Expert competence assessment via a semantic-statistical model of the target subject domain built on OpenAlex concepts*

Oksana Mulesa[1,2,†], Iryna Domanetska[3,†], Yaroslav Khrolenko[3,*,†], Olha Tsyhanok[4,†], Olena Trubitsyna[5,†]

[1] University of Prešov, Prešov, Slovakia
[2] Uzhhorod National University, Uzhhorod, Ukraine
[3] Taras Shevchenko National University of Kyiv, Volodymyrs'ka str. 64/13, Kyiv, 01601, Ukraine
[4] State University of Trade and Economics, Kyoto 19, 02156, Kyiv, Ukraine
[5] Kyiv National Linguistic University, Velyka Vasylkivska 73, 03680, Kyiv, Ukraine

Abstract

This paper proposes a semantic-statistical approach to the quantitative assessment of researchers' competence based on the OpenAlex open knowledge graph. Unlike traditional bibliometric methods, which primarily rely on citation-based and productivity indicators, the proposed model incorporates the semantic structure of the subject domain and the degree of proximity between an expert's research profile and its conceptual core.

The expert evaluation method is based on matching the set of concepts extracted from an expert's publications with a multilayer model of the subject domain and computing an integral competence score as a weighted sum of matches across different levels of semantic distance from the core. The weighting coefficients reduce the influence of peripheral and interdisciplinary matches, thereby increasing the sensitivity of the metric to the conceptual nucleus of the domain.

The approach was validated using the subject domain of "Sentiment Analysis" and real-world data from OpenAlex. The experimental results demonstrate the model's ability to clearly rank candidates according to their thematic relevance while ensuring interpretability of the assessment through the analysis of contributions from different semantic layers. It is shown that the proposed integral indicator more adequately captures both the depth and the breadth of expertise compared to conventional scientometric metrics. The findings confirm the potential of open knowledge graphs as a foundation for building reproducible and scalable systems for scientific competence assessment.

Keywords

competence assessment; semantic-statistical model; OpenAlex concepts; expert ranking; bibliometric indicators

1. Introduction

The growing volume of scientific literature and the increasing complexity of interdisciplinary connections complicate the process of identifying qualified experts for peer review, participation in research projects, and the formation of scientific teams. Various methods for obtaining and processing expert assessments [1–3] take into account the relative competence of specialists when conducting group expert evaluations and deriving a consensus-based, generalized, aggregated assessment. The need to account for competence in small-group expert evaluations has been confirmed experimentally [4]. Traditional approaches based on simple scientometric indicators or expert surveys prove to be insufficiently effective due to their subjectivity, static nature, and

*ITS-2025: Information Technologies and Security, December 11, 2025, Kyiv, Ukraine
* Corresponding author.
† These authors contributed equally.
oksana.mulesa@unipo.sk (O. Mulesa); domanetska@knu.ua (I. Domanetska); yaroslav.khrolenko@knu.ua (Ya. Khrolenko); olzyg@ukr.net (O.Tsyhanok); elenatrubitsyna88@gmail.com (O.Trubitsyna)
0000-0002-6117-5846 (O. Mulesa); 0000-0002-8629-9933 (I. Domanetska); 0009-0004-0641-827X (Ya. Khrolenko); 0000-0002-0009-6562 (O.Tsyhanok); 0009-0003-0539-5718 (O.Trubitsyna)
© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR-WS.org/Vol-4337/paper15.pdf
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-11-26

limited capacity to account for the multidimensional character of scientific activity. Modern open databases provide new opportunities for analyzing the scientific landscape, as they contain not only publication metadata but also links among authors, institutions, and thematic concepts.

A promising research direction is the application of semantic tools provided by scientific publication aggregators for the assessment of expert competence.

The aim of this study is to develop a semantic-statistical model of the target subject domain of expertise based on OpenAlex concepts for the quantitative assessment of researchers' competence. Achieving this objective involves addressing the following tasks: analyzing contemporary methods for expert identification, examining existing approaches to subject domain modeling, substantiating the selection of OpenAlex as the primary data source, and developing a formal model of the subject domain of expertise.

2. Literature review and theoretical foundations matter

2.1. Scientific information aggregators

Traditional methods of assessing competence are based primarily on bibliometric indicators, which allow for a quantitative assessment of scientific productivity and the influence of experts. Classic metrics include the h-index [5], total number of citations, number of publications in high-ranking journals, and journal impact factor. For example, study [6] compares informed peer review with bibliometrics, showing that for the natural sciences, bibliometric methods exceed peer review in terms of objectivity and effectiveness. Similarly, [7] proposes using citation databases to verify the qualifications of experts in court cases by analyzing the publications and citations of British court experts.

Bibliometric methods are often used to select experts for scientific evaluations. Khorasani identifies the competencies of scientometrics specialists, including citation analysis and co-authorship skills [8]. Similarly, [9] analyzes professional standards for bibliometric research evaluation in Europe, emphasizing the use of metrics to assess productivity. However, these methods have limitations: they ignore thematic relevance, the dynamics of scientific domains, and interdisciplinary aspects [10-12]. For example, the h-index does not take into account the semantic proximity of works to a specific topic of expertise.

Expert surveys and peer reviews remain subjective alternatives. In a studies [13, 14] a model of competence is proposed as a continuum from dispositions (knowledge, beliefs) to situational skills and performance. This model is tested through surveys, but suffers from bias and limited scalability.

2.2. Thematic modeling and co-authorship network analysis

The problem of identifying and assessing the competence of experts in a particular subject area is becoming particularly relevant in the context of the rapid growth of scientific information and the increasing complexity of interdisciplinary connections. In response to these challenges, two main directions of automated expert search have emerged: methods based on thematic modeling and methods of co-authorship network analysis.

Thematic modeling methods belong to the class of unsupervised machine learning algorithms designed to identify hidden (latent) thematic structures in large arrays of text documents. In the context of expert analysis, these methods are used to build a researcher's thematic profile based on an analysis of their publications. The fundamental model that became the starting point for most subsequent research is Latent Dirichlet Allocation (LDA) [15], which represents each document as a random mixture of hidden topics, where each topic is characterized by a distribution of words.

However, classic LDA has significant limitations when working with short texts, where word distribution is too sparse for accurate detection of topic mixtures. This has stimulated the development of modified versions of the model, in particular GSDMM (Gibbs Sampling Dirichlet Multinomial Mixture) [16], which assumes that a single short document relates to only one topic. This approach significantly improves the quality of short text clustering and makes it suitable for analyzing a variety of sources where potential experts can express themselves not only in scientific publications.

The second main direction is represented by methods based on the analysis of co-authorship networks. This approach considers the scientific community as a social graph, where the vertices are researchers and the edges are the connections between them, most often joint publications. The fundamental idea is that expertise is not an isolated property of an individual, but can spread through a network through collaboration. Semantics can refine the level of detail at which the expertise of different researchers is determined by explicitly expressing the connections between topics and different subtopics using taxonomy. Such "topic-subtopic" links allow expertise to be extrapolated based on the assumption that expertise in subtopics also indicates expertise in the topic itself.

The realization that not all links in a network are equally significant for all topics has led to the development of hybrid models that integrate semantic analysis with network analysis. The SEMEF (Semantic Expertise Extrapolation Framework) approach [17] uses subject area taxonomy to refine an expert's profile, allowing their expertise to be extrapolated to broader categories based on their publications in narrower subtopics. This makes it possible to find hidden experts who are not direct co-authors of the person conducting the search, but are a few steps away in the collaboration graph.

An analysis of expert search methods shows that modern research is increasingly leaning toward hybrid approaches that combine thematic modeling with network structure analysis. However, the effectiveness of these methods directly depends on how accurately and meaningfully the target subject domain within which the search is conducted is formalized [18].

2.3. Semantic models of subject domain

Modeling a subject domain is a fundamental task of knowledge engineering, which consists in creating a formalized representation of a particular sphere of reality or field of knowledge. In the context of expertise tasks, such models serve as a reference coordinate system against which the researcher's profile is evaluated.

The simplest form of knowledge organization is taxonomy, which is a hierarchical classification system where concepts are arranged according to generic and specific relationships. Taxonomies are attractive because they are easy to understand and use, but they have significant drawbacks: they are static, which means that they are slow to update and cannot keep up with the emergence of new interdisciplinary fields, and they are limited in their expressiveness because they only record hierarchical relationships [19].

The next level of semantic modeling is thesauri, which, in addition to hierarchy, record associative relationships between concepts. An even more expressive means of representing knowledge is ontologies, which formally specify the conceptualization of a subject area by defining classes of objects, their attributes, and arbitrary types of relationships between them. The advantages of ontologies are high accuracy and formalization of knowledge, the possibility of logical deduction of new knowledge, but their widespread use is hampered by the high labor intensity of development and the complexity of keeping them up to date [19].

With the development of natural language processing methods, an alternative approach has emerged based on representing the subject area as a semantic space, where concepts and

documents are presented as vectors in a multidimensional space, and semantic proximity is determined by the geometric distance between vectors [20]. The advantages of semantic spaces are automatic detection of semantics without the involvement of experts and high scalability, but vector representations are difficult for humans to interpret.

Knowledge graphs are an integrated solution that combines the advantages of ontologies with those of semantic spaces. In a knowledge graph, entities are nodes, and the connections between them are edges. For scientometric analysis tasks, open knowledge graphs specializing in scientific information, such as OpenAlex, are of particular interest [21].

3. OpenAlex as a data source for model building

3.1. Scientific information aggregators

In the modern scientometric landscape, there are a number of authoritative aggregators of scientific information. Traditionally, the dominant positions are occupied by commercial proprietary databases Scopus from Elsevier and Web of Science (WoS) Core Collection from Clarivate Analytics. Although these databases are distinguished by a high degree of curation and metadata quality, they have significant limitations: their licensing agreements prevent the free distribution of data, making reproducible bibliometric studies virtually unattainable.

In response to these limitations, several new data sources have emerged in recent years, including Dimensions, Google Scholar, and Microsoft Academic Graph (MAG), which was shut down in early 2022. Research shows that Google Scholar and Microsoft Academic had significantly higher coverage than Scopus and WoS, but Google Scholar has limited functionality for large-scale data extraction.

It is in this context that OpenAlex emerges as a promising open alternative, designed to combine the advantages of different sources and provide free access to a comprehensive catalog of the global scientific system.

3.2. Description of OpenAlex as an open knowledge graph

OpenAlex was launched on 2022 by the non-profit organization OurResearch as a replacement for the discontinued Microsoft Academic Graph. OpenAlex is positioned as a fully open catalog of the global research system, offering an alternative to industry standards for scientific knowledge bases.

Data comes from a wide range of open sources, including Crossref, PubMed, PubMed Central, Unpaywall, ROR, ORCID, and others. As of mid-2024, OpenAlex contains over 240 million scientific works.

OpenAlex operates with eight basic types of entities that form the structure of the knowledge graph: works (scientific documents: journal articles, books, conference proceedings, datasets, dissertations); authors (researchers who create scientific works, with integration with ORCID); sources (journals and repositories); institutions (organizations to which authors are affiliated, using ROR); topics (thematic concepts automatically assigned to works); concepts (a hierarchically organized system of thematic tags); publishers; funders [22].

3.3. Advantages of OpenAlex as a data source for scientometric research

OpenAlex offers a number of significant advantages over traditional commercial databases. Openness is a fundamental characteristic of OpenAlex. The entire dataset is available under a free CC0 license, allowing unrestricted use without licensing barriers. This creates the potential to support the transformation of research practices towards reproducible bibliometrics. A notable

example is the decision by Sorbonne University to cancel its subscription to Web of Science in favor of using open tools, including OpenAlex.

High coverage is a significant advantage of OpenAlex. Research in various subject areas confirms that OpenAlex indexes almost all journals covered by Scopus and Web of Science, plus additional sources from countries outside the Western world [23, 24]. An analysis of nursing literature found that OpenAlex has significantly better representation of non-English works .

Regular updates ensure that the data is up to date. OpenAlex is updated hourly, with new data added or modified with each release. A full database dump is updated monthly.

The availability of thematic concepts is critical to this study. OpenAlex has a built-in system for automatically assigning thematic tags based on analysis of titles, abstracts, source names, and citations. This system includes approximately 65,000 hierarchically organized concepts with 19 top-level root concepts and six levels of descendants.

Integration with persistent identifiers improves data quality. OpenAlex actively uses ORCID to identify authors and ROR to identify institutions. Studies show that OpenAlex records more ORCID identifiers per article than Web of Science and Scopus.

3.4. Analysis of the structure of OpenAlex thematic concepts

The central element of OpenAlex for modeling subject areas of expertise is its system of thematic concepts [25]. Unlike traditional subject headings, OpenAlex concepts reflect the dynamic, multidimensional nature of scientific knowledge. Concepts are abstract ideas discussed in scientific works.

Each work is assigned concepts based on an automatic classifier trained on the Microsoft Academic Graph corpus. For each "work-concept" pair, a score is available that reflects the classifier's confidence in assigning that concept. It is important to note that when a child concept is assigned, all of its parent concepts are also automatically assigned, up to the root concept, and for some of them the score may be 0.0.

In addition to concepts, OpenAlex also uses a Topic system—about 4,500 topics assigned to works using a model that takes into account the title, abstract, source name, and citations. Each topic has a description generated by artificial intelligence, keywords, and belongs to a specific subfield, field, and domain [26].

4. General concept of the semantic-statistical model

Expert competence is not just the number of publications or citations (bibliometrics), but the degree of mastery of key concepts, processes, relationships, and issues in a specific subject area.

If the target subject domain is not modeled (not described as a set of key concepts, their hierarchy, semantics, etc.), then the assessment boils down to subjective judgments or rough metrics (for example, "AI expert" — but what exactly: RL, NLP, computer vision, ethics?).

Furthermore, without a model of the subject domain of expertise, you cannot say whether an expert is "deeply knowledgeable," "partially competent," or "has superficial knowledge."

Domain modeling creates an objective knowledge map against which experts can be ranked. Without it, the assessment becomes subjective or relies solely on general metrics (h-index, number of publications) that do not take into account the specifics of the topic.

Scientific fields evolve rapidly (for example, in 2020–2025, new subdomains appeared in AI: diffusion models, multimodal LLMs, alignment, agentic AI, etc.). A static list of terms quickly becomes outdated. The use of aggregator systems allows us to automatically capture the current state of the domain, in our case, based on global OpenAlex data. Without such modeling, the assessment of competence will be either too narrow (outdated terms) or too broad (an "AI" expert may be an expert in a completely different field).

4.1. Formal representation of the model

The development of a semantic-statistical model of the subject domain of expertise is based on the need for a formalized representation of the structure of scientific knowledge, which would enable further matt with researcher profiles for a quantitative assessment of their competence. The proposed approach integrates a semantic dimension, which reflects the meaningful connections between concepts, with a statistical dimension, which records the significance of individual concepts within a given scientific domain [27].

Formally, the proposed semantic-statistical model of the target subject domain of expertise is defined as a triple :

M = (C, R, Q)

where:

1. C – a set of OpenAlex concepts representing the conceptual composition of the subject area;
2. R – a set of semantic relations between concepts that reflect structural connections within a subject area;
3. Q – a set of frequency values that quantitatively express the significance of each concept within a given subject area.

The set of concepts C is a basic element of the model. OpenAlex concepts are the fundamental building blocks of the proposed model. As noted by the developers of OpenAlex, concepts are abstract ideas discussed in scientific works. In total, OpenAlex indexes about 65,000 concepts, organized hierarchically with 19 root concepts and six levels of descendants.

A two-stage procedure is proposed for forming set C within a specific subject area of expertise. The first stage involves identifying the core of the subject area—a set of key concepts that directly determine its content. This core can be formed by experts or automatically based on the analysis of a representative sample of publications that a priori belong to the field under study. At the second stage, the core is expanded by semantically related concepts, which allows for a more complete representation of the subject area.

Semantic relations R is the structural framework of the model. The set of relations R captures the semantic connections between concepts included in the domain model. In the proposed approach, these relations are not postulated a priori, but are derived from the structure of the OpenAlex knowledge graph itself and from the analysis of publications in the domain.

The model proposes the use of two main types of semantic relations. Hierarchical relations directly follow from the structure of the OpenAlex concept tree. Each concept has defined ancestors—concepts of a higher level of generalization—and descendants—concepts of a lower level that detail its content.

Associative relations capture meaningful connections between concepts that are not hierarchical but reflect their common use in scientific publications. In the proposed model, these relations are defined as indirect associations through a set of works classified by OpenAlex concepts. Two concepts are considered associatively related if there is a statistically significant number of scientific works to which the classifier has assigned both concepts. This approach allows us to identify hidden interdisciplinary connections, which echoes the ideas of expertise extrapolation proposed in the SEMEF approach.

Frequency values Q is a measure of concept significance. The Q set provides a quantitative assessment of the importance of each concept within a given subject area. These weights reflect how central or representative a given concept is to the field of knowledge under study.

To calculate the weight of a concept, it is proposed to use information about the frequency of its use in the publications included in the sample, taking into account the weight coefficients (scores) that the OpenAlex automatic classifier assigns to each "publication-concept" pair.

The synthesis of three components—concepts C, relations R, and frequencies Q—forms a comprehensive semantic-statistical model of the subject area. Concepts C define the conceptual component of the model. Relationships R structure these concepts, creating a network where hierarchical relationships reflect generic-specific relationships, and associative relationships reflect real co-occurrence patterns in scientific literature. Frequencies H give this structure a quantitative dimension, allowing us to determine which concepts are central to a given domain and which are peripheral.

4.2. Iterative construction of a semantic-statistical model of the subject domain

The model is constructed as an iterative extension of the set of concepts describing the object of expertise by analyzing the co-occurrence of terms in relevant scientific publications. At each level, new concepts statistically related to the previous ones are identified, filtered by frequency, and added to the model until thematic saturation or threshold conditions are reached.

Generalized algorithm for model construction:

1. Determination of a set of concepts describing the subject area of expertise:

$$C^0=\{c_n^0\}, n=\overline{1,N} \quad (1)$$

where: N — the number of concepts describing the subject area of expertise,

c_n^0 — n-th concept from the set C^0.

2. Formation of a sample of publications in which at least one concept from C^0 is mentioned:

$$W^{C^0}=\{w_i^{C^0}\}, i=\overline{1,|W^{C^0}|} \quad (2)$$

Each work $w_i^{C^0}$ is characterized by its own set of concepts:

$$C^{(w_i^{C^0})}=\{c_l^{(w_i^{C^0})}\}, l=\overline{1,|C^{(w_i^{C^0})}|} \quad (3)$$

At the same time, repeated inclusion of concepts from C^0 is excluded:

$$C^{(w_i^{C^0})} \cap C^0 \neq \varnothing \quad (4)$$

3. For each concept $c_n^0 \in C^0$, the set of works in which it is mentioned is determined:

$$W^{(c_n^0)}=\{w_i \in W^{C^0} \mid c_n^0 \in C^{(w_i)}\} \quad (5)$$

and the frequency of its occurrence is calculated:

$$q_n^0=|W^{(c_n^0)}| \quad (6)$$

4. Based on the set W^{C^0}, a set of unique first-level concepts is formed, excluding concepts from C^0:

$$C^1=\{c_r^1\}=\left(\bigcup_{i=1}^{|W^{C^0}|} C^{(w_i^{C^0})}\right)\setminus C^0, r=1,|C^1| \quad (7)$$

For each concept $c_r^1 \in C^1$, the frequency of its use is determined by the number of works marked by the aggregator with this concept:

$$q_r^1=|W^{(c_r^1)}|, W^{(c_r^1)}=\{w_i \in W^{C^0} \mid c_r^1 \in C^{(w_i)}\} \quad (8)$$

Concepts with a frequency below the threshold level $L^1=0{,}03 \cdot |W^{C^0}|$ are excluded from consideration.

5. Similarly, iterative expansion of the model is performed, namely, based on the set C^1, a new sample of works is formed:

$$W^{C^1}=\{w_i^{C^1}\}, i=\overline{1,|W^{C^1}|} \quad (9)$$

and a new set of second-level concepts:

$$C^2=\{c_r^2\}=(\bigcup_{i=1}^{|W^{C^1}|}C^{(w_i^{C^1})})\setminus(C^0\cup C^1),r=\overline{1,|C^2|} \quad (10)$$

The frequency is calculated for each concept $c_r^2\in C^2$:

$$q_r^2=|W^{(c_r^2)}|,W^{(c_r^2)}=\{w_i\in W^{C^1}\mid c_r^2\in C^{(w_i)}\} \quad (11)$$

Concepts that do not satisfy the threshold level L^2 are excluded from further consideration.

The process is repeated iteratively until the maximum depth of the model is reached M

$$C^1\to C^2\to\cdots\to C^M \quad (12)$$

The maximum depth of the model is determined either by the exhaustion of new concepts or by reaching the threshold level of co-occurrence frequency.

The application of the algorithm allows us to build a multi-level model of the object of expertise, in which scientific concepts are arranged according to their degree of thematic proximity to the initial core — the basic set of terms describing the subject of evaluation. Each level of the model is formed on the basis of statistical analysis of the co-occurrence of terms in a global array of scientific publications, which ensures a gradual expansion of the context around the object of expertise. Thus, the model performs the function of localizing relevant knowledge within the general scientific space, structured according to the principle of semantic distance. At more distant levels of the model, concepts that go beyond the primary subject area naturally appear. This is because scientific publications that use a common theoretical apparatus can implement various applied directions belonging to other fields of knowledge. In addition, the model includes generalized concepts that serve as semantic linking elements, through which thematic branching into related disciplines is possible. This property of ensures the interdisciplinary permeability of the model, allowing not only to localize relevant knowledge, but also to identify potential points of intersection between different scientific directions.

Figure 1 presents a graphical interpretation of the constructed model of the subject area "Sentiment analysis."

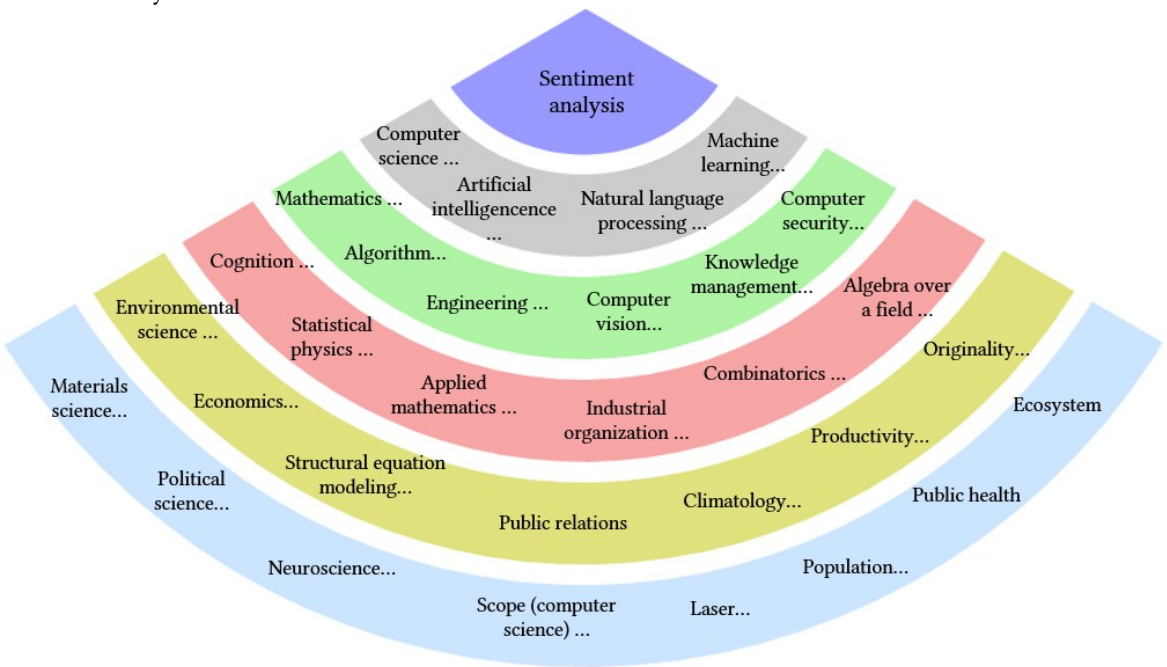


Fig. 1 Graphical interpretation of the model of the subject area "Sentiment analysis"

The constructed semantic-statistical model of the object of expertise serves as the basis for the next stage — assessing the thematic competence of potential experts. In this context, the scientific

profile of an expert is presented as a complete list of their publications, each of which is associated with a corresponding set of concepts.

The assessment is carried out by comparing the concepts of the expert's scientific works with the concepts included in the model of the object of expertise. On this basis, an integral competence indicator is calculated, which reflects the degree of thematic relevance of the expert. This indicator is determined as the weighted sum of quantitative matches between the concepts of both sets, taking into account their level of semantic proximity to the conceptual core of the model.

5. Method of expert evaluation based on a semantic-statistical model. Expert evaluation algorithm

The assessment of an expert's competence is built through the prism of a semantic-statistical model of a given subject area of expertise [2,3,28,29]. The generalized algorithm includes the following steps.

Step 1. Formation of an expert profile

For each expert E , a set of their scientific works is determined:

$$W(E)=\{w_j\}, j=\overline{1,|W(E)|} \tag{13}$$

Each work w_j is described by a set of concepts:

$$p^{(w_j)}=\{p_1,p_2,\ldots\} \tag{14}$$

The expert profile $P(E)$ is a combined set of unique concepts from all of their works:

$$P(E)=\bigcup_{j=1}^{|W(E)|} p^{(w_j)} \tag{15}$$

Step 2. Model of the object of expertise

The model is represented by a multi-level structure of concepts:

$$C^0,C^1,\ldots,C^L \tag{16}$$

where:

C^0 — the core of the model (basic concepts),

C^k — concepts statistically related to the level of C^{k-1} .

For each level k , a weight coefficient α_k is specified, which decreases as k increases:

$$\alpha_k=\frac{1}{2^k},\alpha_0>\alpha_1>\cdots>\alpha_L \tag{17}$$

Step 3. Calculating the coincidence function

For each level C^k , the coincidence function is determined:

$$f(E,C^k)=\sum_{c\in C^k} 1[c\in P(E)]\cdot q_E(c) \tag{18}$$

where:

$1[c\in P(E)]$ — indicator of the presence of the concept c in the expert's profile,

$q_E(c)$ is the number of the expert's works that are protected by the concept c .

Step 4. Integral assessment of competence

Generalized assessment of the expert:

$$A(E)=\sum_{k=0}^{L}\alpha_k\cdot f(E,C^k) \tag{19}$$

This indicator takes into account matches at all levels of the model, giving greater weight to concepts closer to the core.

Step 5. Normalization of assessment

Normalization is performed for comparison between experts:

$$A_{norm}(E)=\frac{A(E)}{max_{E'}A(E')} \quad (20)$$

where the maximum is determined by the assessment of all candidates E' in the group.

The normalized assessment $A_{norm}(E)\in[0;1]$ allows the expert's competence to be interpreted on a relative scale.

High $A_{norm}(E)\geq 0.7$ values indicate a high thematic relevance of the expert's profile to the subject of the competition.

Average $0.4\leq A_{norm}(E)<0.7$ values reflect partial relevance, which may be acceptable in interdisciplinary competitions.

Low $A_{norm}(E)<0.4$ values signal the remoteness of the expert's scientific activity from the subject of the expertise.

Thus, the proposed algorithm for assessing an expert's thematic competence is based on a multilevel semantic-statistical model that allows comparing the scientific profiles of candidates with a structured space of concepts relevant to the object of expertise.

6. Experimental Validation of the Method

Objective of the experiment: Validation of the developed semantic-statistical model for the quantitative assessment of the thematic relevance of an expert's scientific profile.

Object of evaluation: The scientific competence of three candidates with different levels of research output and engagement in the target domain.

6.1. Construction of a multilayer model

1. Selecting the core (layer 0). The starting concept of the domain is Sentiment analysis (in your data, it belongs to "layer 0" of the model).
2. Formation of the domain corpus. OpenAlex collects a set of works that have been assigned a core concept (or core concepts). For each work, the set of assigned concepts and their weights are known. The concepts of the works form a set of concepts of the next level.
3. Iterative expansion of the model (layers 1...d) until the maximum depth (6) is reached.

General statistics on the number of concepts at each level of the model are presented in Table 1. Analysis of the resulting model shows an "explosion" in the number of concepts and subsequent narrowing. The typical pattern is as follows:

- layer 0: 1 concept (core);
- layer 1: dozens of concepts (general "umbrellas" such as Computer science / AI / NLP / ML);
- layers 2–3: maximum number of concepts (massive branching into methods, subtasks, related subdomains);
- Layers 4–6: the number of concepts decreases/stabilizes: the model enters the "long tail" zone (small, narrow, and very rare concepts).

Table 1
Frequency of Special Characters

Model level	1	2	3	4	5	6
Number of concepts	57	620	435	212	97	52

Interpretation: at levels 2–3, the co-occurrence system provides the largest "combinatorial" number of thematic connections; then threshold filtering and semantic drift come into play (some

branches are cut off as weakly related to the core).

If we analyze the semantic quality of the layers, the coverage decreases with depth. By coverage, we mean the proportion of works in the domain corpus in which the concept is "present." For layer 1, there are concepts with very high coverage (they almost always accompany the core because they are general disciplinary "containers"). For deeper layers, the median coverage becomes very small: this is a sign that layers 3+ mainly describe fragments of the subject (subtasks/methods/contexts) rather than the "core" of the domain. The "core" of the domain is a conceptual framework that is present in a significant part of the works and defines the methodological and disciplinary basis of the subject. Layers 3+ are branches of this framework into specific tasks, application areas, and specializations.

Practical conclusion for the methodology: competence assessment must be sensitive to the layer, otherwise "random" tail matches will begin to inflate the scores. In the method proposed by the authors, this is implemented through layer weighting coefficients. α_k.

6.2 Expert evaluation

Three candidates are compared.

The results of the final ranking are presented in Table 2, which contains statistics on how many concepts from each level of the model for the subject of expertise "Sentiment analysis" were found in the candidates' works.

For each candidate, the following was used: the total number of works in the OpenAlex profile (1) and the subset of works related to the subject of the examination (2). In the sample, respectively: 903/100, 174/49, and 5/1 works (total/by subject).

The normalized assessment values already show a clear gradation at the level of integral metrics: the first candidate has the highest thematic saturation of the profile relative to the model; the second has approximately "half" of the leader; the third has a significantly lower relevance.

The key advantage of the layer diagram is evident in the qualitative analysis of where thematic overlap is concentrated. For Bhattacharyya and Li, layer 2 dominates (57.75 and 25 points, respectively), while for Tajrin, layer 1 dominates (7.5 points), and the contribution of layer 2 is minimal (1.5 points). This means that the first two profiles have broad support at the level of "average proximity" to the model core, while the third has support mainly at the level of the immediate context, without sufficient deployment into deeper/broader conceptual sublevels. Below are interpretations for each expert (including the number of works and "significant" concepts by layer).Expert 1 (Pushpak Bhattacharyya). Represents a high level of competence. The main contribution to the overall assessment is made by layer 2: 57.75 points, or about 61.9% of the total; the second most significant is layer 1 (23.5 points; ≈25.2%). The other layers together account for about 13% (layer 3 — 8.125; layer 4 — 1.6875; layer 5 — 1.28125). This distribution is typical for an expert whose thematic profile not only "rests" on the core, but is also well represented in concepts at a medium distance from the core (layer 2), which often corresponds to a stable research program with many subtopics and methodological branches. High scores at levels 2 and 3 simultaneously indicate semantic breadth (many related subconcepts) and structural connectivity (the presence of concepts from subsequent layers), but due to reduced weights, distant layers do not "overlap" core relevance. Integrally, Bhattacharyya is the top-ranked candidate in the group. The integral score is 93.34375, normalized to 1.0.

Expert 2 (Shoushan Li). Contribution profile of layers: layer 2 dominates (25 points; ≈52.4%), followed by layer 1 (17.5; ≈36.6%). Layers 3–5 make a relatively small contribution (3.125; 0.75; 0.375), which is to be expected given their weights. Compared to Bhattacharyya, the candidate has significantly fewer "significant" concepts in layers 2–5, which reduces the integral score by almost half: integral score 47.75, normalized 0.512. At the same time, it is important that the same

dominant structure is preserved (the advantage of layer 2 over layer 1), i.e., the profile is not a random "touch" on the topic, but a systematic presence in the relevant conceptual environment. Therefore, Li can be interpreted as a relevant mid-level expert, suitable for tasks that require competence in a significant part of the thematic context, but with less breadth/scope compared to the leader.

Table 2
Final ranking of candidates

No	Candidate characteristics	Pushpak Bhattacharyya	Shoushan Li	Jannatul Tajrin
1	Total number of works	903	174	5
2	Number of works by subject of expertise	100	49	1
3	Number of "significant" concepts in layer 0	1	1	1
4	Number of points for layer 0	1	1	1
5	Number of "significant" concepts in layer 1	47	35	15
6	Number of points for layer 1	23.5	17.5	7.5
7	Number of "significant" concepts in layer 2	462	10	6
8	Number of points for layer 2	57.75	25	1.5
9	Number of "significant" concepts in layer 3	195	25	1
10	Number of points for layer 3	8.125	3.125	0.125
11	Number of "significant" concepts in layer 4	10	12	1
12	Number of points for layer 4	1.6875	0.75	0.1875
13	Number of "significant" concepts in layer 5	205	12	2
14	Number of points per ball 5	1.28125	0.375	0.03125
15	Total points	93.343750	47.75	10.375
16	Normalized assessment value	1	0.512	0.111

Expert 3 (Jannatul Tajrin). Unlike the first two candidates, the contribution is concentrated in layer 1 (7.5 points; ≈72.3% of the total), while layer 2 accounts for only 1.5 points (≈14.5%). This pattern is consistent with a situation where the profile has a local, narrow match with the immediate environment of the model core (layer 1), but hardly expands into a broader conceptual space (layers 2–5). The reason is further confirmed by the small number of publications: with only one relevant work available, a limited set of concepts is formed, and the model "sees" weak support outside the immediate context. Tajrin receives an integral score of 10.375 (0.111 in normalized representation), which corresponds to low thematic relevance within the selected subject area compared to other candidates.

The experiment yielded the following results:

1. The experiment confirmed the model's ability to clearly distinguish between experts by level. The model successfully distinguishes between the "depth" of expertise (number of works on the topic) and "breadth" (coverage of semantically related layers), which is important for the formation of interdisciplinary research groups.

2. The ranking is stable and explainable. Not only do the total scores separate the candidates, but the distribution of contributions across layers also shows where the relevance lies (core/near-core context/broader thematic space).

3. The weighting coefficient scheme for layersα_k ensures control of interdisciplinary "noise." Even if many concepts appear in distant layers (as in Bhattacharyya), their contribution remains limited, and the overall assessment is determined mainly by layers close to the core. In other

words, this approach allows us to neutralize the influence of random publications on the periphery of the model, focusing the evaluation on the conceptual core.

4. The use of OpenAlex data allows avoiding the subjectivity inherent in traditional survey methods. The model automatically adapts to the dynamic scientific environment , taking into account the real patterns of concept sharing in millions of publications.

5. $A(E)$The results confirm the hypothesis that the integral indicator proposed in the work more objectively reflects competence than a simple count of works, since it takes into account the semantic structure of the subject of expertise and the author's scientific achievements.

The experiment demonstrates that a multi-level conceptual comparison based on OpenAlex allows not only to obtain a numerical assessment of competence, but also to provide an interpreted explanation of the result through the structure of matches by semantic "distance" to the core of the model.

7. Conclusion

This study proposes and experimentally verifies the validity and effectiveness of a semantic-statistical model of the subject area of expertise for the quantitative assessment of scientists' competence. A number of scientific and practical results have been obtained.

The choice of OpenAlex as the basic data source is justified. It is shown that openness, high coverage, regular updates, and the presence of a hierarchical system of concepts make OpenAlex an optimal basis for building reproducible scientometric models, free from the licensing restrictions of commercial databases.

A formal semantic-statistical model of the subject area has been developed. The model is presented in the form of a triple $M=(C,R,Q)$, where$C$ is a set of concepts,R are semantic relations (hierarchical and associative), and$Q$ are frequency values of concept significance. An iterative algorithm for constructing a multilevel model structure based on the analysis of concept co-occurrence in publications is proposed, which allows automatically identifying not only the core of the discipline, but also its peripheral and interdisciplinary connections.

A method for evaluating experts based on comparing their publication profile with a multilevel model is proposed. The integral indicator of competence$A(E)$ takes into account not only the presence of common concepts, but also their semantic "distance" from the core of the model through a system of weighting coefficientsα_k . This allows us to quantitatively measure both the "depth" of immersion in the topic and the "breadth" of coverage of the adjacent conceptual space.

The method's feasibility has been experimentally confirmed. Testing on the "Sentiment analysis" domain using real OpenAlex data demonstrated that:

- the model's ability to clearly rank experts by relevance;
- interpretability of results. Analysis of contributions by layers allows for a qualitative explanation of where the researcher's expertise is concentrated (at the core, in the immediate context, or on the periphery);
- the effectiveness of the weighting scheme for leveling the impact of random publications and interdisciplinary "noise," which focuses the assessment on the conceptual core of competence.

The results confirm the hypothesis that the proposed approach provides a more objective and meaningful assessment of competence compared to traditional bibliometric metrics, as it takes into account the semantic structure of scientific knowledge.

Prospects for further research. Further work will focus on automating the determination of threshold levels for filtering conceptsL^k . Scientific fields differ significantly in terms of the density of their terminology, so a static threshold is ineffective. Solving this problem involves a transition to adaptive filtering parameters that take into account the internal structure and dynamics of a

particular scientific domain. This will increase the sensitivity and selectivity of the model for a wide range of disciplines.

Declaration on Generative AI

During the preparation of this work, the authors used chat-GPT-4 in order to: Grammar and spelling check, Citation management. After using these tools, the authors reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] Andriichuk, O., Tsyganok, V., Kadenko, S., & Porplenko , Y. (2020) Experimental Research of Impact of Order of Pairwise Alternative Comparisons upon Credibility of Expert Session Results Proceedings of the 2nd IEEE International Conference on System Analysis & Intelligent Computing, (SAIC), 05-09 October, 2020, Kyiv, Ukraine, pp. 1-5. https://doi.org/10.1109/SAIC51296.2020.9239126

[2] Kadenko S., Tsyganok V., Szadoczki Z., Bozoki S. An update on combinatorial method for aggregation of expert judgments in AHP. *Production*, 2021, 31, pp. 1–17. https://doi.org/10.1590/0103-6513.20210045

[3] Tsyganok V.V. Providing sufficient strict individual rankings' consistency level while group decision-making with feedback. *Journal of Modelling in Management*. 2013. vol. 8, issue 3. P. 339-347. https://doi.org/10.1108/JM2-11-2010-0050

[4] Tsyganok, Vitaliy V., Kadenko, Sergii V., Andriichuk, Oleh V. Significance of expert competence consideration in group decision making using AHP. *International Journal of Production Research*, 2012, 50(17), pp. 4785–4792. https://doi.org/10.1080/00207543.2012.657967

[5] Hirsch, J.E.: An index to quantify an individual's scientific research output. Proc. Natl. Acad. Sci. U.S.A. 102, 16569–16572 (2005). https://doi.org/10.1073/pnas.0507655102

[6] Abramo, G., D'Angelo, C.: Evaluating research: From informed peer review to bibliometrics. Scientometrics 87, 499–514 (2011). https://doi.org/10.1007/s11192-011-0352-7.

[7] Jones, A.W.: Use of Bibliometric Methods to Verify the Credentials of Expert Witnesses in Forensic Science and Legal Medicine and the Dilemma of Multiple-authorship. Perspect. Legal Forensic Sci. 1(1), 10002 (2024). https://doi.org/10.35534/plfs.2024.10002.

[8] Khorasani, M., Salami, M., Soheili, F., Delghandi, F.: Job competency criteria for scientometrics specialists: Interviews with faculty members and job owners. Scientometrics Research Journal 9(1), 189–210 (2023). https://doi.org/10.22070/rsci.2021.13890.1478.

[9] Jappe, A.: Professional standards in bibliometric research evaluation? A meta-evaluation of European assessment practice 2005-2019. PLoS One 15(4), e0231735 (2020). https://doi.org/10.1371/journal.pone.0231735.

[10] Ninkov, A., Frank, J.R., Maggio, L.A.: Bibliometrics: methods for studying academic publishing. Perspectives on medical education 11(3), 173–176 (2022).

[11] Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., Rafols, I.: The Leiden Manifesto for research metrics. Nature 520, 429–431 (2015). https://doi.org/10.1038/520429a.

[12] Belter, C.W.: Bibliometric indicators: opportunities and limits. J Med Libr Assoc 103, 219–221 (2015). https://doi.org/10.3163/1536-5050.103.4.014.

[13] Blömeke, S., Gustafsson, J.-E., Shavelson, R.: Approaches to Competence Measurement in Higher Education. Zeitschrift für Psychologie 223, 1–2 (2015). https://doi.org/10.1027/2151-2604/a000193.

[14] Zlatkin-Troitschanskaia, O., Shavelson, R.J., Bartnik, T.N.: What have we learned about modelling and assessing mathematics teachers‘ competence?–Insights from a literature review. ZDM – Mathematics Education 57(6), 1093–1108 (2025). https://doi.org/10.1007/s11858-

025-01689-w.
[15] Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation. The Journal of Machine Learning Research 3, 993–1022 (2003).
[16] 12. Yin, J., Wang, J.: A Dirichlet Multinomial Mixture Model-based Approach for Short Text Clustering. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 233–242. ACM (2014). https://doi.org/10.1145/2623330.2623715.
[17] Deng, H., King, I., Lyu, M.R.: Formal Models for Expert Finding on DBLP Bibliography Data. In: 2008 Eighth IEEE International Conference on Data Mining, pp. 163–172. IEEE (2008). https://doi.org/10.1109/ICDM.2008.145.
[18] Liu, Y., Tang, W., Liu, Z., Ding, L., Tang, A.: High-quality domain expert finding method in CQA based on multi-granularity semantic analysis and interest drift. Information Sciences 596, 395–413 (2022). https://doi.org/10.1016/j.ins.2022.02.039.
[19] Currás, E.: Ontologies, Taxonomies and Thesauri in Systems Science and Systematics. Chandos Publishing (2010).
[20] Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781 (2013).
[21] Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G.D., Gutierrez, C., Kirrane, S., Gayo, J.E.L., Navigli, R., Neumaier, S., Ngomo, A.N., Polleres, A., Rashid, S.M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., Zimmermann, A.: Knowledge Graphs. ACM Computing Surveys 54(4), 1–37 (2021). https://doi.org/10.1145/3447772.
[22] Alperin, J.P., Portenoy, J., Demes, K., Larivière, V., Haustein, S.: An analysis of the suitability of OpenAlex for bibliometric analyses. arXiv:2404.17663 (2024). https://doi.org/10.48550/arXiv.2404.17663.
[23] Culbert, J., Hobert, A., Jahn, N., Haupka, N., Schmidt, M., Donner, P., Mayr, P.: Reference coverage analysis of OpenAlex compared to Web of Science and Scopus. arXiv:2401.16359 (2024). https://doi.org/10.48550/arXiv.2401.16359.
[24] Haupka, N., Culbert, J.H., Schniedermann, A., Jahn, N., Mayr, P.: Analysis of the publication and document types in OpenAlex, Web of Science, Scopus, Pubmed and Semantic Scholar. arXiv:2406.15154 (2024). http://arxiv.org/abs/2406.15154.
[25] Priem, J., Piwowar, H., Orr, R.: OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv:2205.01833 (2022). https://doi.org/10.48550/arXiv.2205.01833.
[26] OpenAlex: Work object. OpenAlex technical documentation (2024). https://docs.openalex.org/api-entities/works/work-object.
[27] Domanetska, I.M., Krolenko, Ya.O.: Features of determining the optimal composition of an expert group based on a semantic-statistical approach. Management of complex systems development 64, 194–200 (2025). https://dx.doi.org/10.32347/2412-9933.2025.64.194-200.
[28] Vitaliy Tsyganok, Yaroslav Khrolenko, Iryna Domanetska, Olha Tsyhanok, Olena Trubitsyna. An Automated Approach to Forming a Jury of Student Research Competitions Based on Scientometric and Altmetric Indicators. XXIV International Scientific and Practical Conference "Information Technologies and Security" (ITS 2024). Kyiv, Ukraine, December 19, 2024. CEUR WS Proceedings (ceur-ws.org). - Vol-4068. – pp.176-189. https://ceur-ws.org/Vol-4068/paper14.pdf.
[29] Tsyganok, V., Khrollenko, Y., Tsyhanok, V., Patuk, S., Trubitsyna, O. Strategic Planning Technology: Application Aspects of Artificial Intelligence Linguistic Generative Models. 10th International Scientific Conference "Information Technology and Implementation" (IT&I-2023), Kyiv, 20-21 November 2023. CEUR Workshop Proceedings, 2023, vol.3624, pp. 392 –402. https://doi.org/10.55056/ceur-ws.org/vol-3624/paper_32.pdf.