

Redes de Coautoria Científica nas Instituições de Ensino Superior do Tocantins: Uma Análise Topológica e Estrutural

João Pedro Noronha de M. Souza¹, Tanilson Dias dos Santos²

¹Gerência de Sistemas – Secretaria Estadual de Educação do Tocantins (SEDUC-TO)
Palmas – TO – Brasil

²Curso de Ciência da Computação – Universidade Federal do Tocantins (UFT)
Palmas – TO – Brasil

joao.souza@educ.to.gov.br, tanilson.dias@uft.edu.br

Abstract. *This paper analyzes scientific co-authorship networks of five Higher Education Institutions (HEIs) in Tocantins, Brazil – UFT, Unitins, IFTO, UFNT, and CEULP – covering 1998–2024 (16,600 papers; 32,783 authors), using Graph Theory and Complex Network analysis. Networks were modeled as weighted undirected graphs; topological metrics were validated against Erdős–Rényi and Configuration Model null models. Degree distributions are best described by a power-law model under the AIC criterion ($\alpha \in [2.17, 3.10]$) across all networks, which also display Small-World properties; formal verification via the Clauset, Shalizi and Newman (2009) procedure remains prioritised future work. UFT dominates in scale and cohesion (giant component: 85.85%). An original contribution is the analysis of UFNT – split from UFT in 2021 – as a measurable case of scientific emancipation, with only nine exclusive publications after three years, providing a quantifiable baseline for monitoring institutional autonomy. The proposed pipeline (dual null models + AIC + robustness tests) is replicable for other peripheral Brazilian regions.*

Resumo. *Este artigo analisa as redes de coautoria científica de cinco IES do Tocantins – UFT, Unitins, IFTO, UFNT e CEULP – cobrindo 1998–2024 (16.600 trabalhos; 32.783 autores), com Teoria dos Grafos e Redes Complexas. As redes foram modeladas como grafos não dirigidos ponderados, e métricas topológicas foram validadas por dois modelos nulos (Erdős–Rényi e Modelo de Configuração). Em todas as redes, a distribuição de graus é mais bem descrita por um modelo de lei de potência segundo o critério AIC ($\alpha \in [2,17; 3,10]$) e todas exibem propriedades de mundo pequeno; a verificação estatística formal pelo procedimento de Clauset, Shalizi e Newman (2009) constitui trabalho futuro prioritário. A UFT domina em escala e coesão (componente gigante: 85,85%). Uma contribuição original é a análise da UFNT – desmembrada da UFT em 2021 – como caso de emancipação científica mensurável: após três anos de operação independente, registrou apenas 9 publicações exclusivas, fornecendo linha de base quantificável para monitorar sua autonomia institucional. O pipeline proposto (dois modelos nulos + AIC + robustez) é replicável para outras regiões brasileiras periféricas.*

1. Introdução

Redes de colaboração científica são representações formais das relações de coautoria entre pesquisadores: os autores correspondem a vértices do grafo e a publicação conjunta de um trabalho gera uma aresta entre os pesquisadores envolvidos. Esse formalismo, consolidado por [Newman 2001] e [Barabási et al. 2002], permite aplicar as ferramentas da Teoria dos Grafos ao estudo da ciência, identificando padrões estruturais como a distribuição de graus em lei de potência – característica das chamadas *redes livres de escala*, em que poucos pesquisadores concentram um número desproporcionalmente grande de coautores [Barabási and Albert 1999] – e as propriedades de *mundo pequeno*, que descrevem redes com alta tendência de agrupamento local e distâncias médias entre vértices surpreendentemente curtas [Watts and Strogatz 1998]. No contexto brasileiro, estudos sobre redes de colaboração realizados a partir da Plataforma Lattes documentaram heterogeneidade institucional acentuada e a formação de centros disciplinares de grande influência [Mena-Chalco et al. 2014]. Investigações em escala regional indicam que instituições menores tendem a apresentar menor coesão interna e componentes gigantes mais restritas [Maciel et al. 2013, Brito et al. 2020], enquanto eventos externos, como a pandemia de COVID-19, produzem efeitos detectáveis sobre o volume e os padrões de colaboração [Raynaud et al. 2021].

O estado do Tocantins, integrante da região Norte do Brasil, possui sistema de ensino superior em expansão, com cinco IES identificadas na plataforma OpenAlex: a Universidade Federal do Tocantins (UFT), a Universidade Estadual do Tocantins (Unitins), o Instituto Federal do Tocantins (IFTO), a Universidade Federal do Norte do Tocantins (UFNT) e o Centro Universitário Luterano de Palmas (CEULP). IES de regiões periféricas como o Tocantins apresentam dinâmicas estruturais distintas das grandes universidades do Sudeste – menor densidade de colaboração, maior dependência de coautores externos e componentes gigantes mais restritas – fenômenos que escapam à análise de grandes redes nacionais agregadas [Mena-Chalco et al. 2014].

Um caso particularmente expressivo é o da UFNT: desmembrada da UFT em 2021, essa universidade recentemente criada constitui um fenômeno de *emancipação científica mensurável em rede*. A evolução de sua dependência estrutural da UFT pode ser acompanhada quantitativamente por métricas de grafos, que podem ter implicações diretas para políticas de fomento à pesquisa em instituições recém-criadas.

A análise sistemática das redes de coautoria das cinco IES tocantinenses permanece como uma lacuna na literatura; estudos anteriores [Souza and dos Santos 2025] cobriram apenas parte das IES, sem validação por modelos nulos ou testes de robustez.

O presente trabalho preenche essa lacuna ao propor um *pipeline metodológico replicável* (validação por dois modelos nulos complementares (ER e CM), seleção de distribuições por AIC e testes de robustez estrutural por múltiplos critérios de centralidade) estruturado de forma modular para ser adaptável a estudos de coautoria regionais em outros contextos brasileiros com perfil semelhante. Os objetivos específicos abrangem a caracterização da produção científica por instituição, a análise topológica das redes individuais e da rede união, a identificação de pesquisadores centrais e áreas de concentração, e a validação estatística das propriedades topológicas observadas. Código-fonte e dados

estão disponíveis no repositório do projeto¹.

2. Trabalhos Relacionados

O trabalho de [Barabási and Albert 1999] demonstrou que redes livres de escala emergem por conexão preferencial; [Barabási et al. 2002] documentaram essa propriedade especificamente em redes de coautoria científica. As propriedades de mundo pequeno foram formalizadas por [Watts and Strogatz 1998] e quantificadas pelo índice ω de [Humphries and Gurney 2008]. Tais formalismos e propriedades são de interesse para a análise de redes de coautoria em geral.

No Brasil, [Mena-Chalco et al. 2014] caracterizaram as redes de coautoria de 1.131.912 pesquisadores do Lattes em oito grandes áreas do conhecimento (1990–2010), documentando heterogeneidade significativa entre áreas via componente gigante, grau médio, agrupamento e assortatividade. [Newman 2001], por sua vez, reportou $\alpha \approx 2,5$ para redes de coautoria em Física, valor que serve de referência para situar os expoentes aqui obtidos no espectro da literatura. [Maciel et al. 2013] e [Brito et al. 2020] reforçam que tamanho e coesão das redes variam conforme o contexto regional e institucional. Estes trabalhos nacionais têm como objeto de estudo redes de coautoria de pesquisadores brasileiros; tais dados são importantes para comparar resultados desta pesquisa com resultados de outras da literatura.

[Raynaud et al. 2021] documentaram impacto da pandemia sobre a produção científica. O resultado demonstra a influência de fatores externos sobre a produção científica.

Nenhum desses estudos combina, em uma única análise regional, dois modelos nulos distintos (ER e CM), seleção de distribuições por AIC e testes de robustez por múltiplos critérios de centralidade: o pipeline proposto aqui é estruturado de forma modular (cada etapa pode ser substituída ou estendida independentemente) tornando-o adaptável aos diversos contextos regionais brasileiros com dados equivalentes. O presente trabalho distingue-se ainda por abranger cinco IES de uma região periférica do Norte do Brasil e analisar um caso singular de emancipação institucional mensurável em rede.

3. Metodologia

O design metodológico deste trabalho é orientado pelo objetivo de caracterizar estruturalmente o ecossistema científico de uma região periférica com validação estatística, aspecto que demandou um processo iterativo, em lugar de uma pipeline linear. A Figura 1 ilustra as sete etapas desse processo. Os laços de retorno (representados pelas setas tracejadas) foram necessários neste projeto por razões concretas: ambiguidades na desambiguação de autores só foram detectadas na fase de análise de métricas, exigindo retorno ao pré-processamento para refinar o dicionário de autores; e parâmetros de modelagem (principalmente limiares de peso das arestas) precisaram de ajuste após a visualização das redes geradas.

Os dados foram extraídos da plataforma OpenAlex [Priem et al. 2022] em formato CSV; decisão metodológica que combina duas vantagens: contornar o limite de requisições da API gratuita e acessar uma base de cobertura especialmente ampla para

¹<https://github.com/jpnoronhaa/networks-tocantins>

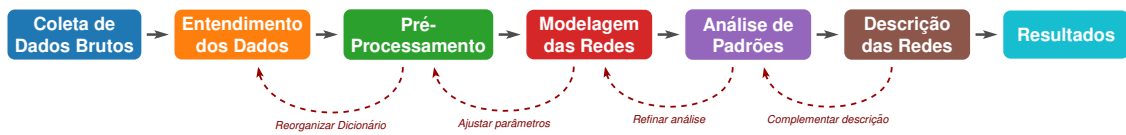


Figura 1. Fluxograma do processo metodológico adotado nesta pesquisa.

IES brasileiras de menor porte, frequentemente sub-representadas em Scopus e Web of Science. Das 18 IES cadastradas no e-MEC para o Tocantins, cinco foram localizadas na plataforma: UFT, Unitins, IFTO, UFNT e CEULP. O recorte temporal abrange 1998 a 2024, com coleta realizada em 29 de maio de 2025. O pré-processamento removeu registros sem autores e entradas duplicadas; a desambiguação de autores e a normalização de nomes institucionais foram realizadas pelos próprios mecanismos do OpenAlex.

A implementação utilizou Python 3.12 com `networkx` 3.4.2, `pandas` 2.2.3 e `scipy` 1.15.2. As visualizações foram geradas no Gephi 0.10.1 [Bastian et al. 2009] com layout Force Atlas 2 [Jacomy et al. 2014] e detecção de comunidades por Louvain [Blondel et al. 2008]; o tamanho dos vértices é proporcional ao grau.

As redes foram modeladas como grafos não dirigidos ponderados $G = (V, E, w)$, onde vértices representam autores, arestas indicam coautoria e os pesos $w(u, v)$ correspondem ao número de publicações conjuntas. Construiu-se ainda a rede união do Tocantins e séries de redes anuais cumulativas para 1998–2024.

Para cada rede extraíram-se: número de vértices e arestas; percentual da maior componente conexa (MCC); grau médio; coeficiente de agrupamento médio (\bar{C}); menor caminho médio (\bar{D}) e assortatividade (r). As métricas reportadas na Tabela 3 foram calculadas sobre a versão binária (não ponderada) do grafo, conforme prática padrão em estudos de coautoria; métricas ponderadas (força de vértice, *weighted clustering coefficient*) constituem extensão a ser explorada em trabalho futuro. A distribuição de graus foi ajustada a cinco modelos: Lei de Potência, Exponencial, Lognormal, Weibull e Poisson. A seleção foi feita por AIC [Akaike 1974]. As métricas \bar{C} e \bar{D} foram comparadas com modelos nulos Erdős–Rényi (ER) e Configuração (CM) [Maslov and Sneppen 2002], e o índice de Small-Worldness calculado como $\omega = (\bar{C}/\bar{C}_{\text{nulo}}) \cdot (\bar{D}_{\text{nulo}}/\bar{D})$ de acordo com [Humphries and Gurney 2008] – definição distinta da métrica ω de [Telesford et al. 2011], que normaliza em relação a um lattice regular. Testes de robustez avaliaram a degradação da MCC sob remoção sequencial de vértices por betweenness, closeness, eigenvector, grau e aleatoriedade.

4. Resultados e Discussão

As cinco redes analisadas revelam um ecossistema científico regional de geometria complexa: coeso em torno da UFT, internamente fragmentado nas IES de menor porte e ainda em constituição na UFNT. Esta seção percorre as camadas dessa topologia buscando, além de descrever os indicadores, interpretar as causas institucionais e geográficas que os geram.

4.1. Produção Científica por Instituição

A Tabela 1 apresenta os volumes de trabalhos e autores por IES. A UFT concentra 13.481 trabalhos e 24.706 autores, correspondendo a 71,32% das publicações exclusi-

vas da região (Tabela 2). O total de trabalhos únicos – 16.600 – evidencia sobreposição interinstitucional limitada: nenhuma publicação envolveu as cinco IES simultaneamente.

Tabela 1. Número de Trabalhos e Autores por IES (1998–2024).

Instituição	Trabalhos	Autores
UFT	13.481	24.706
Unitins	2.087	6.311
IFTO	1.983	5.619
UFNT	558	1.566
CEULP	291	872
Total – Tocantins	16.600	32.783

Tabela 2. Trabalhos Exclusivos por IES.

Descrição	Trabalhos	%
Exclusivos da UFT	11.839	71,32
Exclusivos do IFTO	1.496	9,01
Exclusivos da Unitins	1.355	8,16
Exclusivos da CEULP	191	1,15
Exclusivos da UFNT	9	0,05
Envolvendo as 5 IES	0	0,00

A Figura 2 mostra crescimento consistente da produção após 2010. O pico de 2020 na UFT e no CEULP está alinhado ao aumento documentado de publicações relacionadas à pandemia de COVID-19 [Raynaud et al. 2021]; IFTO, Unitins e UFNT atingiram seus volumes máximos em 2022–2023. O diagrama de Venn, Figura 3, ilustra a pequena sobreposição entre as três IES de maior produção, indicando que a colaboração interinstitucional, embora presente, não é o padrão dominante.

4.2. Métricas Topológicas

A Tabela 3 caracteriza as métricas topológicas das redes cumulativas de 2024. A UFT lidera em todas as dimensões: 85,85% da componente gigante, grau médio de 13,30 e 164.249 arestas. As demais IES apresentam componentes gigantes inferiores a 50% – Unitins (38,92%), IFTO (46,63%), CEULP (46,49%) e UFNT (36,81%) –, indicando que parcela expressiva de seus pesquisadores permanece em subgrafos desconectados.

Os coeficientes de agrupamento médio ($\bar{C} \geq 0,78$) são elevados em todas as redes, sinalizando a formação de triângulos colaborativos que refletem a organização interna dos grupos e PPGs. A assortatividade é positiva em todos os casos; os valores mais altos na Unitins ($r = 0,874$) e no CEULP ($r = 0,852$) indicam colaboração mais estratificada, com pouca permeabilidade entre grupos de portes diferentes.

4.3. Modelos Nulos e Small-Worldness

A Tabela 4 compara a MCC da UFT com os modelos nulos. O coeficiente de agrupamento real ($\bar{C} = 0,7889$) é 1.127 vezes maior que no modelo ER e 95 vezes maior que no

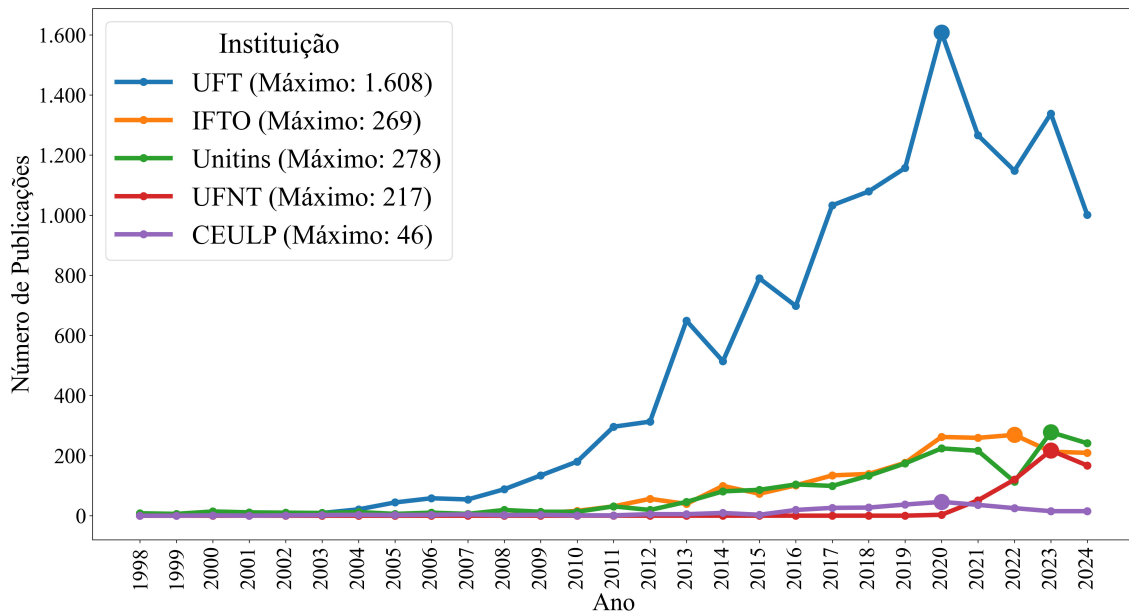


Figura 2. Evolução da Produção Científica das IES Tocantinenses (1998–2024).

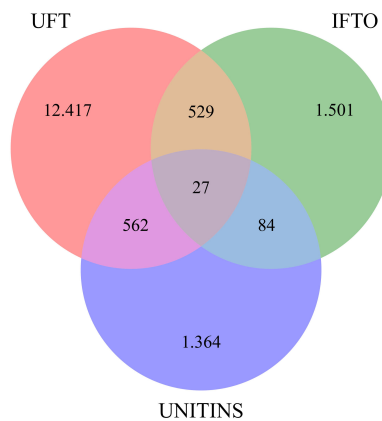


Figura 3. Publicações compartilhadas entre UFT, IFTO e Unitins.

Tabela 3. Métricas Topológicas das Redes de Coautoria (2024).

Métrica	UFT	Unitins	IFTO	UFNT	CEULP	Tocantins
Vértices	24.706	6.311	5.619	872	1.566	32.783
Arestas	164.249	32.185	22.970	3.040	7.290	205.349
MCC (%)	85,85	38,92	46,63	36,81	46,49	81,11
Grau médio	13,30	10,20	8,18	6,97	9,31	12,53
\bar{C}	0,7889	0,8499	0,7955	0,8260	0,8071	0,7927
\bar{D} (MCC)	5,59	6,59	8,82	5,43	6,87	5,89
Assortatividade	0,271	0,874	0,564	0,184	0,852	0,318

CM, descartando origem aleatória. O caminho médio real (5,586) supera o dos modelos em apenas 1,6 e 2,0 saltos, respectivamente – margem que não compromete a eficiência de troca de informação. Os índices $\omega_{ER} = 786,67$ e $\omega_{CM} = 60,92$ confirmam as propriedades de mundo pequeno da rede da UFT. O menor ω frente ao CM é esperado: por

preservar a sequência de graus, o CM produz redes aleatórias já mais agrupadas que as do ER, constituindo um teste mais rigoroso; mesmo assim, $\omega_{CM} \gg 1$ indica organização colaborativa além do que a distribuição de graus produziria.

Resultados análogos foram obtidos para Unitins ($\omega_{ER} = 86,53$; $\omega_{CM} = 17,05$) e para a rede do Tocantins ($\omega_{ER} = 1089,14$; $\omega_{CM} = 82,25$).

Tabela 4. MCC da UFT vs. Modelos Nulos.

Métrica	Rede Real	Erdős-Rényi	Config. Model
Vértices / Arestas (MCC)	21.211 / 156.576	=	=
\bar{C}	0,7889	$0,0007 \pm 0,0000$	$0,0083 \pm 0,0002$
\bar{D}	5,586	$3,968 \pm 0,000$	$3,581 \pm 0,002$
ω (vs. ER/CM)	786,67 / 60,92	-	-

4.4. Distribuição de Graus

O ajuste por AIC indicou Lei de Potência como modelo de melhor ajuste relativo em todas as redes (Peso AIC = 1,00), conforme Tabela 5.

Os expoentes $\alpha \in (2,17; 3,10)$ situam-se no intervalo tipicamente reportado para redes de coautoria na literatura [Barabási et al. 2002, Newman 2001]. Para referência, [Newman 2001] reportou $\alpha \approx 2,5$ em redes de coautoria em Física; o expoente $\alpha = 2,17$ da UFT, o menor do conjunto, indica distribuição de cauda mais pesada e maior concentração de colaborações nos *hubs*. Expoentes mais altos como os da UFNT ($\alpha = 3,10$) e Unitins ($\alpha = 2,98$) descrevem ecossistemas menos hierarquizados.

Uma limitação metodológica deve ser assinalada: o AIC compara ajuste relativo entre modelos, mas não substitui o teste de plausibilidade de [Clauset et al. 2009], cuja aplicação constitui trabalho futuro prioritário (ver Subseção 4.9).

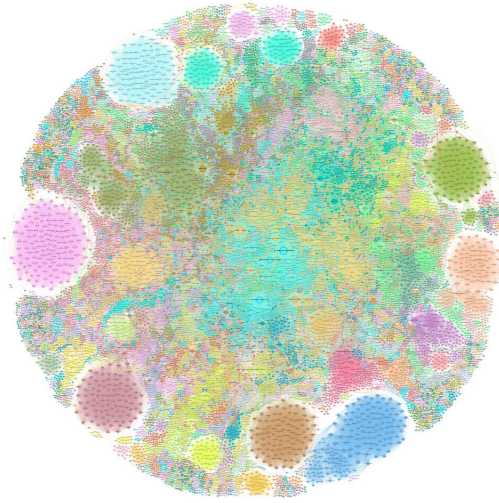
Tabela 5. Parâmetros da Lei de Potência (Distribuição de Graus).

IES	α	x_{min}	AIC
UFT	2,17	8	79.910,58
Unitins	2,98	12	7.927,82
IFTO	2,57	7	16.442,30
UFNT	3,10	5	2.018,90
CEULP	2,81	6	4.311,20
Tocantins	2,26	9	92.444,90

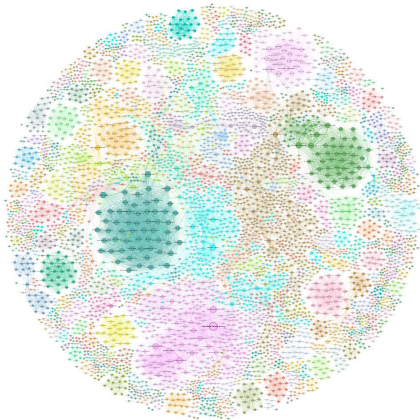
4.5. Visualização e Análise das Redes

As Figuras 4 e 5 apresentam as redes de cada IES geradas no Gephi, com tamanho dos vértices proporcional ao grau e cores identificando comunidades de Louvain.

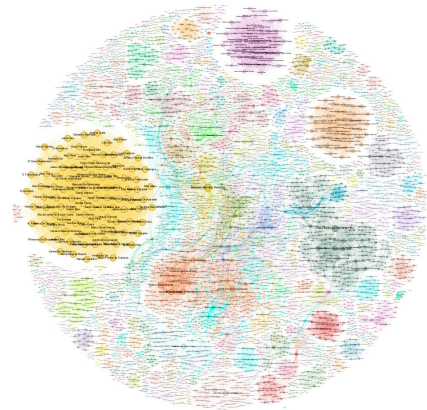
A UFT constitui o núcleo estrutural do ecossistema científico tocantinense. Sua rede, ilustrada na Figura 4a, é a maior e mais coesa do conjunto: componente gigante de 85,85% e Small-Worldness 786 vezes acima do esperado por um modelo aleatório, reflexo de décadas de construção institucional e de PPGs consolidados. Os *hubs* de maior grau pertencem a Agronomia, Engenharia Florestal e Biotecnologia, enquanto os de maior



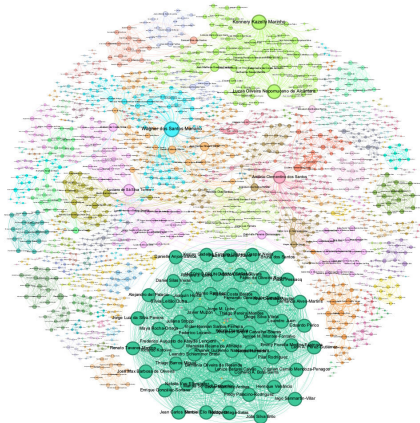
(a) Rede de coautoria da UFT.



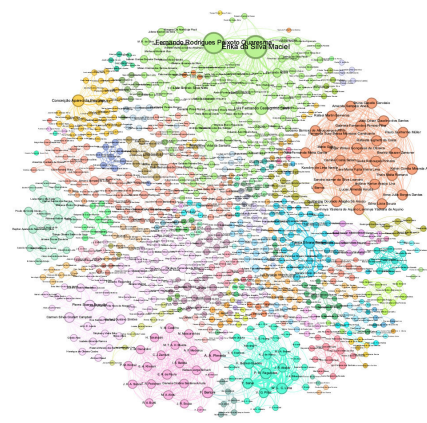
(b) Rede de coautoria da Unitins.



(c) Rede de coautoria do IFTO.



(d) Rede de coautoria da UFNT.



(e) Rede de coautoria do CEULP.

Figura 4. Redes de coautoria das IES do Tocantins em 2024.

betweenness e closeness são de Medicina e Engenharia Ambiental, perfis que exercem funções estruturais distintas na rede, analisados em detalhe na Tabela 6.

A Unitins ocupa posição singular no conjunto: apesar de ser a segunda IES mais antiga da região, sua rede apresenta a menor componente gigante (38,92%) e a maior assortatividade ($r = 0,874$). Esse contraste sinaliza que a produção científica ocorre em núcleos colaborativos internamente coesos, mas pouco conectados entre si – padrão que reflete, em parte, a ausência de PPGs transversais capazes de cruzar as fronteiras entre grupos de pesquisa. Três dos pesquisadores de maior centralidade têm afiliação principal em outras IES: UFT, Universidade do Estado do Pará (UEPA) e Universidade Federal do Maranhão (UFMA). Este padrão pode indicar certa dependência de coautores externos para manter a coesão da rede.

O perfil estrutural da rede do IFTO é marcado pelo maior caminho médio da componente gigante ($\bar{D} = 8,82$), valor muito superior ao das demais IES. Com campi distribuídos em diversas cidades do estado, as colaborações entre pesquisadores de unidades distintas tendem a exigir mais intermediários, alongando os caminhos na rede. Esse resultado é consistente com a literatura sobre instituições policampi, nas quais a dispersão geográfica dificulta a formação de atalhos estruturais [Maciel et al. 2013]. A análise de betweenness indica que pesquisadores que conectam diferentes campi do IFTO ocupam posições especialmente críticas para a coesão da instituição.

A UFNT é o caso mais singular deste estudo. Desmembrada da UFT em 2021, acumulou apenas 9 publicações exclusivas até 2024 – número que reflete diretamente o horizonte temporal disponível: três anos são insuficientes para que um corpo docente em formação consolide linhas de pesquisa próprias. A vasta maioria das colaborações envolve coautores da UFT, e o menor grau médio (6,97) e a menor componente gigante (36,81%) confirmam que a rede autônoma da UFNT está em constituição – um retrato de estágio inicial, não de deficiência estrutural permanente.

Como única IES privada do conjunto, o CEULP opera em condições institucionais distintas das demais, e sua rede reflete essa especificidade. Apresenta alta assortatividade ($r = 0,852$) e componente gigante de 46,49%, sugerindo colaboração concentrada em grupos coesos com pouca permeabilidade entre si – padrão que, assim como na Unitins, pode estar associado à menor quantidade de PPGs com linhas de pesquisa transversais. O pico de produção em 2020 destaca-se de forma particularmente acentuada nos dados do CEULP – mais do que em qualquer outra IES excetuando-se a UFT – e está associado ao aumento de publicações sobre a pandemia de COVID-19. A análise de centralidade revela que os pesquisadores de maior grau e betweenness do CEULP atuam principalmente nas áreas de Medicina e Enfermagem, áreas intensificadas no contexto pandêmico.

4.6. Análise de Centralidade

A Tabela 6 apresenta os pesquisadores de maior centralidade na rede da UFT – coincidente com a rede do Tocantins pela sobreposição estrutural entre ambas. Pesquisadores de maior grau pertencem a Ciências Agrárias e Biológicas, enquanto betweenness e closeness mais altos recaem sobre pesquisadores de Medicina, indicando que volume de colaborações e posição de ponte entre comunidades correspondem a perfis distintos.

Tabela 6. Pesquisadores Mais Centrais – Rede UFT/Tocantins (2024).

Métrica	Pesquisador	Valor	Área
Grau	Pesquisador A	491	Agronomia
	Pesquisador B	399	Eng. Florestal
	Pesquisador C	380	Biotecnologia
Betweenness	Pesquisador D	0,0514	Medicina
	Pesquisador E	0,0395	Eng. Ambiental
	Pesquisador F	0,0343	Eng. Alimentos
Closeness	Pesquisador D	0,2411	Medicina
	Pesquisador E	0,2369	Eng. Ambiental
	Pesquisador F	0,2353	Eng. Alimentos

4.7. Rede do Tocantins

A rede união do Tocantins, ilustrada na Figura 5, tem sua topologia dominada pela UFT, que responde por 71,32% dos trabalhos únicos. Distribuição de graus com $\alpha = 2,26$, grau máximo 498 e Small-Worldness de 1.089 em relação ao modelo ER confirmam que o ecossistema científico regional combina coesão local com acessibilidade global: cada pesquisador está, em média, a menos de seis colaborações de distância de qualquer outro na maior componente, que reúne 81,11% dos 32.783 autores cadastrados.

4.8. Testes de Robustez

Os testes de robustez avaliaram a resistência de cada rede à remoção sequencial de vértices por cinco critérios. Em todas as IES, a ordem decrescente de eficácia para fragmentar a MCC é: betweenness > closeness \approx eigenvector > grau > remoção aleatória.

Na rede da UFT, a remoção dos 1% de vértices com maior betweenness fragmenta mais do que a equivalente por grau, confirmando que a coesão depende criticamente de *brokers* e não apenas dos *hubs* de maior volume. A remoção aleatória confirma a maior resiliência da UFT: a componente gigante permanece acima de 70% após extração de 5% dos vértices, comportamento característico de redes com distribuição de graus heterogênea.

Nas redes de menor porte (UFNT e CEULP), a degradação é mais acentuada sob qualquer critério, reflexo da menor redundância de caminhos. Do ponto de vista de política científica, os perfis de Medicina e Engenharia Ambiental – que lideram betweenness e closeness na rede da UFT (Tabela 6) – são os mais críticos para a conectividade intercomunitária: seu afastamento teria impacto estrutural desproporcional ao seu número absoluto.

4.9. Limitações Metodológicas

Cinco limitações merecem registro explícito. A primeira diz respeito ao critério de seleção de modelos: o AIC identifica o melhor ajuste relativo entre os modelos concorrentes, mas não atesta plausibilidade estatística em sentido absoluto. As afirmações sobre distribuição de graus devem, portanto, ser lidas como expressão do melhor ajuste relativo segundo esse critério, e não como confirmação definitiva de comportamento livre de escala. A aplicação do procedimento baseado em máxima verossimilhança combinado com KS-bootstrap permanece como um dos próximos passos da pesquisa.

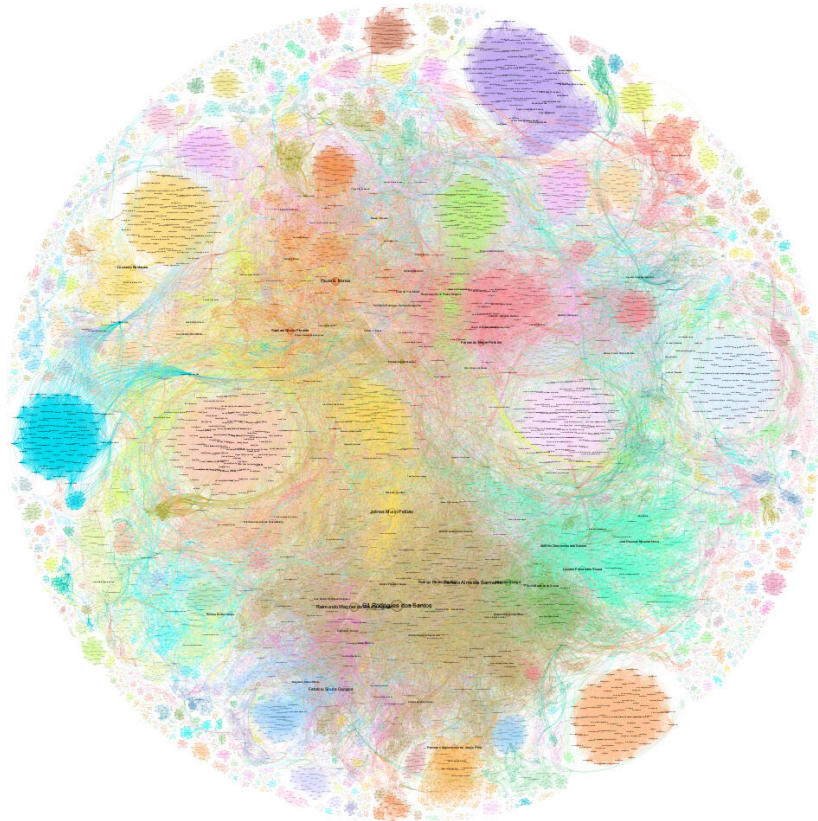


Figura 5. Rede de coautoria do Tocantins em 2024 (32.783 vértices, 205.349 arestas, MCC de 81,11%).

Uma segunda limitação recai sobre as métricas reportadas na Tabela 3, que não capturam a intensidade das colaborações. Medidas ponderadas — como força de vértice e *weighted clustering coefficient* — serão incorporadas em versão futura da análise. Somase a isso a omissão das curvas de robustez por restrição de espaço, o que impede a comparação precisa das velocidades de fragmentação entre as IES segundo cada critério de ataque.

Duas limitações adicionais dizem respeito às fontes de dados e ao recorte temporal. A dependência exclusiva do OpenAlex [Priem et al. 2022] implica que a cobertura das IES de menor porte pode ser incompleta, potencialmente amplificando a assimetria observada entre a UFT e as demais instituições. Por fim, o período disponível para a UFNT restringe-se a 2021–2024, intervalo de três anos insuficiente para a caracterização estrutural estável de uma rede; os indicadores relativos a essa instituição devem, assim, ser interpretados como retrato de um estágio inicial de consolidação.

5. Considerações Finais

As redes de coautoria das cinco IES tocantinenses revelam um ecossistema científico regional organizado em torno de uma assimetria central: a UFT estrutura e conecta esse ecossistema, enquanto as demais instituições operam com fragmentação interna significativa e graus variáveis de dependência em relação à instituição dominante. Essa configuração

provavelmente não é acidental, refletindo a desigualdade histórica no desenvolvimento institucional e na consolidação de programas de pós-graduação na região. A dominância da UFT vai além do tamanho: seu coeficiente de agrupamento médio supera em 95 vezes o do Modelo de Configuração, o que indica que a organização colaborativa emerge de mecanismos institucionais — PPGs, projetos intergrupos, lideranças estáveis — e não apenas da heterogeneidade de graus. A distinção entre perfis de alto grau, como Agronomia e Engenharia Florestal, e perfis de alto *betweenness*, como Medicina e Engenharia Ambiental, é igualmente relevante para a formulação de políticas de pesquisa: trata-se de funções estruturais distintas cujo enfraquecimento teria impactos assimétricos sobre a coesão do ecossistema. As IES intermediárias — Unitins, IFTO e CEULP — partilham o traço de alta coesão local combinada com baixa integração global, sugerindo que o desafio não é criar colaboração, mas conectar grupos que já colaboram internamente sem se integrarem entre si ou com o restante da rede regional.

O caso da UFNT merece atenção especial como fenômeno de emancipação científica mensurável em rede. Desmembrada da UFT em 2021, a universidade acumulou apenas nove publicações exclusivas em três anos, com a maioria de seus artigos envolvendo coautores da instituição de origem. Longe de representar deficiência, esse perfil documenta o estágio inicial de um processo gradual de constituição de identidade científica autônoma. O monitoramento longitudinal das métricas de rede da UFNT constitui um instrumento direto de avaliação de políticas de fomento à pesquisa em instituições recém-criadas, e o presente estudo fornece uma linha de base quantificável contra a qual seus efeitos poderão ser medidos ao longo do tempo.

Do ponto de vista ético, a pesquisa mobiliza dados nominais de pesquisadores a partir de fontes abertas, o que impõe deveres de transparência metodológica e uso responsável das informações. A identificação de atores com alta centralidade expõe indivíduos a avaliações institucionais baseadas em posição de rede — uma forma de métrica que pode ser instrumentalizada de maneira inadequada em processos de financiamento ou avaliação de desempenho. Recomenda-se que análises derivadas deste trabalho explicitem as limitações interpretativas das métricas empregadas e evitem reduções individualizantes de fenômenos que são, por natureza, coletivos e relacionais. A assimetria entre IES documentada aqui deve ser lida como evidência de condicionantes estruturais e históricas, e não como atribuição de mérito ou deficiência a pesquisadores ou gestores específicos.

Como extensões prioritárias, propõe-se a verificação estatística rigorosa da lei de potência por máxima verossimilhança e KS-bootstrap, o cálculo de métricas ponderadas que capturem a intensidade das colaborações, e a análise longitudinal da UFNT como caso de acompanhamento de política científica. A integração com a Plataforma Lattes e com a base Scopus permitiria ampliar a cobertura e reduzir vieses associados à dependência exclusiva do OpenAlex, especialmente para IES de menor porte. Em horizonte mais amplo, a extensão para mineração de redes — incluindo predição de links, detecção de comunidades dinâmicas e rastreamento da difusão de tópicos emergentes — abriria caminho para uma ciência regional mais compreendida em sua dinâmica temporal e temática, e não apenas em sua estrutura estática.

Referências

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.
- Barabási, A., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., and Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3):590–614.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.
- Brito, V. et al. (2020). Redes de coautoria em ciências da saúde no nordeste brasileiro. *Transinformação*.
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703.
- Humphries, M. D. and Gurney, K. (2008). Network ‘small-world-ness’: A quantitative method for determining canonical network equivalence. *PLOS ONE*, 3(4):e0002051.
- Jacomy, M., Venturini, T., Heymann, S., and Bastian, M. (2014). Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PLOS ONE*, 9(6):e98679.
- Maciel, C. O. et al. (2013). Análise da rede de colaboração científica de um departamento de uma universidade brasileira. *Organizações & Sociedade*.
- Maslov, S. and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science*, 296(5569):910–913.
- Mena-Chalco, J. P., Digiampietri, L. A., Lopes, F. M., and Cesar Junior, R. M. (2014). Brazilian bibliometric coauthorship networks. *Journal of the Association for Information Science and Technology*, 65(7):1424–1445.
- Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409.
- Priem, J., Piwowar, H., and Orr, R. (2022). Openalex: A fully-open index of scholarly works, authors, venues, institutions, and concepts.
- Raynaud, M. et al. (2021). Impact of the covid-19 pandemic on publication dynamics and non-covid-19 research production. *BMC Medical Research Methodology*, 21(1):255.
- Souza, J. P. N. d. M. and dos Santos, T. D. (2025). Redes de produção/colaboração científica no estado do Tocantins. *DESAFIOS - Revista Interdisciplinar da Universidade Federal do Tocantins*, 12(1):1–17.
- Telesford, Q. K., Joyce, K. E., Hayasaka, S., Burdette, J. H., and Laurienti, P. J. (2011). The ubiquity of small-world networks. *Brain Connectivity*, 1(5):367–375.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks.
Nature, 393(6684):440–442.