Concentration and Diversity in Collaboration Networks in Electrical Engineering: Evidence from Brazilian Researchers

Fernando A. do Carmo, Gustavo S. Silva, Raimundo C. S. Freire, Antonio F. L. Jacob Jr, and Fábio M. F. Lobato

***Abstract*—Scientific collaboration networks shape research dynamics in fields such as Electrical Engineering, where Brazil's CNPq productivity grant system constitutes a central incentive structure. Building on previous analyses focused on Microelectronics, this study examines collaboration networks across the broader field of Electrical Engineering, encompassing seven subareas: Electrical Materials; Electrical, Magnetic, and Electronic Measurements and Instrumentation; Electrical, Magnetic, and Electronic Circuits; Power Systems; Industrial Electronics, Systems, and Electronic Controls; and Telecommunications. The analysis considers 419 CNPq productivity grant holders and 18,386 publications. The results reveal a hierarchical collaboration structure in which higher grant levels concentrate influence and centrality, while lower levels remain more peripheral despite active collaboration. The network exhibits pronounced geographic concentration in Brazil's Southeast and South regions, as well as marked gender disparities. Compared with prior subfield-focused analyses, Electrical Engineering displays greater thematic heterogeneity and broader inter-level interaction, while maintaining persistent structural inequality. Temporal analysis indicates an increase in average author influence and more compact collaboration clusters over time. These findings suggest that collaboration patterns in Electrical Engineering reflect mechanisms of cumulative advantage associated with hierarchical funding structures.**

Link to graphical and video abstracts, and to code:
https://latamt.ieeer9.org/index.php/transactions/article/view/10313

***Index Terms*—Scientific collaboration networks, CNPq productivity grants, collaboration clusters.**

I. Introduction

COLLABORATION networks constitute a mechanism through which scientific knowledge advances. In fields characterized by high entry barriers and strong resource dependence, collaborative structures are not a byproduct of scientific activity but a strategic response to incentives, funding mechanisms, and institutional constraints [1], [2]. Network-based analyses grounded in graph theory help examine how scientific communities organize themselves and how influence accumulates. [3]–[5]. In Brazil, the National Council for Scientific and Technological Development (CNPq) administers two major fellowship programs that support sustained contributions to Science, Technology, and Innovation (CT&I).

The Research Productivity Fellowship (PQ) and the Technological Development and Innovative Extension Fellowship (DT) follow a hierarchical classification structure with three levels (C, B, and A) for PQ and DT, and a separate Senior Researcher category (PQ-Sr), as established by Normative Resolution 12/2024. CNPq also maintains the Lattes platform, where researchers document publications, collaborations, and institutional affiliations [6]. Despite the availability of comprehensive Lattes data, systematic characterization of fellowship holders at the level of specific subfields remains limited. This gap is relevant given CNPq's recent policy initiatives addressing gender equity and funding effectiveness. The centralized and hierarchical nature of CNPq grants creates conditions associated with the Matthew Effect, a cumulative advantage mechanism whereby early success increases the likelihood of future recognition and network centrality [7], [8]. This funding structure may therefore influence collaboration patterns and the distribution of resources among researchers.

The present study extends the investigation reported in [9], which analyzed collaboration networks within the Microelectronics subarea of Electrical Engineering (EE). In contrast, this work examines EE as an aggregated field, as defined by the CAPES and CNPq classification, encompassing multiple subareas with distinct research practices and collaboration dynamics. Rather than replicating previous results at a larger scale, the goal is to examine how hierarchical funding structures, collaboration networks, and thematic organization interact within a heterogeneous, strategically relevant engineering domain, thereby providing evidence relevant to ongoing CNPq policy discussions on fellowship allocation and evaluation criteria.

EE is examined due to its strategic role in national innovation systems. The field is capital-intensive, requiring costly laboratories and sustained infrastructure investment, and develops strong industry–academic linkages in high-tech sectors (e.g., power systems, automotive, electronics) that tend to be geographically clustered [10]. Rapid technological obso-

The associate editor coordinating the review of this manuscript and approving it for publication was Mario R. Arrieta Paternina (*Corresponding author: Fernando Carmo*).

This study was supported by CNPq - DT-303031/2023-9, PDS-101057/2024-5, and DES-404727/2024-7; by Acordo de Cooperação Técnica N.° 02/2021 (N.° 38328/2020-TJ/MA); by the FINEP (ProAmazonia - 2373/24); and by CAPES (Transformative Agreement).

Fernando Carmo, G. S. Silva, and A. F. L. Jacob Jr are with the State University of Maranhão, São Luís, Brazil (e-mails: fernando.20251001723@aluno.uema.br, gustavo.20241004427@aluno.uema.br, and antoniojunior@professor.uema.br).

R. C. S. Freire is with the Federal University of Campina Grande, Campina Grande, Brazil (e-mail: freire@dee.ufcg.edu.br).

F. M. F. Lobato is with the Federal University of Western Pará, Santarém, Brazil (e-mail: fabio.lobato@ufopa.edu.br).

lescence further increases dependence on stable collaborative arrangements [11]. These characteristics are evident in sectors such as the automotive industry, where electronic components account for 30–40% of vehicle costs and vehicles integrate hundreds of millions of lines of embedded software [12], [13]. Thus, EE serves as a proxy for high-technology development in the Global South, and Latin America in particular, making structural inequalities within the field relevant for understanding constraints on innovation capacity and technology transfer.

The analysis focuses on CNPq productivity fellows in EE as defined by CAPES[1], covering eight subareas: Electrical Materials; Electrical, Magnetic, and Electronic Measurements and Instrumentation; Electrical and Electronic Circuits; Electric Power Systems; Industrial Electronics; Electronic Systems and Controls; and Telecommunications. Data extracted from the Lattes platform comprises 419 fellows and 18,386 publications.

The study pursues three complementary goals: (i) mapping co-authorship networks to characterize collaboration patterns and central actors, (ii) examining their temporal evolution, and (iii) relating collaboration structures to research topics. The contribution of this work does not lie in the proposal of new analytical algorithms, but in the integrated application of established network and topic modeling methods to describe how investment structures, collaboration patterns, and thematic evolution shape research organization and inequality within EE in Brazil.

The remainder of this paper is structured as follows: Section II presents a literature review. The methodological procedures are detailed in Section III. Section IV presents the results and their analysis. Finally, Section V presents the conclusions and future prospects.

II. Related Works

Previous bibliometric and network-based studies have mapped thematic clusters and collaboration patterns across engineering domains [3], [14], demonstrating the relevance of network and topic modeling approaches adopted here.

[15] analyzed interdisciplinary collaborations among PhDs from eight major areas using 263,264 Lattes curricula and over 10 million publications. The results showed that 35.2% of co-authorships cross disciplinary boundaries and 57.6% of researchers participated in at least one interdisciplinary collaboration. Engineering demonstrated a 56.4% interdisciplinary rate. While this work reveals broad collaborative patterns across Brazilian academia, it does not examine the dynamics within specific engineering subareas or characterize the institutional profiles of researchers. The present study addresses these gaps by focusing on EE and incorporating institutional and productivity variables.

[16] examined how co-authorship networks influence the productivity of CNPq fellowship holders. Social network metrics assessed the impact of collaborations on scientific output, highlighting the structural importance of networks in performance evaluation. The analysis remained limited to a single engineering subfield. The current work adopts a similar network-based approach, focusing on CNPq recipients, but expands the scope to encompass all of EE.

[17] used centrality measures (degree, betweenness, closeness) alongside Latent Dirichlet Allocation (LDA) to identify influential researchers and main themes in civil engineering in Nigeria from 663 Scopus publications spanning 1969-2018. Betweenness and closeness centrality were positively associated with citation performance, indicating an active, collaborative environment within the network. Although this study provides insights into collaborative networks and emerging themes, its geographical and disciplinary scope differs substantially from the present focus. The current work extends this approach by employing BERTopic, utilizing a larger sample, and incorporating detailed institutional characterization of Brazilian researchers.

[18] mapped authorship networks among faculty in graduate Accounting programs using data from the Lattes Platform. Analysis of 3,778 publications from 21 programs between 2013 and 2016, using techniques such as degree centrality and betweenness, revealed a low overall network density; however, density increased compared to earlier periods. Institutions such as USP and UFRJ stood out as centers with the highest connectivity and betweenness. Although it contributes to understanding collaborative dynamics in accounting, the study is limited to a single broad area and does not consider aspects such as interdisciplinarity, international collaborations, or detailed institutional characterization.

The current investigation extends this framework by examining CNPq fellowship holders in EE, examining their collaborative networks, temporal evolution, and dominant research topics. This contributes new evidence about the organization of Brazil's engineering research community and its alignment with national funding priorities.

III. Materials and Methods

We adopted the Data Science Trajectory (DST) model [19], which structures the analysis into six phases described in the following subsections. The methodological contribution lies in the integrated application of established network metrics, influence measures, and topic modeling techniques to a comprehensive national dataset, enabling a unified analysis of funding hierarchies, collaboration structures, and thematic dynamics over time.

A. Domain Understanding

This phase focused on defining the analytical scope, identifying gaps in existing research, and specifying the study's goals. Previous investigations examined Brazilian co-authorship networks at broad interdisciplinary scales [15] or within individual engineering subfields [16], whereas a systematic characterization of CNPq productivity fellows across all Electrical Engineering subareas remained unexplored. Based on this context, the study was guided by three goals: (i) mapping the hierarchical and institutional structure of collaboration networks; (ii) characterizing the main research themes and their distribution across subareas; and (iii) analyzing the temporal evolution of collaboration patterns within the field.

[1] https://lattes.cnpq.br/web/dgp/engenharias

The field boundaries followed the CAPES knowledge tree definition of Electrical Engineering, encompassing seven subareas that define the scope of the analysis. This classification delimited the sample to researchers working in Electrical Materials; Electrical, Magnetic, and Electronic Measurements and Instrumentation; Circuits; Power Systems; Industrial Electronics and Controls; and Telecommunications. Within this scope, the methodological design followed the DST model, integrating network analysis and semantic topic modeling to examine relational and thematic dimensions of scientific production.

B. Data Acquisition

Data collection proceeded through multiple filtering stages using the Lattes platform. An institutional API key enabled automated access to the complete curriculum vitae in XML format. A custom Python web crawler collected data for 12,345 researchers holding CNPq productivity fellowships across all knowledge areas defined by CAPES. Each XML curriculum contained structured information about the researcher's academic trajectory. The extraction focused on ten key attributes that characterize both individual profiles and collaborative patterns. Personal identifiers included full name and citation name (the standardized form used in publications). Institutional affiliation data captured the current workplace and the doctorate-granting institution. Research profile information included the declared areas of expertise and a curriculum overview. Geographic data were collected in the Brazilian federal state in which the researcher's institution is located. Publication records contained complete lists of journal articles and book chapters, including titles, co-author names, publication years, and venue information.

The initial dataset of 12,345 researchers underwent filtering to isolate the EE community. Using each researcher's declared areas of expertise, records were matched against the CAPES knowledge tree structure. Only researchers who listed at least one of the seven EE subareas as a primary or secondary area of expertise were retained. This process yielded 419 CNPq productivity fellows. From these 419 curricula, publication data extraction identified 7,891 distinct articles. Each publication record included the full author list, enabling the construction of co-authorship relationships. Publication dates ranged from 1979 to 2023, with 80% appearing after 2000.

C. Data Preparation

The raw data comprised an author dataset with 12,345 researchers and a publication dataset with 1,251,807 publications. Missing values were minimal for core fields; however, approximately 34% of records lacked current institution information. The article dataset showed complete coverage for titles, years, and author lists, although 33% lacked DOI identifiers. Duplicate detection identified 11,295 exact duplicate records and 250,141 publications with identical titles. Author names were normalized using three transformations: conversion to uppercase, removal of leading/trailing whitespace, and regularization of punctuation. This ensured consistent formatting across different representations of the same author.

Publications were processed to identify suitable articles for network construction. Author lists were split, normalized, and mapped to researcher identifiers. Three filtering criteria were applied to ensure network quality: (i) publications must include at least one CNPq EE fellow to maintain focus on the target population, (ii) the first author position must be mapped to a known researcher identifier, ensuring publications had at least one resolvable author, and (iii) publications must list multiple authors, as single-author works provide no collaboration information. These filters reduced the dataset from 1,251,807 to 18,386 articles and 323 authors suitable for network construction. The exclusion of 96 fellows from the final network is a direct result of these topological filters, as researchers without multi-author collaborations involving at least one other CNPq EE fellow or with unresolvable first-author identifiers were removed to maintain structural consistency. For topic modeling, 7,891 publication titles were extracted directly without preprocessing. This preserved original semantic content and terminology, allowing the transformer-based embedding model to leverage its multilingual capabilities and contextual understanding of raw text.

D. Modeling

Two complementary analytical methods were employed to characterize the research community. Network analysis captured the structural organization of collaborations, while topic modeling revealed the thematic content of research output.

1) Co-authorship Network Analysis: The collaboration network was modeled as an undirected, weighted graph using NetworkX. The analysis also employed the Density Index ($\mathcal{D}_{(i,j)}$), based on the works from [20]–[22], designed to capture collaboration intensity more accurately than traditional similarity metrics by accounting for both the mutual dependency between two authors and their total publication volume, as defined by the equation:

$$\mathcal{D}_{(i,j)} = \frac{2P_iP_j}{P_i + P_j} \cdot \ln(\text{count}_{ij} + 1) \tag{1}$$

where P_i represents the percentage of author i's publications co-authored with j, and count_{ij} is the number of joint publications. This formulation captures mutual dependency and productivity while penalizing asymmetric collaborations.

Community detection was performed using the Louvain algorithm in Gephi to optimize modularity. Network characterization employed three metrics: average degree (highly connected researchers), modularity (community structure strength), and PageRank (influential researchers accounting for indirect connections). Temporal analysis partitioned the dataset into five-year windows (1995-1999, 2000-2004, 2005-2009, 2010-2014, 2015-2019, 2020-2023) to track network evolution.

2) Semantic Topic Modeling: For the semantic analysis of research content, BERTopic was selected due to its use of contextualized embeddings that capture semantic relationships and handle technical terminology effectively. The analytical pipeline integrated four sequential stages.

Titles were encoded into 768-dimensional embeddings using the Multilingual MPNet model[2] model from Sentence Transformers, selected for its multilingual support, including Portuguese. UMAP dimensionality reduction was applied to project these embeddings into a 5-dimensional space, configured with 11 neighbors, cosine similarity metric, minimum distance of 0.0, and fixed random state for reproducibility. Cluster identification employed HDBSCAN with a minimum cluster size of 25 documents, Euclidean distance metric, and the excess-of-mass selection method, labeling documents not fitting dense clusters as outliers.

Topic representations were generated through three complementary approaches. KeyBERTInspired extracted semantically relevant terms, while Maximal Marginal Relevance with a diversity parameter of 0.2 provided balanced coverage. A locally-hosted OpenHermes 2.5 Mistral-7B[3] language model generated concise labels of up to five words. The language model operated with full GPU offloading and 512-token context window. After the initial generation, topic representations were refined, and labels underwent post-processing to remove formatting artifacts. The final model identified 47 distinct topics plus an outlier category.

E. Evaluation and Deployment

The results were validated using a qualitative grounded-theory-inspired approach. The inspection was conducted by three domain experts. To evaluate the model's output, the experts inspected the full set of generated topics by analyzing their representative keywords and the top 10 most representative publication titles extracted by the BERTopic algorithm for each cluster. The evaluation criteria focused on (i) semantic coherence, ensuring the grouped terms and publications reflected a recognizable and consistent Electrical Engineering subarea, and (ii) practical relevance to the field's current research landscape. This process ensured the model's accuracy and interpretability from a practical, expert-driven perspective.

To ensure reproducibility and accessibility, all research outputs were organized for deployment. The generated co-authorship networks, centrality indicators, and semantic topics were structured in standardized file formats. All developed scripts and processed data were documented and shared in a public repository to facilitate replication and further research. The findings were also disseminated through a scientific article and technical reports to communicate the results to the academic community.

IV. Results and Discussion

The analysis covered 419 profiles of researchers in the broad field of EE, including its subfields. In total, 18,386 publications were examined. The analysis of research topics was performed using the BERTopic model applied to publication titles from CNPq productivity fellowship recipients. Fig. 1 shows a map of these topics across the EE subareas in a two-dimensional projection.

[2]https://huggingface.co/sentence-transformers/paraphrase-multilingual-mpnet-base-v2

[3]https://huggingface.co/TheBloke/OpenHermes-2.5-Mistral-7B-GGUF

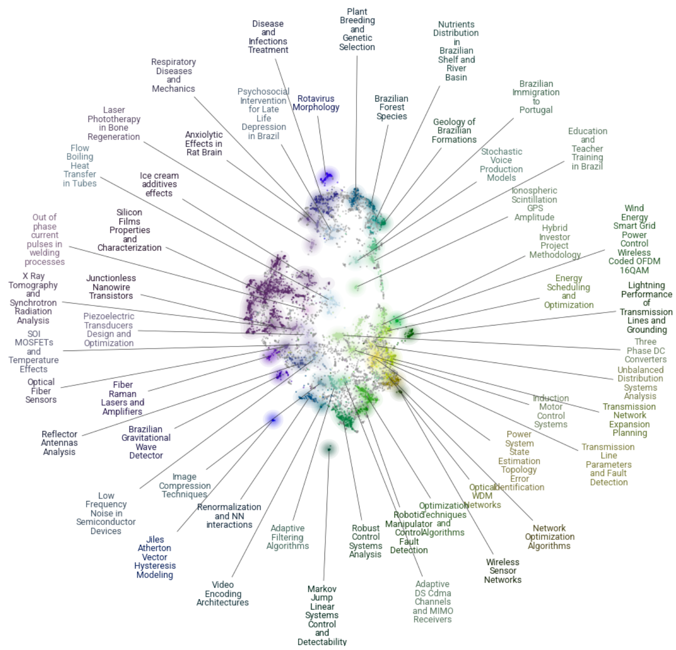

Fig. 1. Map of research topics in EE identified from publication titles.

Each group of points represents semantically similar documents, with colors indicating different topics. The map reveals a diverse thematic structure composed of traditional areas of EE. Three main domains that stand out are Power Systems (including topics on transmission line performance and expansion planning), Control and Automation (covering control systems and adaptive algorithms), and Telecommunications & Electronics (encompassing wireless sensor networks and optical communications).

The analysis highlights the interdisciplinary nature of the topics through groupings composed of traditional areas of EE. Areas in Biomedical Engineering (respiratory diseases and infection treatment), Physics and Space Sciences (Brazilian gravitational-wave detector), and Environmental Sciences (geology and nutrient distribution) are identified. This diversity suggests applications of signal processing, instrumentation, and sensing techniques in multiple academic fields.

The dynamics of the influence of researchers (*PageRank*) and collaboration patterns (connected components) over time (2010-2022) in the co-authorship network are presented in Fig. 2.

The analysis in Fig. 2a) suggests an inverse relationship between the two metrics. In periods of fragmentation or reduction of the largest component (2013–2014 and 2019–2020), there was an increase in the average influence of researchers. With the reduction or division of the largest collaborative group, the average influence of authors increased, especially between 2020 and 2022. This suggests that recent collaborations tend to occur in smaller groups but may be more focused on prominent researchers.

Fig. 2b) shows the monitoring of activity volume through the number of active authors (fellowship recipients with annual publications) and co-authorships (edges in the network). Both metrics peaked between 2014 and 2016, followed by a

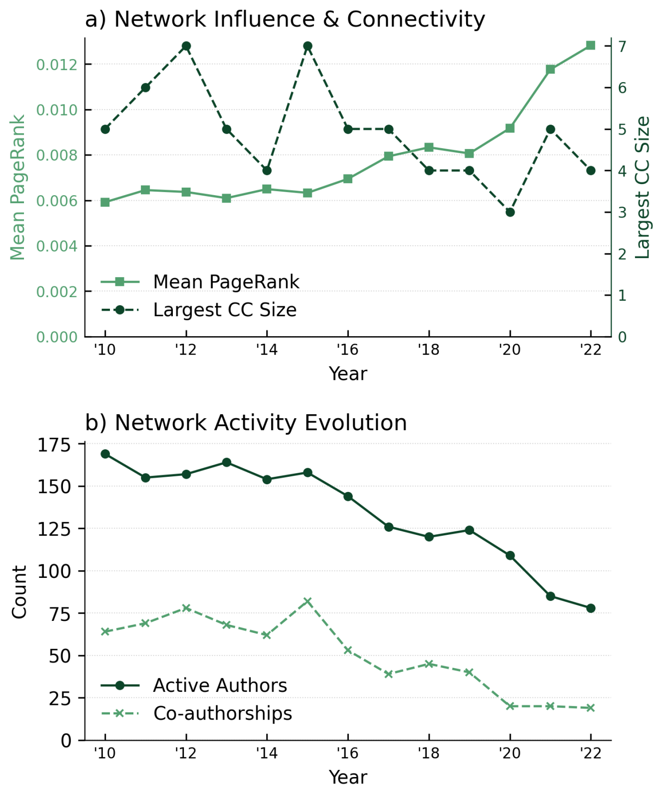

Fig. 2. Temporal dynamics of the co-authorship network from 2010 to 2022. (a) Network influence and connectivity, comparing the mean PageRank score with the size of the largest connected component. (b) Network activity evolution, displaying the annual count of active authors and co-authorship ties.

consistent decline from 2018 to 2022. This reduction does not necessarily reflect a genuine decline in scientific output. At the time of data collection for this study (end of 2024), it is natural that the records for 2021 and 2022 remain incomplete.

Fig. 3 illustrates the main research topics obtained from topic modeling of publication titles using BERTopic.

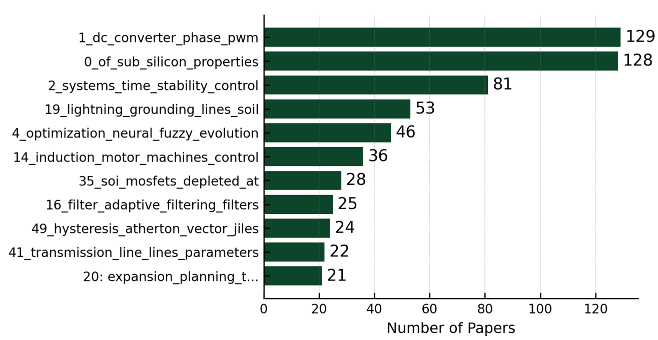

Fig. 3. Frequency of main research topics.

The figure highlights the predominance of topics 1 (DC converters and power electronics) and 0 (semiconductor properties and microelectronics). The presence of specialized topics such as induction machine control (topic 14), MOSFET devices and SOI technology (topic 35), and adaptive filtering (topic 16) indicates the depth of research and the existence of consolidated groups in specific niches.

Table I shows the five most frequent research topics for each year in the last four years of the database.

TABLE I
ANNUAL EVOLUTION OF THE MOST FREQUENT RESEARCH TOPICS

Year	Topic (Top 5)	Count
2019	0_semiconductors_spectroscopy_silicon	32
	1_converters_converter_conversor	13
	3_infection_disease_stercoralis	11
	6_network_networks_bandwidth	10
	18_terapia_reabilitação_tratamento	8
2020	0_semiconductors_spectroscopy_silicon	22
	2_nonlinearities_nonlinear_stability	10
	10_brasileiros_brasileiro_brasileira	8
	18_terapia_reabilitação_tratamento	8
	1_converters_converter_conversor	8
2021	0_semiconductors_spectroscopy_silicon	15
	23_gravitacionais_gravitational_observatory	9
	10_brasileiros_brasileiro_brasileira	5
	18_terapia_reabilitação_tratamento	5
	19_lightning_voltages_voltage	5
2022	0_semiconductors_spectroscopy_silicon	14
	10_brasileiros_brasileiro_brasileira	8
	26_pedagógica_ensino_educacional	7
	1_converters_converter_conversor	6
	12_planning_maintenance_gerenciamento	5

The table indicates that Topic 0 was the most prevalent research topic across all years analyzed, suggesting the persistence of core areas within EE. In addition to microelectronics, power electronics (Topic 1) appears consistently in 2019, 2020, and 2022, indicating sustained research activity within the scientific community. This temporal consistency complements the volume-based results presented above.

The table also reflects the dynamic nature of research trends through distinct temporal peaks. Topic 23 appears exclusively in 2021, possibly corresponding to a specific phase of large-scale research projects or special journal issues. Similarly, Topic 2 is highlighted only in 2020. These temporal fluctuations illustrate how scientific attention can concentrate on specific topics during particular periods.

Fig. 4 presents two scatter plots analyzing community characteristics and topic-specific collaboration patterns across the analyzed articles.

Fig. 4a shows that most topics, including power electronics (Topic 1) and microelectronics (Topic 0), are concentrated between 2.0 and 2.4 authors per article, indicating the predominance of small research teams typically composed of a supervisor and one or two collaborators. Topics with higher average numbers of authors tend to correspond to specialized research areas requiring broader expertise. The lack of a direct relationship between publication volume and team size indicates that high-output research areas tend to maintain lean collaboration structures.

Fig. 4b reveals a positive relationship between the number of unique authors and article volume across topics. Power electronics and microelectronics exhibit the largest research communities, each comprising more than 40 authors. In contrast, control systems (Topic 2) display high dispersion,

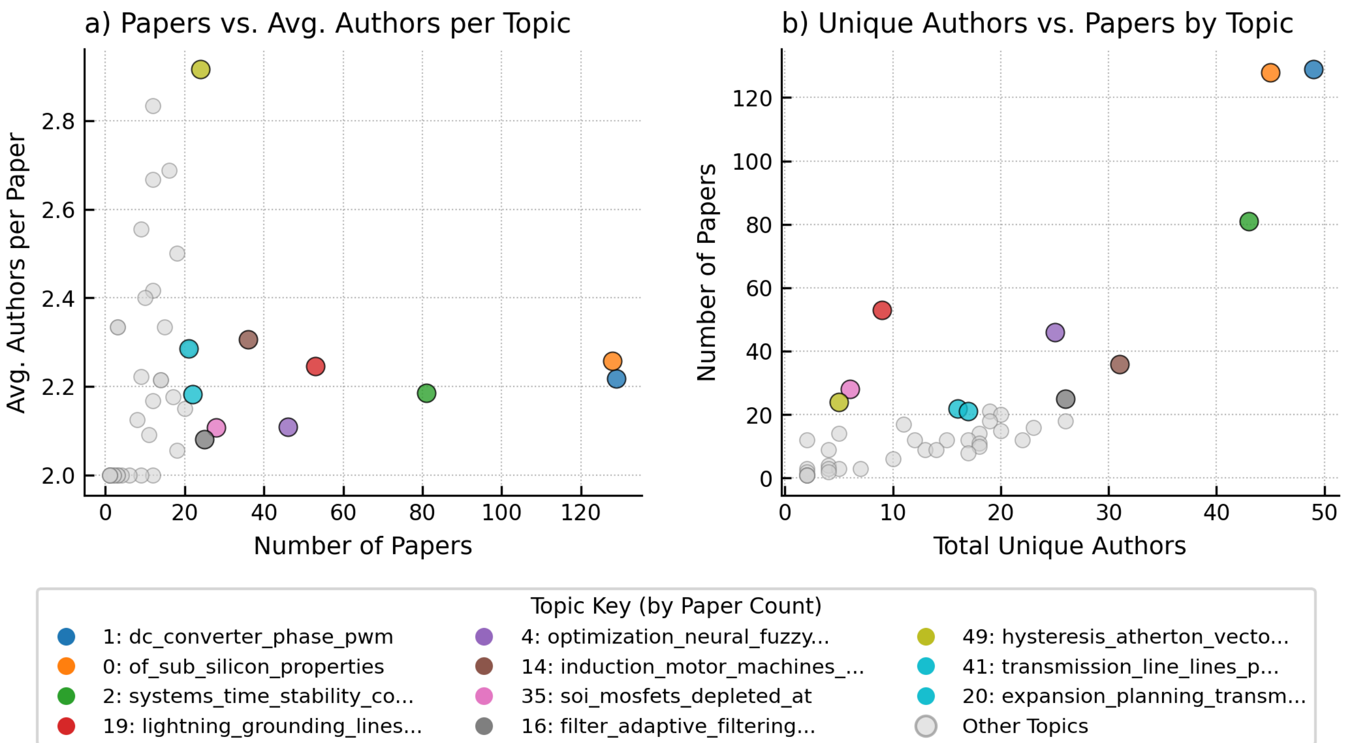

Fig. 4. Collaboration and community dynamics by research topic.

with approximately 80 articles produced by 25 authors, while lightning protection (Topic 19) shows lower productivity, suggesting more sporadic research contributions. The following analysis examines how these topics emerge from collaborative structures among researchers.

To analyze co-authorship patterns, the dataset was segmented into five subnetworks according to CNPq PQ/DT productivity grant levels (1A, 1B, 1C, 1D, and 2). Nodes represent productivity grant holders in the broad field of EE, while edges denote co-authorship relationships. The resulting network comprises 323 nodes and 9,087 edges.

Table II reports the five researchers with the highest influence scores at grant levels 1A and 1B, as measured by the PageRank algorithm. The complete table, covering all grant levels, is available in the supplementary material[4].

An examination of influential researchers by funding level reveals distinct career trajectories and thematic orientations. Level 1A researchers typically have long post-PhD careers and well-established collaborative networks, often linked to leading Brazilian and international institutions. In contrast, researchers at levels 1D and 2 generally obtained their PhDs more recently and tend to focus on emerging research areas, indicating generational renewal within the field. Fig. 5 illustrates these dynamics by showing a subgraph that highlights the most central researchers at each productivity grant level.

The analysis of the co-authorship network, presented in Table II, reveals relevant collaboration patterns among CNPq productivity fellows. Researchers at the highest grant levels, such as 1A and 1B, stand out not only due to their stronger influence, reflected by high PageRank values, but also because their dense interconnections with other groups suggest a consolidated collaborative ecosystem in which the experience of senior researchers contributes to cohesion and leadership in scientific networks.

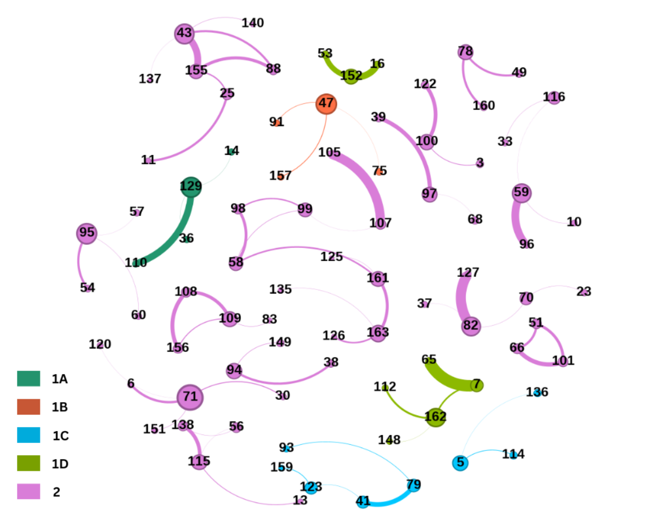

Fig. 5. Collaboration network of the most influential researchers, grouped by productivity grant level.

Table III reports the number of co-authorship links between CNPq productivity grant levels, summarizing inter-level collaboration patterns across the network.

Levels 1C and 2 exhibit the highest intergroup collaboration

[4]https://github.com/gustaph/IEEE-LATAM-Collaboration-Networks-Brazilian-EE-Researchers

TABLE II
MOST INFLUENTIAL RESEARCHERS BY CNPQ PRODUCTIVITY GRANT LEVEL

Grant Level	ID	PageRank	Main Research Areas	Institution	PhD Year
	47	0.011704	Industrial Electronics; Control Systems; Digital Signal Processing	ITA (Brazil)	1993
1A	113	0.006098	Electric Circuits; Measurements and Instrumentation; Telecommunications	USP (Brazil)	1984
	22	0.006098	Electric Circuits; Electrical Materials	INPG (France)	1982
	86	0.006098	Not reported	Not reported	–
	141	0.006098	Not reported	Not reported	–
	129	0.011704	Electric Power Systems	University of Missouri-Columbia (USA)	1981
1B	17	0.006098	Electric Power Systems	UNICAMP (Brazil)	1991
	92	0.006098	Electric Power Systems	UNICAMP (Brazil)	1995
	74	0.006098	Industrial Electronics; Control Systems	Concordia University (Canada)	1998
	90	0.006098	Industrial Electronics; Control Systems; Electric Circuits	UFSC (Brazil)	1994

TABLE III
INTERGROUP COLLABORATION MATRIX OF CNPQ PRODUCTIVITY FELLOWS

Level	1A	1B	1C	1D	2
1A	–	199	1127	330	726
1B	199	–	203	410	712
1C	1127	203	–	543	1363
1D	330	410	543	–	1033
2	726	712	1363	1033	–

volumes, with Level 1C playing a central intermediary role through ties to both senior (1A) and early-career (1D and 2) groups. Levels 1D and 2, despite active collaboration, occupy more peripheral network positions, consistent with the Matthew Effect, whereby researchers in central positions accumulate visibility and access to resources over time. Within the CNPq productivity system, higher grant levels are associated with greater network centrality, while lower levels remain less visible. This suggests that collaboration alone is insufficient to alter network position, as centrality is reinforced by prior recognition and institutional accumulation.

Fig. 6 presents a scatter plot designed to characterize co-authorship pairs based on mutual dependence. Each point represents a unique author pair, where the x-axis measures Author 1's dependence, defined as the percentage of Author 1's publications co-authored with Author 2, and the y-axis represents Author 2's dependence. Point size and color encode the volume of joint publications, distinguishing highly productive collaborations from sporadic ones. Four quadrants correspond to distinct collaboration patterns: Casual Collaborations, Author 1 Dependency, Author 2 Dependency, and Exclusive Pairs.

The distribution reveals distinct collaboration patterns. Large, warm-colored points are concentrated in the "Exclusive Pairs" quadrant, where mutual dependence exceeds 90%, indicating stable, highly productive partnerships. Asymmetric patterns along the main axes, in which one author exhibits near-total dependence while the counterpart exhibits considerably lower dependence, are indicative of mentorship relationships or collaborations involving specialized expertise. The relative sparsity of points in the central region suggests that balanced, moderately interdependent collaborations are less common

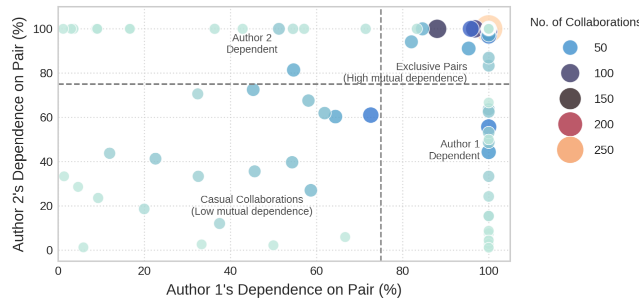

Fig. 6. Characterization of co-authorship pairs by mutual dependence.

than either highly exclusive or casual connected partnerships.

Table IV lists the ten author pairs with the highest density index, which captures the balance between collaboration volume and mutual dependence.

TABLE IV
TOP 10 COLLABORATIONS BY DENSITY INDEX

Author's 1 ID	Author's 2 ID	$\mathcal{D}_{(\mathbf{i},\mathbf{j})}$	Mutual Dependence (%)
48	130	560.95	100.00 / 100.00
82	127	443.68	96.77 / 100.00
105	107	437.97	100.00 / 98.78
7	65	434.87	88.03 / 100.00
59	96	415.80	95.83 / 100.00
110	129	404.34	100.00 / 96.77
84	147	400.73	100.00 / 100.00
8	61	391.20	100.00 / 100.00
2	153	380.67	100.00 / 100.00
12	150	361.09	100.00 / 100.00

The results in Table IV show that the pair (48, 130) exhibits the highest density index, reflecting both complete mutual dependence and a large volume of joint publications. The ranking further demonstrates the discriminative capability of the density index, as it differentiates collaborations with identical or near-identical mutual dependence but distinct publication volumes. For instance, pairs such as (84, 147) and (8, 61), despite exhibiting 100% mutual dependence, present lower density values due to reduced collaboration volume.

The collaborative networks and thematic structures described previously reveal the organizational patterns of research activity in the field. However, the distribution of researchers within these networks is uneven across geographic and demographic dimensions. To examine geographic disparities in the distribution of CNPq productivity grant holders, a spatial analysis (Fig. 7) was conducted using institutional data for researchers with available affiliation information, yielding a sample of 279 scholars.

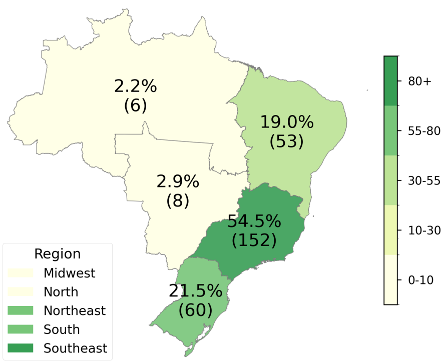

Fig. 7. Researchers' geographical distribution.

The pronounced geographic concentration of researchers in the South and Southeast regions raises the question of whether network inequality primarily reflects the spatial distribution of institutional excellence. These regions host a disproportionate number of Brazil's most established universities, graduate programs, and research infrastructure, which historically attract greater funding, talent, and international collaboration. As a result, geographic centrality may be less a consequence of individual productivity than of long-standing institutional accumulation processes.

However, the observed concentration also suggests a self-reinforcing dynamic in which institutions in these regions benefit from greater visibility and denser collaboration networks, which, in turn, increase their capacity to attract resources and researchers. This feedback loop may limit the integration of researchers from less represented regions, even when collaboration links exist, thereby perpetuating regional disparities within the national research system.

A complementary gender-based analysis (Fig. 8) was performed to characterize representation within the field of EE.

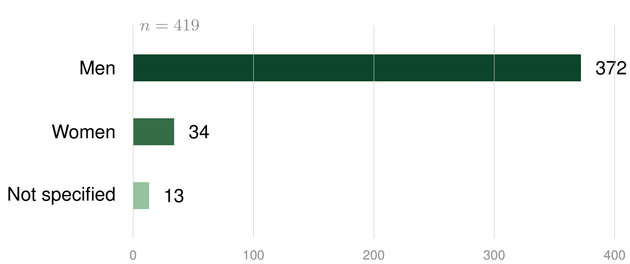

Fig. 8. Gender distribution of CNPq productivity grant holders.

Fig. 8 shows a marked predominance of male researchers, who represent 88.78% of the sample, compared with 8.11% female researchers and 3.10% with not specified gender. This distribution highlights the persistence of gender imbalance in engineering and related STEM fields. The underrepresentation of women at higher CNPq levels mirrors findings for the Engineering macro-area, where male PQ fellows concentrate in upper categories and few women reach top levels, suggesting a glass ceiling within the productivity grant system. Such patterns may result from cumulative disadvantages in access to leadership roles, mentorship networks, and recognition mechanisms that are linked to visibility within collaboration structures. Gender inequality thus appears not only as a demographic imbalance but as a structural feature of hierarchical evaluation systems in Brazilian science.

When considered alongside the co-authorship network analysis, these demographic and geographic patterns help contextualize the observed peripheral positions of Levels 1D and 2. Despite their relatively high number of collaborations, particularly with Level 1C, these groups exhibit lower scientific centrality, suggesting constraints on integration and visibility within the research network.

Research activity and scientific centrality are unevenly distributed across geographic regions, with a strong concentration in the South and Southeast. Gender representation is similarly asymmetric, with a marked predominance of male researchers across all productivity levels.

In addition, hierarchical stratification within the CNPq productivity system is reflected in network structure, as senior levels concentrate influence and centrality, while early-career groups, despite active collaboration, tend to occupy more peripheral positions. These findings indicate that inequality is not limited to a single dimension but is an inherent, multidimensional feature of the observed research networks.

V. Conclusions

This study examined co-authorship networks, thematic organization, and inequality patterns among CNPq productivity fellows in Electrical Engineering in Brazil. The results indicate that collaboration networks are hierarchically structured, with higher grant levels concentrating influence and centrality, while lower levels occupy peripheral positions—consistent with cumulative advantage mechanisms in hierarchical funding systems.

Thematic analysis identified research topics spanning microelectronics, power electronics, telecommunications, control systems, and instrumentation, as well as applications in the health, environmental sciences, and education. Temporal analysis reveals both recurrent themes and year-specific variations reflecting shifts in research focus.

Geographic and demographic analyses indicate uneven distributions within the network. Research activity and centrality are concentrated in Brazil's South and Southeast regions, while gender representation shows lower participation of women, particularly at higher grant tiers. These patterns indicate that inequality in Electrical Engineering manifests across hierarchical, geographic, and demographic dimensions.

Methodologically, this work contributes by integrating network analysis, topic modeling, temporal dynamics, and funding-level information to examine how collaboration structures, funding hierarchies, and thematic organization intersect within a strategic field of Brazilian science and technology. Future research using comparative datasets, extended temporal windows, and additional institutional indicators may complement these findings.

REFERENCES

[1] S. H. Kang, J. Shin, and J. Choi, *How Do Researchers Conduct Scientific Collaboration across Universities?*, ser. HICSS. Hawaii International Conference on System Sciences, 2025, doi: 10.24251/hicss.2025.056.

[2] S. B. d. Jesus and K. d. O. Rodrigues, "Coautoria internacional em artigos científicos: uma análise na produção intelectual dos programas de pós-graduação no campo das engenharias da universidade federal da bahia," *Em Questão*, vol. 31, 2025, doi: 10.1590/1808-5245.31.140393.

[3] M. A. Lopes and R. A. Martins, "Mapping the impacts of industry 4.0 on performance measurement systems," *IEEE Latin America Transactions*, vol. 19, no. 11, pp. 1912–1923, 2021, doi: 10.1109/TLA.2021.9475625.

[4] B. S. Santos, I. Silva, and D. G. Costa, "Symmetry in scientific collaboration networks: A study using temporal graph data science and scientometrics," *Symmetry*, vol. 15, no. 3, p. 601, Feb. 2023, doi: 10.3390/sym15030601.

[5] M. Mamat *et al.*, "Beyond nodes and edges: a bibliometric analysis on graph theory and neuroimaging modalities," *Frontiers in Neuroscience*, vol. 18, Apr. 2024, doi: 10.3389/fnins.2024.1373264.

[6] A. Justino *et al.*, "Análise orientada a dados como auxílio para tomada de decisão em gestão de pesquisa," *Revista da CGU*, vol. 15, no. 27, Jul. 2023, doi: 10.36428/revistadacgu.v15i27.565.

[7] T. Bol, M. de Vaan, and A. van de Rijt, "The matthew effect in science funding," *Proceedings of the National Academy of Sciences*, vol. 115, no. 19, p. 4887–4890, Apr. 2018, doi: 10.1073/pnas.1719557115.

[8] Y. L. Katchanov, Y. V. Markova, and N. A. Shmatko, "Empirical demonstration of the matthew effect in scientific research careers," *Journal of Informetrics*, vol. 17, no. 4, p. 101465, Nov. 2023, doi: 10.1016/j.joi.2023.101465.

[9] F. A. Carmo, G. S. Silva, A. F. L. Jacob, and F. M. F. Lobato, "Concentration and diversity in collaboration networks in microelectronics: A study on cnpq productivity fellows," in *2025 39th Symposium on Microelectronics Technology and Devices (SBMicro)*, vol. 1, 2025, pp. 1–4, doi: 10.1109/SBMicro66945.2025.11197788.

[10] N. Albis *et al.*, "University-industry linkages through the lens of national innovation systems in latin america," *Industry and Innovation*, p. 1–22, Oct. 2025, doi: 10.1080/13662716.2025.2560929.

[11] S. Ma, "Technological obsolescence," *The Review of Financial Studies*, Oct. 2025, doi: 10.1093/rfs/hhaf059.

[12] M. Broy *et al.*, "Engineering automotive software," *Proceedings of the IEEE*, vol. 95, no. 2, p. 356–373, Feb. 2007, doi: 10.1109/jproc.2006.888386.

[13] O. Burkacky, J. Deichmann, G. Doll, and C. Knochenhauer, "Rethinking car software and electronics architecture," McKinsey & Company, Tech. Rep., 2023.

[14] J. Assandre and R. Martins, "Analysis of scientific production on the use of big data analytics in performance measurement systems," *IEEE Latin America Transactions*, vol. 21, no. 3, pp. 367–380, 2023, doi: 10.1109/TLA.2023.10068840.

[15] G. J. P. Junior *et al.*, "On interdisciplinary collaborations in scientific coauthorship networks: the case of the brazilian community," *Scientometrics*, vol. 124, no. 3, p. 2341–2360, Jul. 2020, doi: 10.1007/s11192-020-03605-3.

[16] R. Andrade and L. Rêgo, "A influência da rede de coautoria no nível de produtividade dos bolsistas de produtividade em engenharia de produção," in *Anais do Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*, ser. BraSNAM 2015. Sociedade Brasileira de Computação - SBC, Aug. 2015, doi: 10.5753/brasnam.2015.6786.

[17] I. T. Afolabi *et al.*, "Identifying major civil engineering research influencers and topics using social network analysis," *Cogent Engineering*, vol. 7, no. 1, Jan. 2020, doi: 10.1080/23311916.2020.1835147.

[18] L. M. Welter *et al.*, "Redes de coautoria dos programas brasileiros de pós-graduação em contabilidade," *Contextus: Revista Contemporânea de Economia e Gestão*, vol. 19, no. 10, pp. 146–159, 2021, doi: 10.19094/contextus.2021.61274.

[19] F. Martinez-Plumed, L. Contreras-Ochando, C. Ferri, J. Hernandez-Orallo, M. Kull, N. Lachiche, M. J. Ramirez-Quintana, and P. Flach, "Crisp-dm twenty years later: From data mining processes to data science trajectories," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, p. 3048–3061, Aug. 2021, doi: 10.1109/tkde.2019.2962680.

[20] N. A. Aziz and M. P. Rozing, "Profit (p)-index: The degree to which authors profit from co-authors," *PLOS ONE*, vol. 8, no. 4, p. e59814, 2013, doi: 10.1371/journal.pone.0059814.

[21] M. E. J. Newman, "Scientific collaboration networks. II. shortest paths, weighted networks, and centrality," *Physical Review E*, vol. 64, no. 1, p. 016132, 2001, doi: 10.1103/PhysRevE.64.016132.

[22] A. J. Parish, K. W. Boyack, and J. P. A. Ioannidis, "Dynamics of co-authorship and productivity across different fields of scientific research," *PLOS ONE*, vol. 13, no. 1, p. e0189742, 2018, doi: 10.1371/journal.pone.0189742.

Fernando A. do Carmo received the B.S. degree in Computer Science from the Federal University of Western Pará (UFOPA) in 2025. He is currently pursuing the M.S. degree in Computer and Systems Engineering at the State University of Maranhão (UEMA). His primary research interests include Artificial Intelligence and Data Science, with an emphasis on Text Mining and Natural Language Processing (NLP).

Gustavo S. Silva received the B.S. degree in Computer Engineering from the UEMA, São Luís, in 2024 and is currently pursuing the M.S. degree in Computer and Systems Engineering at the UEMA, São Luís. His primary research interests include Data Science and Artificial Intelligence, with emphasis on Natural Language Processing and AI Engineering.

Raimundo C. S. Freire received the B.S. degree in electrical engineering from the Federal University of Maranhão (UFMA), Brazil, in 1979, the M.S. degree from the Federal University of Paraíba (UFPB), Brazil, in 1982, and the Ph.D. degree from the Institut National Polytechnique de Lorraine, France, in 1988. He conducted postdoctoral research at the École Supérieure des Télécommunications in Paris, France. He is currently a full professor at the Federal University of Campina Grande (UFCG), Brazil. His current research interests include electronic instrumentation and sensors.

Antonio F. L. Jacob Jr received the B.S. degree in computer science from University of Amazon (UNAMA), Pará, Brazil, in 2005 and the M.S. degree in computer science from the Federal University of Pernambuco (UFPE), Pernambuco, Brazil, in 2008. He is currently pursuing a Ph.D. degree in electrical engineering at UFMA, Maranhão, Brazil. Since 2016, he is a Lecturer in the Computer Engineering Department at the UEMA.

Fábio M. F. Lobato received a degree in Computer Engineering (2010), a Master's (2011), and a Ph.D. (2016) from the Federal University of Pará, with a split-time period at the University of Kent. Dr. Lobato conducted postdoctoral research at the University of País Vasco (2018) and the University of São Paulo (2025). He is a CNPq productivity fellow in technological development, with a focus on computational intelligence, natural language processing, social networks, and scientometrics.