



**OBSERVATÓRIO
NACIONAL DE
BLOCKCHAIN**

PRODUÇÃO CIENTÍFICA

Redes de coautoria em
blockchain

**Fevereiro
2026**

*De olho
na web
do **futuro.***

Sumário

Resumo	3
1 Introdução	4
2 Trabalhos Relacionados	6
3 Metodologia	7
3.1 Coleta e pré-processamento de dados	8
3.2 Modelagem da rede de colaborações	9
3.3 Métricas de análise da rede	9
3.4 Métodos de identificação de comunidades	10
4 Resultados	11
4.1 Análise estrutural da rede	11
4.2 Análise de autorias	12
4.3 Análise de instituições	13
4.4 Análise dos métodos de identificação de comunidades	15
4.4.1 Análise estrutural por algoritmo	15
4.4.2 Análise das comunidades geradas por algoritmo	16
5 Conclusões	19
6 Agradecimentos	20
7 Referências	21

Resumo

A tecnologia blockchain, aliada à criação do Bitcoin, impulsionou uma crescente produção acadêmica e a formação de uma rede de colaboração científica sobre o tema. No entanto, existe uma lacuna na literatura de pesquisas que investiguem acerca da estrutura e evolução desta rede no cenário brasileiro. Este trabalho visa preencher esse espaço ao apresentar uma análise da rede de coautoria formada por pesquisadores vinculados a instituições brasileiras que fizeram publicações entre 2008 e 2025. A metodologia envolveu a coleta de metadados das bases OpenAlex e SBC OpenLib, com um total de 1.723 artigos contabilizados, seguida pela modelagem de um grafo ponderado e não-direcionado utilizando ferramentas da biblioteca Igraph. Foram analisadas métricas de centralidade em redes e técnicas para identificação de comunidades explorando os algoritmos Louvain, Fast Greedy e Walktrap. Os resultados revelaram uma rede de baixa densidade, mas com um alto coeficiente de aglomeração local, indicando a formação de núcleos de pesquisa coesos. O algoritmo Louvain apresentou o melhor desempenho em modularidade, identificando com uma maior precisão os grupos mais influentes. A análise destacou ainda a predominância de instituições públicas na liderança da produção nacional.

1. Introdução

A tecnologia blockchain foi proposta para o funcionamento da criptomoeda Bitcoin, em 31 de outubro de 2008, por Satoshi Nakamoto, pseudônimo para uma pessoa ou grupo de pessoas que até hoje é desconhecido. Em um artigo técnico com acesso livre via Internet, misteriosamente atribuído a esse autor anônimo [Nakamoto et al. 2007], foram descritos os fundamentos de blockchain e como o Bitcoin é seguro e descentralizado baseado nessa tecnologia. Atualmente, esse artigo tem mais de 41 mil citações de acordo a plataforma Google Acadêmico¹. A evolução e adesão dessa tecnologia impulsionou a produção acadêmica, com a formação de uma rede de colaborações entre pesquisadores que possibilitou a disseminação de novas descobertas.

No Brasil, as pesquisas sobre blockchain tem crescido notavelmente como mostram os indicadores de produtividade acadêmica do Observatório Blockchain². Em especial, nos anos de 2018 para 2019 a produção acadêmica nacional deu um salto de 131 para 291 artigos anuais, ou seja, mais que o dobro, e ainda há sinais de que essa produção continua crescente. Por isso compreender o comportamento desses autores e a forma como eles se organizam é essencial para identificar onde estão concentrados os núcleos de pesquisa mais influentes e como o conhecimento flui dentro das redes de colaborações. A ciência de redes [Newman 2018] oferece as ferramentas necessárias para o mapeamento e o conhecimento profundo dessas redes, que por apresentarem padrões de conexão não triviais, como a formação de “hubs” e o efeito de mundo pequeno, são caracterizadas como redes complexas [Albert and Barabási 2002].

Todavia, a identificação correta das subestruturas densamente conectadas, conhecidas como comunidades, é um dos principais desafios na análise de redes complexas. O problema de detecção de comunidades - clusterização - deve considerar algumas condições, como: a dispersão e consistência dos dados, a variação no tamanho da rede ao longo do tempo e a escolha do algoritmo que seja mais adequado para particionar o grafo de maneira mais eficiente.

A revisão literária identificou a existência de estudos que mapearam diversas redes de colaboração e as descobertas anteriores já revelaram que essas redes têm uma tendência a serem “livres de escala” e são governadas por ligações preferenciais, nas quais os autores tendem a se conectar com os mais influentes [Barabási et al. 2002]. No contexto nacional, pesquisas recentes seguiram essa abordagem para mapear redes consolidadas, como a

¹ <https://scholar.google.com.br>

² <https://observatorioblockchain.org.br/indicadores>

composta pelos eventos da Sociedade Brasileira de Computação [Carvalho et al. 2025] e encontraram padrões estruturais sobre a produtividade da rede. Complementarmente, a seleção dos métodos mais adequados para identificar comunidades está condicionada às características particulares da rede, sendo necessário análises comparativas para justificar a escolha [Aynaud and Guillaume 2010]. Porém, existe uma lacuna na análise específica sobre o crescimento e evolução dos estudos específicos sobre blockchain no contexto nacional, considerando variadas fontes de dados e comparando o desempenho de métodos de identificação de comunidades para validar os grupos de pesquisas mais prolíficos.

Neste trabalho, propomos uma análise dessa rede de coautorias, considerando o escopo temporal de 2008 à 2025. As etapas para realização desta pesquisa consistiram em: coletar os dados, realizar um pré-processamento, modelar um grafo não-direcionado e ponderado. A partir desse grafo, analisamos métricas de centralidade [Newman 2018] dos nós da rede, i.e., os autores e identificamos comunidades nessa rede aplicando os algoritmos mais conhecidos em redes complexas - Louvain, Fast Greedy e Walktrap - para identificação de comunidades.

Para além da contribuição técnica, o mapeamento da rede de colaborações sobre blockchain apresenta aplicações práticas diretas para a comunidade acadêmica. Os resultados obtidos permitem a identificação dos principais pólos de pesquisa e dos grupos mais ativos, o que serve como fonte valiosa de informação para auxiliar no direcionamento de recursos e bolsas por parte das instituições de ensino e de agências de fomento à pesquisa. Adicionalmente, para os pesquisadores a análise da rede pode facilitar na descoberta de potenciais colaboradores e estimular a integração entre as instituições.

Os resultados mostram que a rede de colaborações em pesquisas sobre blockchain é caracterizada por ter uma baixa densidade (0,0021), mas com um coeficiente de aglomeração local alto (0,8355), o que indica a formação de núcleos mais coesos. Além disso, foi possível identificar que a produção nacional é liderada por instituições públicas em sua maioria, com destaques para a Universidade de São Paulo, Universidade Federal do Rio de Janeiro e a Universidade Federal de Santa Catarina. Com relação à detecção de comunidades, os algoritmos *Louvain*, *Fast Greedy* e *Walktrap* foram utilizados para validar a solidez das estruturas identificadas, utilizando a modularidade para avaliar a qualidade do particionamento. Todos os algoritmos apresentaram valores altos para esta métrica (0,9222 - 0,9183 - 0,8854), um comportamento esperado por se tratar de uma rede muito esparsa, com destaque para *Louvain*, revelando colaboração intensa dentro dos grupos, mas esparsa entre eles.

Este trabalho está organizado nas seguintes seções: a Seção 2 apresenta a revisão da literatura e os trabalhos relacionados a este. A Seção 3 descreve a metodologia utilizada para

realizar esta pesquisa. A Seção 4 expõe os resultados obtidos, incluindo uma análise topológica, um ranking de autores e instituições e a comparação entre os algoritmos implementados. A Seção 5 traz as conclusões e sugestões para trabalhos futuros.

2. Trabalhos Relacionados

A ciência de redes permite entender e identificar as dinâmicas e padrões existentes dentro de uma rede de colaborações. Em [Barabási et al. 2002] são estabelecidas bases importantes ao analisar redes complexas associadas à evolução temporal, seus resultados indicaram que essas redes são “livre de escala” e a estrutura é regida por ligações preferenciais. Isso explicaria a formação de “hubs”, nos quais poucos autores são fortemente conectados, enquanto o restante possui poucas conexões. Estudos recentes demonstram que essas análises permanecem atuais e relevantes, como o apresentado por [Zhang et al. 2025] que mapeia a estrutura e as relações de conectividade entre países pertencentes à Iniciativa Cinturão e Rota (BRI), corroborando a tese de que conhecer a topologia da rede é necessário para compreender o fluxo do conhecimento e a distribuição de influência em comunidades acadêmicas.

No cenário brasileiro, existem diversos trabalhos que focam em estudar as redes de coautoria nacionais utilizando bases bibliométricas como a da plataforma Lattes. Em [Digiampietri et al. 2014, Dias et al. 2019] foram realizadas pesquisas semelhantes focadas nos doutores com currículo cadastrado no Lattes e foram encontradas redes heterogêneas, de baixa densidade e com uma alta dispersão, o que é esperado para redes de grande escala. De modo semelhante, foi apresentado por [Boaventura et al. 2014] um estudo com uma visão mais microscópica, orientado a analisar a evolução temporal das colaborações entre docentes cadastrados na plataforma, revelando uma tendência em formar grupos de pesquisa pequenos, porém duradouros.

Na área da Computação, em [Carvalho et al. 2025] é conduzido um estudo meta-científico sobre o perfil de colaborações da comunidade acadêmica vinculada ao Congresso da Sociedade Brasileira de Computação (CSBC), o maior evento científico dedicado à computação da América Latina, com dados extraídos da biblioteca digital SBC OpenLib (SOL). Duas análises foram feitas: uma global do CSBC, com todos os seus eventos associados, e uma específica do BraSNAM, um evento satélite desse congresso. Os resultados revelaram que o BraSNAM possui uma comunidade com alta proximidade e uma forte tendência à colaboração, o que difere da análise macroscópica do CSBC, que demonstrou haver uma baixa generalização de fenômenos entre seus eventos, com cada um deles apresentando especificidades próprias.

Para uma análise mais aprofundada da estrutura de uma rede de coautorias, os métodos de detecção de comunidades são essenciais para identificar os grupos mais prolíficos da rede.

No entanto, o processo de modelagem deve considerar alguns fatores, como: o tamanho da rede, a evolução temporal das relações e as possíveis inconsistências nos dados. A identificação correta dos autores e grupos chaves, especificamente em redes de colaboração, não depende exclusivamente da escolha do algoritmo, mas da capacidade em lidar com as particularidades e variações nos dados e nas relações entre pesquisadores. As pesquisas mais atuais discutem inclusive a aplicação de abordagens diferentes, como a centralidade fuzzy, para representar de maneira mais precisa os relacionamentos entre os nós [Porreca et al. 2025].

A seleção do método ideal deve ter como base as propriedades topológicas da rede, pois para redes sociais dinâmicas, qualquer modificação, ainda que pequena, pode acarretar em mudanças drásticas nas comunidades detectadas [Aynaud and Guillaume 2010]. Diante disso, [Leão et al. 2017] propõem que antes de iniciar o processo de detecção de comunidades a rede deve passar por um pré-processamento, considerando as características temporais e estruturais, removendo arestas aleatórias, que seriam os denominados “ruídos”, visando uma melhor qualidade dos dados que serão analisados.

Nesse sentido, este trabalho está alinhado com a literatura, pois aplica técnicas de Análise de Redes Sociais para compreender a estrutura da rede de colaborações acadêmicas sobre blockchain, encontrando resultados que corroboram com os estudos supracitados sobre a formação de hubs e a topologia de redes de grande escala [Barabási et al. 2002, Zhang et al. 2025]. Entretanto, apresenta diferenças de [Digiampietri et al. 2014, Boaventura et al. 2014], por utilizar uma base de dados focada em um eixo temático - a tecnologia blockchain - e por integrar base de dados variadas (OpenAlex e SBC OpenLib). Além disso, contribui com as discussões recentes sobre a identificação de grupos de colaboração [Porreca et al. 2025], ao aplicar e comparar os algoritmos de detecção de comunidades em um cenário real, trazendo uma análise detalhada sobre a formação e a maturidade desses grupos de pesquisa no Brasil.

3. Metodologia

Nesta seção, vamos especificar as etapas de coleta, processamento e análise dos dados bibliométricos referentes à produção científica em blockchain. A linguagem de programação Python, executada em notebooks do Google Colab, foi utilizada em todas essas etapas juntamente à biblioteca lgraph [Csardi and Nepusz 2006] e a API do modelo de linguagens

grande (LLM) GPT-4o³. Essas ferramentas foram essenciais para calcular e analisar métricas de redes e comunidades de pesquisadores.

3.1. Coleta e pré-processamento de dados

A base de dados utilizada foi construída com trabalhos publicados entre os anos 2008 e 2025, a partir de metadados extraídos de duas plataformas:

- OpenAlex⁴: catálogo bibliográfico de dados de pesquisa científica totalmente aberto, gratuito e de livre acesso.
- SBC-OpenLib (SOL)⁵: biblioteca digital mantida pela Sociedade Brasileira de Computação (SBC), que permite acesso gratuito à pesquisas da área de Computação.

Prompt Utilizado

role: system

“Você é um especialista em blockchain e sua tarefa é analisar o título e resumo de um trabalho acadêmico para identificar se está ou não no contexto de blockchain.

Será informado o título e o resumo do artigo.

Retorne apenas 1 se o artigo está no contexto de tecnologia blockchain e 0 caso contrário.

role: user

O artigo tem o título: '{title}' e o seguinte resumo: '{abstract}'.

Indique apenas 1 se o artigo utiliza tecnologia blockchain e 0 caso contrário.

A seguir, foi realizado a remoção de duplicatas e a desambiguação de nomes de instituições e de autores. Nesse caso, foi passado como parâmetro para o GPT-4o uma planilha contendo todos os dados organizados no formato CSV, e foi requisitado via *chat* o retorno no formato JSON, conforme o *prompt* mostrado abaixo. Esse último procedimento resultou em um conjunto de dados final com 1.723 artigos.

³ <https://openai.com/api/>

⁴ <https://openalex.org/>

⁵ <https://sol.sbc.org.br/index.php/indice>

Prompt Utilizado

“Você é um assistente especializado em curadoria de dados acadêmicos.

Recebe uma lista de artigos com 'título', 'autores' e 'instituições'.

Seu trabalho é padronizar nomes de autores e instituições, corrigindo ambiguidade e variações.

Retorne no formato JSON com os campos originais corrigidos e padronizados.”

3.2. Modelagem da rede de colaborações

A rede de coautorias foi modelada como um grafo não-direcionado e ponderado $G = (V, E, W)$, a implementação ocorreu com o auxílio da biblioteca Igraph, que possui como uma das principais características a capacidade de lidar com redes grandes. A construção das arestas foi feita de modo que elas conectassem todas as combinações de autores possíveis por artigo. Sendo assim para essa rede, os nós (V) representam os autores e as arestas (E) são as conexões entre cada par de autor.

A definição de peso (W) para cada aresta foi um parâmetro considerado para quantificar a frequência de coautoria, ou seja, o número de vezes que um par de autor colaborou. Esse processo foi implementado com o auxílio da classe Counter da biblioteca collections do Python. A lista de nós e a de arestas ponderadas foi usada pelo Igraph para instanciar o grafo da rede e serviu como base para as todas as análises subsequentes deste trabalho.

3.3. Métricas de análise da rede

Para compreender as características da rede, algumas métricas estatísticas, fundamentadas nos conceitos clássicos de redes complexas [Tabassum et al. 2018, foram calculadas utilizando funções nativas do Igraph. De um ponto de vista global, a densidade foi mensurada para avaliar a conectividade da rede, ou seja, o quão denso ou disperso o grafo pode ser. Ela é determinada pela razão entre o número de arestas que existem e o número máximo de arestas que podem existir no grafo, dado por: $\frac{n(n-1)}{2}$, onde n representa o número total de nós (autores) na rede.

Além disso, o cálculo do coeficiente de aglomeração médio é outro parâmetro global importante, pois ele é uma medida quantitativa que verifica a tendência dos nós em se agrupar, formando clusters ou os chamados triângulos na rede. Basicamente, ele mede a

probabilidade que dois autores tem de colaborar baseado em suas conexões em comum. A análise da capacidade de fragmentação da rede foi calculada por meio da divisão do grafo em componentes conectados, com ênfase no componente gigante que abrange o grupo de pesquisadores mais influentes.

Em um contexto local, o grau dos nós é uma métrica fundamental, pois indica a quantidade de co autores distintos que um autor possui. Portanto, através dessa medida é possível identificar os autores mais prolíficos e ativos dentro da rede, sendo esses os principais responsáveis pela disseminação dos conhecimentos relacionados à tecnologia blockchain.

3.4. Métodos de identificação de comunidades

Os algoritmos de detecção de comunidades possibilitam a partição da rede em subgrafos nos quais os autores apresentam características semelhantes e mantêm colaborações frequentes, além de destacar os núcleos de pesquisa mais relevantes. Diante disso, foi realizada uma análise exploratória inicial na literatura sobre os seguintes algoritmos: Louvain, Leiden, Walktrap, Infomap, Label Propagation e Fast Greedy. A motivação para a escolha dos algoritmos (Louvain, Fast Greedy e Walktrap) aplicados neste trabalho se deu pela necessidade de avaliar métodos com diferentes abordagens e como estes se adequam às com as particularidades da rede analisada. A modularidade foi a métrica utilizada para medir a qualidade do particionamento da rede gerado por cada algoritmo [Newman 2006].

Abaixo destacamos os três algoritmos selecionados:

- **Louvain:** adota uma abordagem “gulosa” baseada na otimização multinível da modularidade da rede [Blondel et al. 2008], dividindo-se em duas fases: a primeira que otimiza localmente, movendo cada nó para a comunidade na qual apresentar o maior aumento da modularidade, enquanto na segunda fase é onde ocorre a agregação da rede, as novas comunidades são transformadas em “super nós” e o peso das arestas é recalculado, por fim o processo se repete até que não haja mais ganho de modularidade. Esse algoritmo frequentemente realiza partições de alta qualidade e é reconhecido pela sua eficiência em trabalhar com redes de grande escala.
- **Fast Greedy:** também utiliza uma heurística aglomerativa para otimizar a modularidade [Clauset et al. 2004], nele cada nó inicia o processo como uma comunidade individual e, interativamente, avalia os pares de comunidades vizinhas, fundindo aqueles que resultarem em um ganho positivo de modularidade. O algoritmo repete esse procedimento, sempre escolhendo a melhor escolha local, até que a rede atinja o estado que apresentou a maior modularidade. Apesar do bom desempenho, como o fast greedy tende a fazer escolhas que são localmente ótimas e não pode voltar atrás nessas decisões, esse algoritmo pode acabar não alcançando um resultado que seja ótimo globalmente.

- **Walktrap:** fundamenta-se na premissa de que caminhadas aleatórias (random walks) tendem a permanecer “presas” por mais tempo dentro de comunidades mais densas [Pons and Latapy 2006]. Sendo assim, o algoritmo define uma medida que calcula a distância entre os nós baseando-se na probabilidade deles se visitarem durante a caminhada e a partir disso utiliza essa métrica para realizar um agrupamento aglomerativo. Ainda que seja diferente dos demais métodos escolhidos, a inclusão do Walktrap nas análises se deve a sua eficiência ao lidar com redes mais esparsas.

4. Resultados

Esta seção apresenta os resultados obtidos a partir da análise da rede de colaborações, especificamente: análise estrutural da rede (Seção 4.1), análise de autorias (Seção 4.2), análise de instituições (Seção 4.3) e análise dos métodos de identificação de comunidade (Seção 4.4).

4.1. Análise estrutural da rede

A primeira etapa da análise dos resultados consiste em caracterizar a estrutura da rede de colaborações. Os nós representam os autores e as arestas indicam a existência de ao menos uma colaboração entre eles. As métricas gerais são importantes para entender a dimensão, a conectividade e as características principais que definem essa rede.

A Tabela 1 reúne as métricas gerais que foram calculadas para o grafo original da rede. A rede possui 3.669 autores (nós) e 14.304 colaborações (arestas), derivadas de 1.723 artigos. A baixa densidade (0,0021), com um valor muito próximo de zero, indica que essa rede é bastante dispersa, ou seja, a conectividade entre os autores é pequena. Esse comportamento é esperado por se tratar de uma rede de grande escala.

Tabela 1. Métricas gerais da Rede de colaborações científicas

Grafo Original	
Nº de nós (autores)	3669
Nº de arestas (colaborações)	14304
Densidade	0,0021
Grau médio	7,7972
Maior grau	115
Menor grau	1
Peso médio	1,2319

Coeficiente de aglomeração médio	0,8355
Nº de componentes conectados	527
Tamanho do componente gigante	1237
Nº de arestas do componente gigante	6525
Densidade do componente gigante	0,0085
Nº de documentos	1723

Outro parâmetro que valida essa constatação é a grande quantidade de componentes conectados (527), pois revela que os autores não colaboram muito entre si e essa área de blockchain ainda é composta por grupos de pesquisas muito isolados.

Ainda que a conectividade da rede não seja alta, o componente gigante encontrado (1.237 nós) apresentou uma densidade superior a da rede completa, o que demonstra que os autores pertencentes a esse grupo central possuem uma maior interconectividade.

No que se refere aos nós, o grau médio (7,7972) aponta que em média um autor possui colaborações com aproximadamente outros 8 autores. No entanto, o cálculo do maior grau (115) e do menor grau (1), expõem valores muito discrepantes, atestando que existem autores bastante centrais (\textit{hubs}), como também confirma a presença de muitos autores com apenas 1 colaboração publicada.

Por fim, o coeficiente de aglomeração médio (0,8355) é uma dos valores mais inusitados dentre as métricas analisadas, pois por ser um valor muito próximo de 1,0, que é o valor máximo, pode-se afirmar que existe uma tendência muito grande que os coautores de um autor também colaborem entre si. Esta propriedade corrobora com a aplicação dos métodos de detecção de comunidades, os quais serão analisados na Seção 4.4.

4.2. Análise de autorias

Como forma de identificar os pesquisadores mais ativos nesta área de pesquisa, foi realizado um levantamento de artigos publicados por autor. A Tabela 2 apresenta o ranqueamento dos 15 autores com maior número de produções publicadas e também demonstra a evolução de suas publicações ao longo do período de 2016 à 2024, anos que concentram a produção deste grupo. A partir destes dados é possível identificar que Joel J. P. C. Rodrigues é o autor mais participativo dessa rede, com um total de 33 artigos publicados, seguido por Alex Borges Vieira (25 publicações) e Glauber Dias Gonçalves (24 publicações), indicando que eles são ativamente engajados nos estudos sobre blockchain.

Tabela 2. Ranking dos 15 autores mais ativos na rede

Autor	Total	2016	2017	2018	2019	2020	2021	2022	2023	2024
Joel J. P. C. Rodrigues	33	0	1	0	2	7	14	1	0	8

Alex Borges Vieira	25	1	0	1	3	4	6	3	4	3
Glauber Dias Gonçalves	24	0	0	0	1	2	5	2	5	9
Emanuel Ferreira Coutinho	24	0	0	0	0	8	4	7	4	1
Antônio Abelém	22	0	0	2	2	2	2	2	6	6
Gabriel Antonio F. Rebello	21	0	0	1	6	6	1	2	2	3
Otto Carlos M. B. Duarte	21	0	0	1	6	6	4	3	1	0
Gustavo F. Camilo	21	0	0	0	5	4	3	4	2	3
Diogo M. F. Mattos	21	0	0	2	3	5	4	3	2	2
Roben Castagna Lunardi	20	0	0	2	4	8	1	2	3	0
Rodrigo da Rosa Righi	20	0	0	2	3	4	7	2	1	1
Pamella Soares	19	0	0	0	0	1	6	7	3	2
Charles C. Miers	19	0	0	0	3	6	4	2	3	1
Avelino F. Zorzo	18	0	0	3	4	6	1	2	2	0
Cristiano André da Costa	17	0	0	1	3	3	6	2	1	1

A análise temporal da evolução das publicações também evidencia alguns padrões:

- Picos de Produtividade: alguns autores apresentaram anos em que tiveram um aumento mais acentuado nas suas produções, como: Joel J. P. C. Rodrigues (14 publicações em 2021), Roben Castagna Lunardi (8 publicações em 2021) e Rodrigo da Rosa Righi (7 publicações em 2022).
- Crescimento Recente: houve alguns pesquisadores que inicialmente não possuíam uma participação muito ativa, mas nos anos mais recentes exibiram um crescimento mais evidente. Por exemplo, Antônio Abelém que até 2022 possuía uma tendência de 2 artigos por ano, atingiu um pico em 2023 e 2024, com 6 publicações em cada ano.
- Produção Constante: autores como Alex Borges Vieira e Gustavo F. Camilo demonstraram uma média de produção constante ao longo dos anos considerados na análise, sem aumentos muito abruptos, mas mantendo uma participação contínua na rede.

Em suma, para esse grupo de autores é possível perceber que as produções se concentram nos anos mais atuais (2023 e 2024), o que indica que houve uma intensificação e amadurecimento recente da pesquisa na área de blockchain por essa parcela dos principais autores da rede. Identificar os autores-chave é importante para entender a composição da rede e como o conhecimento está sendo disseminado, servindo como base para as análises subsequentes das instituições e dos métodos de identificação de comunidades.

4.3. Análise de instituições

Do mesmo modo, uma investigação acerca das instituições mais prolíficas foi realizada, com o intuito de identificar os principais centros de pesquisa e colaborações sobre a tecnologia blockchain. A Tabela 3 traz dados referentes às 15 instituições com maior quantidade de estudos publicados, além de demonstrar a evolução da produção científica dessas instituições, abrangendo o escopo temporal de 2008 à 2025.

Tabela 3. Ranking das 15 instituições mais prolíficas da rede

Instituição	Total	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
Universidade de São Paulo	122	0	0	0	1	1	0	3	2	2	3	7	13	15	21	18	19	13	4
Universidade Federal do Rio de Janeiro	88	0	0	0	0	0	1	0	0	1	2	4	14	17	17	12	10	10	0
Universidade Federal de Santa Catarina	86	0	0	0	0	0	0	0	1	0	0	3	12	18	14	14	12	12	0
Universidade Federal Fluminense	71	0	0	0	0	0	0	1	0	0	1	5	4	15	16	13	8	8	0
Universidade Federal do Ceará	70	0	0	0	0	0	0	0	1	0	0	2	2	11	15	15	16	8	0
Universidade Estadual de Campinas (UNICAMP)	68	0	1	0	0	0	0	1	0	0	2	7	12	9	11	9	12	4	0
Universidade do Vale do Rio dos Sinos	53	0	0	0	0	0	0	0	0	0	0	4	7	10	13	5	10	4	0
Universidade Federal de Pernambuco	47	0	1	0	0	0	0	0	0	0	2	5	5	5	9	8	7	5	0
Universidade Federal de Juiz de Fora	42	0	0	0	0	0	0	0	0	1	0	1	5	7	12	6	4	6	0
Universidade Federal do Rio Grande do Sul	41	0	0	0	0	0	0	0	0	1	0	3	10	3	6	5	7	6	0
Universidade Federal da Bahia	41	0	0	0	0	0	0	0	0	0	0	0	3	6	12	7	5	7	1
Universidade de Brasília	39	2	0	1	0	0	0	0	0	0	3	3	3	5	8	5	4	5	0
Pontifícia Universidade Católica do Rio Grande do Sul	37	0	0	0	0	0	0	0	0	0	2	5	6	9	2	2	9	2	0
Universidade Federal do Pará	36	0	0	0	0	0	0	0	0	1	0	2	6	4	3	3	9	8	0
Fundação Getúlio Vargas	36	0	0	0	0	0	0	0	0	0	0	4	3	6	10	4	4	4	1

Essa tabela destaca a Universidade de São Paulo, com 122 artigos publicados, como a instituição mais ativa nessa rede, seguido pela Universidade Federal do Rio de Janeiro (88 publicações) e a Universidade Federal de Santa Catarina (86 publicações). A partir desse ranking é possível indicar que as instituições públicas e federais são os grandes hubs de pesquisa em blockchain.

A evolução temporal auxilia a entender o interesse sobre esse tema. Os anos iniciais de 2008 até 2016 mostram uma atividade de publicação muito baixa, sendo o ano de 2017 considerado um ponto de inflexão, no qual marcou o início de um aumento acentuado das produções científicas dessas instituições. O pico da atividade científica na área ocorre entre 2019 e 2021, com a USP atingindo o ápice em 2021 com mais de 20 artigos publicados no ano. Após esse período de crescimento, nota-se que houve uma estabilização nos anos subsequentes.

4.4. Análise dos métodos de identificação de comunidade

A análise das métricas de redes sociais na Seção 4.1 relatou um coeficiente de aglomeração médio de 0,8355 o que sugere que essa rede possui uma predisposição muito grande a formar clusters coesos. A partir disso, selecionamos três algoritmos de detecção de comunidades (Louvain, Fast Greedy e Walktrap) com o intuito de definir essas estruturas.

4.4.1 Análise estrutural por algoritmo

A Tabela 4 apresenta alguns parâmetros globais resultantes da aplicação de cada um desses algoritmos na rede para efeito de comparação de desempenho entre eles.

Tabela 4. Métricas gerais da rede após a aplicação de cada um dos métodos de identificação de comunidades

	Louvain	Fast Greedy	Walktrap
Nº de comunidades	546	547	677
Modularidade	0,9222	0,9183	0,8854
Média de nós/membros por comunidade	6,72	6,71	5,42
Maior comunidade	130	161	201
Menor comunidade	2	2	2
Visualização	Link	Link	Link

A modularidade é a métrica que avalia a qualidade da divisão produzida por esses algoritmos e todos eles obtiveram valores muito altos. O Louvain foi o que mais se destacou dentre eles, alcançando o maior valor (0,9222), o Fast Greedy se aproximou desse desempenho (0,9183) e o Walktrap obteve o menor valor (0,8854), mas ainda sim é um resultado bastante expressivo.

A quantidade de comunidades geradas pelo Louvain e o Fast Greedy também é bem similar (546 e 547, respectivamente), bem como a média de membros por comunidades (6,72 e 6,71, cada), enquanto o Walktrap demonstrou uma tendência a fragmentar mais a rede, gerando 677 comunidades, mas com uma média menor de membros (5,42).

4.4.2 Análise das comunidades geradas por algoritmo

Com o objetivo de ter uma compreensão mais detalhada sobre a estrutura e a qualidade dos métodos aplicados na rede, foi realizada uma análise detalhada das 15 maiores comunidades identificadas por cada algoritmo.

Na Tabela 5 está representada as maiores comunidades identificadas pelo Louvain⁶, elas cobrem 586 artigos, o que representa 34,01\% do dataset total. Alguns destaques são: a comunidade 1 que tem como líder Diogo M. F. Mattos e possui o maior tamanho (130 nós), a comunidade 2 (119 nós) liderada por Joel J. P. C. Rodrigues (top 1 autor, com 33 artigos publicados), que apresenta a maior quantidade de coautores (108), a comunidade 5 (103 nós) que tem Glauber Dias Gonçalves como líder (top 3 autor, com 24 artigos publicados), na qual a colaboração mais frequente é com Alex Borges Vieira (top 2 autor, com 25 artigos publicados). A comunidade 11 (62 nós) também se sobressai, pois apresenta a colaboração mais frequente da rede (17 publicações), entre Roben Castagna Lunardi (Top 10 autor, com 20 artigos publicados) e Avelino F. Zorzo (Top 14 autor, com 18 artigos publicados). Demonstrando que os principais grupos de pesquisas estão concentrados em torno dos autores mais centrais.

O Fast Greedy⁷ obteve um resultado muito similar ao Louvain, com uma cobertura de 34,13\% das publicações (588 artigos). No entanto, apesar das semelhanças, as comunidades identificadas apresentam algumas diferenças em suas composições, assim como é possível visualizar na Tabela 6. Os pontos mais relevantes foram: a comunidade 2 (158 nós) liderada por Alex Borges Vieira, que mantém a colaboração mais frequente entre ele e Glauber Dias Gonçalves, a comunidade 3 (141 nós) que o líder é Diogo M. F. Mattos (top 9 autor, com 21 artigos publicados) e a comunidade 4 (119 nós), que é liderada por Joel J. P. C. Rodrigues, essas comunidades validam os núcleos de pesquisa similares aos identificados pelo Louvain.

⁶ <https://tinyurl.com/26mpm3lx>

⁷ <https://tinyurl.com/23alqmyg>

A maior comunidade agora é liderada por Marcos A. Simplicio, com um tamanho de 161 nós, indicando a predisposição deste algoritmo em criar comunidades maiores em comparação ao Louvain.

Tabela 5. Top 15 cobertura de artigos (Louvain): 591/1723 (34.30\%)

ID	Tamanho	Líder	Grau do líder	Colaboração mais frequente	# Colaborações	# Artigos
1	130	Diogo M. F. Mattos	78	Gabriel Antonio F. Rebello <> Gustavo F. Camilo	16	61
2	119	Joel J. P. C. Rodrigues	108	Ashok Kumar Das <> Joel J. P. C. Rodrigues	7	37
3	115	Marcos A. Simplicio	38	Diego Fernandes Gonçalves Martins <> Marco Aurélio Amaral Henriques	9	80
4	111	Célio V. N. de Albuquerque	35	Arthur A. Vianna <> Bruno T. Gondim	4	31
5	103	Glauber Dias Gonçalves	43	Alex Borges Vieira <> Glauber Dias Gonçalves	15	55
6	95	Pamella Soares	30	Jerffeson Souza <> Pamella Soares	13	66
7	78	Rodrigo da Rosa Righi	34	Cristiano André da Costa <> Rodrigo da Rosa Righi	16	38
8	70	Antônio Abelém	26	Antônio Abelém <> Billy Pinheiro	15	48
9	68	Alexander Leff	67	Alexander Leff <> Ann Charlotte Laska	1	1
10	66	Fabiola Greve	26	Fabiola Greve <> Jauberth Abijaude	7	27
11	62	Roben Castagna Lunardi	24	Avelino F. Zorzo <> Roben Castagna Lunardi	17	39
12	62	Wilson Vieira Melo	25	Luiz Fernando Rust da Costa Carmo <> Wilson Vieira Melo	7	32
13	55	Mário A. R. Dantas	15	Jean Everson Martina <> Lucas M. Palma	6	34
14	52	Johnny Marques	30	Juliana Maria Trindade Bezerra <> Celso Massaki Hirata	4	13
15	50	Vladimir Moreira Rocha	22	Arlindo Flávio da Conceição <> Vladimir Moreira Rocha	6	29

Figura 6. Top 15 cobertura de artigos (Fast Greedy): 588/1723 (34.13\%)

ID	Tamanho	Líder	Grau do líder	Colaboração mais frequente	# Colaborações	# Artigos
1	161	Marcos A. Simplicio	34	Antônio Abelém <> Billy Pinheiro	15	112
2	158	Alex Borges Vieira	46	Alex Borges Vieira <> Glauber Dias Gonçalves	15	81
3	141	Diogo M. F. Mattos	80	Gabriel Antonio F. Rebello <> Gustavo F. Camilo	16	64
4	119	Joel J. P. C. Rodrigues	108	Ashok Kumar Das <> Joel J. P. C. Rodrigues	7	37
5	100	Pamella Soares	30	Jerffeson Souza <> Pamella Soares	13	68
6	97	Célio V. N. de Albuquerque	35	Arthur A. Vianna <> Bruno T. Gondim	4	25

7	87	Johnny Marques	41	Jauberth Abijaude <> Péricles de Lima Sobreira	7	24
8	85	Rodrigo da Rosa Righi	35	Cristiano André da Costa <> Rodrigo da Rosa Righi	16	40
9	68	Alexander Leff	67	Alexander Leff <> Ann Charlotte Laska	1	1
10	62	Roben Castagna Lunardi	24	Avelino F. Zorzo <> Roben Castagna Lunardi	17	39
11	55	Mário A. R. Dantas	15	Jean Everson Martina <> Lucas M. Palma	6	34
12	46	Rodrigo Santaella Gonçalves	16	Benedito Cristiano Aparecido Petroni <> Rodrig...	6	28
13	43	Wilson Vieira Melo	25	Luiz Fernando Rust da Costa Carmo <> Wilson Vieira Melo	7	23
14	36	Paulo Henrique Alves	30	Isabella Frajhof <> Paulo Henrique Alves	7	11
15	32	A Lafontaine	31	A Lafontaine <> Adeel Abbas Dhahri	1	1

Figura 7. Top 15 cobertura de artigos (Walktrap): 295/1723 (17.12\%)

ID	Tamanho	Líder	Grau do líder	Colaboração mais frequente	# Colaborações	# Artigos
1	201	Diogo M. F. Mattos	97	Gabriel Antonio F. Rebello<> Gustavo F. Camilo	16	89
2	68	Alexander Leff	67	Alexander Leff<> Ann Charlotte Laska	1	1
3	53	Joel J. P. C. Rodrigues	52	Ashok Kumar Das<> Joel J. P. C. Rodrigues	7	18
4	43	Charles C. Miers	32	Charles C. Miers<> Marco Marques	9	29
5	42	Rodrigo da Rosa Righi	31	Cristiano André da Costa<> Rodrigo da Rosa Righi	16	22
6	38	Pamella Soares	30	Jerffeson Souza<> Pamella Soares	13	25
7	36	Antônio Abelém	23	Antônio Abelém<> Billy Pinheiro	15	30
8	32	A Lafontaine	31	A Lafontaine<> Adeel Abbas Dhahri	1	1
9	31	Paulo Henrique Alves	30	Isabella Frajhof<> Paulo Henrique Alves	7	9
10	30	Ricardo de Andrade Lira Rabêlo	29	Romuer R. V. Silva<> Ricardo de Andrade Lira Rabêlo	4	7
11	29	Jauberth Abijaude	22	Fabiola Greve<> Jauberth Abijaude	7	10
12	28	Vladimir Moreira Rocha	21	Arlindo Flávio da Conceição<> Vladimir Moreira Rocha	6	14
13	27	Roben Castagna Lunardi	24	Avelino F. Zorzo<> Roben Castagna Lunardi	17	21
14	25	Mário A. R. Dantas	15	Alex Sandro Roschidt Pinto<> Bruno Machado Agostinho	4	10
15	25	Rafael T. de Sousa	22	Emanuel Ferreira Jesus<> Vanessa R. L. Chicarino	4	9

O Walktrap⁸ foi o algoritmo que teve os resultados mais diferentes dentre os três, pois fragmentou mais a rede. Na Tabela 7 esses resultados são detalhados e pode-se perceber que a maior comunidade gerada possui 201 nós (comunidade 1, liderada por Diogo M. F. Mattos). Porém, as demais comunidades apresentam tamanhos bem menores (68 e 55 nós, 2º e 3º comunidades). Alguns dos principais líderes já não aparecem neste ranking, o que pode indicar que foram combinados na comunidade 1 ou em outras comunidades menores. O indicativo mais relevante foi a cobertura do dataset, apenas 295 publicações são compreendidas por esse algoritmo, o que representa 17,53\% do total. Como conclusão, o Walktrap foi menos coeso ao identificar os principais grupos de pesquisa, mesmo com a alta modularidade.

5. Conclusões

Este trabalho permitiu conhecer e entender de maneira mais detalhada a estrutura e a dinâmica dos relacionamentos acadêmicos na rede de colaborações científicas na área de blockchain, expondo características significativas, como: a baixa densidade e a forte tendência à formação de clusters, isso somado a existência de grupos de pesquisa consolidados em torno de autores e instituições centrais, revela que apesar da rede possuir grandes hubs, a produção encontra-se concentrada nos núcleos principais, com uma alta colaboração interna, mas pouca conexão entre as comunidades.

A comparação entre os métodos de detecção de comunidades indicou que a depender da escolha do algoritmo, os resultados podem apresentar algumas diferenças consideráveis, com uma maior aglutinação de nós por comunidade, por exemplo. O Louvain foi o algoritmo que mais se destacou devido a maior modularidade e menor tempo de execução, comparado aos demais algoritmos.

Entretanto, algumas limitações foram encontradas durante a realização desta pesquisa. Durante a fase de coleta, ao aplicar a filtragem pelo termo “blockchain”, trabalhos que utilizem nomenclaturas relacionadas ao tema, mas sem referenciar esta palavra-chave, podem ter sido descartados. A restrição de artigos publicados por instituições brasileiras também é um outro desafio enfrentado, pois depende da precisão dos metadados preenchidos nas bases de origem. Ademais, ainda que o uso de LLMs tenha auxiliado no pré-processamento reduzindo as inconsistências, a desambiguação de nomes se trata de um processo complexo e por isso ainda pode haver ruídos nos dados.

Como proposta para trabalhos futuros, é possível aprofundar-se na avaliação das comunidades encontradas, por meio de uma análise com ênfase nos principais temas abordados pelos pesquisadores de cada grupo, utilizando técnicas de extração de tópicos

⁸ <https://tinyurl.com/2yn3hq7m>

aplicadas aos títulos ou resumos (abstracts) das publicações. Essa abordagem possibilitaria compreender não somente as relações entre os autores e sua frequência, mas também os tópicos mais relevantes estudados por eles, objetivando uma perspectiva ainda mais valorosa sobre a evolução da produção científica em blockchain no Brasil.

5. Agradecimentos

Os autores agradecem à Rede Nacional de Pesquisa (RNP) pelo apoio neste trabalho. Todos os dados coletados e processados estão disponíveis para análises via *dashboards* interativos na seção de indicadores Observatório Nacional Blockchain⁹. Essa iniciativa faz parte do Projeto Ilíada executado pela RNP e CPQD no âmbito do MCTI PPI-Softex recursos da Lei 8248/1991.

⁹ <https://observatorioblockchain.org.br/indicadores>

7. Referências

- [Albert and Barabási 2002] Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47.
- [Aynaud and Guillaume 2010] Aynaud, T. and Guillaume, J.-L. (2010). Static community detection algorithms for evolving networks. In *8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, pages 513–519.
- [Barabási et al. 2002] Barabási, A. L., Jeong, H., Nédá, 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(3–4):590–614.
- [Blondel et al. 2008] Blondel, V. D., Guillaume, J., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.
- [Boaventura et al. 2014] Boaventura, M., Boson, K., Silva, A., Veloso, A., and Junior, W. M. (2014). Caracterização temporal das redes de colaboração científica nas universidades brasileiras: Anos 2000–2013. In *Anais do III Brazilian Workshop on Social Network Analysis and Mining*, pages 9–20, Porto Alegre, RS, Brasil. SBC.
- [Carvalho et al. 2025] Carvalho, L. P., Lima Filho, S., Brandão, M. A., Oliveira, J., Santoro, F. M., Silva, M. F. d., and Marinho, (2025). Csbcs (2013–2023) and brasnam (2012–2023): A panoramic meta-scientific study on quantitative and collaborations. *iSys: Revista Brasileira de Sistemas de Informação*, 18(1):3:1–3:44.
- [Clauset et al. 2004] Clauset, A., Newman, M. E. J., and Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, 70(6):066111.
- [Csardi and Nepusz 2006] Csardi, G. and Nepusz, T. (2006). The igraph software. *Complex syst*, 1695:1–9.
- [Dias et al. 2019] Dias, T. M. R., Moita, G. F., and Dias, P. M. (2019). Um estudo sobre a rede de colaboração científica dos pesquisadores brasileiros com currículos cadastrados na plataforma lattes. *Em Questão*, 25(1):63–86.
- [Digiampietri et al. 2014] Digiampietri, L., Alves, C., Trucolo, C., Delgado, K., and Mugnaini, R. (2014). Análise da rede de relacionamentos dos doutores brasileiros. In *Anais do VIII Brazilian e-Science Workshop*, pages 21–28, Porto Alegre, RS, Brasil. SBC.
- [Leão et al. 2017] Leão, J., Brandão, M., de Melo, P., and Laender, A. (2017). Classificação de relações sociais para melhorar a detecção de comunidades. In *Anais do VI Brazilian Workshop on Social Network Analysis and Mining*, pages 647–657, Porto Alegre, RS, Brasil. SBC.
- [Nakamoto et al. 2007] Nakamoto, S., Bit, B., et al. (2007). Bitcoin: A peer-to-peer electronic cash system. 2008.
- [Newman 2018] Newman, M. (2018). *Networks*. Oxford university press.

- [Newman 2006] Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582.
- [Pons and Latapy 2006] Pons, P. and Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, 10(2):191–218.
- [Porreca et al. 2025] Porreca, A., Maturo, F., and Ventre, V. (2025). Fuzzy centrality measures in social network analysis: Theory and application in a university department collaboration network. *International Journal of Approximate Reasoning*, 176:109319.
- [Tabassum et al. 2018] Tabassum, S., Pereira, F. S., Fernandes, S., and Gama, J. (2018). Social network analysis: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(5):e1256.
- [Zhang et al. 2025] Zhang, C., Feng, X., Li, Y., and Wang, Y. (2025). Scientific collaboration network structure and connectedness among the “belt and road” countries. *Journal of the Knowledge Economy*, 16(2):7908–7933.



**OBSERVATÓRIO
NACIONAL DE
BLOCKCHAIN**