DOI:10.59787/2413-5488-2025-51-3-107-119

**[1]Andrii Biloshchytskyi, [1,2]Oleksandr Kuchanskyi, [1]Aidos Mukhatayev, [2]Yurii Andrashko, [1]Adil Faizullin, [1]Aidar Ispussinov**

*[1]Astana IT University, Kazakhstan*
*[2]Uzhhorod National University, Ukraine*

# FORMATION OF MULTI-UNIVERSITY SCIENTIFIC AND EDUCATIONAL COMMUNITIES BASED ON SCIENTOMETRIC ANALYSIS

**Abstract.** The article examines the problem of forming multi-university scientific and educational communities based on scientometric and network analysis methods. A formal model for identifying inter-university collaborations is proposed, which is based on constructing a metric space of scientific publications and further clustering using the Louvain algorithm. The model makes it possible to represent universities and researchers as nodes of a collaboration network, where the edge weights are determined by the number of co-authored publications. The developed approach was applied to data from the open OpenAlex database for the analysis of universities in Kazakhstan. The obtained results revealed the existence of a core of leading institutions (Al-Farabi Kazakh National University, Nazarbayev University, L. N. Gumilyov Eurasian National University, and Astana IT University) that demonstrate the highest intensity of collaboration, while other universities remain at the periphery of the research network. It was also established that national collaborations in Kazakhstan are significantly stronger than international ones, which highlights the need for further internationalization and integration into global scientific networks. The results obtained have both theoretical significance for the development of scientometrics and practical value for shaping science policy and designing strategies to improve the effectiveness of academic collaboration.

**Keywords:** multi-university scientific communities, scientometric analysis, network clustering, scientific collaboration network, citation network, international collaboration.

## Introduction

Improvement of scientific activity, development of innovation, stimulation of fundamental and applied research, and the promotion of the country’s scientific potential worldwide are possible through the formation of multi-university scientific and educational communities. At the same time, an important role is played by the development of models and methods for organizing scientific research aimed at establishing procedures to ensure the formation of a rational composition and structure of multi-university scientific and educational communities. The foundation of such formation should be the evaluation of objective performance indicators and the normative principles of integrating the nation’s science into the global scientific community.

The modern stage of scientific development is characterized by the globalization of knowledge and intensive scientific cooperation between educational and research institutions. Joint studies conducted across several universities make it possible to address complex scientific problems by combining the expertise and resources of different teams. Moreover, the inclusion of representatives from other universities fosters the creation of interdisciplinary projects. As research has shown, international scientific collaborations usually result in publications with higher scientific impact (measured by citation counts) compared to isolated national works. Multi-university scientific and educational communities are defined as stable networks of universities that collaborate in research and education on a regular basis. The development of such communities is particularly important for countries striving to strengthen their scientific potential and integrate into the global scientific space. In Kazakhstan, in particular, the effective use of scientific potential and its internationalization have been identified by the government as critical tasks for reinforcing the national scientific system. State programs (such as the "Kazakhstan-2050" strategy and the State Program for the Development of Education and Science 2020-2025) explicitly emphasize the necessity of enhancing the contribution of science to the country's socio-economic development through globalization of science and international collaboration.

Despite the obvious advantages, the formation of inter-university scientific communities faces several challenges. First, many traditional assessments of universities and researchers have not considered their network interactions, focusing only on internal institutional indicators (Kosztyán et al., 2021). However, institutions may underestimate the importance of collaboration development as a factor of success. Second, there are regional imbalances: global scientific interaction has historically been dominated by Western countries, while researchers from some other regions remain less integrated into global networks (Zhang et al., 2025). For post-Soviet states, a decline in mutual connections was observed after the collapse of the USSR, along with a simultaneous reorientation towards cooperation with Western countries. Kazakhstan, as part of the post-Soviet space, also faces the legacy of limited international integration.

Although over the past decade the country has undertaken numerous reforms in higher education and science (joining the Bologna Process, the "Bolashak" program for training young researchers abroad, the establishment of the new Nazarbayev research university, etc.), its contribution to global science remains moderate (Narbaev et al., 2025; Amirbekova et al., 2025). Kazakhstan's share in international publications is growing, but the country's visibility in global scientific networks is still low. Among the reasons are low R&D intensity (less than 0.2% of GDP devoted to science) and structural barriers hindering the development of scientific cooperation (Narbaev et al., 2025).

Another important stage of transformation in Kazakhstan's higher education has been its internationalization. Internationalization has influenced both publication activity and academic mobility. Kazakhstan has established a requirement for researchers to publish in journals indexed by Scopus and Web of Science, while other Central Asian countries, such as Turkmenistan and Tajikistan, did not have such requirements (Hladchenko & Moed, 2021). Access to international mobility through international grant projects and other funding sources has become an important dimension of post-Soviet changes in higher education, providing new opportunities for collaboration with international colleagues (Narbaev et al., 2025; Kuzhabekova et al., 2022). These changes were also enabled by the introduction of project and program management theory and practice in education (Nurtayeva et al., 2024).

In 2016, Kazakhstan ranked 85th among 218 countries in terms of the number of scientific publications and 108th in terms of citation counts (NASK, 2023). On average, Kazakhstani researchers published about 200 articles per year in Scopus-indexed journals, with an average of 4.7 citations per publication (Kuzhabekova & Ruby, 2018). The study (Biloshchytskyi et al., 2025) analyzed the activities of higher education institutions in the Republic of Kazakhstan, examining the principles of forming their information and educational environment and developing a ranking of universities using a proposed method based on data from the national institutional assessment system (IQAA).

Thus, the dissemination of knowledge is increasingly transcending geographical boundaries, significantly transforming models of scientific collaboration worldwide (Lin et al., 2023). The establishment of quality assurance systems in science and education is a key prerequisite for the effective development of higher education in a country (Biloshchytskyi et al., 2024). The relevance of this study lies in the need to develop formal approaches for building and optimizing multi-university scientific and educational communities, particularly with a focus on developing countries such as Kazakhstan.

The research problem is to formalize the network interaction between universities based on scientometric analysis, to identify existing communities (clusters) of universities, and to propose ways to optimize the structure of such a network in order to enhance the efficiency of scientific collaboration. In other words, it is necessary to answer the question: how can the dynamics of inter-university interaction be formally represented, and what measures can strengthen the integration of universities into the global scientific community? Solving this problem will have both theoretical significance and practical value for science policy (by supporting decision-making in fostering collaborations and reforming the scientific system).

**Methods and Materials**

Let us consider a set of subjects of scientific activity (universities and individual researchers), which are organized and united according to the criterion of a common focus of research topics and shared scientific interests. Such a space of scientific actors will be referred to as a scientific community. Since this space may include researchers affiliated with different universities as well as multiple universities themselves, such a scientific community is inter-university. The development of such an inter-university scientific community takes place through the publication of joint scientific articles, the establishment of collaborative research projects, and similar activities. The connections between researchers in such a community are defined through a scientific collaboration network or a citation network of scientific publications.

Let $W_s$ is a certain inter-university scientific community, $W_s \subset W$, where $W$ is a common educational and scientific space. Let $A = \{ a_1, a_2, \ldots, a_n \}$ is the set of all researchers, n – is the number of researchers. Let $Q = \{ q_1, q_2, \ldots, q_m \}$ is the set of scientific publications authored by researchers from set A, m is the number of scientific publications. The task consists in constructing such subject-specific spaces that include researchers working in similar scientific fields. To achieve this, the following steps are proposed: to define a metric space and establish distances between researchers' scientific publications; to perform cluster analysis of the publications and construct clusters of similar scientific works; to assign names to the clusters and establish correspondences between publications and cluster labels; and, finally, to build inter-university scientific communities based on these clusters, which will consist of researchers.

Let $W^s = \{ \eta_1, \eta_2, \ldots, \eta_r \}$ is the set of inter-university scientific communities. The membership of a researcher $a \in A$ in an inter-university scientific community $\eta \in W^s$ will be denoted by $\mu(a, \eta) \in [0,1]$. Here, the value 0 indicates that the researcher does not belong to the specified inter-university scientific community, while the value 1 means that all of the researcher's publications correspond to this field and the researcher belongs exclusively to this community, $\sum_{j=1}^{r} \mu(a, \eta_j) = 1$,

$$\mu(a, \eta) : A \times W^s \to [0,1]. \quad (1)$$

As a result of applying the method for identifying inter-university scientific communities, we obtain the distribution of each researcher's membership across different inter-university scientific communities. This is because many researchers, as well as all organizations as subjects of scientific activity, publish within several fields. At the very least, they have some publications that do not belong to their primary field. For example, a given researcher may publish the majority of their works in the field of scientometrics but also have several publications in the field of environmental monitoring. Importantly, the total membership of a researcher across all fields is equal to one.

Let us define a metric space for establishing the distance between the vertices of the citation graph of scientific publications $(Q, g)$, which consists of the set of scientific publications Q and a certain distance function g, defined for any pair of elements in this set. The distance g is defined as a mapping from the Cartesian square of the set Q onto the set of real numbers, i.e: g: $Q\times Q\to \mathbf{R}$, where **R** the set of real numbers. The distance between publications g is a non-negative real-valued function $g(q_i, q_j) \geq 0$, $i = \overline{1, m}$, $j = \overline{1, m}$, which is defined for $\forall\ q_i, q_j \in Q$. For function $g(q_i, q_j)$, $i \neq j$ the axioms of identity, symmetry, and the triangle inequality hold:

$g(q_i, q_j) = 0 \Leftrightarrow i = j$,

$g(q_i, q_j) = g(q_j, q_i)$, $\forall\ q_i, q_j \in Q$,

$g(q_i, q_j) \leq g(q_i, q_e) + g(q_e, q_j)$, $\forall q_i, q_j, q_e \in Q$, (2)

In this case, when $g(q_i, q_i) = 0$, $i = \overline{1, m}$, $j = \overline{1, m}$:

$$g(q_i, q_i) \leq g(q_i, q_j) + g(q_j, q_i) = 2g(q_i, q_j). \tag{3}$$

Let (Q,C) be a directed graph that represents the relationships between the scientific publications from the set Q and the citations from the set C, Q is the set of vertices of the graph, C is the set of arcs of the graph, i.e., pairs $(q_i, q_j)$, $q_i, q_j \in Q$, which represents the citation of a publication $q_j$ in publication $q_i$, $i \neq j$. We will represent such a graph in space by a set of points corresponding to the vertices of the graph and arcs connecting them.. Distance $g(q_i, q_j)$ between arbitrary vertices $q_i, q_j \in Q$ can be calculated as the length of the shortest path from vertex $q_i$ to vertex $q_j$, if it exists. A path between vertices $q_i$ and $q_j$, is defined as a sequence of vertices $q_i = q_{i_0}, q_{i_1}, q_{i_2}, \ldots, q_{i_k} = q_j$, $i_0 < i_1 < i_2 < \ldots < i_k$, $k \in \mathbf{N}$, each of which, except for the last one, is connected to the next by an arc $c_{y,y+1}$, $y = \overline{0, k-1}$. For example, two vertices $q_{i_0}, q_{i_1}$ are connected by an arc $c_{0,1}$, the vertices $q_{i_1}, q_{i_2}$ are connected by an arc $c_{1,2}$ and so on. The length of a path is the number of arcs in this path. The shortest path is the one with the smallest length. The weights of the arcs are equal to 1, i.e., the weight function of the graph is defined as (Q,C): $f: C \to \{1\}$ and the length of the path from vertex $q_i$ to vertex $q_j$ is defined as follows:

$$r(q_i, q_j) = \sum_{i=0}^{k-1} f(c_{y,y+1}). \tag{4}$$

If there is no path between vertices $q_i$ and $q_j$ then the length will be considered arbitrarily large, for example $r(q_i, q_j) = m + 1$. The shortest path is found using the Bellman–Ford algorithm with a complexity of $O(QC)$ or Dijkstra's algorithm with a complexity of $O(Q^2)$.

Let a metric space be given (Q,g). It is necessary to partition the set of publications Q into a certain number of subsets that do not intersect with each other. Such subsets of the set Q are called clusters. Let us denote the set of clusters by $Y=\{y_1,y_2,\ldots,y_z\}$, where z is the number of clusters into which the set P is partitioned. To define the set of clusters Y, the following conditions must be satisfied: each publication necessarily belongs to one of the clusters, i.e. $\bigcup_{i=1}^{z} y_i = P$, each publication belongs to a unique cluster, i.e. $y_i \cap y_j = \varnothing,\ \forall i \neq j$, each cluster combines publications that are sufficiently close (in terms of the distance g).

It should be emphasized that when calculating the distances between publications using the approaches described above, so-called isolated publications may appear. An isolated publication is understood as one for which the distance to any other publication from the set P is equal to infinity. According to the definition of a cluster, no other publication can be assigned to the cluster that contains an isolated publication. As a result, clusters consisting of only a single work are formed. The presence of such clusters complicates further analysis and provides no meaningful information about the scientific domains to which these publications belong. Therefore, before performing the clustering procedure, it is advisable to exclude isolated publications from consideration.

For the clustering of such graphs, it is proposed to use the Louvain method (Blondel, 2008), which is based on the maximization of graph modularity. Modularity is defined as the sum of the differences between the fraction of edges within a given subgraph and the square of the fraction of edges with one endpoint belonging to that subgraph. That is, the modularity of the clustering of the publication graph can be defined as follows:

$$\overline{Q} = \sum_{v=1}^{z} (\beta_v - \alpha_v^2), \qquad (5)$$

where $\alpha_v$ – the fraction of citations where either the citing publication or the cited publication belongs to the cluster $y_v$, that is

$$\alpha_v = \frac{\left\| \left\{ q_i \in Q \middle| (q_i, q_j) \in C, q_i \in y_v, q_j \in Q \right\} \cup \left\{ q_i \in Q \middle| (q_j, q_i) \in C, q_i \in y_v, q_j \in Q \right\} \right\|}{\operatorname{card}(C)}, \qquad (6)$$

where $\beta_v$ is the fraction of citations where both the citing publication and the cited publication belong to the cluster $y_v$, that is

$$\beta_v = \frac{\left\| \left\{ q_i \in Q \middle| (q_i, q_j) \in C, q_i \in y_v, q_j \in y_v \right\} \right\|}{\operatorname{card}(C)}, \qquad (7)$$

where card(C) is the number of edges of the graph (Q,C).

We assume that the initial graph (Q,C) has been clustered, for example using the Louvain method, and an initial partition of the set of publications Q into z clusters has been obtained. As a result of the clustering procedure of scientific publications, a set of clusters Y is formed. The cardinality of the set Y may be quite large, which complicates further analysis. One way to address this problem is to aggregate the constructed clusters by merging those that are close to each other and contain a small number of elements. To achieve this, it is necessary to determine the centroid of each constructed cluster.

The centroid of a cluster

$$y_k = \left\{ q_1^k, q_2^k, \ldots, q_{\mu_k}^k \right\},\ k = \overline{1, z}, \qquad (8)$$

is defined as the object within the cluster $\Omega^k$, or which the total distance to all other objects in the cluster is minimal:

$$\Omega^k = \arg\min \left( \sum_{i=1}^{\mu_k} g(q_i^k, q_j^k),\ j = \overline{1, \mu_k} \right), \qquad (9)$$

where $\mu_k$=card($y_k$) is the number of objects that belong to the cluster $y_k$, $k = \overline{1, z}$.

The algorithm for implementing such cluster merging consists of the following steps:

1. We initialize a counter b=0.

2. We find the centroids $\Omega^k$ of each cluster $y_k = \{q_1^k, q_2^k, \ldots, q_{\mu_k}^k\}$.

3. We calculate the distances between the centroids of each cluster by comparing the annotations of the publications corresponding to the centroids.

4. If there exist clusters $y_k$ and $y_l$ such that the distance between their centroids does not exceed the threshold value δ, i.e., the condition $g(\Omega_k, \Omega_l) \le \delta$, then

4.1. We increase the counter by one b=b+1.

4.2. We form a new cluster $y_{z+b} = y_k \cup y_l$.

4.3. Clusters $y_k$ and $y_l$, are removed from further consideration.

4.4. We find the centroid of the cluster $\Omega_{z+b.}$

4.5. We calculate the distances from the centroid of the cluster $y_{z+b}$ to the centroids of the other clusters $y_1$, $y_2$, …,$y_{z+b-1}$.

4.6. Return to Step 4.

5. If the distance between the centroids of the clusters exceeds the threshold value for all clusters, i.e., there do not exist clusters $y_k$ and $y_l$ for which the condition $g(\Omega_k, \Omega_l) \le \delta$, then the execution of the algorithm is terminated.

Thus, it is possible to construct an inter-university scientific community that unites researchers based on the commonality of their scientific activities. Such a community is represented in the form of a scientific collaboration network. The formal models of the metric space and clustering described above were implemented in the Gephi 0.10.1 software environment (Gephi) using the Louvain algorithm for analyzing the network of co-authored publications obtained from the OpenAlex database (OpenAlex). In this way, the theoretical approach to clustering scientific publications found practical application in constructing the collaboration graph of universities and researchers in Kazakhstan.

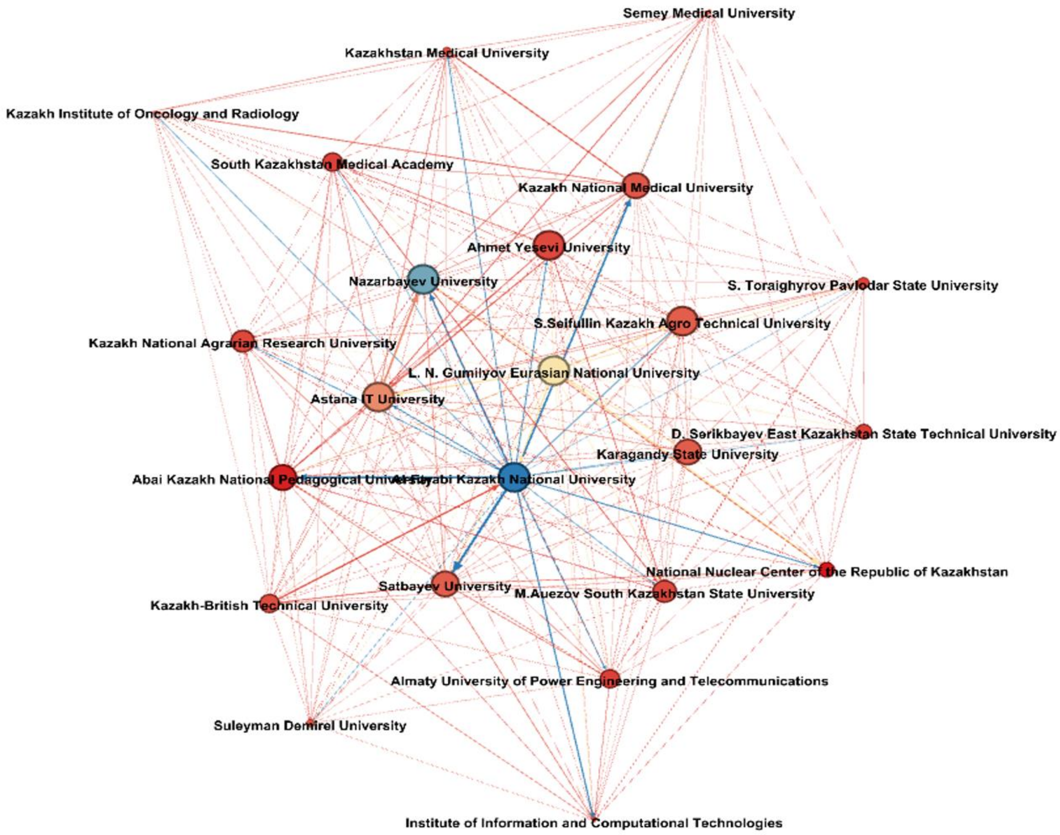
**Results**

The proposed model of a metric space of scientific publications, combined with the Louvain algorithm, made it possible to move from theoretical description to practical implementation of the analysis of inter-university scientific communities. Based on open data from the OpenAlex database, a network of scientific interactions was constructed, where publications were treated as elements of clustering and their links as indicators of scientific collaboration. The implementation of this model in the Gephi software environment enabled not only the formal grouping by scientific domains but also the visualization of actual cooperation structures between universities and individual researchers. The constructed network became the basis for further quantitative and qualitative analysis of the intensity of scientific interaction. According to the proposed model, the vertices of the graph represent universities, while the edge weights are determined by the number of joint publications. This approach allows for a direct interpretation of the formalized community structure based on real data. Within the study of inter-university scientific interaction, data from the open OpenAlex database were used, which aggregates information about scientific publications, authors, organizations, and their interrelations. To obtain relevant information, a sample of records related to universities was extracted, forming a dataset that reflects cooperation at the level of article co-authorship. Particular attention was paid to the identification of affiliations: for each publication, the institution of the authors was verified, which made it possible to establish inter-institutional links.

The collected data were transformed into a network structure, where the nodes represent universities and the edges correspond to instances of joint scientific activity, in particular, co-authored publications. The edge weights were defined proportionally to the number of such joint works, which makes it possible to assess the intensity of interaction between individual institutions. Based on this information, in September 2025, a graph was constructed that reflects the current state of scientific cooperation between universities (Figure 1). For the visualization and analysis of the network, the Gephi software environment was applied, which provided the opportunity to use algorithms for community detection, centrality calculation, and other structural characteristics of the graph. The visual representation made it possible to identify a core of universities with the highest intensity of collaboration, as well as peripheral institutions involved in a smaller number of international projects.

**Figure 1.**

*Graph of inter-university scientific collaboration for universities of the Republic of Kazakhstan*




For the quantitative representation of the results of network construction, an adjacency matrix was created, in which each university is considered as a vertex of the graph, and the values in the cells define the strength of interaction between pairs of institutions. The weight of an edge is interpreted as the number of jointly published scientific works, which makes it possible to compare the level of cooperation among universities. Thus, the matrix reflects not only the fact of cooperation but also the intensity of scientific links. The analysis of the obtained data revealed a clear regional specificity. In particular, universities of the Republic of Kazakhstan, in the vast majority of cases, collaborate primarily with each other, forming a dense internal scientific environment. At the same time, the number of joint publications with foreign institutions is relatively small, indicating weaker integration into international research networks. This is confirmed both by the numerical values in the matrix and by the visual structure of the graph constructed in Gephi, where the most intensive links are formed among domestic universities.

**Table 1.**
*Assessment of the level of inter-university scientific collaboration*

| OpenAlex ID | Label | Weighted degree |
|---|---|---|
| I185571130 | Al-Farabi Kazakh National University | 20180 |
| I60559429 | Nazarbayev University | 17444 |
| I10232997 | L. N. Gumilyov Eurasian National University | 11145 |
| I4210141757 | Astana IT University | 8012 |
| I204275683 | Satbayev University | 6425 |
| I191378831 | Kazakh National Medical University | 6196 |
| I4210120897 | Suleyman Demirel University | 6188 |
| I4210128045 | Abai Kazakh National Pedagogical University | 4298 |
| I4210132247 | National Nuclear Center of the Republic of Kazakhstan | 3992 |
| I4210124329 | S.Seifullin Kazakh Agro Technical University | 3030 |
| /I2800164252 | Karagandy State University | 3014 |
| I171571821 | Kazakh-British Technical University | 2386 |
| I4210128576 | Kazakhstan Medical University | 2206 |
| I4210121010 | Kazakh Institute of Oncology and Radiology | 2172 |
| I2801707353 | Ahmet Yesevi University | 2142 |
| I4210151108 | M.Auezov South Kazakhstan State University | 2138 |
| I4210128817 | Kazakh National Agrarian Research University | 2079 |
| I4210101268 | S. Toraighyrov Pavlodar State University | 1852 |
| I4210117112 | South Kazakhstan Medical Academy | 1795 |
| I247491477 | Almaty University of Power Engineering and Telecommunications | 1779 |
| I4210135055 | Institute of Information and Computational Technologies | 1747 |
| I4210117727 | D. Serikbayev East Kazakhstan State Technical University | 1662 |
| I4210135680 | Semey Medical University | 1601 |
| I4210151480 | Institute of Mathematics and Mathematical Modeling | 1509 |
| I4210131705 | Abylkas Saginov Karaganda Technical University | 1500 |
| I4210143225 | Ministry of Education and Science of the Republic of Kazakhstan | 1491 |
| I4210159690 | Kostanay State University A Baitursynov | 1439 |
| I4210106095 | Aktobe Regional State University named after K.Zhubanov | 1389 |
| I4210114545 | Almaty Technological University | 1363 |
| I4210143508 | International Information Technologies University | 1363 |
| I4210129744 | Korkyt Ata Kyzylorda State University | 1313 |
| I4210108340 | University of International Business | 1271 |
| I4210159258 | Shakarim University | 1267 |
| I175124709 | Süleyman Demirel University | 1266 |
| I4210130920 | West Kazakhstan Marat Ospanov State Medical University | 1239 |
| I3130577743 | Narxoz University | 1213 |
| I4210144657 | Global Health Research Center of Central Asia | 1183 |
| I4210119570 | Kokshetau State University | 1180 |
| I4210111563 | National Academy of Sciences of the Republic of Kazakhstan | 1170 |
| I4210099401 | Institute of Plant Biology and Biotechnology | 1149 |
| I4210134323 | Maqsut Narikbayev University | 1141 |
| I4210101819 | M. Kh Dulati Taraz State University | 1140 |
| I4210096915 | Sarsen Amanzholov East Kazakhstan University | 1115 |

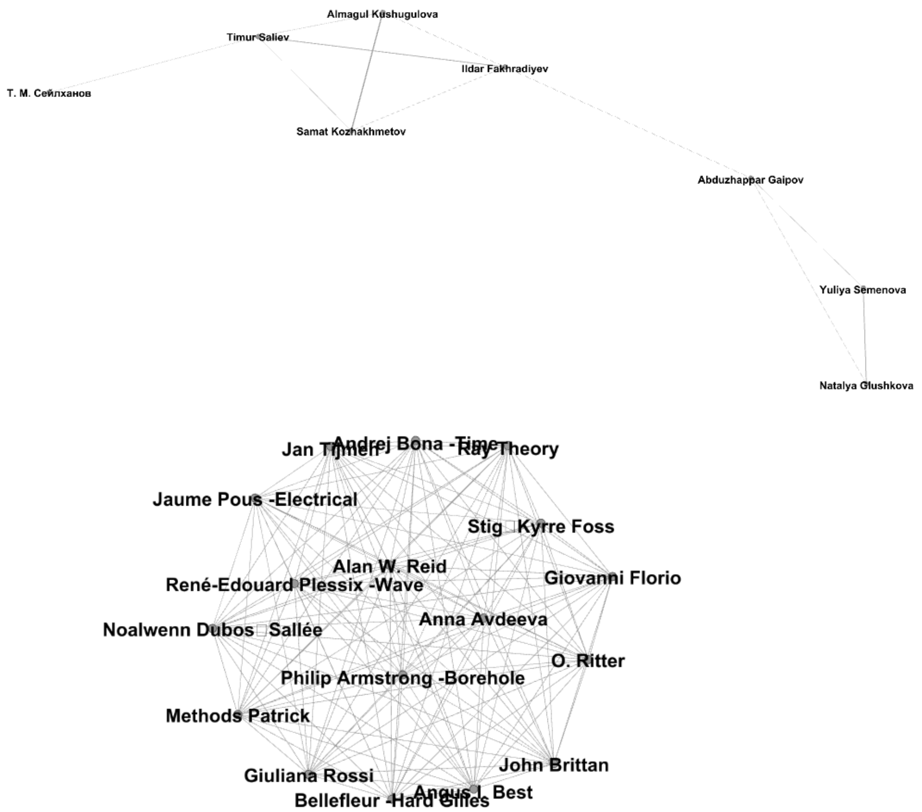
| | | |
|---|---|---|
| I4210115547 | Karaganda Medical University | 1104 |
| I3130777500 | Turan University | 1098 |
| I4210101002 | National Center for Biotechnology | 1093 |
| I4210157209 | Caspian University | 1054 |
| I4210158929 | National Research Center for Maternal and Child Health | 1042 |
| I2801153214 | Fesenkov Astrophysical Institute | 1002 |

A similar approach was applied to construct a collaboration network among individual researchers affiliated with universities and research institutions of the Republic of Kazakhstan. In total, the sample included 67,907 researchers, for each of whom the presence or absence of co-authored publications with colleagues was recorded. As a result, a graph was constructed in which the vertices represent researchers, and the edges reflect instances of co-authorship. The weight of the edges was determined by the number of joint scientific works, allowing for a quantitative assessment of the intensity of individual scientific interaction.

The analysis of the structure of this graph revealed the existence of several clearly defined clusters characterized by a high density of internal connections. Figure 2 presents two clusters with components where the strength of collaboration exceeds the threshold value of 200 joint publications, which made it possible to focus on the most significant research communities. It was found that these clusters largely correlate with specific scientific fields in which co-authorship is most intensive.

**Figure 2.**

*Collaboration graph among individual researchers of the Republic of Kazakhstan, where the strength of collaboration exceeds the threshold value of 200 joint scientific publications.*



The analysis also revealed an interesting anomaly in the structure of the network. Among the clusters, one stands out clearly, demonstrating exceptionally close collaboration between Kazakhstani universities and foreign researchers in the field of geology. In contrast, a large number of smaller clusters are formed predominantly by researchers from Kazakhstan and reflect specific national research directions in medicine, information technology, and other applied fields. Such a structure indicates the simultaneous existence of strong international ties in selected areas and fragmentation of domestic scientific communities in several disciplines. Thus, the collaboration network at the level of individual researchers demonstrates a high degree of internal coherence within specialized scientific fields and complements the picture obtained from the analysis of inter-university interactions. It clearly shows that the intensity of cooperation is largely determined by researchers' affiliation with scientific schools and fields of knowledge.

**Discussions**

The formation of inter-university communities is essential for ensuring the synergy of scientific resources and competencies. No single university or individual researcher possesses the full spectrum of knowledge and expertise required to solve complex interdisciplinary problems. Such an integration of researchers and universities makes it possible to combine different schools and methodologies, enhance the quality and novelty of research, and avoid duplication of scientific results and efforts. An important outcome of establishing international scientific communities is the increased global competitiveness of universities, the growth in the number of joint publications in high-ranking journals, the rise in citation levels, and higher impact indices. This, in turn, makes universities and researchers more visible on the global scientific and educational stage. Another effect of forming and identifying inter-university scientific communities is knowledge exchange and the preparation of highly qualified personnel, fostering the mobility of students and faculty, increasing the likelihood of funding for new projects and programs, and ensuring the sustainability of collaboration. All these factors allow for a focus on creating scientific trends and long-term development strategies. Thus, inter-university scientific communities serve as a mechanism for integration into the global scientific space, enhancing research quality, creating conditions for innovation, and ensuring the sustainable development of both individual universities and national scientific systems as a whole.

The obtained results confirm that the formation of inter-university scientific communities is a key mechanism for integrating national science into the global scientific space. The analysis of the university collaboration network in Kazakhstan revealed the existence of a core of institutions with a high level of joint publications (Al-Farabi Kazakh National University, Nazarbayev University, L. N. Gumilyov Eurasian National University, and Astana IT University), while a number of universities occupy a peripheral position and are characterized by weaker integration. Such a structure is also typical for other developing countries, where scientific activity tends to be concentrated in a limited number of leading centers.

A comparison with international experience shows that enhancing the effectiveness of multi-university communities requires systemic measures to support interdisciplinary projects, international grant programs, and the promotion of researcher mobility. For example, in the Baltic countries, integration into European research networks has significantly increased the visibility of universities in international databases. For Kazakhstan, an important task is to expand cooperation beyond the internal system, since the current network structure indicates the predominance of national collaborations over international ones.

It is also important to emphasize the methodological aspect: the use of scientometric approaches and clustering algorithms (in particular, the Louvain method) made it possible to identify hidden patterns in the structure of collaboration that are difficult to detect using traditional statistical methods. At the same time, a limitation is the dependence on OpenAlex data, which do not always fully reflect local journals and conference proceedings. This opens up prospects for further improvement of the model by combining multiple data sources and integrating alternative metrics of scientific impact.

**Conclusion**

In this study, a formal model was developed for the identification and analysis of inter-university scientific communities based on the network structure of co-authored publications. The application of this model to OpenAlex data made it possible to construct a map of scientific interactions among Kazakhstani universities and to identify key centers of scientific activity. It was found that the highest intensity of collaboration is demonstrated by the country's leading universities, while international integration remains limited.

The practical value of the obtained results lies in the possibility of applying network analysis methods for the development of science policy aimed at supporting strategic partnerships, enhancing the competitiveness of universities, and increasing their visibility in the global academic space.

Among the limitations of this study, it is important to note the use of only one database (OpenAlex) and the focus on universities in Kazakhstan, which reduces the possibility of direct comparison with other countries. Future research should be directed toward:

- expanding the analysis to the level of thematic domains and individual scientific schools;
- assessing the dynamics of community formation over time;
- modeling scenarios of international integration, taking into account potential growth in science funding and participation in international projects.

Thus, the proposed approach demonstrates its potential as a tool for strategic planning in higher education and science, allowing not only the identification of current research networks but also the design of strategies for their further development.

**Conflict of Interest Statement**

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

**Funding Information**

This paper was written in the framework of the state order to implement the science program according to the budget program 217 "Development of Science", IRN No. AP19678627 with the topic: "Development of the information technology for the formation of multiuniversity scientific and educational communities based on the scientometric analysis theory".

**Author Contributions**

Andrii Biloshchytskyi – Conceptualization, Resources, Supervision; Oleksandr Kuchanskyi – Writing, Original draft preparation; Aidos Mukhatayev – Methodology, Writing-Reviewing; Yurii Andrashko – Editing, Writing-Reviewing; Adil Faizullin – Investigation, Project administration, Formal analysis; Aidar Ispussinov – Visualization, Investigation.

## References

Kosztyán, Z. T., Fehérvölgyi, B., Csizmadia, T., & Kerekes, K. (2021). Investigating collaborative and mobility networks: Reflections on the core missions of universities. Scientometrics, 126, 3551–3564. https://doi.org/10.1007/s11192-021-03865-7

Zhang, L., Gong, J., Yang, Y., & Li, J. (2025). Cities in global and sustainable scientific collaboration network: Unveiling cooperation centers. Humanities and Social Sciences Communications, 12, 1406. https://doi.org/10.1057/s41599-025-05667-1

Narbaev, T., Amirbekova, D., & Bakdaulet, A. (2025). A decade of transformation in higher education and science in Kazakhstan: A literature and scientometric review of national projects and research trends. Publications, 13(3), 35. https://doi.org/10.3390/publications13030035

Amirbekova, D., Narbaev, T., & Kussaiyn, M. (2022). The research environment in a developing economy: Reforms, patterns, and challenges in Kazakhstan. Publications, 10(4), 37. https://doi.org/10.3390/publications10040037

Hladchenko, M., & Moed, H. F. (2021). The effect of publication traditions and requirements in research assessment and funding policies upon the use of national journals in 28 post-socialist countries. Journal of Informetrics, 15(4), 101190. https://doi.org/10.1016/j.joi.2021.101190

Kuzhabekova, A., Ispambetova, B., Baigazina, A., & Sparks, J. (2022). A critical perspective on short-term international mobility of faculty: An experience from Kazakhstan. Journal of Studies in International Education, 26(4), 454–471. https://doi.org/10.1177/10283153211023613

Nurtayeva, D., Kredina, A., Kireyeva, A., Satybaldin, A., & Ainakul, N. (2024). The role of digital technologies in higher education institutions: The case of Kazakhstan. Problems and Perspectives in Management, 22(1), 562–577. https://doi.org/10.21511/ppm.22(1).2024.44

National Academy of Sciences of Kazakhstan. (2023). National report on science. Government of Kazakhstan. https://www.gov.kz/uploads/2023/11/17/7ce8a403d55b7f2b1cb4ab194337832f_original.5055287.pdf

Kuzhabekova, A., & Ruby, A. (2018). Raising research productivity in a post-Soviet higher education system: A case from Central Asia. European Education, 50(3), 266–282. https://doi.org/10.1080/10564934.2017.1348445

Biloshchytskyi, A., Makhatayev, A., Kuchanskyi, A., Omirbayev, S., Andrachko, Y., & Biloshchytska, S. (2025). Method for assessing quality assurance in higher education institutions. TEM Journal. 14 (1), 372-386. https://doi.org/10.18421/TEM141-33

Lin, Y., Frey, C. B., & Wu, L. (2023). Remote collaboration fuses fewer breakthrough ideas. Nature, 623(7989), 987–991. https://doi.org/10.1038/s41586-023-06767-1

Biloshchytskyi, A., Omirbayev, S., Mukhatayev, A., Kuchanskyi, O., Hlebena, M., Andrashko, Y., Mussabayev, N., & Faizullin, A. (2024). Structural models of forming an integrated information and educational system “quality management of higher and postgraduate education”. Front. Educ. 9:1291831. doi: 10.3389/feduc.2024.1291831

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Gephi. The open graph visualization platform. https://gephi.org/

OpenAlex. The open catalog of the global research system. https://openalex.org/

**Information about authors:**

***Andrii Biloshchytskyi*** - Doctor of Technical Sciences, Professor, Vice-Rector for Science and Innovation, Astana IT University, Kazakhstan, email: a.b@astanait.edu.kz, ORCID 0000-0001-9548-1959

***Oleksandr Kuchanskyi*** - Doctor of Technical Sciences, Professor, School of Artificial Intelligence and Data Science, Astana IT University, Kazakhstan, email: kuczanski@gmail.com, ORCID 0000-0003-1277-8031

***Aidos Mukhatayev*** **-** Candidate of Pedagogical Sciences, Professor, School of General Education Disciplines Astana IT University, Kazakhstan, email: mukhatayev.aidos@gmail.com, ORCID 0000-0002-8667-3200

***Yurii Andrashko*** - Candidate of Technical Sciences, Associate Professor of Department of System Analysis and Optimization Theory, Uzhhorod National University, Ukraine, email: yurii.andrashko@uzhnu.edu.u, ORCID 0000-0003-2306-8377

***Adil Faizullin*** - PhD, Department of Quality Assurance, Astana IT University, Kazakhstan, email: faizullin.adil@gmail.com, ORCID 0000-0001-5644-9841 ***(corresponding author)***

***Aidar Ispussinov*** **-** Vice-Rector for Digitalization, Astana IT University, Kazakhstan, email: a.manas@astanait.edu.kz, ORCID 0009-0001-4134-776X.