Making Invisible Colleges Visible:
Multidimensional Scaling and Interactive 3D Exploration of the Vilnius University Co-authorship Network

Martynas Sabaliauskas, Karolis Pempė
E-mail: martynas.sabaliauskas@mif.vu.lt
Vilnius University, Institute of Data Science and Digital Technologies

Abstract

Publication counts and the h-index describe individual researchers but ignore who works with whom. Co-authorship networks reveal whether the actual research groups – the *invisible colleges* [1] – follow the formal structure of an institution. We reconstructed the co-authorship network of Vilnius University from the national repository eLABa and explored it in two complementary ways: **static SMACOF maps** that reveal the global cluster structure, and an **interactive 3D explorer** where anyone can find a researcher and follow their collaborations.

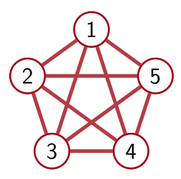
From repository to weighted graph

1. **Organizational tree** — recursive walk over 3 levels → **4 103** staff, emeriti & affiliates, 15 units
2. **Harvest eLABa records** — cached, rate-limited requests → one query per person
3. **Extract & deduplicate** — authors matched to staff list → **75 671** unique publications
4. **Fractional tie weights** — symmetric weight matrix → **3 221** authors, **20 512** ties
5. **Static files ≤ 10 MB** — SMACOF maps & 3D explorer → no server, no database

Tie strength by fractional counting [2]:

$$a_{ij} = \sum_{p \in P_{ij}} \frac{1}{N_p}$$

P_{ij}: joint papers of i and j; N_p: number of *all* authors of p, external ones included. A two-author paper adds 0.50, a five-author paper only 0.20 per pair, so consortium papers do not dominate.



$+\frac{1}{5} = 0.20$ for each of 10 pairs

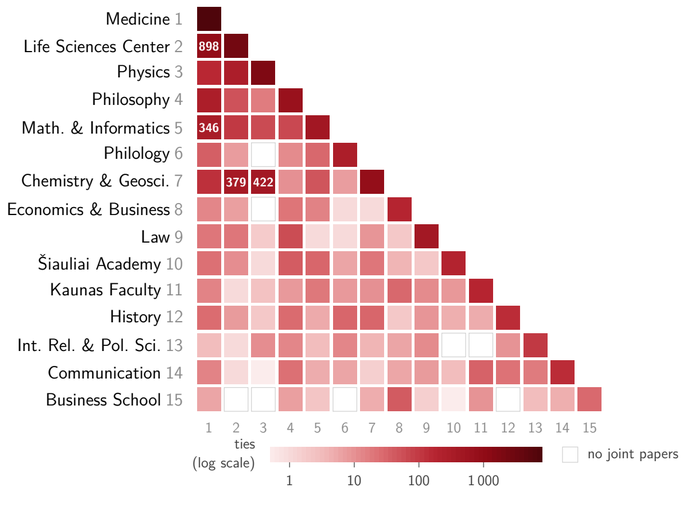
A sparse, clustered small world

Authors (vertices)	3 221
Co-authorship ties (edges)	20 512
Density	0.004
Degree: median / max	8 / 165
Average clustering coefficient	0.428
Largest component	2 868 (89.0 %)
Mean shortest path (largest comp.)	4.13

Sparse, strongly clustered, short paths: a small world, as is typical of scientific collaboration [3, 4].

Do collaborations follow faculties?

- **82.8 %** of ties join authors sharing a faculty; faculty partition: $Q = 0.694$ ($Q > 0.3$: significant community structure [6]).
- Louvain [7], blind to organizational data, reaches a moderately higher $Q = 0.813$, but its largest community (306 authors) is far smaller than the largest faculty (893): large faculties split into smaller research groups.



Ties between faculty-level units (columns = row numbers; ties of multi-unit authors split proportionally). Numbers: the four strongest interfaculty axes.

Static view: SMACOF maps

Metric MDS places the authors in $\mathbb{R}^m$ ($m = 2, 3$) by minimizing the raw stress

$$\sigma(X) = \sum_{i<j} \big(\delta_{ij} - d_{ij}(X)\big)^2$$

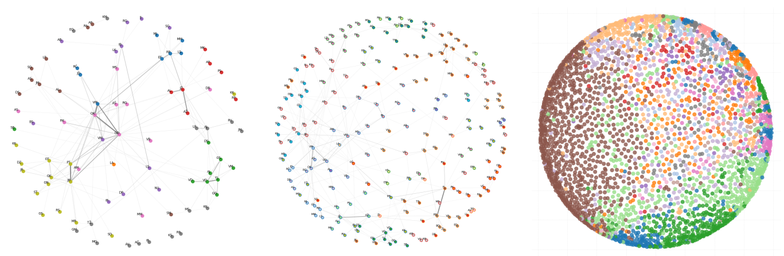
with the SMACOF majorization algorithm [5]. Tie strengths become dissimilarities

$$\delta_{ij} = \frac{1}{\ln(1 + a_{ij}) + 1} \in (0, 1],$$

- defined for all pairs: $\delta_{ij} = 1$ without joint papers;
- bounded, and damping a few very strong ties (max $a_{ij} = 159.8$ vs. median 0.25).

A PCA initial configuration makes every map reproducible.

Clusters sharpen with scale



one institute — 91 authors | one faculty — 248 authors | university — ≈2 900 authors

Colours: research groups; groups or departments; faculties. With more points, individual collaborations weigh less in the global stress and the organizational structure emerges.

Try it: interactive 3D explorer



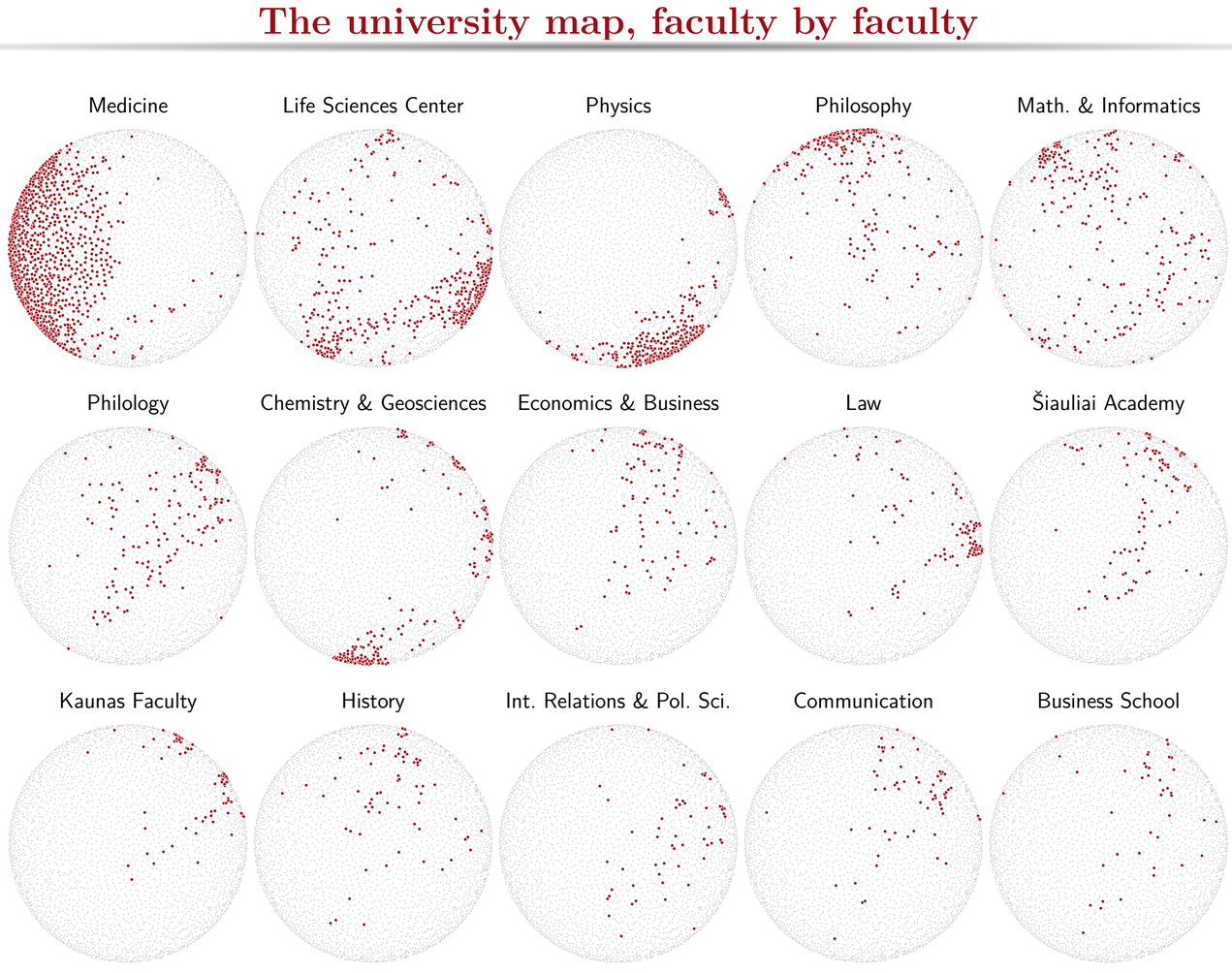
- Force-directed 3D layout in WebGL (Three.js): runs in any modern browser, no installation, no server-side code.
- Sphere volume ∝ publications; striped spheres for the 107 authors in several faculties.
- Filters, search and author cards: positions, publications, co-authors ranked by tie strength; views shareable by URL.

Scan & explore!

Opens the whole-university network with the Faculty of Mathematics and Informatics selected; tick more faculties in the right-hand panel.

ctrlon.github.io/is_final_prod_vu

The university map, faculty by faculty



2D SMACOF map of about 2 900 authors with at least one co-author at the university; each panel highlights one faculty-level unit (red) against all others (grey). Medicine, Life Sciences, Physics, Chemistry and Philosophy hold distinct regions at the rim, whereas most smaller humanities and social-science units scatter towards the centre.

Conclusions

1. A largely automated pipeline turns a national repository into a university-wide weighted co-authorship graph.
2. Collaboration largely follows faculty boundaries; large faculties contain smaller research communities.
3. The views are complementary: SMACOF maps for the global structure (deterministic, interpretable distances), the 3D explorer for individual researchers.

Next: temporal evolution, external collaborators and betweenness centrality to find bridging researchers.

Acknowledgements

This research was funded by the Research Council of Lithuania (LMTLT), agreement No S-SV-26-165.

References

[1] Crane, D. (1972). *Invisible Colleges: Diffusion of Knowledge in Scientific Communities.* Univ. of Chicago Press.
[2] Perianes-Rodriguez, A., Waltman, L., van Eck, N. J. (2016). Constructing bibliometric networks: a comparison between full and fractional counting. *J. Informetrics* 10(4), 1178–1195.
[3] Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *PNAS* 101(suppl. 1), 5200–5205.
[4] Watts, D. J., Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442.
[5] de Leeuw, J., Mair, P. (2009). Multidimensional scaling using majorization: SMACOF in R. *J. Stat. Softw.* 31(3), 1–30.
[6] Newman, M. E. J., Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev. E* 69, 026113.
[7] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech.* P10008.