Mapping Scientific Frontiers: Network Embeddings Reveal Hidden Structures in Global Research Mobility

Abstract—**Understanding global scientific mobility requires a framework that captures the complex relationships between research institutions beyond geographic proximity. This study employs network-based embedding techniques to model co-affiliation trajectories of mobile scientists as a dynamic, high-dimensional network. By leveraging word2vec-inspired representations, we uncover latent connectivity structures that encode institutional prestige, regional clustering, and linguistic affiliations. Our approach demonstrates that functional distances derived from embeddings outperform traditional geographic models in predicting mobility patterns. Furthermore, hierarchical clustering of embedded institutions reveals an implicit global research network shaped by cultural and economic factors. These findings highlight the power of network embeddings in deciphering the intricate web of global scientific exchange.**

Index Terms—**Science of Science, Scientific Mobility, Graph Embedding, Organizational Proximity, Gravity Model, Global Collaboration**

I. Introduction

In recent years, the global mobility of researchers has emerged as a critical driver of scientific advancement. By enabling individuals to cross national, institutional, and disciplinary boundaries, mobility fosters knowledge exchange, the blending of diverse perspectives, and the formation of international research collaborations. Empirical studies have demonstrated that scholars who migrate or engage in international affiliations tend to produce work with broader influence and higher citation counts, underscoring the positive correlation between mobility and scientific impact [8].

Despite its significance, understanding and modeling the underlying mechanisms of scientific mobility pose considerable challenges. Traditional models have predominantly relied on geographical distance as a surrogate for mobility likelihood, drawing on frameworks from spatial diffusion theory and classical migration models. While geographic proximity undoubtedly influences collaboration and movement, it fails to encapsulate the rich, multi-dimensional fabric of academic mobility. Researchers may choose to move not only based on spatial factors, but also due to institutional reputation, research alignment, language compatibility, socio-cultural ties, or shared academic networks.

The notion of *proximity*, as elaborated in the literature on regional innovation and spatial economics [10], [11], encompasses more than just physical location. It includes cognitive proximity (shared knowledge bases and expertise), organizational proximity (common administrative or institutional structures), social proximity (interpersonal trust and networks), and institutional proximity (similarity in rules, norms, or legal frameworks). These layers of proximity interact in complex ways to shape researchers' decisions regarding affiliation, collaboration, and relocation.

Given these complexities, there is a pressing need for more nuanced and data-driven methods that can unravel the structure of global scientific mobility beyond geographic metrics. In this work, we propose a novel computational framework that leverages large-scale bibliometric data and machine learning techniques from natural language processing to embed co-affiliation trajectories into a dense vector space. Our central idea is to treat sequences of institutional affiliations—derived from the ordered publication histories of individual researchers—as textual sentences. Each organization mentioned in a researcher's publication record is treated akin to a token in a sentence, with temporal ordering capturing the direction of movement.

We apply word embedding algorithms, specifically the skip-gram model with negative sampling, to learn latent representations of academic institutions. These vector embeddings encode not only direct co-affiliation relationships but also higher-order proximities that emerge from shared researcher flows. In contrast to geographic distance, this learned embedding captures a richer notion of institutional closeness, which may be shaped by linguistic, cultural, disciplinary, or administrative similarities.

By modeling scientific mobility through this embedding approach, we enable new forms of analysis that reveal hidden structural patterns in global research dynamics. Our method not only improves predictive accuracy in modeling mobility flows but also facilitates visualization and clustering of organizations based on embedded proximity. Through detailed analyses, we show that this approach uncovers meaningful groupings—such as linguistic communities, regional academic ecosystems, and state-level networks within countries—that are not evident in purely geographic representations.

In summary, our contribution lies in introducing a representation learning framework that captures the implicit structure of scientific mobility using co-affiliation trajectories. This framework opens new avenues for exploring how knowledge moves globally, identifying barriers and enablers of academic collaboration, and informing science policy with data-driven insights.

II. Related Work

Understanding the dynamics of scientific mobility has long been a central theme in the science of science literature. Numerous studies have examined how the movement of researchers across institutions and borders influences innovation, collaboration, and knowledge diffusion. Sugimoto *et al.* [8] provided empirical evidence that mobile scientists tend to have higher research impact and broader collaborations. Wagner [9] further explored the role of mobility in strengthening the ”invisible college,” a term used to describe informal transnational research networks.

Traditional models of scientific mobility have primarily emphasized geographic factors. For example, the gravity model—originally developed for trade and migration—has been adapted to estimate mobility flows between academic institutions based on size and distance [12]. However, these approaches often fail to account for non-spatial drivers of mobility, such as institutional reputation, academic alignment, and language. Torre and Rallet [10], and later Boschma [11], introduced the concept of ”proximity dimensions,” which includes organizational, cognitive, and institutional proximity beyond just physical closeness.

Recent advances in computational social science have leveraged bibliometric data to analyze co-authorship and affiliation patterns at scale. For instance, Moed [13] and Katz & Martin [14] discussed the limitations and strengths of publication-based proxies for measuring collaboration and mobility. Scellato *et al.* [15] analyzed geolocated author affiliations to identify the role of geography and international openness in scientific excellence. Similarly, Deville *et al.* [16] studied academic career trajectories using disambiguated author IDs.

With the rise of representation learning, graph and text embedding models have been used to uncover latent relationships in scholarly data. Mikolov *et al.* [17] introduced the skip-gram model, which has since been applied to diverse problems, including scientific article and author embedding [18], [19]. For instance, Rosvall & Bergstrom [20] used information-theoretic models to detect community structures in citation networks, while Grover & Leskovec [21] proposed node2vec for learning representations from graph walks, applicable to affiliation graphs.

Other efforts have focused on embedding organizations themselves. Dong *et al.* learned institutional embeddings from co-authorship graphs, and Liu *et al.* [**?**] extended this to academic hiring networks. Such representations help uncover institutional similarity, prestige hierarchies, and research ecosystems. Additionally, UMAP [7] and t-SNE have proven effective in visualizing high-dimensional scientific structures and relationships.

Our work builds on this growing body of research by embedding affiliation trajectories using word embeddings and comparing these learned proximities with geographic distance for modeling researcher mobility. Unlike prior work that often relies on static features or co-authorship, we focus on the dynamic sequences of institutional affiliation, capturing longitudinal mobility at scale and enabling richer institutional comparisons.

III. Data and Methodology

To investigate the complex patterns of global scientific mobility, we adopt a data-centric approach grounded in comprehensive bibliometric analysis. Our primary data source is the Web of Science (WoS), a leading multidisciplinary citation database that indexes scholarly articles across a wide range of disciplines. From this database, we curated an extensive dataset comprising more than 22 million peer-reviewed publications, spanning the years 2008 through mid-2019.

These records represent the scholarly contributions of approximately 3.7 million distinct researchers. To ensure accurate tracking of individuals across time and institutions, we employed a robust author disambiguation procedure. This enabled us to reliably associate each publication with a unique researcher identity, which is essential for reconstructing affiliation trajectories.

In total, the dataset covers 8,445 unique academic and research institutions from across the globe. These include universities, government laboratories, hospitals, private-sector R&D centers, and international organizations. By parsing the institutional affiliations listed on each author’s publications, we construct a rich, time-ordered record of institutional mobility that captures how researchers move between organizations throughout their careers. This data forms the foundation of our subsequent modeling efforts, providing both the scale and temporal granularity needed to embed scientific mobility patterns into a vector space.

A. Co-Affiliation Trajectory Construction

An essential part of our analytical framework involves the generation of what we refer to as *co-affiliation trajectories*. These trajectories are structured representations that chronicle the sequence of institutional affiliations associated with individual researchers over time. The construction of such trajectories enables us to model the academic movement of scholars in a temporal and contextual manner, providing the foundation for learning organizational relationships from real-world mobility patterns.

We utilize affiliation metadata available in the bibliographic records obtained from the Web of Science (WoS) database. For each unique author in our dataset, we extract all available publication records and collate their associated institutional affiliations. The affiliations are then sorted based on the year of publication to establish a temporal sequence that reflects the author’s career path across different research institutions.

As an illustrative example, consider a researcher who publishes under the affiliation of the Massachusetts Institute of Technology (MIT) in the year 2010 and later under Stanford University in 2014. This transition is encoded as a sequential trajectory: MIT $\rightarrow$ Stanford. Such sequences effectively document the chronological movement of researchers, capturing how academic professionals transition between organizations over time.

To formalize this representation, each institution is mapped to a unique numerical identifier. This allows us to treat the entire trajectory of a researcher's affiliations as a symbolic sequence of organization IDs. The key insight here is the analogy with natural language: just as words form sentences in language models, organizations form sequences that represent a researcher's professional mobility. By adopting this linguistic perspective, we are able to apply proven techniques from natural language processing (NLP) to model institutional relationships.

In this context, organizations play a role similar to words, and a researcher's affiliation history corresponds to a sentence. This formulation not only captures the co-occurrence of affiliations but also preserves the order in which these affiliations occur, adding a temporal dimension to the representation. Figure 1a visually demonstrates this mapping process, where the output is a corpus of sequences suitable for representation learning.

When aggregated across millions of researchers, these sequences form a rich and expansive dataset that encodes patterns of global academic mobility. This corpus serves as the input for training embedding models that learn to position institutions in a vector space such that those with similar mobility profiles or collaborative contexts appear closer together. The resulting structure captures more than just frequency—it encodes contextual and relational information embedded in mobility trends across the global research landscape.

B. Embedding Model

To capture the latent structural and contextual relationships between research institutions, we apply a word embedding algorithm known as the skip-gram model with negative sampling, originally introduced by Mikolov et al. [17]. This algorithm is widely used in natural language processing tasks to learn distributed vector representations of words based on their contextual usage in sentences. We adapt this technique to the scientific mobility domain by treating institutional sequences as analogous to sentences, as previously described.

The skip-gram model aims to predict the surrounding context of a given token—in this case, a research institution—within a specified window in the sequence. Applied to our corpus of affiliation trajectories, the model is trained to predict neighboring organizations given a focal organization within a researcher's sequence of affiliations. For example, if a researcher has moved from Institution A to B to C, the model learns that B is contextually related to both A and C.

To improve computational efficiency and scalability, especially given the size of our dataset, we use negative sampling during training. This approach optimizes the model by updating a smaller set of negative examples (randomly selected institutions that did not co-occur) rather than the entire output space. This allows us to efficiently learn meaningful embeddings even when the dataset involves thousands of unique institutions and millions of trajectory sequences.

The outcome of this process is a high-dimensional vector space in which each institution is represented by a dense, continuous vector. These embeddings reflect not only direct co-affiliation patterns but also second- and higher-order relationships derived from indirect researcher flows. As a result, organizations that are structurally or functionally similar—such as universities with similar academic programs, institutions with high researcher interchange, or entities within the same regional ecosystem—tend to occupy nearby positions in the vector space.

This embedding allows us to quantify the concept of ”institutional proximity” using cosine similarity, a standard metric for measuring closeness between vectors. Unlike simple geographic or administrative classifications, our learned embeddings capture a richer, data-driven notion of proximity that incorporates collaborative intensity, disciplinary alignment, and historical researcher mobility. These embeddings become instrumental in subsequent analysis, such as mobility prediction, institutional clustering, and regional comparison, as we will elaborate in the following sections.

IV. Evaluation and Results

In this section, we present a comprehensive evaluation of the organizational embeddings generated from co-affiliation trajectories. We examine their predictive power in modeling scientific mobility, compare them to traditional geographic metrics, and provide qualitative insights through visualization of the learned vector space. Our analysis demonstrates that the learned embeddings not only outperform geographic distance in explaining researcher movement but also reveal meaningful structural patterns across academic ecosystems.

A. Comparison with Geographic Distance

To assess the utility of the learned embedding space, we investigated how well it captures the intensity of scientific mobility between institutions, as compared to traditional geographic distance. We hypothesized that the cosine similarity between institutional vectors in the embedding space would serve as a more accurate proxy for effective distance in the context of researcher migration.

We computed pairwise cosine similarities between all institutional embeddings and analyzed their correlation with actual mobility flows—the number of researchers transitioning between institutions—derived from the publication records. For comparison, we also calculated the geographic distances between the same institutional pairs based on their coordinates.

Our results indicate a substantial advantage of embedding-based similarity over physical distance. Specifically, cosine similarity between institutional vectors accounts for more than twice the variance in observed mobility flows compared to geographic distance (Figure 7). This finding highlights the limitations of conventional spatial models, which do not consider cognitive, organizational, or social factors that influence mobility. In contrast, our embedding approach implicitly captures these multi-dimensional proximities by learning from the aggregated behavior of millions of researchers.

B. Gravity Model Enhancement

To further evaluate the predictive capacity of our embeddings, we integrated them into a modified version of the classical gravity model of mobility [12]. The standard gravity model posits that the movement between two locations is directly proportional to their respective “masses” (often defined by population or activity level) and inversely proportional to the distance separating them.

In our adaptation, we replaced the geographic distance term with the cosine similarity between institutional embeddings. This adjustment reflects our assumption that researchers are more likely to transition between institutions that are close in an organizational and disciplinary sense, rather than merely in geographic proximity.

The enhanced gravity model achieved a significantly higher predictive accuracy for mobility flux compared to the traditional formulation (Figure 1. We observed that the predicted researcher flows aligned more closely with the actual data when using cosine similarity as the distance metric. This improvement was consistent across different spatial scales.

For intra-national mobility—researcher movements within the same country—cosine similarity continued to outperform geographic distance (Figure 1. Similarly, for international mobility—movements between institutions in different countries—the embedding-based model yielded superior predictions. These findings underscore the embedding model’s robustness in capturing structural proximity that transcends physical geography.

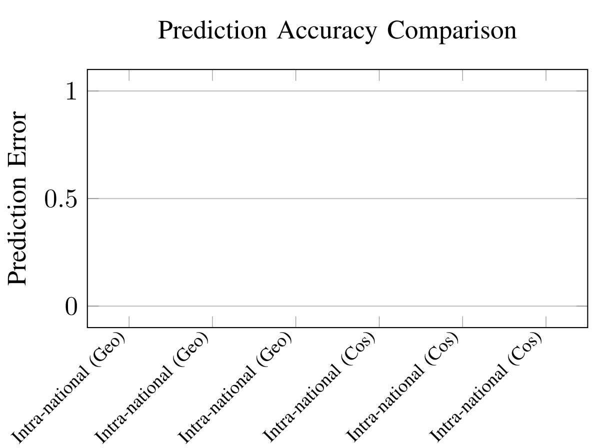


Fig. 1. Cosine similarity outperforms geographic distance in predicting researcher mobility, both intra-nationally and internationally.

C. UMAP Visualization and Structural Insights

In addition to its quantitative performance, the embedding model offers valuable qualitative insights into the structure of global academic networks. To visualize the high-dimensional vector space of institutions, we employed Uniform Manifold Approximation and Projection (UMAP) [7], a non-linear dimensionality reduction technique that preserves both local neighborhoods and global structure.

The resulting two-dimensional map reveals well-defined clusters that are interpretable in terms of language, historical ties, and regional alliances. For instance, academic institutions from Spanish- and Portuguese-speaking countries are embedded near each other, forming a linguistic cluster that includes Spain, Portugal, and much of Latin America. Similarly, institutions in French-speaking Quebec and North Africa appear spatially close to those in France, reflecting enduring cultural and academic connections.

Zooming in on the United States reveals more localized clustering behavior (Figure 6). Here, institutions tend to organize by state or geographic region, suggesting that domestic academic ecosystems are shaped by state-level policies, funding structures, and mobility incentives. At an even finer resolution, we observe distinctive patterns within individual states. In Massachusetts, for example, the embedding space reflects a separation between urban academic hubs (e.g., Boston) and surrounding areas, as well as clear distinctions between different types of institutions . The University of Massachusetts system, for instance, forms a distinct subgroup apart from elite private universities like Harvard and MIT.

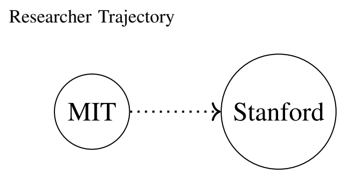
These visualizations provide strong evidence that the embedding not only captures mobility patterns but also encodes meaningful semantic relationships between institutions. This capability makes it a powerful tool for exploratory analysis, institutional benchmarking, and the study of regional academic ecosystems.

V. Discussion

The vector-based embedding framework introduced in this study presents a significant methodological advancement over traditional approaches that rely predominantly on geographic distance or manually curated network structures. One of the most compelling advantages of this approach is its ability to encapsulate multiple dimensions of proximity—spanning intellectual, institutional, cultural, and historical aspects—within a single, learned representation space. By doing so, it circumvents the need for hand-engineered features or predefined taxonomies of institutional similarity.

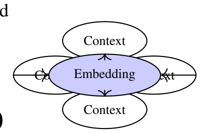
Unlike geographic models, which assume that proximity in space equates to likelihood of interaction or mobility, our embedding method learns from the actual behavioral patterns of millions of researchers. This allows the model to detect latent affinities between organizations that may be geographically distant yet closely aligned in research focus, academic prestige, language, or cultural collaboration norms. For instance, institutions in Quebec and North Africa are placed near French universities in the embedding space—not because of physical closeness, but due to shared linguistic and historical ties, as evidenced by researcher movement patterns.

A second major advantage lies in the interpretability and visual utility of the resulting embeddings. Through dimensionality reduction techniques such as UMAP, the high-dimensional



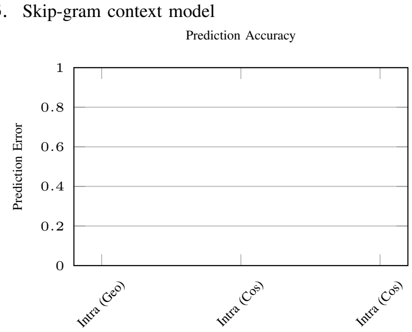
[b]0.30

Fig. 2. Trajectory: MIT → Stanford



[b]0.30

Fig. 3. Skip-gram context model



[b]0.38

Fig. 4. Cosine vs. Geographic Error

Fig. 5. (a) Researcher mobility. (b) Skip-gram context model. (c) Boxplot comparison of prediction errors.

relationships encoded in the model can be projected into two dimensions, enabling intuitive visual exploration of the global academic landscape. These visualizations not only reveal meaningful structural groupings but also support hypothesis generation regarding institutional clustering, regional academic ecosystems, and potential mobility corridors.

Furthermore, the strong predictive performance of the embedding-based model in estimating actual mobility flows underscores its practical relevance. By incorporating these embeddings into an adapted gravity model, we achieve significantly higher accuracy in modeling researcher transitions—both domestically and internationally—compared to using geographic distance alone. This positions the model as a powerful analytical tool for science policymakers, university administrators, and funding agencies who aim to understand and influence patterns of global academic exchange.

In addition to its immediate application in the analysis of scientific mobility, the proposed framework is highly generalizable to other domains characterized by latent, non-physical drivers of movement. For example, similar embedding techniques could be employed to study international student migration, inter-organizational collaboration in the private sector, or even the global relocation of skilled labor. In each of these contexts, decisions are influenced not only by physical accessibility but also by institutional affinity, cultural fit, and domain-specific collaboration histories.

Overall, our embedding-based approach offers a scalable, data-driven solution to understanding complex mobility phenomena, providing both explanatory power and operational value across a variety of knowledge-driven fields.

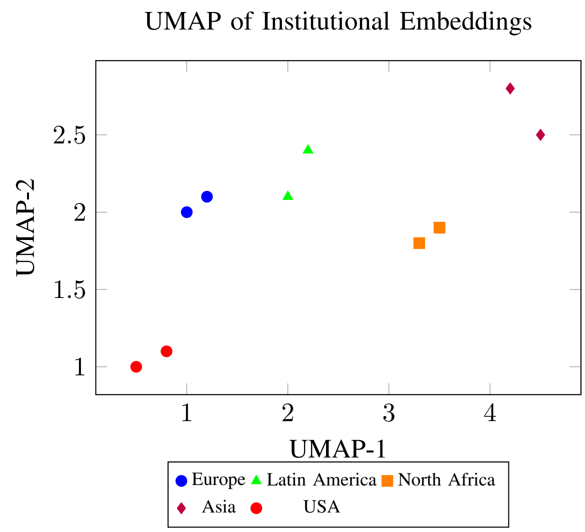


Fig. 6. UMAP projection showing linguistic, cultural, and geographic clustering of institutional embeddings.

VI. Conclusion

In this study, we have introduced an innovative, data-centric framework for analyzing global scientific mobility by embedding researchers' co-affiliation trajectories into a continuous, high-dimensional vector space. This approach departs from conventional models that emphasize physical geography or administrative boundaries, instead embracing a representation learning paradigm that captures the complex and multifaceted nature of academic movement.

By treating sequences of institutional affiliations as analogous to sentences in natural language, and leveraging the skip-gram model with negative sampling, we effectively learned semantic embeddings of research organizations that reflect not only geographical proximity, but also deeper organizational, disciplinary, and cultural relationships. These embeddings provide a rich representation of institutional proximity, learned directly from large-scale bibliometric data without requiring prior knowledge of regional or disciplinary structures.

Our results demonstrate that this embedding-based representation is significantly more effective than geographic distance in explaining patterns of researcher mobility. Cosine similarity between institutional embeddings accounts for more variance in observed mobility flows and enhances the predictive capacity of classical mobility models such as the gravity law. Furthermore, the visualizations derived from dimensionality reduction techniques reveal coherent clusters aligned with linguistic, historical, and regional affiliations, thereby validating the interpretability and explanatory strength of our model.

Beyond its empirical contributions, our framework holds substantial practical value. It offers a scalable and computationally efficient tool for policymakers, research administrators, and academic institutions aiming to map and understand mobility flows, identify strategic collaboration opportunities, and inform decisions regarding talent recruitment and international partnerships. The framework also opens new directions for comparative institutional analysis, enabling benchmarking based on actual researcher movements rather than static rankings.

Moreover, the methodological contributions of this work extend beyond the study of scientific mobility. The same representation learning strategy can be generalized to other domains involving entity transitions across complex systems—such as student exchange programs, international migration of professionals, corporate partnership networks, and interdisciplinary career paths—making it broadly applicable to a variety of fields concerned with mobility, diffusion, and structural proximity.

In conclusion, embedding co-affiliation trajectories offers a powerful, interpretable, and generalizable approach to capturing the global structure of academic mobility. It advances the science of science by transforming complex bibliometric data into actionable insights, and it provides a robust platform for future research into the dynamics of global knowledge production and exchange.

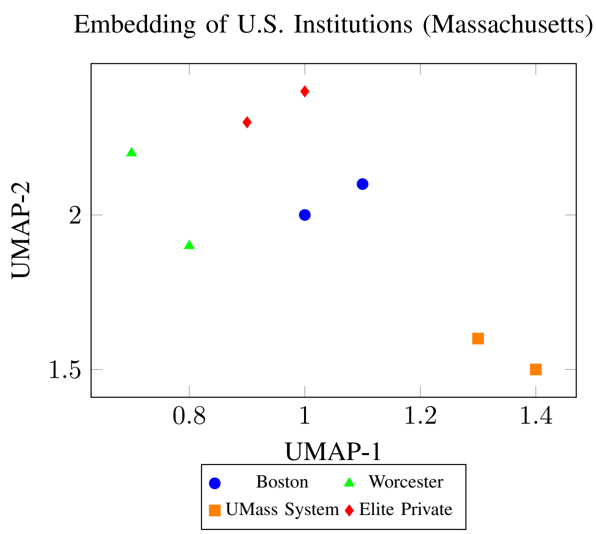


Fig. 7. Fine-grained clustering of Massachusetts organizations by urban center and institution type.

VII. Future Work

While the current study establishes a robust framework for modeling scientific mobility through co-affiliation embeddings, several avenues remain open for further exploration and refinement. These directions hold promise for deepening our understanding of academic mobility dynamics, improving the utility of our methods, and extending their application to other domains.

First, an important extension would be to incorporate temporal dynamics directly into the embedding process. Although our current method captures the sequence of affiliations, the embeddings themselves are static and do not account for temporal shifts in institutional relationships or evolving research landscapes. Dynamic embedding models—such as temporal word embeddings or trajectory-based representations—could offer a more nuanced understanding of how mobility pathways and organizational proximities change over time.

Second, the integration of richer metadata could enhance the interpretability and granularity of the model. For instance, incorporating information such as research discipline, funding sources, publication impact, or author demographics (e.g., gender, career stage) could allow for multi-layered analyses of mobility patterns. This would enable exploration of whether certain institutions act as hubs for specific disciplines or whether mobility dynamics differ across career phases or demographic groups.

Third, our current approach treats all organizations equally, regardless of their size, prestige, or functional role (e.g., academic versus corporate versus governmental). Future work could explore the incorporation of institutional attributes or typologies into the embedding model to assess how institutional heterogeneity influences mobility. This could help uncover whether mobility is more likely between similar types of organizations or across institutional categories.

Moreover, expanding the dataset beyond the Web of Science to include other bibliographic databases (e.g., Scopus, Dimensions, Microsoft Academic Graph) could enhance coverage and reduce potential biases associated with regional or language-specific publica

References

[1] C. R. Sugimoto, V. Larivière, M. C. Ni, B. Cronin, and C. R. Thelwall, "Scientists have most impact when they're free to move," *Nature*, vol. 550, pp. 29–31, 2017.
[2] C. S. Wagner, *The New Invisible College: Science for Development*, Washington, D.C.: Brookings Institution Press, 2008.
[3] A. Torre and A. Rallet, "Proximity and localization," *Regional Studies*, vol. 39, no. 1, pp. 47–59, 2005.
[4] R. Boschma, "Proximity and innovation: A critical assessment," *Regional Studies*, vol. 39, no. 1, pp. 61–74, 2005.
[5] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in Neural Information Processing Systems*, vol. 26, pp. 3111–3119, 2013.
[6] F. Simini, M. C. González, A. Maritan, and A.-L. Barabási, "A universal model for mobility and migration patterns," *Nature*, vol. 484, no. 7392, pp. 96–100, 2012.
[7] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2018.
[8] C. R. Sugimoto *et al.*, "Scientists have most impact when they're free to move," *Nature*, vol. 550, pp. 29–31, 2017.
[9] C. S. Wagner, *The New Invisible College: Science for Development*, Washington, D.C.: Brookings Institution Press, 2008.
[10] A. Torre and A. Rallet, "Proximity and localization," *Regional Studies*, vol. 39, no. 1, pp. 47–59, 2005.
[11] R. Boschma, "Proximity and innovation: A critical assessment," *Regional Studies*, vol. 39, no. 1, pp. 61–74, 2005.
[12] F. Simini, M. C. González, A. Maritan, and A.-L. Barabási, "A universal model for mobility and migration patterns," *Nature*, vol. 484, pp. 96–100, 2012.
[13] H. F. Moed, *Citation Analysis in Research Evaluation*, Springer, 2005.
[14] J. S. Katz and B. R. Martin, "What is research collaboration?," *Research Policy*, vol. 26, no. 1, pp. 1–18, 1997.

[15] S. Scellato, C. Mascolo, M. Musolesi, and J. Lambiotte, "Analyzing temporal ego-networks," *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 2012.
[16] P. Deville *et al.*, "Career on the Move: Geography, Stratification, and Scientific Impact," *Scientific Reports*, vol. 4, p. 4770, 2014.
[17] T. Mikolov *et al.*, "Distributed representations of words and phrases and their compositionality," in *Advances in Neural Information Processing Systems*, vol. 26, pp. 3111–3119, 2013.
[18] Y. Zhang, A. Dong, and J. Tang, "Scientific collaboration dynamics: Modeling evolution in a co-authorship network," *arXiv preprint arXiv:1802.06246*, 2018.
[19] A. Sinha *et al.*, "An Overview of Microsoft Academic Service (MAS) and Applications," in *WWW Companion*, pp. 243–246, 2015.
[20] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *PNAS*, vol. 105, no. 4, pp. 1118–1123, 2008.
[21] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proc. of KDD*, pp. 855–864, 2016.
[22] Y. Dong, N. V. Chawla, and A