

Mapping Power and Collaboration in Urban Data Governance: A Complex Network Analysis of Social Actors in Brazil

Isabella Cavalcanti¹, Marcelo Tramontano²

^{1,2} Institute of Architecture and Urbanism (IAU), University of São Paulo, São Carlos, SP, Brazil

Abstract. *Urban data governance is increasingly shaped by sociotechnical infrastructures that may reproduce structural inequalities, particularly in Global South contexts. This paper examines how social actors are structured and interconnected in urban data governance in Brazil through the GHZ Network, combining Social Network Analysis (SNA) and critical data studies. Based on documentary research and relational sampling, we constructed a network of 35 core actors and 809 connected entities, modeling collaborations, funding relations, and data exchanges. Results reveal a sparse yet highly connected network, dependent on a small number of intermediary actors—primarily civil society organizations—while public sector actors exhibit low centrality. The analysis also highlights territorial concentration in Brazil’s South–Southeast and the influence of international funders, pointing to governance asymmetries. The study contributes a replicable framework for mapping sociotechnical networks and provides insights into how network structures relate to power and inequality in urban data governance.*

1. Introduction

Digital transformation has become central to public administration, promoting data-driven practices in urban governance. In Brazil, this process has been institutionalized through policy frameworks such as the Brazilian Digital Transformation Strategy (Brasil, 2018; 2022), which supports the development of digital public services and the strategic use of data in decision-making, including urban management.

While digital systems may improve decision-making, they may also reproduce or intensify inequalities, particularly when governance lacks transparency or inclusiveness (Broussard, 2018; Perez, 2019; D’Ignazio & Klein, 2020). These dynamics are especially relevant in Global South contexts, where structural conditions differ from those typically addressed in the literature.

Despite growing research on digital government, the relational dynamics shaping urban data governance remain underexplored. This paper asks: how are social actors structured and interconnected in urban data governance in Brazil, and what do these relationships reveal about power, collaboration, and inequality?

The analysis combines systems and complexity perspectives with critical feminist and decolonial approaches. This study contributes by proposing a framework for modeling sociotechnical networks based on documentary data and relational sampling, offering an empirical analysis of the Brazilian case and a critical interpretation of how network structures relate to broader sociopolitical dynamics.

This paper contributes in three interconnected ways. Methodologically, it proposes a framework for modeling sociotechnical networks based on documentary data and relational sampling. Analytically, it provides an empirical account of how power, collaboration, and inequality are structured in urban data governance in Brazil. Interpretively, it advances a critical reading that connects network structures to broader sociopolitical dynamics. These contributions support the identification of opportunities for more distributed and inclusive data governance practices.

2. Background

This study approaches the city as a complex sociotechnical system, drawing on systems and complexity perspectives (Bertalanffy, 2010; Morin, 2015; Batty, 2013). To analyze urban data governance, we adopt an interdisciplinary approach combining urban studies and complex network analysis, which provides the conceptual and methodological basis for the GHZ Network.

Urban data governance is embedded in sociotechnical and historical dynamics, where digital technologies may both support decision-making and reinforce inequalities (Batty, 2013; Kitchin & Dodge, 2014; Townsend, 2015; Noble, 2018; Eubanks, 2018). This is particularly relevant in Global South contexts, where data infrastructures intersect with existing structural asymmetries.

To examine these dynamics, we draw on critical perspectives that conceptualize data as embedded in power relations. In particular, Data Feminism (D'Ignazio & Klein, 2020) highlights how data practices can reproduce inclusion and exclusion, emphasizing the need for more transparent and participatory governance. In this study, data governance is understood as the set of practices and institutional arrangements shaping the production, circulation, and use of data (Obendiek, 2023).

Methodologically, the analysis is grounded in complex network theory, using concepts such as centrality, modularity, and connectivity to examine relational structures among actors. This approach supports the identification of coordination patterns, intermediation roles, and structural clustering, enabling a combined quantitative and interpretive analysis of urban data governance.

Rather than aiming at statistical generalization, the study adopts an analytical perspective focused on identifying relational patterns and sociotechnical dynamics. The findings are intended to contribute to theoretical and methodological discussions on urban data governance, particularly in Global South contexts.

3. Methodology

The GHZ Network is an experimental sociotechnical artifact that combines network modeling with the development of a digital platform for exploring interactions among actors in urban data governance. This paper focuses on the network modeling component, framed by systems thinking, complexity theory, and critical perspectives that understand data governance as a socially situated process.

The study adopts an exploratory approach to map and model a sociotechnical network of actors involved in urban data governance. The network is used as an

analytical instrument to examine structures of interaction, collaboration, and data production among institutional actors.

Data collection followed a structured protocol: (1) identification of an initial set of actors based on thematic relevance and prior involvement in urban data governance initiatives in Brazil; (2) systematic consultation of official sources, including institutional websites, reports, and public documents; (3) extraction of explicitly documented relationships between actors; (4) classification of relationships into predefined categories (funding, partnerships, data provision, and institutional networks); and (5) iterative expansion through link-tracing procedures until no substantially new actors were identified.

A relationship (edge) was recorded whenever an explicit reference to interaction between two actors was identified in at least one official source, such as the description of a joint project, funding support, data collaboration, or participation in a shared initiative. All relationships were treated as binary (unweighted), meaning that the presence of a documented interaction was sufficient to establish a connection, regardless of its intensity or frequency. This decision reflects both the nature of the available data and the exploratory objective of identifying structural patterns rather than measuring interaction strength.

To enhance data reliability, the collection and coding process followed consistent inclusion criteria based on explicit textual evidence. The dataset was iteratively reviewed during construction, including cross-checking of actor identities, removal of duplicates, and verification of relationship classifications. While no formal inter-coder reliability metrics were applied, the combination of systematic coding rules and collaborative review aimed to ensure internal consistency and traceability of the network construction process.

The database was exported as node and edge lists and analyzed using Python libraries, including NetworkX, NumPy, PyVis, and CSV tools. While the broader GHZ Network includes the development of an interactive platform, this paper focuses on the analytical potential of the network model.

The qualitative interpretation of network patterns was conducted collaboratively by the authors, based on an iterative reading of network metrics, visualizations, and contextual information about the actors. Rather than following a formal coding scheme, the interpretive process was guided by theoretically informed analytical lenses, particularly critical data studies and Data Feminism, allowing the identification of patterns related to power, intermediation, and inequality. Disagreements in interpretation were resolved through discussion, ensuring coherence in the analytical narrative.

4. Network Modeling (Rede GHZ)

4.1. Identification of social actors

The network modeling process began with the identification of an initial set of 35 social actors selected based on their documented involvement in urban data governance initiatives in Brazil, particularly in relation to data production, analysis, or use in urban policy contexts between 2018 and 2025. This initial set reflects a purposive sampling

strategy, combining thematic relevance and visibility in publicly available sources. From this core, additional actors and relationships were identified through documentary link-tracing procedures (Bowen, 2009; McLevey, Scott, and Carrington, 2024).

Actors were included based on their documented use of digital technologies and data-related practices in the context of urban governance. In the case of civil society organizations, inclusion criteria required demonstrated engagement in the production, modeling, or use of data to support social agendas, as well as a clear relationship with urban territories, either through direct action, advocacy, or influence on public policy processes. For public sector actors, inclusion was based on their institutional relevance in data governance at the national level and their role in the formulation or implementation of urban public policies.

Edges between actors were established based on the identification of explicit relationships documented in publicly available sources. A connection was recorded only when the relationship was explicitly documented in at least one official source. Reciprocal confirmation was not required, meaning that a single documented reference was sufficient to establish an edge. This criterion prioritizes traceability over symmetry in the representation of relationships. These references included mentions of project partnerships, funding relationships, data provision, or participation in shared initiatives or networks.

At the end of this process, 35 social actors were identified, forming the initial core of the network, as illustrated in Figure 1. This set constitutes a purposive and relationally expanded sample suitable for network-based analysis.



Figure 1. Main social actors in the GHZ Network, showing the core actors identified. Source: Authors.

Rather than an exhaustive representation of the field, this sample should be understood as analytically bounded, reflecting both the applied selection criteria and the scope of the study. Future research may incorporate additional actors, expanding the network and further refining the analysis.

4.2. Network structure

In addition to the 35 core actors defined as the primary nodes of the network, a further 809 actors connected to these core entities were incorporated into the database.

In alignment with the analytical objectives of the study, a set of attributes was defined for each actor, including the type of relationship established with other actors. Relationships were categorized into four types: (i) funding, (ii) project partnerships, (iii) data provision, and (iv) participation in institutional networks or coalitions. These categories provide analytical depth by clarifying how and why connections between actors are formed.

Relationships were modeled as undirected, as the analysis focuses on the existence of connections rather than their directionality. This choice reflects both the nature of the data sources and the analytical emphasis on structural connectivity. While this simplification does not capture asymmetries in direction or intensity, it supports a consistent representation of relational structures across heterogeneous data sources.

This categorization process can be understood as a form of qualitative coding, in which documented interactions were systematically classified into analytically defined categories. The coding scheme was applied consistently across all relationships, based on explicit textual evidence rather than interpretive inference, ensuring comparability and supporting the integration of qualitative and quantitative analysis. Relationships were modeled as undirected, as the analysis focuses on the existence of connections rather than their directionality, reflecting both the nature of the data sources and the emphasis on structural connectivity.

In this initial modeling, all edges were treated as unweighted, meaning that each identified relationship was assigned the same value. This decision reflects the exploratory nature of the study, prioritizing the identification of relational structures over the measurement of interaction intensity. Future developments may incorporate weighted edges to account for variations in the strength or frequency of relationships. A summarized diagram of these attributes is presented in Figure 2, offering a visual overview of the relational dimensions captured in the network.

Another attribute included was the temporal duration of relationships. While some connections result from short-term, project-based interactions, others persist over longer periods and may be continuously renewed.

Given the interest in examining the influence of international actors on local data governance processes, an additional attribute was introduced to capture the origin of actors, distinguishing between national and international entities.

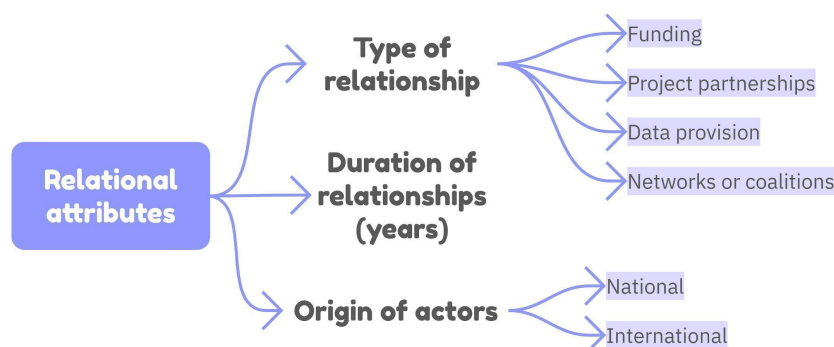


Figure 2. Relational attributes of actors in the GHZ Network. Source: Authors.

4.3. Data model

The relational data model was designed to systematically organize information on social actors and their interactions, ensuring consistency and supporting subsequent network analysis. The database consists of three main structures:

Actors: including all social actors with attributes such as sector, thematic focus, and territorial scope.

Relationships: capturing interactions between actors, categorized by type, temporal duration, and origin (national or international).

Attributes: detailing node and edge properties, enabling both quantitative metrics and qualitative interpretation.

Data normalization procedures were applied to remove duplicates, standardize institution names, and maintain consistent representation of bidirectional relationships. This ensures the network accurately reflects observed interactions and supports rigorous analyses.

To enhance internal consistency, the dataset was reviewed during the modeling process to verify the coherence of actor attributes and relationships, as well as to eliminate duplications and inconsistencies. While no formal inter-coder reliability metrics were applied, the combination of systematic data structuring and collaborative interpretation contributed to the analytical robustness of the network within the scope of this exploratory study.

The database was exported as node and edge lists for complex network analysis using Python libraries, including NetworkX, NumPy, Pathlib, PyVis, and CSV. A visual representation of the data model is shown in Figure 3, illustrating the main entities and their connections.

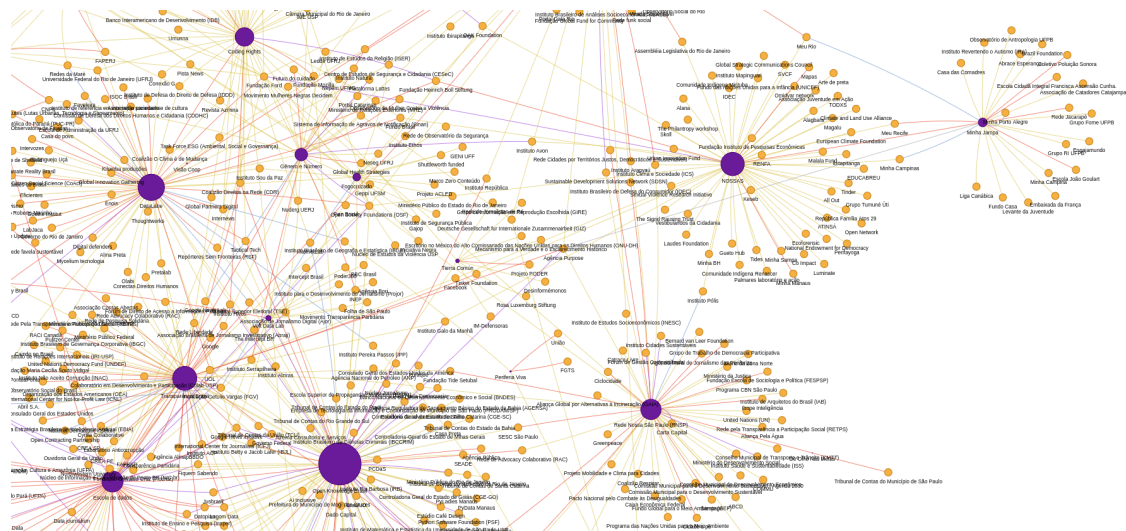


Figure 3. Relational data model of the GHZ Network, showing the main entities and their relationships. Source: Authors.

5. Network Analysis

5.1. Quantitative analysis

Following the methodological framework, the GHZ Network was analyzed using a Social Network Analysis (SNA) approach, examining not only actors but also the relationships, interdependencies, and information flows connecting them. Drawing on Data Feminism (D'Ignazio & Klein, 2020), these relationships are understood as historically and politically constructed, shaped by power asymmetries that produce exclusions and hierarchies. This perspective supports an integrated analysis of both network structure and its social, political, and institutional implications.

The quantitative analysis was organized into four axes: (i) overall structure and connectivity, (ii) centrality metrics, (iii) degree distribution and global patterns, and (iv) community detection. The GHZ Network comprises 844 nodes (35 core actors and 809 connected entities) and 1,017 edges, with low density (0.003), indicating limited direct connections among actors. Nevertheless, two main components were identified, with the largest containing 805 nodes, suggesting a structurally integrated network capable of sustaining information flows despite its sparsity.

Degree centrality identifies Casa Fluminense as a key hub, while the Government of the State of Rio de Janeiro shows low connectivity (0.001), indicating a gap between formal authority and network coordination. Betweenness centrality highlights Casa Fluminense, Open Knowledge Brasil, and DataLabe as key intermediaries enabling communication across network segments. Closeness centrality indicates that Open Society Foundations and Fundação Ford occupy structurally advantageous positions for information dissemination. The nodes with the highest degree centrality are presented in Figure 4.

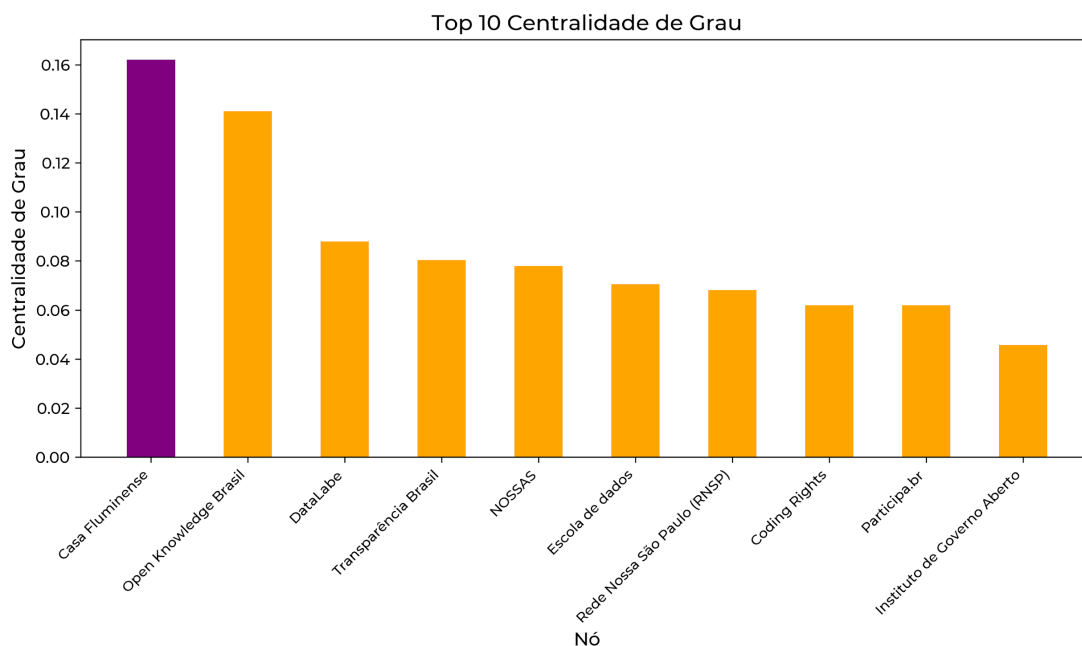


Figure 4. Nodes with the highest degree centrality in the GHZ Network. Source: Authors.

The degree distribution confirms structural asymmetries: despite a relatively low average degree (2.514), high variance (75.13) indicates a concentration of connections among a few nodes, while the majority of actors exhibit limited participation. The negative assortativity coefficient (-0.523) reflects a disassortative network, in which hubs tend to connect with low-degree nodes, consolidating their coordinating function. Community detection revealed 15 distinct subgroups, illustrated in Figure 5, with high modularity (0.737), suggesting well-defined cooperation clusters within the network.

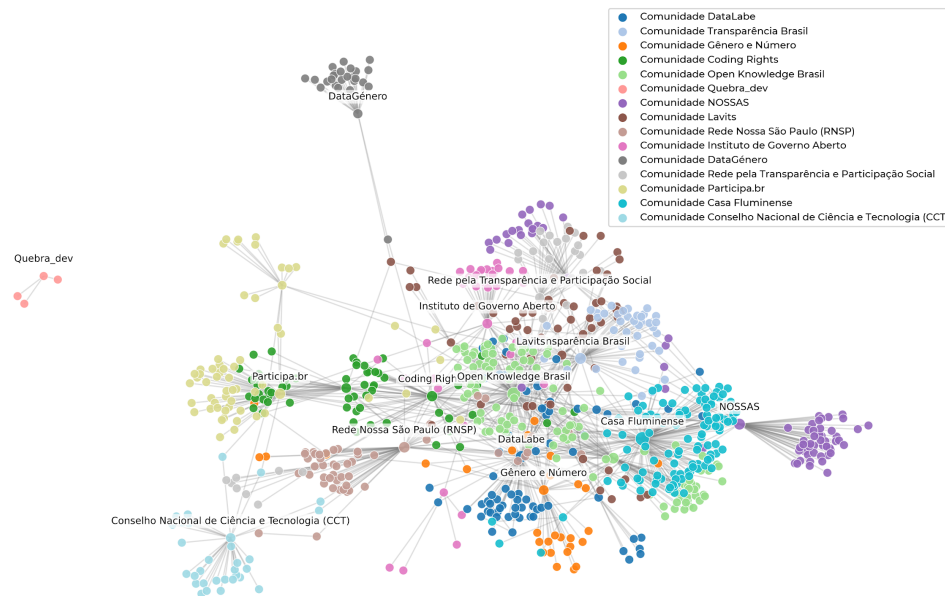


Figure 5. Identified communities within the GHZ Network, illustrating structural clusters and subgroup organization. Source: Authors.

To assess the robustness of the observed network structure, we conducted a percolation-based analysis simulating both random failures and targeted attacks on highly connected nodes. The procedure consisted of iteratively removing increasing fractions of nodes and measuring the size of the largest connected component. To account for stochastic variability, the random failure scenario was repeated 50 times, and results were aggregated using mean values and standard deviations. The results, presented in Figure 6, indicate that while the network is relatively resilient to random failures, it is highly vulnerable to targeted removal of hubs, with rapid fragmentation occurring as key intermediary actors are removed. This behavior reinforces the structural dependence on a small number of central nodes and supports the interpretation of the network as both cohesive and fragile.

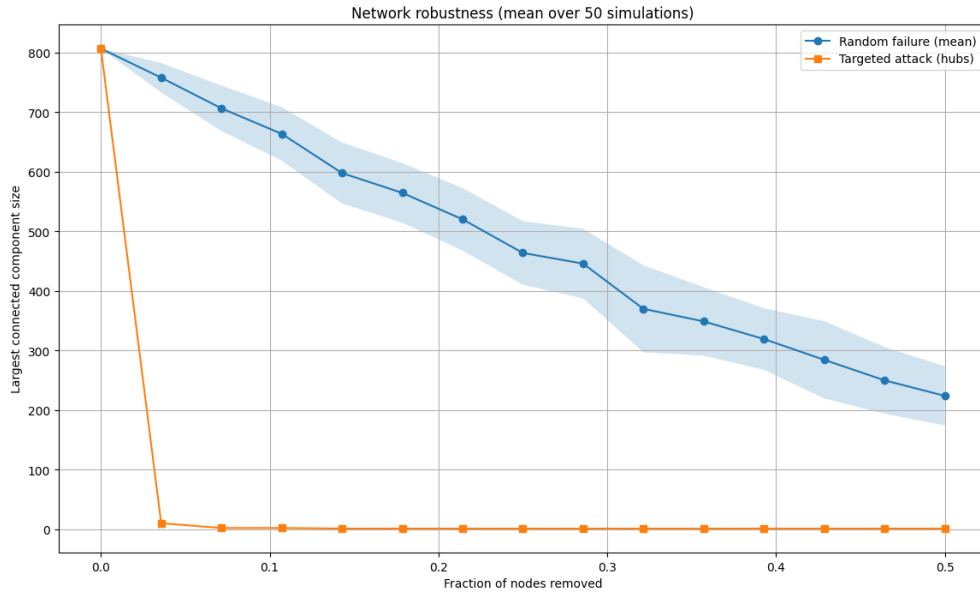


Figure 6. Network robustness under random failures and targeted attacks, showing the largest connected component vs. fraction of removed nodes (mean $\pm$ standard deviation, 50 simulations). Source: Authors.

5.2. Qualitative interpretation

The qualitative analysis complements the quantitative findings by shifting the focus from network structure to the meanings of relationships and their political implications. This interpretation combines insights from network analysis with historical, institutional, and social contexts, providing a deeper understanding of how relational configurations shape urban data governance.

Actor activity shows a clear concentration in Brazil's South–Southeast region, particularly in Rio de Janeiro and São Paulo. Although no areas are entirely unconnected, the georeferenced analysis (Figure 7) reveals regional gaps and potential opportunities for new connections. This highlights the importance of considering not only ties, but also the spatial distribution of data governance relations.

Funders play a central role in shaping network dynamics. More than half of the resources come from international organizations, such as Open Society Foundations, as well as foreign governmental bodies, revealing power asymmetries that influence agendas, methodologies, and data usage. From a critical perspective, these patterns raise questions regarding justice, transparency, and local autonomy, exposing tensions between data production, community care, and resource distribution. The predominance of international actors highlights the need for governance models that promote accountability, consent, and collective benefit, preventing the reproduction of structural inequalities.

In summary, the qualitative analysis contextualizes and interprets the structural patterns identified quantitatively, linking network metrics to social, political, and institutional processes. The integration of numerical analysis and critical interpretation provides a situated and comprehensive understanding of the GHZ Network, highlighting

hubs, bridges, and clusters that shape urban data governance in Brazil, as well as the tensions and opportunities emerging from historical and territorial asymmetries.

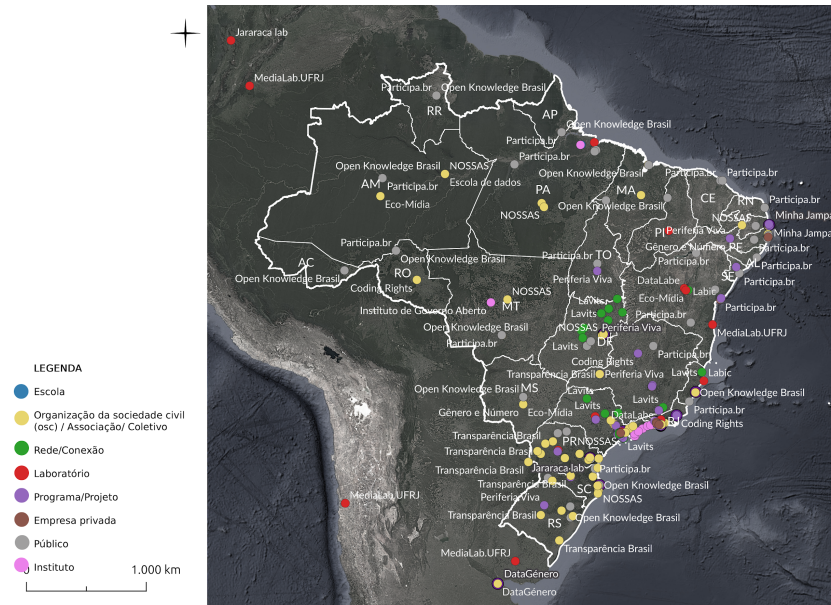


Figure 7. Territorial visualization of the GHZ Network's activity: Brazil and neighboring countries. Source: Authors.

6. Discussion

The GHZ Network reveals a low-density yet highly connected structure, in which cohesion depends on a limited number of central actors functioning as hubs and intermediaries. Civil society organizations play a key role in sustaining coordination, suggesting that urban data governance operates through distributed sociotechnical arrangements rather than exclusively state-centered structures.

The concentration of connections indicates dependence on key actors, generating both cohesion and vulnerability, as confirmed by the robustness analysis. The network remains relatively stable under random failures but fragments rapidly when highly connected nodes are removed, reinforcing the structural importance of intermediary actors.

The presence of international funders, such as the Open Society Foundations and the Ford Foundation, highlights external influences that may shape agendas and resource distribution, raising questions about data autonomy and governance asymmetries. Additionally, the concentration of actors in Brazil's South-Southeast reflects persistent regional inequalities in access to data infrastructures and participation.

These findings should be interpreted within the limits of the adopted methodology. As a documentary-based and relationally sampled network, the results reflect patterns observable in publicly documented interactions rather than a comprehensive representation of the entire governance ecosystem. Consequently, interpretations regarding power, collaboration, and inequality combine empirical

network measures with theoretically informed analysis and should not be understood as directly generalizable to other contexts

7. Final considerations

This paper presented the GHZ Network as an approach for analyzing urban data governance through the integration of Social Network Analysis and critical perspectives. The results highlight an ecosystem structured by strategic actors, with strong participation from civil society, international influence, and territorial inequalities.

This work contributes by proposing a sociotechnical framework for analyzing urban data governance that integrates quantitative network modeling with situated qualitative interpretation. Rather than offering a purely descriptive mapping, the GHZ Network functions as an analytical device that makes visible how power, collaboration, and inequality are structured through relational configurations among actors.

From a practical perspective, the framework can support public institutions, civil society organizations, and researchers in identifying strategic actors, coordination gaps, and underrepresented territories. It also provides a basis for designing more inclusive and collaborative data governance arrangements, particularly in contexts marked by institutional fragmentation and social inequality, such as Brazil.

This study presents important limitations that must be considered when interpreting its findings. First, the network was constructed based on documentary data and relational sampling, which means that the inclusion of actors depends on their visibility in publicly available sources. As a result, the network may underrepresent less institutionalized, informal, or locally embedded actors. Second, the selection of the 35 core actors and their subsequent expansion to 809 connected entities reflects an analytically guided rather than exhaustive sampling strategy, which constrains the generalizability of the results.

The findings should be interpreted as indicative of relational patterns derived from documented interactions rather than as a definitive representation of the entire urban data governance ecosystem. As such, conclusions regarding power, collaboration, and inequality reflect patterns observable within the constructed network and should be understood in light of the study's methodological scope and data limitations.

Additionally, the findings should be understood as indicative of relational patterns within a specific empirical configuration rather than as representative of urban data governance in Brazil as a whole or in other contexts. The interpretation of power, collaboration, and inequality is therefore bounded by the scope of the dataset and the methodological choices made in constructing the network.

Finally, while the integration of quantitative metrics and qualitative interpretation enables a richer analysis, the interpretive dimension remains situated and dependent on the analytical framework adopted. These limitations do not invalidate the findings but rather define the conditions under which they can be interpreted and extended.

This study highlights that urban data governance should be understood not only as a technical domain but as a sociotechnical and political process structured by

relationships among actors. By making these relationships visible, the GHZ Network contributes to a more nuanced understanding of how data-driven governance operates and where opportunities for more inclusive and coordinated arrangements may emerge.

Acknowledgment

We gratefully acknowledge the financial support provided by CAPES for this research.

References

- Amrute, S. (2016). *Encoding race, encoding class: Indian IT workers in Berlin*. Duke University Press.
- Batty, M. (2013). *The new science of cities*. MIT Press.
- Bertalanffy, L. von. (2010). *Teoria geral dos sistemas: Fundamentos, desenvolvimento e aplicações* (F. M. Guimarães, Trad.). Editora Vozes. (1ª ed. 1968).
- Brasil. (2018). *Estratégia brasileira para a transformação digital – E-Digital: Ciclo 2018–2022*. Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC).
- Brasil. (2022). *Estratégia brasileira para a transformação digital (E-Digital): Ciclo 2022–2026*. Centro de Gestão e Estudos Estratégicos (CGEE) & Ministério da Ciência, Tecnologia e Inovações (MCTI).
- Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT Press.
- Bowen, G. A. (2009). *Document analysis as a qualitative research method*. Qualitative Research Journal, 9(2), 27–40.
- D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Kitchin, R., & Dodge, M. (2014). *Code/space: Software and everyday life*. MIT Press.
- Morin, E. (2015). *Introdução ao pensamento complexo* (E. Lisboa, Trad.). Sulina. (1ª ed. 1990).
- Morin, E., & Le Moigne, J.-L. (2000). *A inteligência da complexidade* (N. Maria, Trad.). Fundação Peirópolis. (1ª ed. 1999).
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Obendiek, A. S. (2023). *Data governance: Value orders and jurisdictional conflicts*. Oxford University Press.
- Perez, C. C. (2019). *Invisible women: Exposing data bias in a world designed for men*. Random House.
- Townsend, A. (2015). *Cities of data: Examining the new urban science*. Public Culture, 27(2), 201–212.