um caminho direcionado de qualquer vértice u para qualquer outro vértice v . Se existir apenas um caminho entre dois vértices, independentemente da direção, o grafo é fracamente conexo. A conectividade também pode se referir ao número mínimo de vértices (ou arestas) que precisa ser removido para que o grafo se torne desconexo. Isso é chamado de conectividade de vértices ou conectividade de arestas, respectivamente. A conectividade de vértices avalia a vulnerabilidade do grafo a falhas ou ataques,

determinando quantos vértices devem ser removidos para que o grafo se fragmente em componentes desconexos. Essa propriedade está intimamente relacionada à topologia do grafo, já que a forma como os vértices e arestas estão organizados influencia diretamente a resiliência da rede. A topologia em grafos refere-se à forma como os vértices e as arestas de um grafo estão organizados e interconectados. Descreve a estrutura ou o padrão das conexões dentro do grafo, independentemente das propriedades geométricas, como a posição dos vértices no espaço. A análise da topologia de grafos é essencial para entender a organização e o comportamento de redes complexas em diversas áreas, como redes sociais, redes de transporte e redes biológicas. Na literatura, o termo “redes complexas” refere-se a grafos com uma estrutura topológica não trivial, compostos por vértices conectados por arestas (BARABÁSI et al., 2002).

A detecção de comunidades e a clusterização são técnicas da análise de redes que permitem a identificação de padrões estruturais e agrupamentos dentro de um grafo. Em muitos contextos, como redes sociais, redes biológicas e redes de colaboração científica, os vértices não se conectam de maneira aleatória; ao contrário, eles tendem a formar grupos altamente interconectados, denominados clusters ou comunidades. Formam-se subconjuntos de vértices que possuem maior densidade de conexões entre si do que com o restante da rede. Essa estrutura indica grupos de entidades que compartilham características comuns, como interesses acadêmicos similares em redes de coautoria ou grupos sociais coesos em redes de relacionamento.

A análise topológica de uma rede complexa pode ser feita pela utilização de parâmetros para a compreensão da estrutura e do comportamento dessas redes. Estes parâmetros são baseados nas definições de Teoria dos Grafos, explicados anteriormente, e a distribuição de grau é um desses. A distribuição de graus é representada por uma função de probabilidade $P(k)$, onde k é o grau de um vértice, e $P(k)$ é a fração ou a proporção de vértices que têm exatamente k conexões. O caminho mínimo médio, outro parâmetro, é a média dos comprimentos dos caminhos mínimos entre todos os pares de vértices em um grafo. O caminho mínimo médio fornece uma medida da “distância média” entre os vértices de um grafo, refletindo o quão compacta ou dispersa é a rede em termos de conectividade. O diâmetro de uma rede é a maior distância entre qualquer par de vértices no grafo e pode ser calculado pelo maior valor dentre os caminhos mínimos entre todos os vértices do grafo, caso seja fortemente conexo. Esses parâmetros combinados fornecem um quadro da estrutura da rede de colaboração científica e são essenciais para compreender as dinâmicas de cooperação e os padrões de produção científica no estado do Tocantins.

MÉTODOS E MATERIAIS

A pesquisa realizada neste trabalho é de natureza exploratória e descritiva, tendo como objetivo principal a análise das redes de produção científica no estado do Tocantins a partir de dados coletados. A escolha por este tipo de pesquisa se deve à necessidade de mapear, descrever e interpretar as colaborações acadêmicas e institucionais no estado.

Na seção anterior, foram apresentados os fundamentos teóricos que embasam a análise, com uma revisão da Teoria dos Grafos. A seção “Métodos e Materiais” detalha o processo de coleta de dados, a modelagem dos grafos e os métodos de análise empregados. Em seguida, seção “Dados e Discussões”, a apresentação e discussão dos resultados obtidos. Por fim, as conclusões abordarão a relevância dos achados para a comunidade acadêmica e o potencial impacto no fomento à pesquisa no estado.

Para modelar as redes relacionadas à produção científica no estado tocaninense devemos ter os dados da produção científica, como por exemplo os nomes de autores, título dos trabalhos, tópicos, instituições, data de publicação, número de citações, etc. Há certa dificuldade técnica na coleta desses dados. Revistas científicas possuem modelos de trabalhos diversos, por exemplo, o que pode dificultar nos tipos de dados que podem ser coletados. A plataforma OpenAlex possui uma base de dados de produções científicas, (PRIEM et al., 2022), e foi utilizada para mitigar os problemas relacionados à coleta dos dados brutos. Em primeiro momento foram escolhidas 5 instituições relevantes do estado, 3 públicas e 2 privadas, para analisar suas produções: Universidade Federal do Tocantins (UFT), Universidade Estadual do Tocantins (Unitins), Instituto Federal do Tocantins (IFTO), Universidade Luterana do Brasil (Ulbra) e Centro Universitário Católica do Tocantins (UniCatólica). Os dados do UniCatólica não foram encontrados na plataforma OpenAlex. Os dados da UFT, Unitins, IFTO e Ulbra puderam ser coletados. Entretanto, a Ulbra não é uma instituição exclusivamente tocaninense e seus dados não representam unicamente as produções feitas no estado. Por esse motivo, apenas as redes das 3 instituições públicas, UFT, Unitins e IFTO, foram modeladas.

Os processos de consumir a API e o arquivo CSV (formatos de coleta oferecidos pelo OpenAlex), modelar os grafos, analisar os parâmetros e construir os gráficos foram realizados com o uso da linguagem de programação Python (versão 3.12.3) em um ambiente Jupyter (versão 6.5.4), empregando as bibliotecas scipy (versão 1.13.0), pandas (versão 2.0.3), numpy (versão 1.24.3), matplotlib (versão 3.8.4) e networkx (versão 3.1). Os gráficos foram produzidos com o estilo específico da biblioteca Science Plots, (GARRETT et al., 2021). Para definir a visualização das redes, que foram modeladas de forma adequada utilizando a biblioteca NetworkX, foi empregada a ferramenta Gephi (versão 0.10.0), um software especializado na exploração, visualização e análise de grafos complexos. O uso do Gephi permitiu a criação de representações visuais

detalhadas das redes de colaboração científica, facilitando a compreensão das estruturas de relacionamento entre os pesquisadores e instituições. Além disso, o software oferece uma variedade de algoritmos de layout, métricas de centralidade e ferramentas de análise de comunidades, proporcionando uma visualização interativa e flexível dos dados, essencial para a interpretação das interações e padrões presentes nas redes modeladas.

RESULTADOS E DISCUSSÃO

Durante o desenvolvimento desta pesquisa, algumas dificuldades técnicas e metodológicas foram identificadas. Uma das principais barreiras foi a coleta de dados referente à produção científica nas instituições do estado do Tocantins. O uso da plataforma OpenAlex facilitou esse processo, mas ainda assim houve desafios. Um exemplo foi a falta de dados completos para certas instituições, como o Centro Universitário Católica do Tocantins (UniCatólica), cujas informações não estavam disponíveis na plataforma. Além disso, a Ulbra, uma das instituições inicialmente consideradas, apresentou dados que não são exclusivamente tocaninenses, o que comprometeu sua inclusão na análise final.

A Tabela 1 apresenta a quantidade de trabalhos relacionados às instituições e a quantidade de autores da rede, de 1998 a 2024. A documentação do OpenAlex informa que o seu tipo de dado para instituições são universidades e outras organizações às quais os autores afirmam estar afiliados. Os trabalhos são documentos acadêmicos como artigos de periódicos, livros, conjuntos de dados e teses. Esse número de autores não representa somente aqueles que estão vinculados às instituições, mas sim os que produziram ciência associados de alguma forma às instituições.

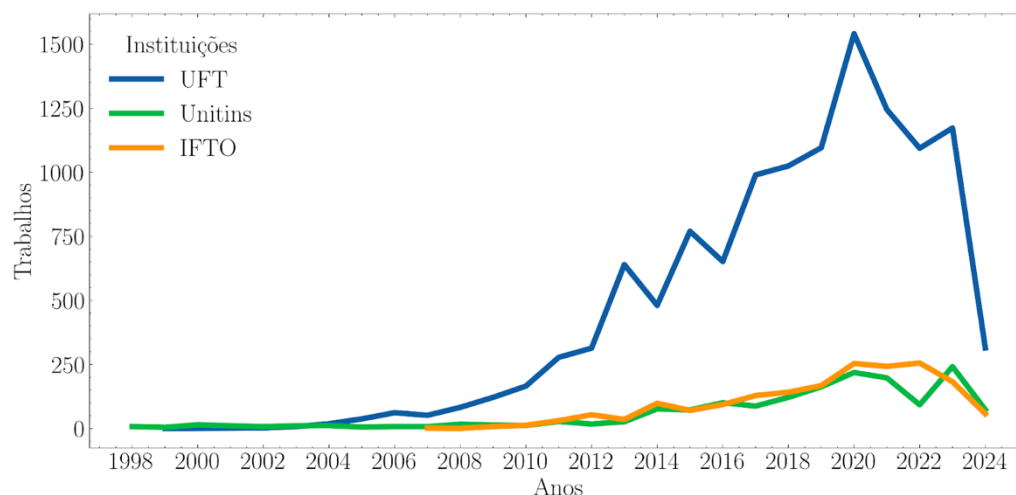
Tabela 1 – Número de trabalhos e autores das redes

Instituição	UFT	Unitins	IFTO
Quantidade de Trabalhos	12.170	1.658	1.842
Quantidade de Autores	21.557	4.564	5.238

A Figura 2 é um gráfico que ilustra as progressões das publicações entre 1998 a 2024. Há um crescimento depois do ano de 2010 em todas as instituições. No ano de 2020 houve um aumento significativo em relação aos anos anteriores e seu patamar não foi atingido novamente nos anos posteriores, com exceção da Unitins que foi a única instituição que em 2023 obteve um número maior de publicações em relação ao ano de 2020. É razoável dizer que a Pandemia de COVID-19 influenciou a produção científica do período. Outros trabalhos

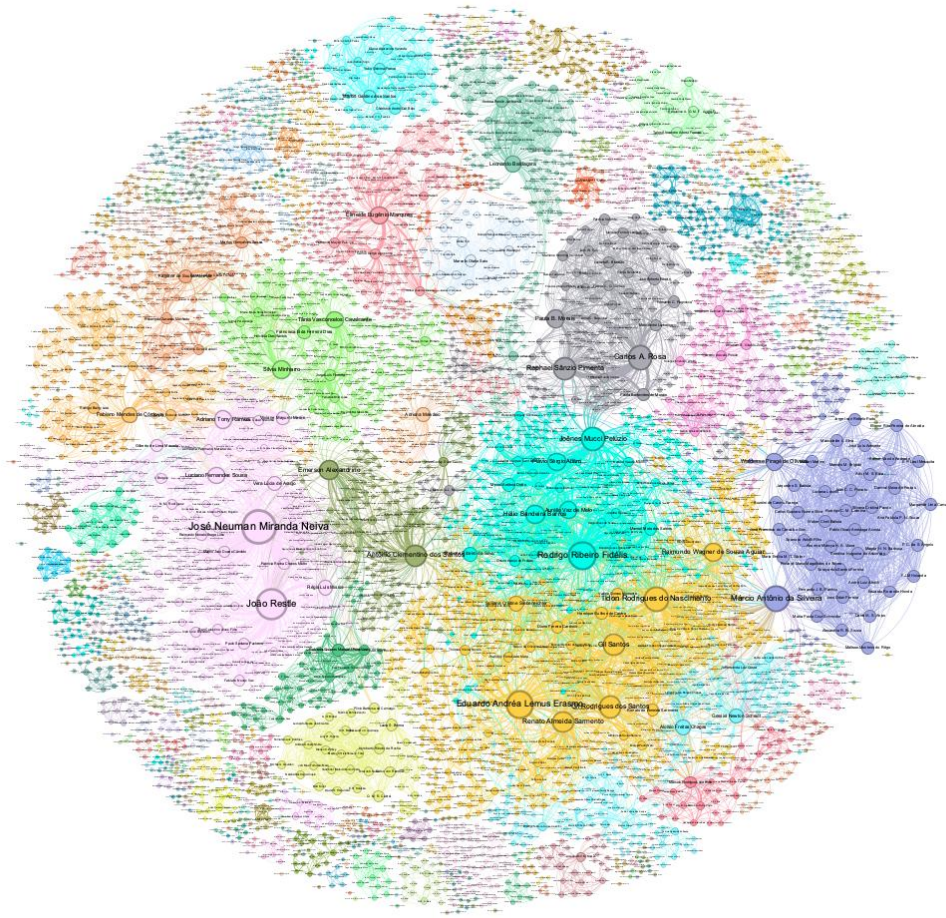
exploram mais detalhadamente esse caso, como, por exemplo, a mudança na dinâmica da produção sobre temas não relacionados ao vírus dentro das Ciências da Saúde durante o período pandêmico, (RAYNAUD et al., 2021).

Figura 2 – Linha do tempo de publicações científicas



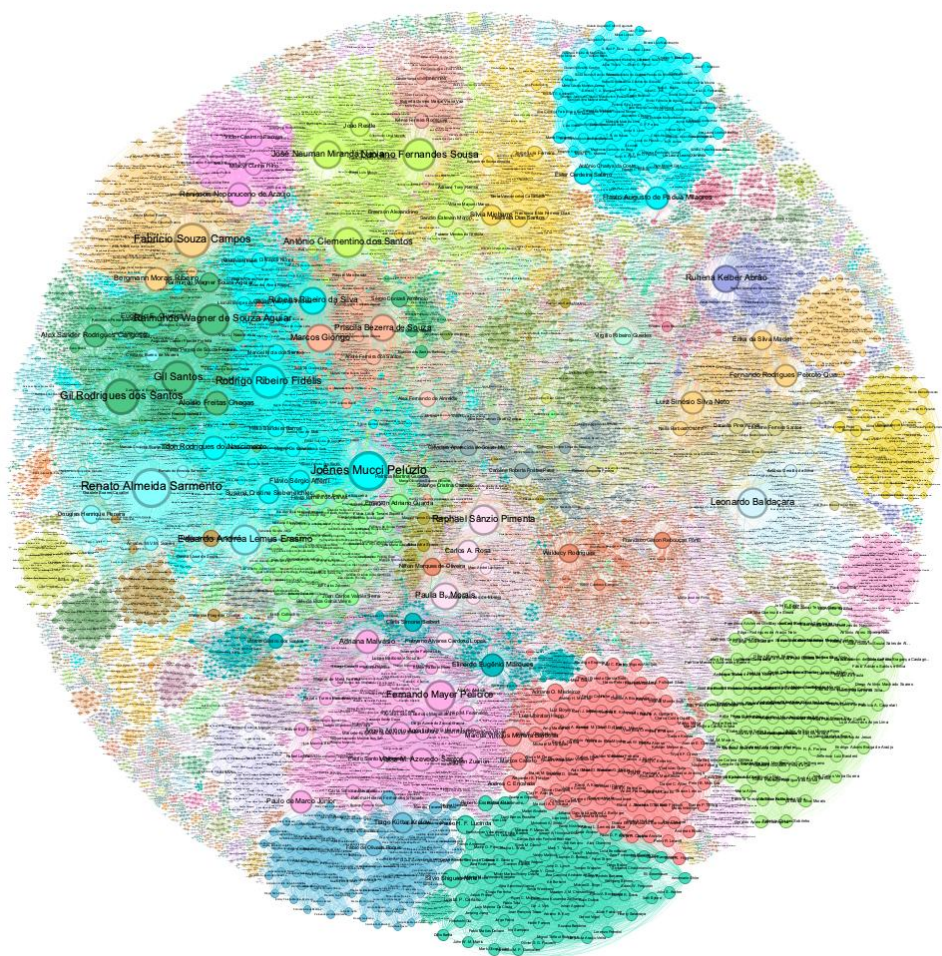
As Figuras 3 e 4 ilustram as redes de coautoria da UFT nos anos de 2014 e 2024, respectivamente. Em 2014, a rede contava com 4.486 vértices e 16.313 arestas, além de 559 comunidades identificadas. Já em 2024, observa-se um crescimento da rede, que passou a contar com 21.557 vértices, 131.610 arestas e 1.069 comunidades. Outro aspecto importante foi a expansão da maior componente conexa da rede. Em 2014, essa componente abrangia 60,19% dos vértices totais, enquanto em 2024, esse valor subiu para 86,32%, indicando uma maior integração entre os autores e uma conectividade mais robusta na rede de coautoria da UFT.

Figura 3 – Rede de coautoria da UFT em 2014



Perceba que na Figura 3, cada agrupamento de vértices, nesta pesquisa denotada por cluster, está destacada por uma cor diferente. Dessa forma, clusters de cores diferentes representam autores que possuem no mínimo uma publicação em comum no dataset analisado. Outro fato de importante observação é que o tamanho do vértice é proporcional a quantidade de colaborações encontradas no momento da coleta. De forma análoga a espessura da aresta é proporcional à quantidade de colaborações existentes entre dois autores. Observe que na Figura 3, se destacam os vértices dos pesquisadores José Neuman Miranda Neiva, Eduardo Andreia Lemos..., Márcio Antonio da Silveira, Rodrigo Ribeiro Fidélis, entre outros. Como estes vértices estão em maior destaque no cluster, isso pode indicar que estes autores são de alguma forma liderança em seu cluster as áreas de pesquisa. Obviamente clusters maiores indicam a interação de um grupo de pesquisa maior, enquanto cluster menores indicam que os autores possuem colaborações mais restritas em suas produções.

Figura 4 – Rede de coautoria da UFT em 2024



A Figura 4 ilustra a evolução da rede de coautoria da UFT em um intervalo de dez anos com relação aos dados ilustrados na Figura 3. É possível notar que o grafo de colaboração construído é bem mais denso, onde o número de comunidade da rede de colaboração é quase o dobro da analisada em 2014. Com relação ao número de vértices/autores na colaboração, é possível notar que o número de autores na rede subiu quase 5 vezes. Além disso, nota-se que há diversos cluster com vértices com tamanhos relativamente homogêneos, o que indica que há uma produção e interação constante/contínua entre os vértices que compõe o cluster, isso indica que o crescimento em produção ocorre em conjunto entre os membros do cluster. Outros clusters ainda apresentam características de possuir um vértice dominante, onde o grau dele é bem maior que os demais, visualmente notado pelo tamanho maior dentre os outros elementos do cluster ao qual pertence.

As Figuras 5 e 6 ilustram as redes de coautoria da Unitins e do IFTO no ano de 2024, respectivamente. A rede de coautoria da Unitins é composta por 4.564

vértices e 18.465 arestas, com sua maior componente conexa englobando 45,92% dos vértices. Já a rede do IFTO apresenta 5.238 vértices e 21.845 arestas, sendo que sua maior componente conexa abrange 39,77% dos vértices.

Figura 5 – Rede de coautoria da Unitins em 2024

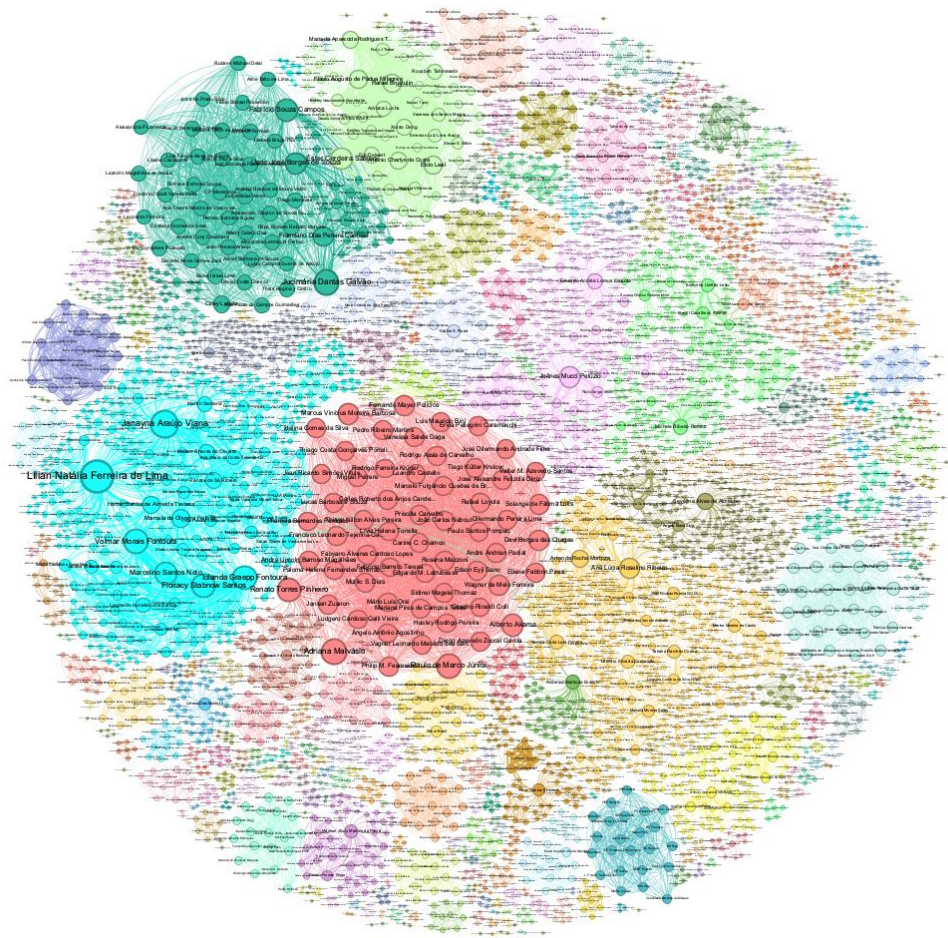
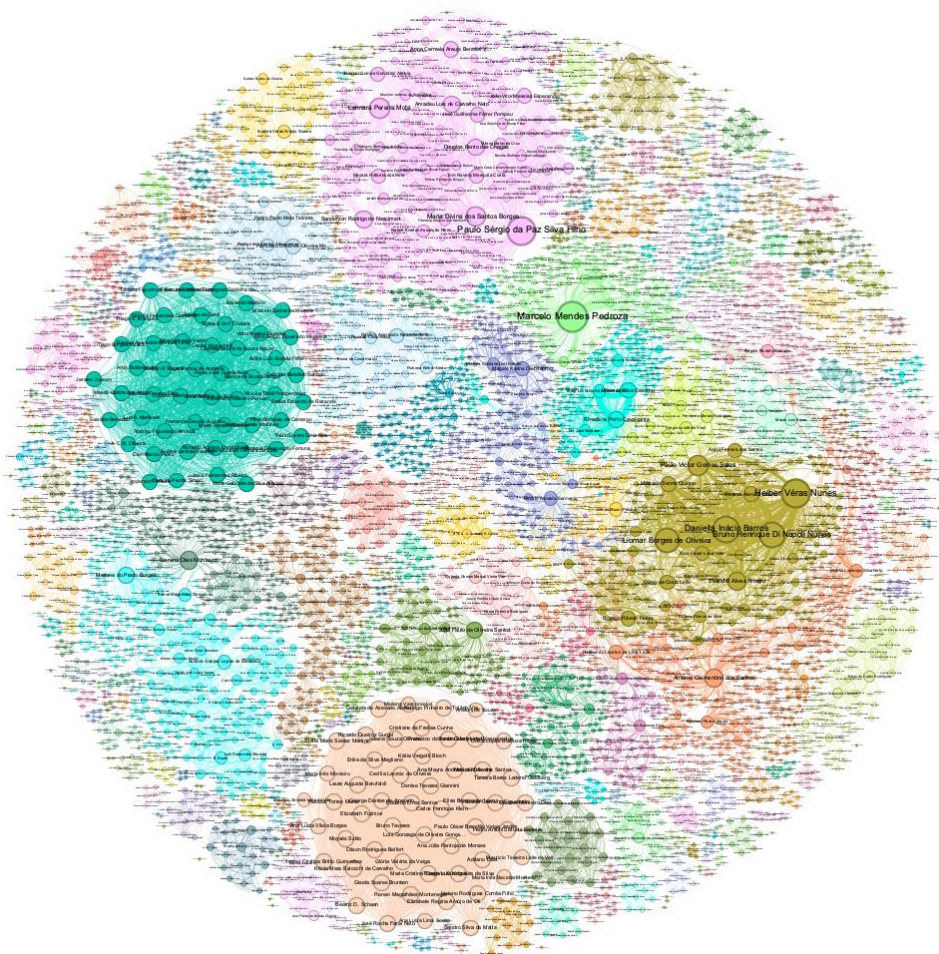


Figura 6 – Rede de coautoria do IFTO em 2024



CONSIDERAÇÕES FINAIS

A aplicação da Teoria dos Grafos na modelagem de redes científicas demonstrou ser uma ferramenta poderosa para entender as interações e colaborações no contexto acadêmico do estado do Tocantins. A análise das redes de coautoria, representadas por grafos, permitiu visualizar o crescimento e a evolução das colaborações entre pesquisadores e instituições ao longo do tempo. O uso de ferramentas como o NetworkX para preparação de dados e o Gephi para visualização e análise aprofundada das redes. A evolução da rede de coautoria da UFT entre 2014 e 2024 demonstra um aumento significativo na integração dos autores, com um crescimento no número de vértices, arestas e comunidades. Esse fenômeno reflete a expansão da produção científica e a intensificação das colaborações na instituição. Da mesma forma, as redes da Unitins e do IFTO, analisadas em 2024, mostram estruturas robustas de colaboração, embora ainda com componentes desconexas, o que indica áreas para possível fortalecimento de interações futuras.

Os desafios enfrentados na coleta de dados, como a falta de informações completas de algumas instituições, e a necessidade de adequar formatos de diferentes fontes, destacam a importância de uma infraestrutura de dados mais integrada e acessível para facilitar a pesquisa e o mapeamento da produção científica. Trabalhos futuros podem abordar outros tipos de redes relacionadas a produção científica, como as redes de citações, tópicos, palavras-chave etc.

Esses resultados não apenas contribuem para uma compreensão mais profunda das redes de colaboração científica no Tocantins, mas também oferecem subsídios importantes para a formulação de políticas de incentivo à pesquisa. O mapeamento das colaborações e o diagnóstico da conectividade entre instituições e pesquisadores podem guiar iniciativas de fomento que promovam maior integração e produtividade, potencializando o desenvolvimento científico na região.

Agradecimentos

O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq – Brasil.

Referências Bibliográficas

- BARABÁSI, Albert-Laszlo et al. **Evolution of the social network of scientific collaborations**. Physica A: Statistical mechanics and its applications, v. 311, n. 3-4, p. 590-614, 2002.
- BASTIAN, Mathieu; HEYMANN, Sebastien; JACOMY, Mathieu. **Gephi: an open source software for exploring and manipulating networks**. In: Proceedings of the international AAAI conference on web and social media. 2009. p. 361-362.
- BIGGS, N.; LLOYD, E. K.; WILSON, R. J. **Graph Theory**, 1736-1936. [S.l.]: Oxford University Press, 1986.
- BONDY, John Adrian; MURTY, Uppaluri Siva Ramachandra. **Graph theory**. Springer Publishing Company, Incorporated, 2008.
- ERDŐS, Paul et al. **On the evolution of random graphs**. Publ. math. inst. hung. acad. sci, v. 5, n. 1, p. 17-60, 1960.
- NASCIMENTO, Sabrina do; BEUREN, Ilse Maria. **Redes sociais na produção científica dos programas de pós-graduação de ciências contábeis do Brasil**. Revista de Administração Contemporânea, v. 15, p. 47-66, 2011.
- SZWARCFITER, Jayme Luiz. **Teoria computacional de grafos: Os Algoritmos**. Elsevier Brasil, 2018.
- ESTRADA, Ernesto. **The structure of complex networks: theory and applications**. American Chemical Society, 2012.
- PRIEM, Jason; PIWOWAR, Heather; ORR, Richard. **OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts**. arXiv preprint arXiv:2205.01833, 2022.

GARRETT, John et al. **garrettj403/SciencePlots**: 2.1. 1. Zenodo, 2023.

RAYNAUD, Marc et al. **Impact of the COVID-19 pandemic on publication dynamics and non-COVID-19 research production**. BMC medical research methodology, v. 21, p. 1-10, 2021.