Network analysis of conferences: Mapping the backbone of ESANN topics

Arya Nair and Jens Christian Claussen

University of Birmingham, School of Computer Science, Edgbaston, UK

Abstract. Academic conferences are central to knowledge dissemination and community formation. The *maps of science* approach [Boyack, Klavans and Börner 2005] introduced the framework of visualization of sets of text documents, based on distances calculated from (dis)similarity between the text documents. Can a similar framework be applied to conferences, to visualize subtopics and their fine structure and time development? Here, we focus on the abstracts of the ESANN conference series to trace topic and collaboration dynamics from 2010 to 2025, based on Topic Modelling of 2,000 abstracts used SPECTER Embeddings, DBSCAN/HDBSCAN Clustering, UMAP Visualisation, and Temporal Drift Analysis. A co-authorship network of 3,500 authors and 7,500 ties was examined through Centrality Measures, Clustering Coefficients, Louvain Community Detection, and Largest Connected Component Analysis. Our findings reveal continuity in Data Mining and Graph Learning, rapid growth in Deep Learning for Natural Language Processing and Medical Imaging, and a decline of Feature Selection and Spectral Clustering. The collaboration network shows a fragmented core-periphery structure reliant on a few hubs and brokers, reflecting both continuity and disruption in Machine Learning research.

Academic conferences are essential for knowledge sharing, idea exchange, and building research collaborations. They often serve as catalysts for emerging scientific trends, acting as focal points where researchers present novel findings, establish collaborations, and set the research agenda in their respective fields [1, 2, 3]. The study of conferences is therefore a valuable methodological lens for understanding how scientific domains evolve and how communities of practice interact within them. ESANN's longevity and focus make it an ideal case study for understanding how topics emerge, transform, or decline, and how collaborative communities form and persist.

Foundational studies [4] revealed that scientific collaboration networks typically exhibit small-world properties, high clustering, and skewed degree distributions, with a minority of highly connected individuals exerting disproportionate influence. Subsequent research has shown that such networks densify over time, reflecting the rise of interdisciplinary research and the expansion of team sizes [5, 6]. To analyse collaboration structures, community detection algorithms such as Louvain and its refinement Leiden [7] have become standard, uncovering modular groupings of authors that often correspond to disciplinary or institutional communities. Nevertheless, the reliance on co-authorship as a proxy for collaboration carries limitations. As [3] emphasise, collaboration does not always equate to intellectual influence or knowledge transfer. Moreover, most studies privilege large, multidisciplinary datasets, overlooking the conference scale where the bulk of machine learning research is disseminated [8].

In this study, approx. 2000 ESANN (2010–2025) proceedings PDFs were parsed, and text cleaning (tokenisation, stopword removal, lowercasing and punctuation stripping, name disambiguation) was applied to ensure consistency. Text documents were converted into embeddings (numerical vectors that capture semantic meaning) using Sentence-BERT to identify research themes, DBSCAN and HDBSCAN for clustering and UMAP for projection into 2D.

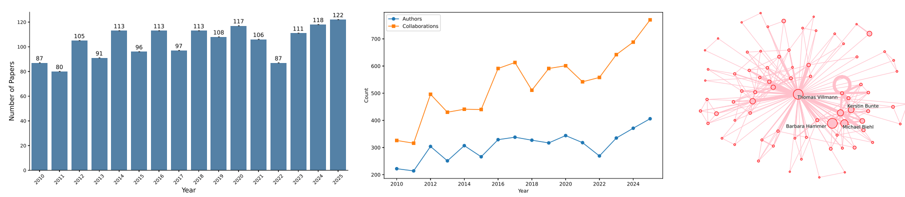

Fig. 1: Number of ESANN papers (left), authors (middle), and collaboration network of a frequent ESANN contributor (right).

Temporal Dynamics and Stability: Growth of Participation, Collaborations and Fragmentation. – Longitudinal analysis (Fig. 1) shows that both the number of authors and collaborations have increased steadily over time, with collaborations rising at a faster pace than author participation, indicating a shift towards collective research, consistent with global patterns [6]. This reflects the broader trend of science becoming more collective, where multi-institutional teams now dominate knowledge production [3, 6]. The steeper collaboration curve suggests research networks are becoming denser and more interconnected, echoing Newman's findings [9] on the evolving structure of scientific collaboration.

Fig. 2: ESANN Topic Evolution from 2010 to 2025

Topic Evolution and Transitions. – Figs. 2 and 3 present the longitudinal evolution of research topics at ESANN. Themes as Data Mining and Big Data Analytics and Graph Learning, demonstrate continuity, reflecting their foundational role in the field [10]. In contrast, earlier areas like Feature Selection and Spectral Methods show signs of decline. This confirms broader observations that traditional approaches, while once central, have gradually given way to more flexible and scalable techniques better suited for modern data environments [3]. Their shrinking presence suggests that such topics become methodological background rather than active frontiers of research. The most striking trend is the emergence of new, high-impact streams. Deep Learning for NLP and Medical Imaging rise sharply in recent years, reflecting the transformative influence of breakthroughs in neural architectures such as convolutional and transformer models [11]. These shifts illustrate how disruptive innovations diffuse across disciplines, spawning new communities while reshaping existing ones.

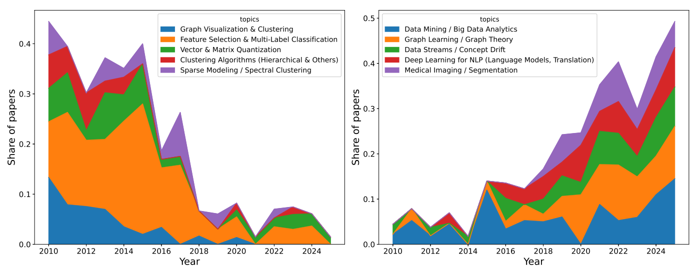

Fig. 3: Declining and growing Topics at ESANN

Topic-Based Community Labelling. – The collaboration network, visualised in Fig. 4 illustrates how research communities within the field have become spatially organised around distinct thematic clusters. Each coloured community corresponds to a cohesive research topic, highlighting both disciplinary silos and points of interdisciplinary convergence. For example, communities centred on Data Mining and Big Data Analytics and Graph Learning appear densely interconnected, reflecting the increasing overlap between large-scale data processing and graph-based approaches in recent years [11]. – In contrast, more peripheral clusters such as Feature Selection and Multi-Label Classification or Spectral Methods are smaller and less connected, signalling their gradual decline in influence [10]. Interestingly, emergent areas such as Medical Imaging and Generative Models bridge traditionally distinct communities, suggesting that methodological innovations in machine learning are diffusing into applied domains such as healthcare. This pattern reinforces long-standing findings in the sociology of science: collaboration networks tend to self-organise around intellectual traditions, while simultaneously enabling the diffusion of novel ideas across disciplinary boundaries [2, 4]. In particular, the prominence of neural-network-related clusters reflects the broader paradigm shift brought by deep learning, mirroring the exponential rise in publications in NLP and medical image analysis [12].

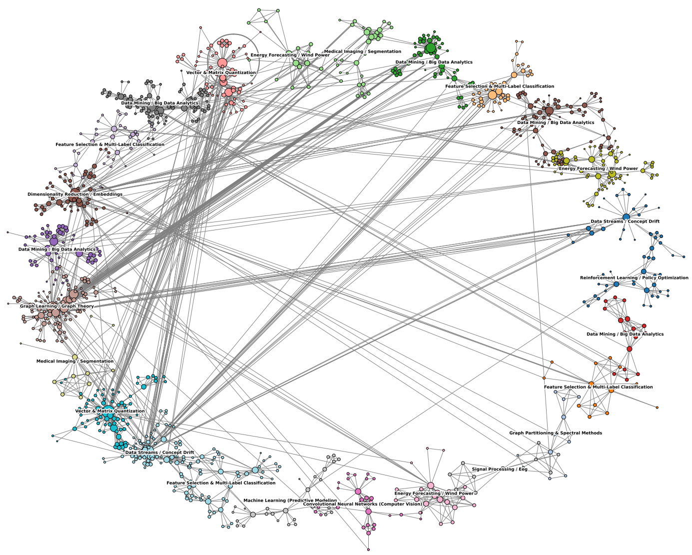

Fig. 4: Collaboration Network structured by communities

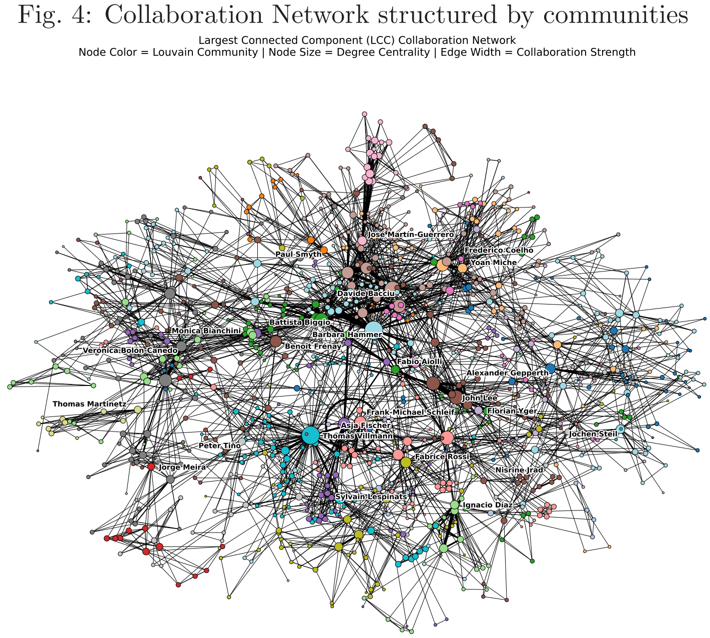

Fig. 5: Largest connected component (LCC) of the overall collaboration network

Role of Influential Authors and Brokers. – The Louvain algorithm partitioned the largest connected component (LCC) into a series of medium-sized communities, with the largest comprising 103 authors (Fig. 3). This modular arrangement indicates that the collaboration network at ESANN is pluralistic and decentralised, rather than dominated by a single cohesive group. Such a pattern aligns with the view of scientific collaboration as a constellation of overlapping research communities, each oriented around thematic or methodological foci [2, 4]. Within these clusters, certain individuals emerge as local hubs, occupying structurally central positions within their own communities. Figures such as Felipe França, Davide Bacciu, and Benoît Frénay appear as large nodes in the visualisation, signifying their role as leaders who concentrate collaboration ties and maintain cohesion locally. Their presence indicates uneven distribution of influence in scientific networks: a small subset of researchers sustains the intellectual and social fabric of each community [3, 13]. Equally important are bridge authors, including Barbara Hammer, Carlos Alzate, and Johan Suykens, sitting at boundaries between communities. Their positions correspond to high betweenness centrality, marking them as crucial brokers of knowledge flow across otherwise disconnected subgroups. Without such connectors, the LCC would fragment into isolated modules, as their ties are the thin yet essential threads binding together a broader research field. In addition to hubs and brokers, the network also features a long tail of peripheral authors, many of whom maintain only one or two collaborations. These actors represent either niche specialists, newcomers entering the field, or researchers whose participation is limited to isolated projects. Their presence highlights the heterogeneity of engagement: while some actors sustain central positions and long-term influence, others remain marginal to the wider collaborative [14, 15]. Taken together, the LCC displays a fragile core-periphery structure. Communities are robust internally due to dense local ties, but the global structure remains vulnerable.

Centrality analyses underscored the multi-dimensional influence of key figures. Barbara Hammer combined high degree, betweenness, and eigenvector centrality, marking her both as a hub and as a broker. Thomas Villmann's coauthor network highlights diverse collaborations with 71 distinct researchers (Fig. 1). Davide Bacciu showed high closeness centrality, reflecting efficient access across the network, while Luca Oneto and Davide Anguita ranked highly in eigenvector centrality, highlighting their embeddedness in influential clusters. These findings confirm that ESANN's cohesion relies heavily on a small group of hubs and brokers actors who sustain intra-community ties and connect distant clusters. However, this also makes the structure fragile: removing these brokers would fragment the largest connected component further.

Interpretation within Collaboration Theory. – The findings resonate with core–periphery models of scientific collaboration [16] where cohesive cores coexist with fragmented peripheries. The dominance of a few central authors reflects Price's principle [13] of cumulative advantage, in which established figures consolidate influence over time. However, the network lacks the hallmarks of a small-world structure [17]: clustering remained low and path lengths long, suggesting that ESANN encourages broad connectivity rather than tightly knit clusters. This diverges from Crane's "invisible college" hypothesis [14], depicting ESANN as a fabric of loosely connected groups, bridged by key individuals.

Conclusions. – Exploring the intellectual landscape of ESANN, our analysis shows a field that is both expanding and fragmenting. The results from topic modelling demonstrated that themes such as data mining and graph learning provide continuity, acting as the backbone of the conference, while more recent breakthroughs in deep learning for NLP and medical imaging have surged to prominence. At the same time, once-central areas like feature selection and spectral clustering have steadily declined, reflecting the shifting methodological priorities of the community. This suggests that ESANN is not just a venue for presenting technical results but a living social system that mirrors the broader dynamics of machine learning. It reveals how intellectual progress is sustained by both continuity and disruption: established methods provide stability, while disruptive innovations like deep learning reshape the landscape and drive new waves of collaboration. The ESANN community embodies the paradoxes of contemporary science: rapid growth alongside fragmentation, continuity alongside disruption, and dense local ties alongside fragile global connections. Our study highlights the importance of examining conferences as spaces where knowledge and collaboration co-evolve. It also opens pathways for future research that could extend the dataset to include other venues, track citations as well as co-authorships, and apply dynamic models to capture the unfolding process of community change. Understanding these dynamics is vital not only for evolution of a single conference but also for appreciating how research communities at large adapt, renew, and sustain themselves in the face of shifting paradigms.

References

[1] Boyack, K.W., Klavans, R., and Börner, K. (2005) Scientometrics 64, 351

[2] Börner, K., Sanyal, S. & Vespignani, A. (2010) Ann Rev Inf Sci Tech 41, 537–607

[3] Glänzel, W. & Schubert, A. (2005) In H.F. Moed, W. Glänzel & U. Schmoch (Eds.), Handbook of Quantitative Science and Technology Research (pp. 257–276), Springer

[4] Newman, M.E.J. (2001) Proc. Nat. Acad. Sci. 98, 404–409

[5] Barabási, A. L., et al. (2002) Physica A 311, 590–614

[6] Wagner, C.S. & Leydesdorff, L. (2005) Research Policy 34, 1608–1618

[7] Traag, V. A. , Waltman, L, van Eck, N. J (2019) Sci. Rep. 9, 5233

[8] Vrettas, G. & Sanderson, M. (2015) J Assoc Inf Sci Tech 66, 2674–2684

[9] Newman, M.E.J. (2010) Networks: An Introduction. Oxford University Press

[10] Blondel, V.D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. (2008) J Stat Mech P10008

[11] LeCun, Y., Bengio, Y. & Hinton, G. (2015) Nature 521, 436-444

[12] Grootendorst, M. (2022) arXiv:2203.05794

[13] Price, D.J. de S. (1963) Little science, big science. Columbia University Press

[14] Crane, D. (1972) Invisible colleges. University of Chicago Press

[15] Melin, G. (2000) Research Policy 29, 31-40

[16] Borgatti, S.P. & Everett, M.G. (2000) Social Networks 21, 375-395

[17] Watts, D.J. & Strogatz, S.H. (1998) Nature 393, 440-442