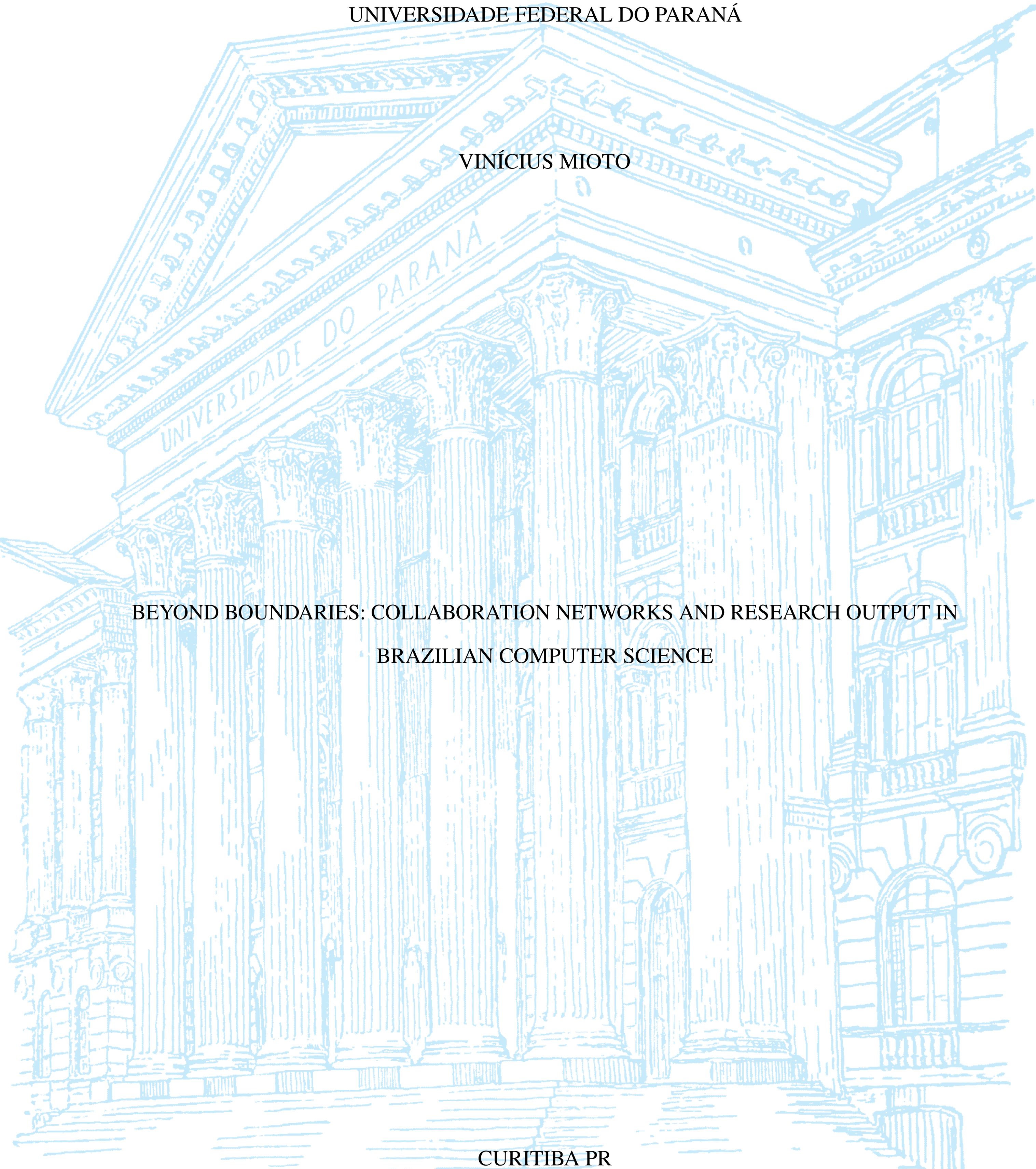
UNIVERSIDADE FEDERAL DO PARANÁ



VINÍCIUS MIOTO

BEYOND BOUNDARIES: COLLABORATION NETWORKS AND RESEARCH OUTPUT IN BRAZILIAN COMPUTER SCIENCE

CURITIBA PR

2025

VINÍCIUS MIOTO

BEYOND BOUNDARIES: COLLABORATION NETWORKS AND RESEARCH OUTPUT IN BRAZILIAN COMPUTER SCIENCE

Trabalho apresentado como requisito parcial à conclusão do Curso de Bacharelado em Ciência da Computação pela Universidade Federal do Paraná, setor de Ciências Exatas.

Área de concentração: *Computação*.

Orientador: André Luís Vignatti.

CURITIBA PR

2025

Ficha catalográfica

Substituir o arquivo `0-iniciais/catalografica.pdf` pela ficha catalográfica fornecida pela Biblioteca da UFPR (PDF em formato A4).

Instruções para obter a ficha catalográfica e fazer o depósito legal da tese/dissertação (contribuição de André Hochuli, abril 2019. Links atualizados Wellton Costa, Nov 2022):

1. Estas instruções se aplicam a dissertações de mestrado e teses de doutorado. Trabalhos de conclusão de curso de graduação e textos de qualificação não precisam segui-las.
2. Verificar se está usando a versão mais recente do modelo do PPGInf e atualizar, se for necessário (`https://gitlab.c3sl.ufpr.br/maziero/tese`).
3. conferir o *checklist* de formato do Sistema de Bibliotecas da UFPR, em `https://bibliotecas.ufpr.br/servicos/normalizacao/`
4. Enviar e-mail para "`referencia.bct@ufpr.br`" com o arquivo PDF da dissertação/tese, solicitando a respectiva ficha catalográfica.
5. Ao receber a ficha, inseri-la em seu documento (substituir o arquivo `0-iniciais/catalografica.pdf` do diretório do modelo).
6. Emitir a Certidão Negativa (CND) de débito junto a biblioteca, em `https://bibliotecas.ufpr.br/servicos/certidao-negativa/`
7. Avisar a secretaria do PPGInf que você está pronto para o depósito. Eles irão mudar sua titulação no SIGA, o que irá liberar uma opção no SIGA pra você fazer o depósito legal.
8. Acesse o SIGA (`http://www.prppg.ufpr.br/siga`) e preencha com cuidado os dados solicitados para o depósito da tese.
9. Aguarde a confirmação da Biblioteca.
10. Após a aprovação do pedido, informe a secretaria do PPGInf que a dissertação/tese foi depositada pela biblioteca. Será então liberado no SIGA um link para a confirmação dos dados para a emissão do diploma.

Ficha de aprovação

Substituir o arquivo 0-iniciais/aprovacao.pdf pela ficha de aprovação fornecida pela secretaria do programa, em formato PDF A4.

ACKNOWLEDGEMENTS

Em primeiro lugar, gostaria de expressar minha profunda gratidão aos meus pais por todo o suporte e apoio incondicional para que eu pudesse concluir essa etapa. Por meio de suas vivências, meus pais me ensinaram a importância de valorizar a educação, algo que, por circunstâncias da vida, eles não tiveram a oportunidade de desfrutar. Também agradeço aos meus amigos, principalmente à minha companheira Ana, por estarem ao meu lado, tanto nos dias de luta, quanto nos dias de glória.

"Se eu vi mais longe, foi porque estava sobre os ombros de gigantes". Portanto, expresso total admiração aos profissionais de educação que fizeram parte da minha jornada acadêmica. Destaco os professores André Vignatti, Carlos Maziero e Fabio Petrillo. Esses e outros profissionais são fontes de inspiração e conhecimento, pelos quais serei eternamente grato. Com uma sensibilidade especial, souberam reconhecer as particularidades e desafios enfrentados pelos alunos, oferecendo acolhimento e apoio. Essa dedicação, que vai além das obrigações acadêmicas, reflete um compromisso genuíno com a formação humana e com o futuro de cada estudante. Suas lições transcendem a sala de aula e continuarão a guiar minha trajetória.

RESUMO

Este trabalho analisa a colaboração em pesquisas de Ciência da Computação no Brasil utilizando métodos de ciência das redes e bibliometria. Apesar da alta produção científica, os padrões limitados de colaboração podem explicar o baixo impacto em citações. Investigamos: a extensão das colaborações internacionais entre as subáreas brasileiras e como os níveis de colaboração variam entre elas. Nossos resultados revelam que áreas teóricas favorecem parcerias internacionais, enquanto áreas aplicadas tendem a preferir colaborações domésticas. Identificamos poucos pesquisadores "ponte" que atuam como elementos chave na estrutura das redes de alto impacto. O estudo mapeia essas redes e demonstra como parcerias internacionais estratégicas se correlacionam com o impacto da pesquisa, oferecendo insights para o aprimoramento de políticas de colaboração e avaliação.

Palavras-chave: Colaboração. Coautoria. Ciência das Redes. Pesquisa. Ciência da Computação

ABSTRACT

This work analyzes Brazilian computer science research collaboration using network science and bibliometric methods. Despite high publication output, Brazil's limited collaboration patterns may explain its low citation impact. We examine: the extent of international collaborations across Brazilian subfields, and how collaboration levels vary among computer science subfields. Our findings reveal that theoretical areas favor international partnerships while applied areas prefer domestic collaborations. We identified few "bridge" researchers serving as key structural elements in high-impact networks. The study maps these networks and demonstrates how strategic international partnerships correlate with research impact, providing insights for improving collaboration and evaluation policies.

Keywords: Collaboration. Co-authorship. Network Science. Research. Computer Science

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

There is widespread agreement that collaborations, domestic or international, serve as a key metric for assessing the quality and relevance of postgraduate programs and research proposals. Researchers are subject to evaluations that inspect publication counts, citation metrics, journal impact factors, and collaborative networks, with international partnerships often conferring enhanced prestige and visibility (Ramos, 2017). Consequently, the existing evaluation system may exert pressure on researchers to pursue collaborations with authors from other subfields or other countries, potentially leading to shifts in academic research priorities. A notable example can be found in the domain of computer science, which has experienced remarkable growth and transformation in recent decades (Bird et al., 2009; Wainer et al., 2009; Madaan and Jolad, 2014).

Despite significant global growth in computer science research, regional disparities persist in research output, citation impact, and international participation. Emerging research communities, such as Brazil's, clearly demonstrate this disparity. While Brazil holds the 12th spot in global computer science publication output, an examination of its domestic and international collaborations suggests a possible link to reduced citation impact when compared to leading countries. In this context, understanding the collaborative structures that underpin successful research ecosystems becomes essential for countries seeking to enhance their global scientific position.

Collaboration networks provide valuable insights into knowledge production dynamics and can help identify strategic pathways for strengthening research communities. This analytical approach is particularly relevant for Brazil, where computer science research has exhibited consistent growth but continues to face challenges in achieving worldwide recognition proportionate to its publication volume. By examining these networks, we can better understand the structural dynamics of research collaborations, additionally, we might get insights of scientific trends and their impact in Brazilian computer science.

The present study examines the collaborative dynamics within Brazilian computer science research through network analysis and bibliometric methods. Our investigation addresses three key motivating questions:

- **Q1:** To what extent do Brazilian computer science engage in collaborative research with international researchers and institutions?
- **Q2:** How do collaboration levels vary among the different subfields of Brazilian Computer Science research?
- **Q3:** How do the structural characteristics of collaboration networks vary across different subfields in Brazilian Computer Science research?

To address the proposed motivating questions, we conduct a comprehensive analysis using data from OpenAlex, focusing on publication trends, citation metrics, and co-authorship relations from 2015 to 2024. Our findings reveal notable variations in collaboration intensity among different subfields, reflecting the diverse research cultures. Based on our analysis, we offer three significant contributions to the understanding of research collaboration in computer science:

- We quantify international collaboration rates by classifying publications as internationally collaborative when at least one co-author is affiliated with a non-Brazilian institution.

Our analysis compiles 11 computer science subfields, presenting detailed collaboration metrics, partner country distributions, and citation performance across international and domestic publications;

- We mapped interdisciplinary collaboration patterns by assigning each author a primary subfield based on publication history and categorizing publications as single-subfield or multi-subfield. Our approach enables a nuanced examination of interdisciplinary engagement, revealing publication rates, citation averages, and the frequency of authors publishing across different computer science subfields;
- By leveraging network analysis and visualization techniques, we noticed that recurrent co-authorship relations are unusual for highly cited works, and we also identify that these collaboration networks feature few "bridge" researchers who connect different research communities and countries, facilitating knowledge transfer.

The remaining of this work is structured as follows: Section 2 reviews prior work on collaboration networks in computer science research. Section 3 outlines our methodology, including data collection and analytical techniques. Section 4 presents an overview of the global research landscape in computer science. Sections 5, 6, and 7 address our motivating questions by analyzing international collaboration patterns, domestic subfield differences, and network dynamics, respectively. Finally, Section 8 concludes the work and proposes directions for future research.

2 RELATED WORK

Early network science studies established key structural features of scientific collaboration networks. (Newman, 2001a,b) observed power-law distributions in research productivity across various fields and noted that theoretical disciplines have fewer collaborators than experimental ones. (Barabási et al., 2002) demonstrated scale-free networks with preferential attachment in mathematics and neuroscience. (Elmacioglu and Lee, 2005) confirmed the "six degrees of separation" phenomenon in computer science using DBLP[1] data, yet (Madaan and Jolad, 2014) revealed increasing collaboration patterns and author counts per publication over time.

2.1 INTERNATIONAL COLLABORATION PATTERNS

International scientific collaboration research has evolved methodologically over time. (Luukkonen et al., 1993) established fundamental metrics distinguishing between absolute and relative measures. (Guan and Ma, 2004) applied these metrics across major countries, highlighting USA dominance in computer science publications. (Niu and Qiu, 2014) found that most of China's international co-authorships involved scientifically advanced nations.

In the Latin American context, (Wainer et al., 2009) positioned Brazil as a regional leader still lagging behind global scientific powers, while (Delgado-Garcia et al., 2014) documented strong regional ties, particularly between Brazil-Chile and Argentina-Brazil. (Pessoa Junior et al., 2022) showed how geographic proximity influences Brazilian collaboration networks and reflects regional economic disparities. Recently, (Okamura, 2023) identified a global "Shrinking World" collaboration phenomenon using OpenAlex data, noting post-2019 divergence between the US and China.

2.2 BRAZILIAN COMPUTER SCIENCE COLLABORATION NETWORKS

Several important studies have analyzed Brazilian scientific networks, primarily using the Lattes Platform curriculum database[2]. (Mena-Chalco et al., 2012) developed methodologies for automatic co-authorship identification and network characterization across 176,114 researcher profiles. (Digiampietri et al., 2012) examined co-authorship evolution among computer science professors, revealing increasing collaboration within and between graduate programs.

(Boaventura et al., 2014) found that networks across major Brazilian universities (2000-2013) displayed similar densification patterns, with small-world characteristics and an inverse relationship between internal collaboration and program quality. (Dias et al., 2018) showed that higher-level CNPq grant recipients demonstrate more balanced publication patterns between conferences and journals.

Regarding interdisciplinary engagement, (Júnior et al., 2018) analyzed 245,583 Brazilian researchers with doctorate degrees, revealing that 37.4% of co-authorship relationships cross disciplinary boundaries. (Carvalho et al., 2023) examined a decade of Brazilian Computer Science Society Congress interactions, finding distinct community characteristics but a fragmented collaboration landscape with only 39.36% of researchers in the giant network component.

[1]https://dblp.org
[2]https://lattes.cnpq.br

2.3 COMPUTER SCIENCE SUBFIELD COLLABORATION PATTERNS

Network analysis reveals significant variation in collaboration patterns across computer science specializations. Using DBLP data, (Bird et al., 2009) found high collaboration in Data Mining and Software Engineering contrasted with more isolated patterns in Theory and Cryptography. (Biryukov and Dong, 2010) tracked community evolution from 1970, showing smaller, less connected collaborations in Algorithms, Theory, and Programming Languages versus high collaboration intensity in Computational Biology and Web research.

(Franceschet, 2011) characterized computer science collaboration as moderate compared to other fields, with higher collaboration intensity in conference papers reflecting the field's conference-centric culture. For measuring collaboration quality, (Brandão et al., 2016) proposed "tieness" as a metric for co-authorship strength, showing that high-quality publications exhibit stronger collaborative ties.

When examining citation patterns, (Druszcz and Vignatti, 2024) identified significant disparities within Brazilian computer science subfields, with Computer Vision receiving substantially more citations than Algorithms or Formal Methods. This research highlighted limitations in applying uniform citation metrics across diverse specializations and proposed normalization techniques for fairer comparisons.

3 METHODOLOGY

We employ OpenAlex (Priem et al., 2022), an open-access scholarly index, for our analysis. It provides comprehensive metadata and, as shown by (Culbert et al., 2024), offers reference coverage and citation rates comparable to proprietary databases like Web of Science and Scopus. Furthermore, (Velez-Estevez et al., 2023) highlight the superior metadata of OpenAlex, including higher ORCID identifier rates, comprehensive institutional affiliations that support robust geographical analysis, and systematic classification of publications within disciplinary subfields.

Our initial exploratory analysis utilized the OpenAlex web GUI[1] to filter for computer science articles and book chapters published between 2015 and 2024. This yielded a dataset with key metrics, including publication count, citation data, open-access percentage, and rankings of countries, universities, subfields, and others. After initial exploration, we required more detailed data. While OpenAlex provides a full database snapshot, we chose to utilize their API for granular data retrieval, storing the results as CSV files. To effectively use the API, we studied the underlying database model, as illustrated by the simplified entity-relationship diagram in Figure 3.1.

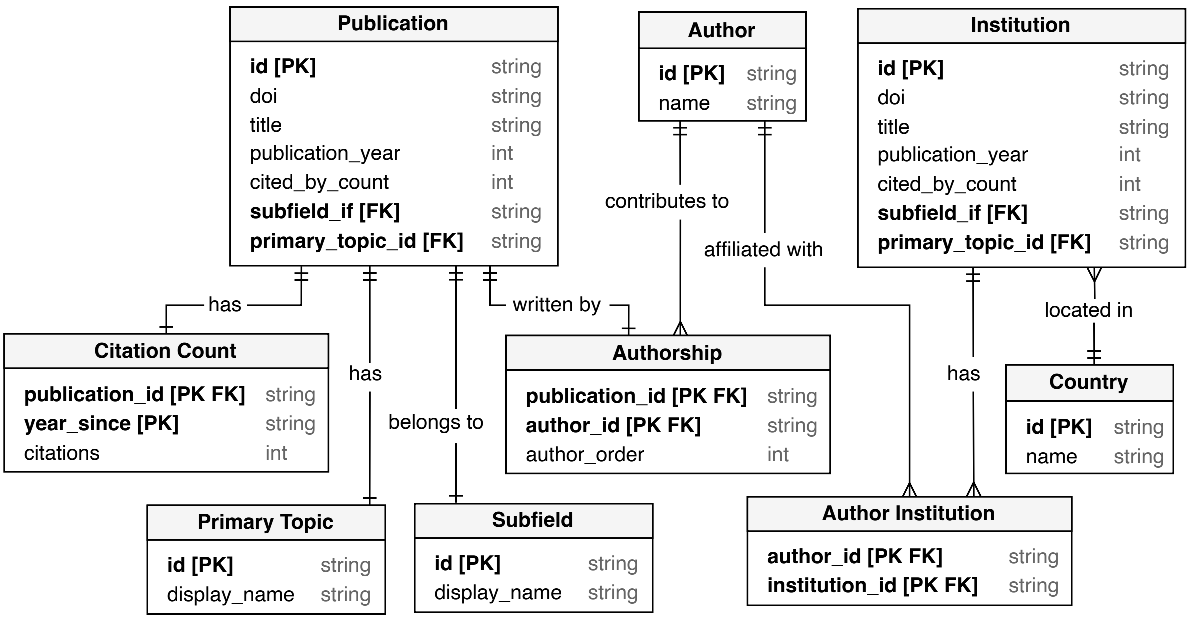


Figure 3.1: Simplified entity–relationship diagram from OpenAlex database. Adapted from OpenAlex API Documentation.

Figure 3.2 illustrates our complete *Extract*, *Transform*, and *Load* (ETL) pipeline. The *Extract* phase involved retrieving data from the OpenAlex API using Python scripts (version 3.13.0). We highlight that the API provides two levels of metadata: *API Response Metadata*, comprising overall statistics such as total record count, citation counts, and database response time; and *Publication Metadata*, which contains detailed information for each publication (e.g., title, authors, publication year, and country affiliations). This dual-level structure allowed us to efficiently gather both aggregate metrics and granular publication details.

[1]https://openalex.org/

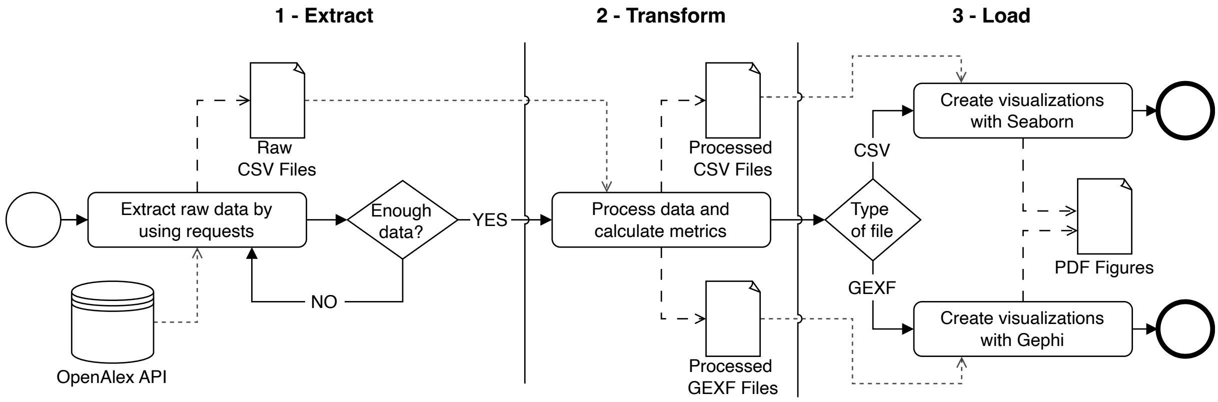


Figure 3.2: Complete pipeline for data collection and analysis.

For simpler metrics such as publication counts for each country and/or subfield, we utilized only the API Response Metadata. One must observe that by using exclusively the Response Metadata, we cannot apply specific filters or any data cleaning. Therefore, when it comes to detailed analyses, specifically for Brazilian publications, we retrieved the comprehensive Publication Metadata, which in turn provides much more information. Due to API request limitations, we implemented a batching strategy, handling pagination constraints (maximum 25 records per request) and organizing extraction by logical groups (e.g., one subfield at a time, one country at a time). The output of this phase is raw data stored in CSV.

In the second phase, *Transform*, we consolidated the batched data through concatenation and performed data cleaning to create processed files, for instance, we removed duplicated entries and works without DOI. For publication metrics comparisons we retained the CSV file format. For collaboration network analysis, we constructed co-authorship networks, where the nodes represent researchers and edges represent collaborative relationships, weighted by collaboration frequency. We exported these networks as GEXF files, by using the Python NetworkX library (version 3.4.2).

In the *Visualization* phase, we used two approaches: Seaborn (0.13.2) for CSV data plots and Gephi (0.10.0) for GEXF network visualizations. Both approaches produced PDF figures for publication. All code and documentation are available on our public GitHub repository[2].

[2]https://github.com/viniciusmioto/beyond_boundaries

4 GLOBAL RESEARCH LANDSCAPE

Understanding the global landscape is essential before analyzing Brazilian collaboration patterns. Therefore, in this section, we provide a global overview of computer science research from 2015 to 2024, by using exclusively the *API Response Metadata* from OpenAlex (see Section 3).

According to Table 4.1, Brazil ranks 12th globally in publication output with 76,184 publications and 447,919 citations. This positions Brazil in the middle tier of global research productivity, trailing significantly behind leaders such as China and the United States, yet surpassing other Latin American countries that failed to reach the top 15. However, Brazil's citation ratio of 5.88 falls considerably short of high-impact nations like Australia, Great Britain, and the United States, all of which exceed 20 citations per publication. More notably, this citation gap persists even when comparing Brazil to countries with similar publication volumes, such as Italy (12.67), Spain (12.52), and South Korea (13.37). This pattern reveals a clear disconnect between Brazil's publication volume and its citation impact, suggesting that while Brazilian researchers are productive in terms of output, their work may have limited international visibility or influence within the global research community.

	Code	Country	Total Publications	Citations	Ratio
1	CN	China	694,103	8,280,834	11.93
2	US	United States of America	474,474	9,590,230	20.21
3	IN	India	311,644	2,224,750	7.14
4	ID	Indonesia	266,047	755,078	2.84
5	DE	Germany	141,044	1,960,782	13.90
6	GB	Great Britain	140,019	3,026,576	21.62
7	JP	Japan	103,265	799,048	7.74
8	FR	France	92,131	1,011,787	10.98
9	CA	Canada	84,076	1,562,176	18.58
10	RU	Russian Federation	83,214	363,210	4.36
11	IT	Italy	80,647	1,021,944	12.67
12	BR	Brazil	76,184	447,919	5.88
13	ES	Spain	75,433	944,337	12.52
14	KR	South Korea	74,421	994,882	13.37
15	AU	Australia	68,502	1,561,906	22.80

Table 4.1: Research output and citation count (articles and book chapters only) metrics across 15 countries, showing publication volume, cumulative citations, and citation ratios. Data from 2015 to 2024, inclusive. Source: OpenAlex.

Figure 4.1 compares the research priorities of Brazil, China, the United States, and India, revealing unique patterns in Brazilian computer science research profile. Specifically, Brazil shows a clear specialization in Information Systems, maintaining 35-40% of its publications in this area throughout the analyzed period. This is significantly higher than the US (15%) and China, which started with 20% and ended with 10% in this subfield.

Artificial Intelligence (AI) is Brazilian second most productive area (20-25%), aligning with global trends but with a more moderate focus than the US and China, where AI output

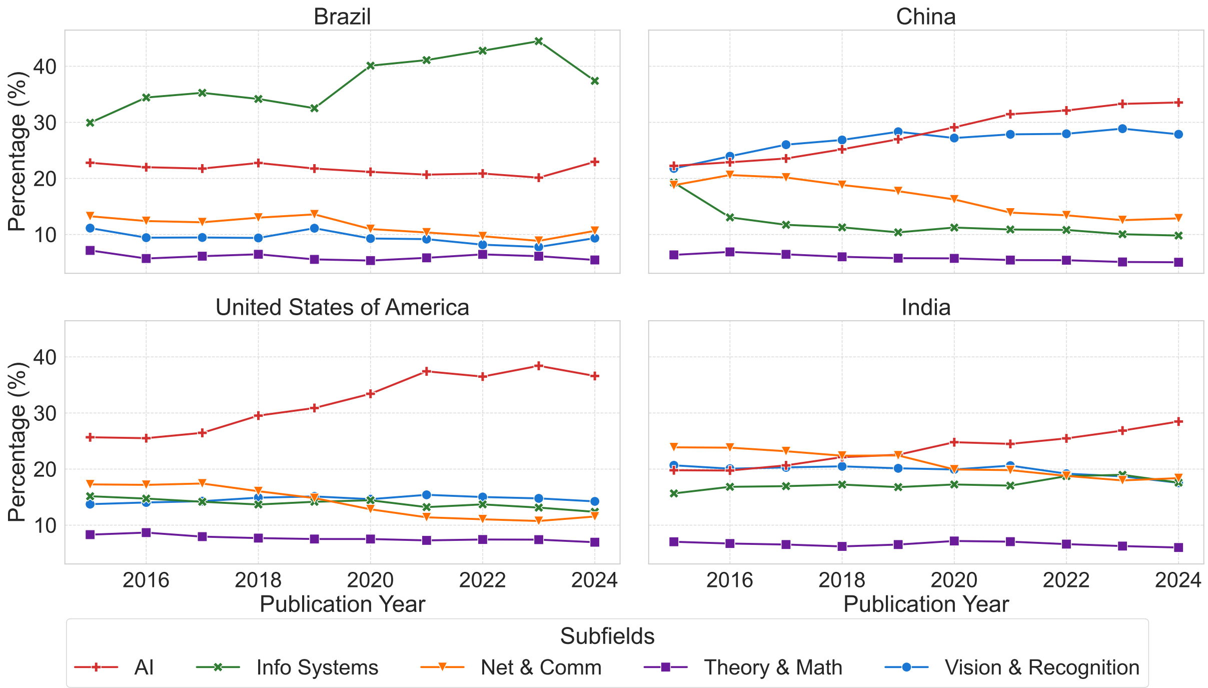


Figure 4.1: Distribution of publications in the top five computer science subfields for Brazil, China, the U.S., and India (2015–2024), relative to the total publications per year for each country.

has increased sharply from 30% to nearly 40% between 2015 and 2024. AI is the leading subfield in India, demonstrating a similar growth trend, albeit with less relative disparity. Brazil exhibits lower activity in Computer Vision & Pattern Recognition compared to the other countries. Computational Theory & Mathematics remains a relatively small area (less than 10%) across all four countries.

The stability of Brazilian research distribution across subfields from 2015 to 2024 contrasts sharply with other nations, especially the significant shift of China towards Artificial Intelligence. While data from 2023-2024 shows a slight decrease in Information Systems and increase in AI, the overall stability of Brazil may suggest entrenched institutional structures and funding mechanisms that favor continuity rather than realignment. This stability fosters deep expertise in specific domains, but may hinder adaptation to evolving global research priorities.

5 INTERNATIONAL COLLABORATION LEVELS (Q1)

To address our first research question, we examined international collaboration patterns by categorizing publications as internationally collaborative when they included at least one co-author affiliated with a non-Brazilian institution. As shown in Table 5.1, international collaboration rates vary substantially across computer science subfields. Computational Theory & Mathematics demonstrates the highest collaboration rate at 37.47%, followed by Computer Networks & Communications at 32.66%. In contrast, Information Systems, despite being the subfield with the largest publication volume, exhibits the lowest international collaboration rate at only 16.74%. Human-Computer Interaction represents another subfield with limited international engagement, with fewer than one-quarter of its publications involving foreign co-authors. When examining the complete dataset, we find that three-quarters of all publications (75.26%) remain exclusively domestic collaborations, indicating that international research partnerships, while present, is an unextreme mode of scientific collaboration in Brazilian computer science.

Subfield	**Domestic-only**		**International**	
	Publications	**%**	**Publications**	**%**
Theory & Math	2,471	62.54	1,480	37.46
Networks & Comm	5,309	67.34	2,575	32.66
Graphics & CAD	251	68.21	117	31.79
Hardware & Arch	811	69.38	358	30.62
AI	10,009	70.68	4,152	29.32
Vision & Recognition	4,470	71.93	1,744	28.07
Signal Processing	1,543	72.20	594	27.80
Software	583	72.24	224	27.76
CS Apps	2,260	74.69	766	25.31
HCI	1,443	78.42	397	21.58
Info Systems	22,135	83.26	4,450	16.74
Total	**51,285**	**75.26**	**16,857**	**24.74**

Table 5.1: International vs domestic publication in Brazilian computer science, with average citation counts.

Although this study does not primarily focus on citation analysis or research impact evaluation, we observed a striking disparity in citation performance between collaboration types. International publications receive more than three times the average citations of domestic-only publications (12.56 versus 3.95 citations). This citation advantage is particularly pronounced in Networks & Communications, Artificial Intelligence, and Information Systems, as evidenced when comparing domestic metrics in Table 5.2 with international metrics in Table 5.3.

However, citation distributions typically exhibit long-tailed characteristics, as noted by (Druszcz and Vignatti, 2024), where a small number of highly-cited publications skew the average upward while the majority of papers receive relatively few citations. This skewed distribution complicates straightforward interpretation of average citation counts. To better understand and validate this apparent citation gap between domestic and international collaborations, we expanded our analysis to include additional metrics that provide a more nuanced view of citation patterns (Tables 5.2 and 5.3).

Domestic	Avg.C	Std.D	Med	Q1	Q3	Max.C	%(=0)	%(<5)
Theory & Math	5.92	14.73	1.0	0.0	6.0	214	37.52	36.75
Networks & Comm	4.89	18.06	1.0	0.0	4.0	886	44.81	34.22
Graphics & CAD	3.00	7.65	1.0	0.0	3.0	81	45.82	38.65
Hardware & Arch	3.39	6.22	1.0	0.0	4.0	56	38.35	43.65
AI	5.51	16.81	1.0	0.0	5.0	480	41.83	36.14
Vision & Recognition	6.20	24.69	1.0	0.0	5.0	944	36.91	38.28
Signal Processing	5.60	15.26	2.0	0.0	6.0	354	32.40	42.26
Software	4.67	9.89	2.0	0.0	5.0	158	34.99	41.17
CS Apps	4.60	15.82	1.0	0.0	4.0	366	38.67	42.92
HCI	4.24	10.19	1.0	0.0	4.0	121	42.00	38.88
Info Systems	2.16	9.32	0.0	0.0	1.0	745	60.48	30.61
Total	**3.95**	**14.68**	**1.0**	**0.0**	**3.0**	**944**	**49.02**	**34.52**

Table 5.2: Detailed Citation Metrics for Domestic-only Collaboration by Subfield.

International	Avg.C	Std.D	Med	Q1	Q3	Max.C	%(=0)	%(<5)
Theory & Math	13.87	101.91	4.0	1.0	11.0	3,587	20.74	37.50
Networks & Comm	14.30	35.22	4.0	1.0	14.0	540	22.83	33.05
Graphics & CAD	7.82	19.45	2.0	0.0	7.0	131	29.06	42.74
Hardware & Arch	6.88	13.44	3.0	0.0	8.0	145	25.70	41.06
AI	14.11	46.82	4.0	1.0	13.0	1,730	24.04	33.16
Vision & Recognition	14.43	33.77	4.0	1.0	13.0	508	22.88	34.52
Signal Processing	12.96	32.99	5.0	1.0	12.0	519	19.19	33.67
Software	11.17	20.07	3.0	1.0	15.0	145	19.20	37.95
CS Apps	10.21	41.64	2.0	0.0	8.0	885	31.46	37.99
HCI	12.35	83.41	2.0	0.0	7.0	1,616	30.48	38.29
Info Systems	9.95	200.14	1.0	0.0	4.0	13,248	45.48	32.45
Total	**12.56**	**112.38**	**3.0**	**0.0**	**10.0**	**13,248**	**29.43**	**34.13**

Table 5.3: Detailed Citation Metrics for International Collaboration by Subfield.

The median values further illustrate this gap, with domestic papers generally receiving just one citation compared to three citations for international work. The impact extends beyond averages and medians, while nearly half of domestic publications (49.02%) receive zero citations, only 29.43% of international collaborations remain on zero. For the publications with less than five citations the results are more similar between the collaboration types. When it comes to the quartiles, domestic type got zero citation for all subfields in the first quartile. Contrariwise, this does not happen with Graphics & CAD, Hardware & Architecture, CS Applications, HCI, and Information Systems. The third quartile shows greater advantage for all subfields in the international side.

Despite their superior overall performance, international collaborations show much higher standard deviation results, this causes a wider range of outcomes, from breakthrough papers achieving exceptional visibility to others with modest impact. Figure 5.1 displays the distribution of citation counts for the two types of collaborations in a logarithmic scale for both axes, and it confirms the long-tailed characteristic for both types of collaborations. Analyzing the distribution plot, we notice that international collaborations show a more extended tail in the citation distribution, with some publications achieving exceptionally high citation counts, and this matches with the metrics of the previous Tables 5.2 and 5.3.

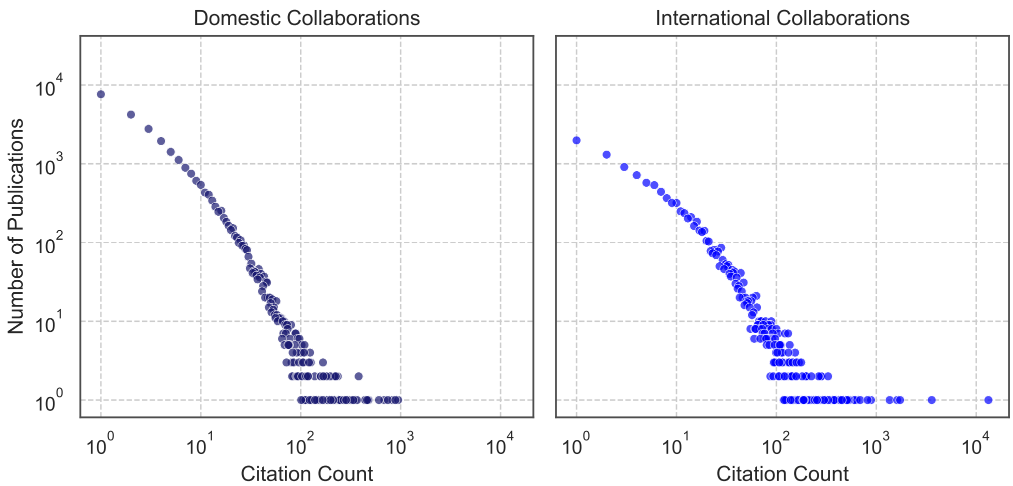


Figure 5.1: Citation distribution for domestic and international collaboration types.

Figure 5.2 depicts the intensity variation by partner country and subfield. The United States is the primary collaborator of Brazil across all subfields, with strong ties in Computational Theory & Mathematics (20.9%), Graphics & Computer-Aided Design (20.7%), and Hardware & Architecture (20.3%). Portugal emerges as the second most significant collaborator, likely facilitated by shared linguistic and cultural ties. Other European countries like Spain, Germany, and France exhibit specialized collaborations, with France notably engaging in Hardware & Architecture and Theory & Mathematics.

To examine researcher mobility patterns between Brazilian and foreign institutions, we tracked changes in institutional affiliations of the authors as documented in their publication records. Our methodology involved grouping publications by individual authors and organizing them chronologically to identify shifts in institutional ties over time. We classified a mobility event as occurring when an author's primary institutional affiliation changed from Brazilian to foreign institutions (*Brazilian to Foreign*) or vice versa (*Foreign to Brazilian*). For instance, if an author published with a Brazilian institution in 2016 and subsequently appeared with a foreign institutional affiliation in 2021, we recorded this as a Brazilian to Foreign movement in 2021.

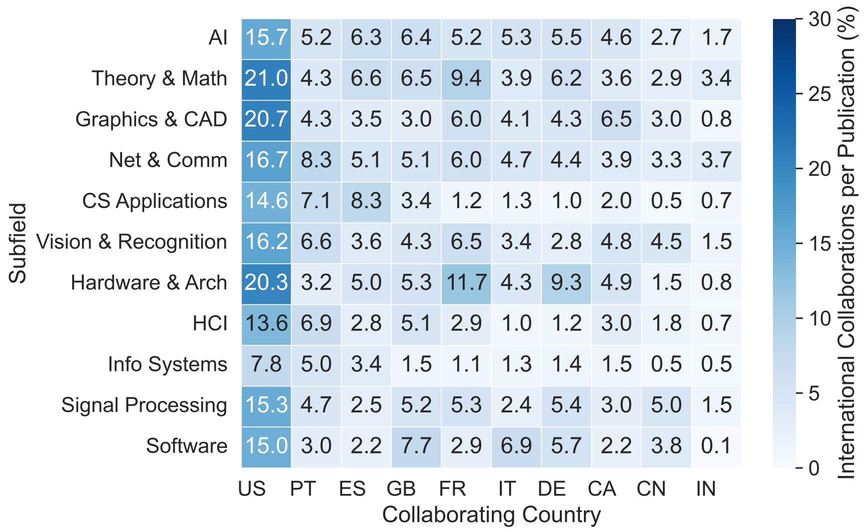


Figure 5.2: International collaboration heatmap for Brazilian CS publications by subfield and partner country. Values represent the percentage of international collaborations normalized by total publications in each subfield.

Figure 5.3 shows that more researchers moved out of Brazil than into it during the last decade. In 2021, for instance, around 240 authors moved from Brazilian to foreign institutions, while about 170 moved in the opposite direction. This outward trend is consistent across most of the years analyzed. These numbers may reflect several factors, such as graduate students seeking master's or PhD degrees abroad, participation in exchange programs, or researchers looking for better opportunities overseas. On the other hand, authors that have experienced such abroad opportunities, might come back after it, so they also corroborate to the Foreign to Brazilian movement. Still, the smaller gap in recent years could suggest a growing ability of Brazilian institutions to attract or retain talent. It is worth noting that our analysis covers only a ten-year period, which may miss longer-term trends or temporary disruptions. Additionally, our dataset includes publications where at least one author is affiliated with a Brazilian institution. A broader timeframe and deeper contextual data would help paint a more complete picture.

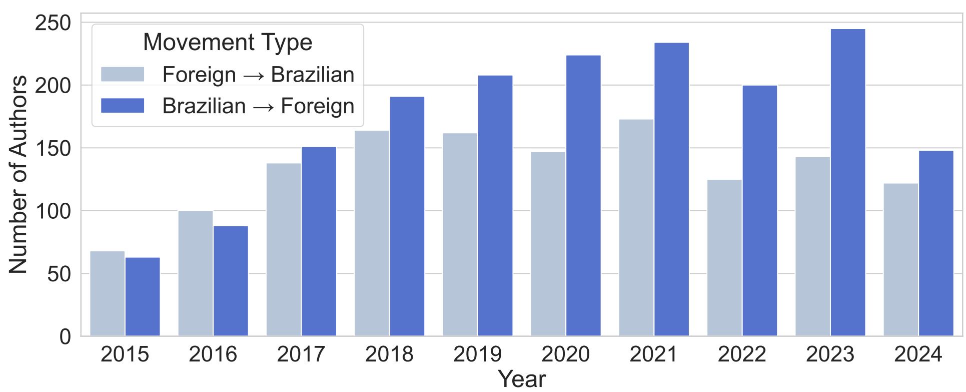


Figure 5.3: International movements over the years.

Figure 5.4 shows the temporal perspective of the international collaboration rates across the top five subfields in terms of publication volume (non-cumulative). The overall upward trajectory across all five subfields in the most recent years (2022-2024) indicates a growing trend toward international research partnerships in Brazilian Computer Science. Theory & Math maintained the highest level of international engagement, with a notable increase to more than

45% in 2024. Networks & Communications demonstrates stable international collaboration rates of approximately 30-35% over the period. Information Systems consistently had the lowest international collaboration rate of approximately 15-20%, though it shows a slight upward trend in recent years.

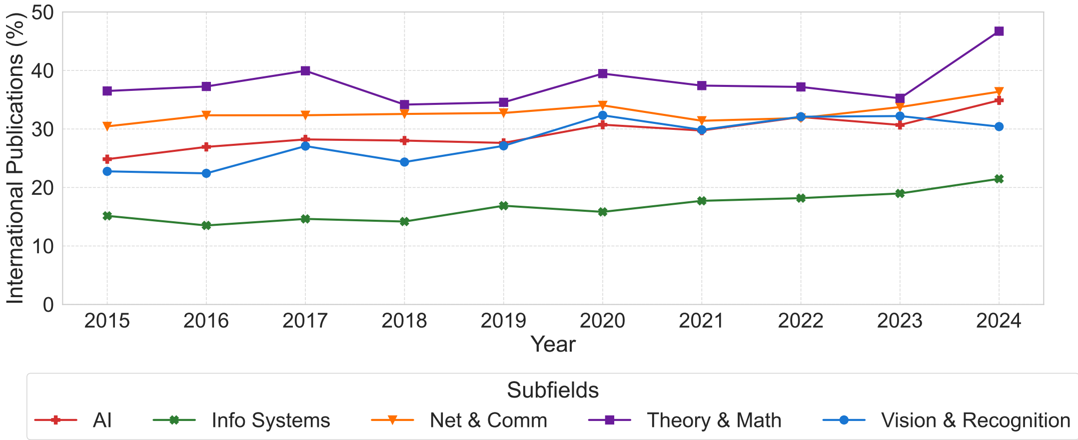


Figure 5.4: International collaboration trends for the top five Brazilian CS subfields by publication volume.

[Q1] Main Observations

Subfield variation: International collaboration varies significantly by subfield. Theory & Math leads international collaboration (37%), while Info Systems lags (16%).
Citation advantage: International papers receive more citations (12.56 vs 3.95) and have lower zero-citation rates (29% vs 49%).
Partner countries: The United States dominates as primary partner across all subfields; Portugal and other European countries serve as key secondary collaborators.
International flow: More computer science researchers moved out of Brazil than into it during the last decade.
International trend: All major subfields show increasing international collaboration (2022-2024), with Theory & Mathematics reaching more than 45% in 2024.

6 COLLABORATION DIFFERENCES ACROSS SUBFIELDS (Q2)

Our analysis reveals distinct patterns of interdisciplinary collaboration across Brazilian computer science subfields. We first classified each publication as either *single-subfield* (all authors from the same primary subfield) or *multi-subfield* (at least one author from a different primary subfield). To enable this classification, we assigned each author a primary subfield based on their publication history. In cases where an author had an equal number of publications in multiple subfields, we applied tie-breaking rules by sorting the publications within the tied subfields by year and number of citations (both descending). We found 5,069 authors that had fewer than five publications or which the tie-breaking strategy was not enough to classify them, so we label their primary subfield as "unknown". We excluded the publications in which all authors had an unknown primary subfield.

As seen in Table 6.1, Software leads in multi-subfield publications with 57%, followed by Signal Processing (54%), and Hardware & Architecture (51%). However, these three subfields have a small amount of total publications. In contrast, Information Systems shows the lowest interdisciplinary collaboration rate (13%) despite having the largest publication volume, with Theory & Math (19%) and AI (27%) also favoring single-subfield collaboration. Multi-subfield publications generally achieve higher average citations (6.12) than single-subfield works (3.53), suggesting a broad advantage for interdisciplinary research.

Subfield	**Single-subfield**			**Multi-subfield**		
	Publications	%	Avg. Cit.	Publications	%	Avg. Cit.
Software	183	42.96	3.82	243	57.04	4.81
Signal Processing	563	46.34	4.54	652	53.66	4.85
Hardware & Arch	295	48.60	3.18	312	51.40	3.36
Graphics & CAD	108	54.55	2.96	90	45.45	3.44
CS Apps	1,116	57.59	3.30	822	42.41	5.36
HCI	690	58.03	2.83	499	41.97	7.79
Vision & Recognition	2,266	59.76	5.76	1,526	40.24	7.62
Net & Comm	2,970	65.03	5.69	1,597	34.97	6.38
AI	6,252	72.07	5.09	2,423	27.93	7.08
Theory & Math	1,959	82.87	5.70	405	17.13	5.99
Info Systems	15,604	87.76	1.91	2,177	12.24	4.80
Total	**32,006**	**74.86**	**3.53**	**10,746**	**25.14**	**6.12**

Table 6.1: Summary of single and multi-subfield publications across subfields with corresponding percentages and average citations counts.

Similarly to the previous section, the citation patterns follow a long-tailed distribution, as we can see on Figure 6.1. Despite the multi-subfield category showing higher citation averages, the difference here is much more subtle when we compare to the domestic versus international partition from the Section 5. This suggests that while interdisciplinary collaboration may enhance visibility and impact, its benefits are less pronounced than those resulting from international partnerships.

Figure 6.2 shows divergent trajectories in interdisciplinary engagement for the five largest subfields by publication volume (non-cumulative). Although all five of these high-volume

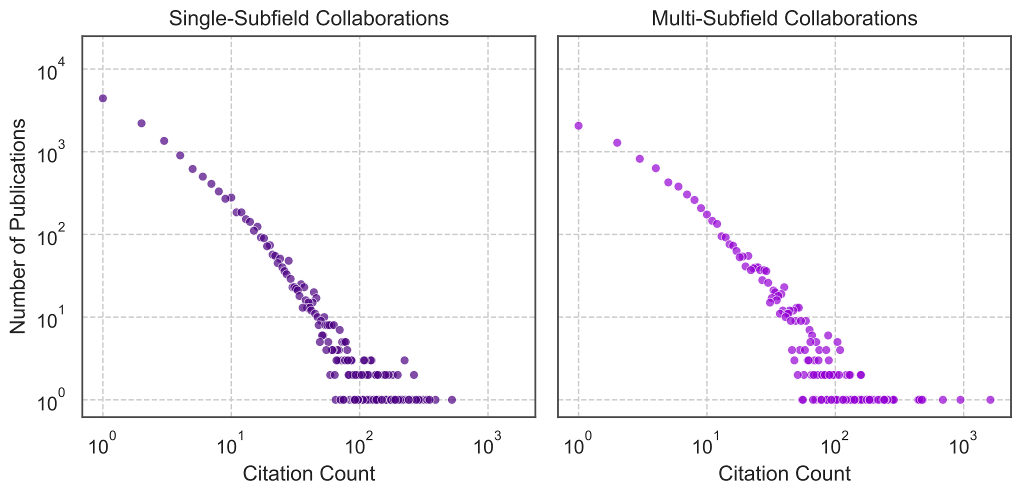


Figure 6.1: Citation distribution for single-subfield and multi-subfield collaboration types.

subfields rank among those with the lowest proportions of multi-subfield output in Table 6.1, in Vision & Recognition, the share of multi-subfield publications rises from roughly 45% in 2015 to almost 60% by 2024, indicating that researchers in this area increasingly collaborate with colleagues from other specializations. A similar upward trend appears in Networks & Communications and Artificial Intelligence, where multi-subfield publications account for approximately 50% and 40%, respectively. By contrast, Information Systems and Computational Theory & Mathematics exhibit more erratic patterns in multi-subfield collaboration. Information Systems reached less than twenty percent, and Theory & Mathematics approximately thirty percent in the last year.

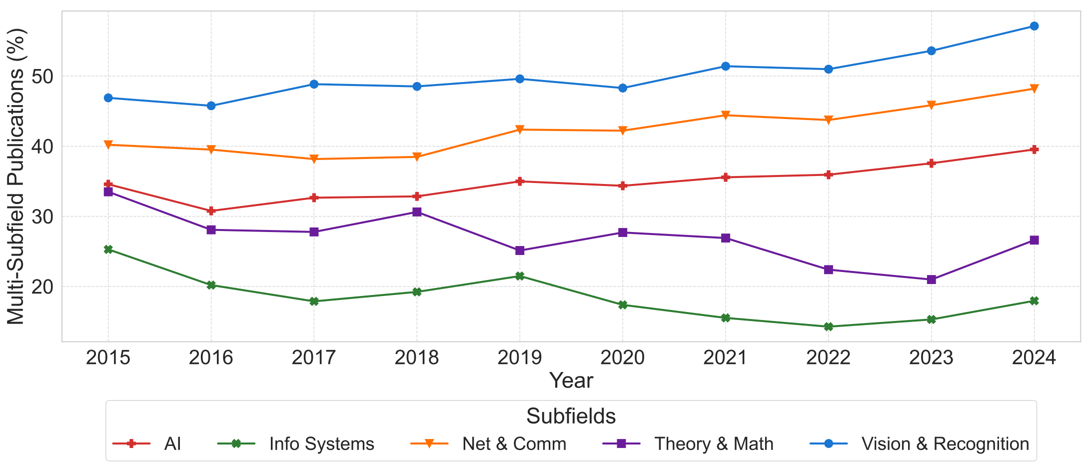


Figure 6.2: Multi-subfield collaboration trends for the top five Brazilian CS subfields by publication volume.

Table 6.2 exhibits which secondary and tertiary subfields authors publish in most frequently beyond their primary specialization, along with the percentage of authors who have published across multiple subfields. Hardware & Architecture researchers appear as the most interdisciplinary authors (19.25%), frequently publishing with Networks & Communications and Artificial Intelligence colleagues. Although Information Systems has the highest number of researchers, it demonstrates the lowest cross-disciplinary engagement, with only 3.62% of its authors publishing in AI and Networks & Communications.

Primary Subfield	Authors	Cross-field %	2nd Subfield	3rd Subfield
Hardware & Arch	1,392	19.25	Net & Comm	AI
Signal Processing	2,478	14.77	AI	Vision & Recognition
Net & Comm	10,696	13.77	Info Systems	AI
Vision & Recognition	9,802	13.50	AI	Net & Comm
Software	868	12.44	Info Systems	AI
AI	22,909	9.75	Info Systems	Vision & Recognition
CS Apps	4,788	9.73	Info Systems	AI
HCI	3,039	9.51	Vision & Recognition	Info Systems
Graphics & CAD	537	7.45	Vision & Recognition	AI
Theory & Math	7,151	6.70	AI	Net & Comm
Info Systems	60,121	3.62	AI	Net & Comm

Table 6.2: Subfields researchers most frequently publish in beyond their primary specialization and the percentage of authors that have published in other subfields.

Comparing Tables 6.1 and 6.2 we noticed a subtlety picture of cross-disciplinary engagement among the subfields. While Software and Hardware & Architecture have higher proportions of multi-subfield publications, their overall publication volumes and number of authors are relatively small. In contrast, Artificial Intelligence and Information Systems exhibits a lower percentage of multi-subfield publications, which might suggest a more focused, within-field publication pattern. On the other hand, Table 6.2 shows that both AI and Information Systems frequently appear as a secondary or tertiary subfield for many authors from other specializations. This juxtaposition reveals that while AI and Information Systems primarily publish within their own domains, they simultaneously serve as critical collaborative partners for researchers from other, less prolific subfields.

Table 6.3 presents a summary of authorship metrics across various computer science subfields in Brazilian publications. The average number of authors per paper is relatively consistent across subfields, ranging from 3.22 in Graphics & CAD to 3.95 in Signal Processing, with an overall average of 3.72. Despite these similar averages, the data reveals substantial variation in collaboration practices. Subfields such as Vision & Recognition, AI, and Theory & Mathematics show extremely high maximum author counts (exceeding 100), suggesting the presence of large collaborative projects. In contrast, subfields like Graphics & CAD and Software report more modest maximums (10 and 23, respectively). The percentage of single-authored papers also varies widely, with Information Systems (11.17%), Theory & Mathematics (10.63%), and HCI (10.33%) showing the highest rates of solo authorship, potentially reflecting more theoretical or conceptual contributions. On the other hand, Hardware & Architecture and Signal Processing have the lowest rates of single authorship, under 2.2%, indicating stronger norms of team collaboration.

Subfield	Avg Authors	Std.Dev	Median	Max Authors	%Single Auth
Signal Processing	3.95	2.30	4.00	45	2.15
Vision & Recognition	3.92	2.55	3.00	+100	4.60
Hardware & Arch	3.91	2.09	4.00	28	1.71
Net & Comm	3.81	2.49	3.00	56	6.52
Info Systems	3.73	2.56	3.00	58	11.17
AI	3.62	2.70	3.00	+100	8.64
Theory & Math	3.60	3.32	3.00	+100	10.63
CS Applications	3.60	2.13	3.00	41	8.10
HCI	3.55	2.30	3.00	42	10.33
Software	3.52	1.93	3.00	23	5.20
Graphics & CAD	3.22	1.69	3.00	10	5.43
Total	**3.72**	**2.58**	**3.00**	**+100**	**8.77**

Table 6.3: Summary of Author Metrics per Subfield CS for Brazilian Publications.

[Q2] Main Observations

Subfield variation: Software, Signal Processing, and Hardware & Architecture lead in multi-subfield publications, while Info Systems and Theory & Math are more insular.
Citation advantage: Multi-subfield papers have higher average citations (6.12 vs. 3.53), suggesting broader impact.
Interdisciplinary trend: High-volume subfields like Vision & Recognition, AI, and Networks & Communications increased their multi-subfield output over time.
Cross-subfield authors: Hardware & Arch and Signal Processing researchers are more likely to publish outside their primary area. AI is the most common secondary and tertiary subfield in general.
Authorship patterns: Some subfields have larger teams and fewer solo papers (e.g., Signal Processing and Vision & Recognition), while others like Theory & Math and Info Systems have more single-authored work.

7 NETWORK STRUCTURE AND DYNAMICS (Q3)

Understanding the underlying structure of co-authorship networks can help us to identify the dynamics within a research ecosystem. Thus, this section delves into the topological characteristics of Brazilian computer science collaborations, employing network science metrics to uncover insights into their organization and the roles of individual researchers. Table 7.1 provides the metrics of the structural properties of collaboration networks within different subfields.

Network	**Publications**	N	E	F	C	N_{LC}
Info Systems	26,585	65,349	196,087	0.83	0.80	26,543
AI	14,161	28,636	119,515	0.64	0.77	17,211
Net & Comm	7,884	14,701	48,756	0.64	0.78	8,763
Vision & Recognition	6,214	13,776	49,678	0.69	0.82	7,641
Theory & Math	3,951	8,568	46,513	0.87	0.80	3,132
CS Apps	3,026	7,023	17,872	0.85	0.78	2,705
Signal Processing	2,137	4,399	13,340	0.89	0.82	1,409
HCI	1,840	4,516	12,116	0.93	0.78	1,174
Hardware & Arch	1,169	2,276	7,084	0.65	0.83	1,345
Software	807	1,736	4,221	0.86	0.80	630
Graphics & CAD	368	925	1,709	0.98	0.79	105
Full network	**68,142**	**128,847**	**492,524**	**0.64**	**0.78**	**76,770**

Table 7.1: Network metrics for Brazilian computer science subfields. N: number of nodes; E: number of edges; F: fragmentation; C: average clustering coefficient; and N_{LC}: number of nodes in the largest connected component.

The full network, encompassing all publications, has $N = 119,228$ nodes (N) and 465,163 edges (E), with a fragmentation (F) of 0.64 and an average clustering coefficient (C) of 0.79. This suggests a relatively connected network with a high propensity for collaborators of an author to also be collaborators with each other. Comparing individual subfields, Information Systems exhibits the largest network with 57,243 nodes and 179,780 edges, it also has highest value for the largest connected component. Yet it has a high fragmentation of 0.84, indicating a less interconnected structure compared to its size. Conversely, subfields like Networks & Communications and AI show lower fragmentation values (0.60 and 0.62 respectively), suggesting more integrated collaboration patterns within these areas despite having fewer nodes and edges.

To examine collaborations in highly cited publications, we filtered all the publications with at least 40 citations, we choose this number because it gives us approximately 10% of the nodes of the full network (6,811). Furthermore, to investigate recurrent collaborations in such highly cited network, we filtered the edges where the weight is at least 2 (authors have published two or more works together), and removed the isolated nodes. Table 7.2 shows the metrics for the highly cited publications network and for the respective recurrent collaboration network. We highlight the huge drop in the number of edges and the clustering coefficient value. Furthermore, the recurrent highly cited network has an extremely high fragmentation as expected, due to the edge weight filter (Borgatti, 2006). These metrics shows that recurrent collaborations are unusual in highly cited works

Figure 7.1 displays this network colored by the country (affiliation) of the author, with the Fruchterman-Reingold layout algorithm to arrange nodes based on their connectivity patterns.

Network	**Publications**	N	E	F	C	N_{LC}
Highly Cited	1,775	6,811	57,515	0.85	0.90	2,573
Recurrent Highly Cited	1,775	1,063	2,237	0.97	0.65	137

Table 7.2: Network metrics for specific collaboration networks. N: number of nodes; E: number of edges; F: fragmentation; C: average clustering coefficient; and N_{LC}: number of nodes in the largest connected component.

The network features a structure with several small clusters and few larger, connected components, which underlines the fragmentation and clustering metrics of Table 7.2. The largest component consists predominantly of Brazilian authors (gray nodes) forming a central hub, surrounded by international collaborators. This clustering pattern suggests that while many researchers collaborate within smaller, tightly knit groups, fewer engage in broader international networks. Notably, authors from the United States (dark blue) and European countries such as Italy (light blue), Germany (purple), and France (green) appear frequently in co-authorship relationships with Brazilian researchers. These international collaborations, however, appear concentrated around specific sub-networks rather than being evenly distributed throughout the network.

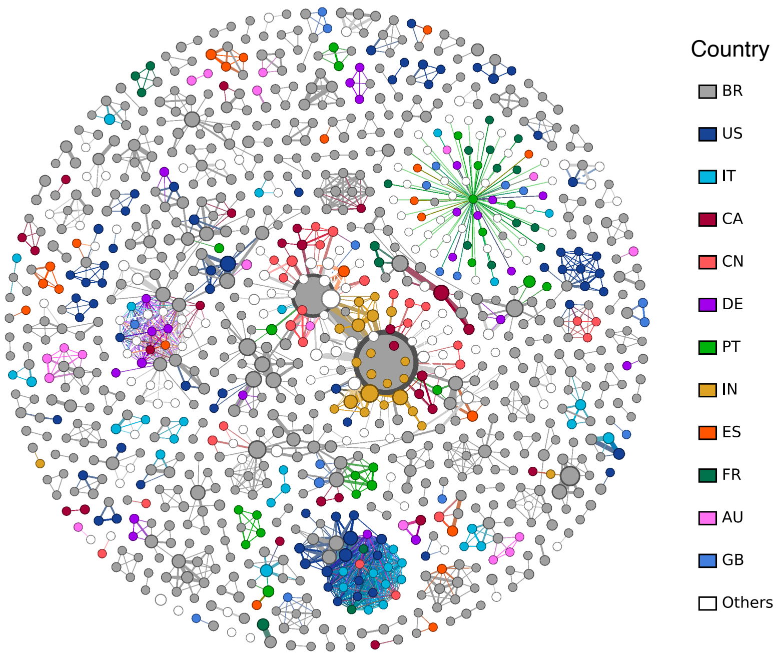


Figure 7.1: Network visualization of international co-authorship recurrence patterns in Brazilian computer science publications with at least 40 citations. Nodes represent authors and edges co-authorship recurrence.

The analysis of collaboration networks and co-authorship rates (see Section 5) shows a strong preference for domestic partnerships, with 75% of publications involving only Brazilian researchers. However, a few hubs have strong presence of international authors (Figure 7.1), localized pockets of global collaboration. It's also important to note that recurrent co-authorship may come from extended versions of earlier work or book chapters derived from smaller publications. Overall, while these patterns point to a solid national research base, they also suggest opportunities to strengthen international collaboration.

Figure 7.2 shows the primary subfield of the authors in the recurrent highly cited network. The visualization exhibits distinct clustering patterns, with clear modularity for Information Systems, Theory & Mathematics, and Artificial Intelligence. This indicates that despite cross-disciplinary efforts, highly cited research primarily emerges from within-field collaborations.

The exception lies in bridge nodes, researchers connecting otherwise separate communities, who facilitate knowledge transfer across disciplines, particularly between AI and Computer Vision, and between Networks and Information Systems.

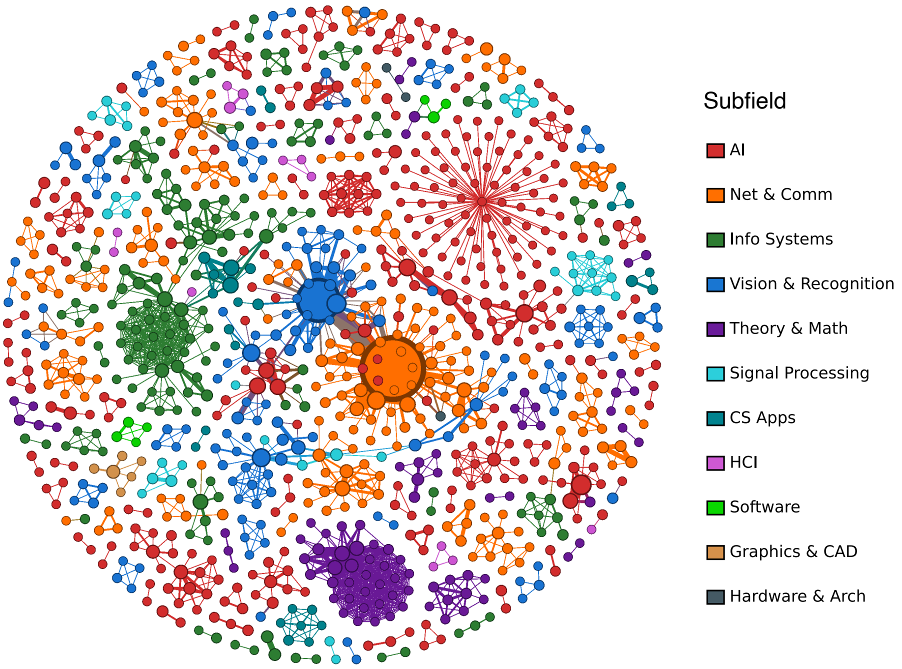


Figure 7.2: Network visualization of subfield co-authorship recurrence patterns in Brazilian computer science publications with at least 40 citations. Nodes represent authors and edges co-authorship recurrence.

In a collaboration network, betweenness centrality measures how often a researcher acts as a "bridge" connecting separate groups in a collaboration network (Freeman, 1978). Betweenness centralization quantifies how reliant a network is on a few key "connectors" to link its members (Bird et al., 2009). Figure 7.3 shows Hardware & Architecture consistently maintaining the highest centralization, indicating influential researchers who bridge previously separate research groups. This subfield also experienced dramatic centralization growth until 2022 before declining sharply. AI demonstrated stable, moderate centralization throughout the period, while Software exhibited significant fluctuations suggesting community reorganization. Computer Vision & Patter Recognition shows an increasing peak until 2019, followed by a more stable period, indicating that some researchers have emerged and dominated a big portion of collaborations in the first period. The overall centralization increase until 2022 followed by decline potentially signals a maturing research ecosystem.

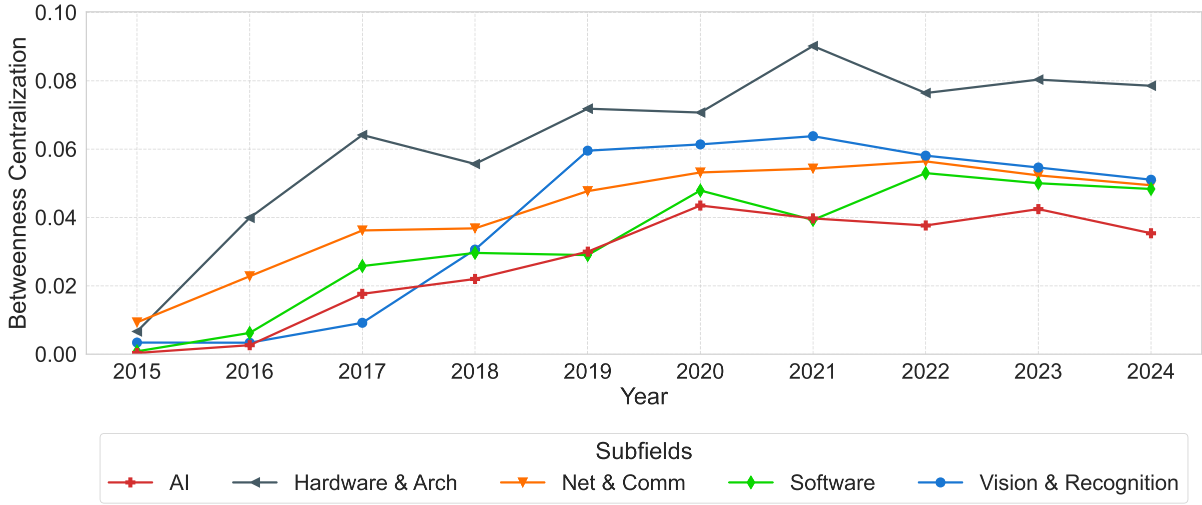


Figure 7.3: Betweenness centralization of top 5 most centralized CS subfields in Brazil from 2015 to 2024.

[Q3] Main Observations

Fragmentation variations: Some areas (e.g., Info Systems $F = 0.83$, HCI $F = 0.93$) consist of many small, disconnected author clusters, whereas Net & Comm ($F = 0.64$) and AI ($F = 0.64$) are more interconnected.
Similar clustering: Most of the subfields have a clustering coefficient around 0.78–0.83, indicating that co-authors of an author tend to collaborate with each other.
Highly cited networks: Brazilian authors form central hubs, and international authors are not evenly distributed over the network, suggesting that international partnerships concentrate around a few key researchers.
Interdisciplinary collaborations: Cross-disciplinary edges exist but largely rely on a handful of "bridge" researchers who connect otherwise separate communities.

8 CONCLUSIONS AND FUTURE DIRECTIONS

This study analyzes the dynamics of collaboration in Brazilian computer science research, using bibliometric data from OpenAlex and network analysis. Regarding our first motivating question, we found that approximately 75% of Brazilian computer science publications involve exclusively domestic partnerships, yet publications originating from international collaborations tend to have a higher citation volume. International collaboration is highest in Computational Theory & Mathematics (37%) and Computer Graphics & CAD (32%), and lowest in Human-Computer Interaction (22%) and Information Systems (17%). The US is the primary international partner, with European countries (particularly France, Germany, and Spain) showing more specialized collaboration patterns.

For the second question, our analysis shows substantial variation in cross-disciplinary collaboration among subfields. Hardware & Architecture researchers exhibit the widest collaborative scope, whereas Computational Theory & Mathematics and Information Systems maintains specialized research ecosystems. In terms of citations, the multi-subfield collaborations usually achieve higher numbers, however this advantage is much more modest than the difference that we saw between international vs domestic. Information Systems, Artificial Intelligence and Computer Networks & Communications respectively have the largest number of authors. Additionally, AI is the subfield that appears more often as a secondary and tertiary subfield for the authors.

Regarding the third question, network centralization metrics indicate bridging researchers are key for collaborations and knowledge transfer, despite Brazilian computer science communities remaining primarily organized along traditional subfield lines. These findings have significant implications for research policy and funding. The focused initiatives to increase cross-border partnerships, particularly in applied areas like Information Systems and AI, might enhance the global impact of Brazilian research. Likewise, supporting key bridging researchers is vital for fostering interdisciplinary innovation.

Finally, our methodology provides a framework for future studies. These could extend the analysis in three key areas: i) integrate additional research databases (Scopus, Google Scholar, DBLP, etc.) and adopt a wider temporal range to validate findings and improve coverage; ii) explore the impact of institutional factors (e.g., funding, geography, organization) on collaboration patterns; iii) employ other measures of impact and quality beyond research output, citation counts, and averages. These methodological enhancements may disclose previously unrecognized value patterns and foster more equitable evaluation of research contributions across disciplines.

REFERENCES

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.

Bird, C., Devanbu, P., Barr, E., Filkov, V., Nash, A., and Su, Z. (2009). *Structure and Dynamics of Research Collaboration in Computer Science*, pages 826–837. Society for Industrial and Applied Mathematics.

Biryukov, M. and Dong, C. (2010). Analysis of computer science communities based on dblp. In Lalmas, M., Jose, J., Rauber, A., Sebastiani, F., and Frommholz, I., editors, *Research and Advanced Technology for Digital Libraries*, pages 228–235, Berlin, Heidelberg. Springer Berlin Heidelberg.

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.

Borgatti, S. P. (2006). Identifying sets of key players in a social network. *Computational & Mathematical Organization Theory*, 12(1):21–34.

Brandão, M., Diniz, M., and Moro, M. (2016). Using topological properties to measure the strength of co-authorship ties. In *Anais do V Brazilian Workshop on Social Network Analysis and Mining*, pages 199–210, Porto Alegre, RS, Brasil. SBC.

Carvalho, L., Filho, S., Brandão, M., Oliveira, J., Santoro, F., and Silva, M. (2023). Uma década de interações entre eventos e pesquisadores do csbc: Um estudo meta-científico. In *Anais do XII Brazilian Workshop on Social Network Analysis and Mining*, pages 139–150, Porto Alegre, RS, Brasil. SBC.

Culbert, J., Hobert, A., Jahn, N., Haupka, N., Schmidt, M., Donner, P., and Mayr, P. (2024). Reference coverage analysis of openalex compared to web of science and scopus.

Delgado-Garcia, J. F., Laender, A. H., and Meira, W. (2014). Analyzing the coauthorship networks of latin american computer science research groups. In *2014 9th Latin American Web Congress*, pages 77–81.

Dias, T. M. R., Moreira, T. H. J., and Dias, P. M. (2018). Caracterização e análise das redes de colaboração científica dos bolsistas de produtividade em pesquisa do cnpq. In *Anais do VII Brazilian Workshop on Social Network Analysis and Mining*, Porto Alegre, RS, Brasil. SBC.

Digiampietri, L., Mena-Chalco, J., Silva, G., Oliveira, L., Malheiros, A., and Meira, D. (2012). Dinâmica das relações de coautoria nos programas de pós-graduação em computação no brasil. In *Anais do I Brazilian Workshop on Social Network Analysis and Mining*, pages 105–116, Porto Alegre, RS, Brasil. SBC.

Druszcz, F. and Vignatti, A. (2024). Citation analysis disparity between sub-areas of brazilian computer science. In *Anais do XIII Brazilian Workshop on Social Network Analysis and Mining*, pages 24–34, Porto Alegre, RS, Brasil. SBC.

Elmacioglu, E. and Lee, D. (2005). On six degrees of separation in dblp-db and more. *SIGMOD Rec.*, 34(2):33–40.

Franceschet, M. (2011). Collaboration in computer science: A network science approach. *Journal of the American Society for Information Science and Technology*, 62(10):1992–2012.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215–239.

Guan, J. and Ma, N. (2004). A comparative study of research performance in computer science. *Scientometrics*, 61(3):339–359.

Júnior, G. P., Dias, T. M. R., Silva, T. H. P., and Laender, A. H. F. (2018). Importância das colaborações interdisciplinares nas redes de coautoria científica. In *Anais do VII Brazilian Workshop on Social Network Analysis and Mining*, Porto Alegre, RS, Brasil. SBC.

Luukkonen, T., Tijssen, R. J. W., Persson, O., and Sivertsen, G. (1993). The measurement of international scientific collaboration. *Scientometrics*, 28(1):15–36.

Madaan, G. and Jolad, S. (2014). Evolution of scientific collaboration networks. In *2014 IEEE International Conference on Big Data (Big Data)*, pages 7–13.

Mena-Chalco, J., Digiampietri, L., and Cesar-Jr, R. (2012). Caracterizando as redes de coautoria de currículos lattes. In *Anais do I Brazilian Workshop on Social Network Analysis and Mining*, pages 72–83, Porto Alegre, RS, Brasil. SBC.

Newman, M. E. J. (2001a). Scientific collaboration networks. i. network construction and fundamental results. *Phys. Rev. E*, 64:016131.

Newman, M. E. J. (2001b). Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64:016132.

Niu, F. and Qiu, J. (2014). Network structure, distribution and the growth of chinese international research collaboration. *Scientometrics*, 98(2):1221–1233.

Okamura, K. (2023). A half-century of global collaboration in science and the “shrinking world”. *Quantitative Science Studies*, 4(4):938–959.

Pessoa Junior, G. J., M. R. Dias, T., H. P. Silva, T., and H. F. Laender, A. (2022). A geographic analysis of the interdisciplinary collaborations in the brazilian scientific community. *Journal of the Brazilian Computer Society*, 28(1):1–12.

Priem, J., Piwowar, H., and Orr, R. (2022). Openalex: A fully-open index of scholarly works, authors, venues, institutions, and concepts.

Ramos, M. Y. (2017). Internacionalização da pós-graduação no brasil: lógica e mecanismos1. *Educação e pesquisa*, 44.

Velez-Estevez, A., Perez, I., García-Sánchez, P., Moral-Munoz, J., and Cobo, M. (2023). New trends in bibliometric apis: A comparative analysis. *Information Processing & Management*, 60(4):103385.

Wainer, J., Xavier, E. C., and Bezerra, F. (2009). Scientific production in computer science: A comparative study of brazil and other countries. *Scientometrics*, 81(2):535–547.