Information Processing and Management 63 (2026) 104424

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/ipm

Knowledge precedence networks: Mining progression patterns of scientific discoveries beyond prerequisites

Shibing Xiang [a,b], Bing Liu [e,f], Xin Jiang [c,d], Zhengan Huang [b], Yifang Ma [a,*]

[a] Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, Guangdong, China
[b] Pengcheng Laboratory, Shenzhen, Guangdong, China
[c] Institute of Artificial Intelligence & Beijing Advanced Innovation Center for Future Blockchain and Privacy Computing & LMIB, Beihang University, Beijing, China
[d] Hangzhou International Innovation Institute, Beihang University, Hangzhou, China
[e] LMIB & School of Mathematical Sciences, Beihang University, Beijing, China
[f] Zhengzhou Aerotropolis Institute of Artificial Intelligence, Zhengzhou, Henan, China

ARTICLE INFO

Dataset link: https://github.com/xiangshb/KPN-Mining

Keywords:
Knowledge precedence network
Progression pattern
Scientific discovery
Career path
Learning path navigation
Science of science

ABSTRACT

Understanding how knowledge evolves through scientists' career paths is essential for advancing education and innovation. This study constructs Knowledge Precedence Networks (KPNs) to uncover scientific progression patterns in real-world practice across 19 disciplines, analyzing the research trajectories of 4,969,403 scientists and 80 million publications from the OpenAlex dataset. We propose the **CoCiTCD** method, which integrates **Co-Ci**ting networks with **T**emporal **C**ommunity **D**etection to capture knowledge progression structures by identifying research communities, selecting representative concepts, and deriving temporal concept pairs. KPNs across Mathematics, Computer Science, and Engineering emphasize the critical role of foundational concepts in supporting advanced topics. For example, Algorithms bridge Mathematics and Computer Science, driving advancements in Artificial Intelligence and Data Science. We evaluate the alignment between KPNs for 303 concepts and theoretical prerequisite relations annotated by large language models, revealing how scientists engage with knowledge over time. The KPN attains a recall of 25.77% in best case, complemented by the citation-based KCN reaching 26.6%. This consistently low alignment indicates that empirical real-world topic transitions frequently diverge from theoretical prerequisite orderings. Furthermore, an AUC of 0.76 on our sample variational ROC curve underscores the robustness of our KPN approach in capturing the nuanced, innovative nature of knowledge progression. The KPNs provide valuable insights for research planning, learning path design, interdisciplinary collaboration, and understanding the hierarchical knowledge structure, thereby contributing to the Science of Science by uncovering real patterns of knowledge progression across disciplines.

1. Introduction

Concepts are the fundamental units of knowledge representation (Omerovic et al., 2001). The systematic organization of scientific concepts (Sun et al., 2022) is crucial for managing the rapid expansion of scientific knowledge (Bornmann et al., 2021), particularly in an era of exponential information overload (Bawden & Robinson, 2020). Since acquiring new knowledge typically builds upon prior understanding, foundational concepts often serve as prerequisites for mastering advanced topics (Nafa et al., 2022). The

* Correspondence to: No. 1088 Xueyuan Blvd, Nanshan District, Shenzhen, Guangdong Province, 518055, PR China.
E-mail address: mayf@sustech.edu.cn (Y. Ma).

https://doi.org/10.1016/j.ipm.2025.104424
Received 30 June 2025; Received in revised form 11 September 2025; Accepted 26 September 2025
Available online 13 October 2025

prerequisite relation (Xiao et al., 2021) captures the learning dependencies between concepts across diverse educational contexts, including Skills (Kumar et al., 2023), Courses (Stavrinides & Zuev, 2023), Mathematical ontologies (Kirillovich et al., 2021), Wikipedia concepts (Alzetta et al., 2019), Textbooks (Jia et al., 2021) and Massive Open Online Course (MOOCs) (Manrique et al., 2018). Synonyms such as the precedence relation (Xiao et al., 2021), surmise relation (Doignon & Falmagne, 2016; Wang et al., 2023), or propaedeutic relation (Adorni et al., 2019) refer to the same relationship.

The knowledge network with prerequisite relations among scientific concepts constitutes the Concept Prerequisite Network (CPN), which is crucial for both educational applications (Lu et al., 2019) and scientific advancement (Sun et al., 2022). For educators, CPNs facilitate educational organization such as course design (Manrique et al., 2018), curriculum planning (Sun et al., 2022) and learning topics ordering (Kumar et al., 2023). For students, CPNs serve as cognitive tutors (Ritter et al., 2007), offering personalized learning paths (Gan et al., 2022; Kumar et al., 2023), as well as recommending exercises (He et al., 2022) and knowledge concepts (Alatrash et al., 2024). Additionally, CPNs also facilitate knowledge acquisition (Li et al., 2019), and support intelligent systems such as educational QA (Yang, Yang, et al., 2024).

Many studies have been proposed to learn annotated prerequisite relations among concepts (Tang et al., 2023), which can be broadly categorized into expert queries (Dowling et al., 1996), statistical approaches (Bart & Read, 1984; Scheines et al., 2014), machine learning (Xiao et al., 2022), deep learning (Li et al., 2020; Sun et al., 2022) and large language models (LLMs) (Tang et al., 2023; Yang, Yang, et al., 2024) based methods. Similar works, such as those on knowledge evolution patterns (Yang, Wu, & Lyu, 2024) and knowledge graph (KG) reasoning (Ma et al., 2024), also provide valuable insights. Existing prerequisite learning studies predominantly rely on manually annotated ground truth for evaluation, using protocols such as PREAP (Alzetta et al., 2024) or expert querying (Dowling et al., 1996; Koppen & Doignon, 1990), while unsupervised prerequisite inference methods are rarely used as ground truth. However, manual annotation is challenging due to its labor-intensive nature, the scarcity of prerequisite relations (Liang et al., 2017), and expert blind spot issues (Scheines et al., 2014). In addition, the concepts studied are often domain-specific (Lu et al., 2019) and rarely address cross-disciplinary precedence relations as they occur in real-world scientific practice. To overcome these limitations, we leverage five LLMs to automatically annotate concept prerequisite pairs.

Despite extensive modeling of prerequisite relations in MOOCs and textbooks, few studies have explored how scientific discoveries with temporal patterns are manifested in actual research practice, as the research trajectories inherently reveal the temporal order in which scientists engage with research topics. A research trajectory is the sequence of an author's published papers, with each paper capturing key scientific discoveries represented by its main topics or Wikipedia concepts. Thus, it remains unclear how scientific concepts are temporally organized within researchers' careers, and our understanding of knowledge progression within or across disciplines is limited. Existing CPNs primarily connect concepts via semantic proximity, forming static graphs that reflect idealized and rigid learning dependencies while overlooking behavioral and temporal evidence (e.g., timestamped adoption, author sequencing, usage trajectories). As a result, these structures map prerequisite learning order rather than the innovative, non-linear progression patterns in real scientific discovery. For example, whereas a CPN might require "Statistical learning" to precede "Deep learning", scientists might explore "Deep learning" first and later reinforce their research with "Statistical learning". To address this gap, we introduce the Knowledge Precedence Network (KPN), which models the temporal advancement of knowledge in scientists' career trajectories based on actual research practices. Compared to CPNs, KPNs provide an empirically grounded view of how innovation-related concept transitions unfold over time, which makes KPNs better suited to reflect the dynamic and innovative nature of real scientific discovery. Accordingly, we employ CPNs solely as an external benchmark to assess whether KPNs recover logical prerequisite relations.

The proposed KPNs further enable us to investigate several key questions: Do scientists' research trajectories follow the prerequisite structures encoded in CPNs? To what extent do KPNs derived from groups of scientists align with these theoretical patterns? And how does this alignment differ between concepts within a single discipline versus those spanning multiple disciplines? Addressing these questions is crucial for understanding whether the CPNs that guide educational content organization also underpin the real-world scientific progress and career development. Importantly, while uncovering empirical precedence patterns offers novel macro-level insights, our preliminary evidence of low theoretical prerequisite alignment indicates that many observed transitions may reflect strategic exploration, opportunistic collaboration, funding shifts, or methodological convergence rather than canonical knowledge dependencies. This gap introduces a methodological boundary: KPNs alone are insufficient for tasks demanding high-precision pedagogical scaffolding such as syllabus ordering, mastery modeling.

Our study leverages extensive publication records of about five million scientists (4,969,403) from the OpenAlex dataset (Priem et al., 2022), and further propose a network-based framework to construct KPNs across 19 disciplines, providing an evidence-based perspective on knowledge progression. We first construct author-specific co-citing network (CCN), where research works are linked through co-citing relations and clustered into distinct research communities. We then identify temporal concept pairs by filtering community pairs that satisfy the empirical transition periods. The resulting KPNs capture hierarchical structures, and provide insights into how knowledge advances throughout academic careers. To further assess the temporal progression patterns, we compare the resulting KPNs with theoretical CPNs, illustrating the degree of alignment between real scientific research and conventional learning pathways. This exploration highlights the progression patterns of scientific discovery, supporting the interpretation that earlier topical foci tend to precede later diversification.

Our work differs from previous studies by examining both disciplinary and cross-disciplinary KPNs of scientific discoveries as they emerge in actual research trajectories. We also assess how closely scientists' topic learning and research progression follow the prerequisite structures found in traditional educational settings. By systematically analyzing research trajectories across 19 disciplines and benchmarking against LLM-annotated CPNs, we offer new insights into the alignment and divergence between theoretical and practical knowledge structures. This approach not only overcomes the limitations of domain specificity and annotation bottlenecks but also bridges the gap between theoretical models and real-world scientific discovery. To the best of our knowledge, this is the first study to extract cross-disciplinary KPNs spanning 19 fields. Unlike existing studies that focus on theoretical learning dependencies in CPNs, our work focuses on practical progression patterns represented by KPNs in real-world scientific discovery. This approach highlights the dynamic and innovative nature of how knowledge evolves in practice, rather than its theoretical structure. Furthermore, prior works rely on supervised machine learning with labeled, domain-specific prerequisite datasets, our method uniquely uncovers patterns from the behavioral evidence of scientists across 19 disciplines. We also annotate concept pairs with prerequisite relations as benchmarks using five LLMs, which are utilized to analyze deviations of the actual KPNs and significantly improve efficiency compared to traditional human annotations. By leveraging this evidence-based framework, we provide insights to support informed decision-making, such as mentoring junior researchers, planning research and learning pathways across diverse scientific domains.

Research Objectives. (1) We aim to uncover the KPNs across 19 disciplines by analyzing the research trajectories of over a million scientists. (2) We analyze the hierarchical KPN structures to reveal the temporal progression patterns of scientific discovery in researchers' career paths. (3) By comparing these real-world KPNs with theoretical CPNs, we explore to what extent these KPNs align with the theoretical CPNs. Ultimately, this work provides valuable insights into the hierarchical knowledge progression patterns while exploring their practical applications in education and research planning.

Our key contributions include:

1. The first to utilize extensive scientific research trajectories to extract precedence relationships.
2. Developing a network-based CoCiTCD framework to analyze the KPN across 19 disciplines.
3. Uncovering the real-world knowledge transition patterns or evolution dynamics in scientific career paths.
4. Providing insights into the hierarchical structure and research dependencies of scientific concepts.
5. Extending the contextual mutual prerequisite formulation and releasing the LLM-annotated SciConPreq dataset.
6. Revealing the low alignment between KPN in research trajectories and the theoretical CPN.

Roadmap. In Section 2, we provide a brief review of related works on concept prerequisite relation learning. In Section 3, we detail the datasets, methods for creating CCNs, community detection, representative concepts, temporal concept pairs. In Section 4, we empirically evaluate KPNs via joint hyperparameter selection, prerequisite alignment, cross-discipline and career-stage precision, WCR/semantic/resolution ablations, citation-based KCN comparison, and structural analysis of low recall. In Section 5, we present and analyze the results of the real-world KPNs concerning Mathematics, Computer science and Engineering. In Section 6, we show how KPNs enable personalized, behavior-grounded learning path navigation, derive theoretical and practical implications for scientific progression, and analyze limitations. Finally, we provide a further discussion in Section 7.

2. Related works

Our work lies in the field of knowledge discovery, which aims to uncover topic or concept transition networks from the research trajectories of millions of scientists, essentially mining the temporal precedence relations between concepts. To the best of our knowledge, there is limited existing work that directly addresses this problem. The network-based approach to analyzing scholarly behavior is supported by established methodologies in the field. Barnett and Park (Barnett & Park, 2023; Park & Barnett, 2024) successfully employed network analysis techniques to reveal patterns in elite scholarly behavior, demonstrating the effectiveness of network-based methods for elucidating the structure and evolution of scholarly networks. Beyond behavioral network analyses, KG embeddings (Li, Huang, et al., 2025) provide vector representation of entities, which is useful for modeling concept prerequisite relations. Recent advances like AdaGCN (Li, Chen, et al., 2025) derives relation–entity interaction embeddings and employs a self-gated attention mechanism with graph convolutional network (GCN) to adaptively aggregate or suppress neighbor messages for KG completion. APKGC (Jian et al., 2025) enhances multimodal KG completion robustness via adaptive noise sampling across textual and visual modalities and an attention penalty mitigating overly peaked attention. Although these models address static triple completion rather than empirical temporal concept precedence, these mechanisms are informative to prerequisite relation learning. Related studies, such as Sun et al. (2024), extract concept prerequisite relations from concept learning paths. However, our approach fundamentally differs in both data sources and the nature of the relations extracted. Specifically, we mine concept transitions or precedence relations from the career trajectories, rather than from educational or learning sequences. While Yang (2025) employs cosine similarity calculations between journal distribution vectors to quantify topic-switching patterns and demonstrates that frequent topic transitions drive scientific innovation, our KPN methodology constructs explicit temporal networks using co-citing analysis to map concept precedence relationships in practice. Although their methodology was not directly utilized in our work, we found their conclusions particularly insightful, we discussed their findings to strengthen our conclusions. Unlike prerequisite relations, which reflect learning dependencies, the transitions between concepts observed in scientific research trajectories do not necessarily follow strict prerequisite logic, as the evolution of research topics is often nonlinear and exploratory. Related studies cover a range of concept relation extraction and learning methods, including co-occurrence relations (Wang et al., 2022), causal relations (Dominici et al., 2025; Scheines et al., 2014), prerequisite relations (Tang et al., 2023), and temporal patterns in knowledge evolution (Huang et al., 2024). Among these, prerequisite relations have attracted significant attention and have been extensively explored.

Recent methodologies for studying knowledge evolution (Huang et al., 2024; Yang & Hu, 2025) highlighted how knowledge emerges, diverges, and fades over time. For example, Shao et al. (2022) trace the evolution of AI by combining bibliometric analysis to identify key trends. However, their approach primarily catalogs topics as they evolve, without addressing finer-grained precedence or dependencies among knowledge elements. Yang, Wu, and Lyu (2024) show with ERGMs on biomedical EFCNs that citation formation is dominated by preferential attachment and homophily, whereas clustering-based transitivity is weak. KPNs advance EFCNs by constructing large-scale temporal KPN across 19 disciplines, revealing gaps between actual empirical career trajectories and ideal prerequisite hierarchies. EvoPath (Liu, Cheng, et al., 2025) uses LLMs to evolve schema-level meta-paths among heterogeneous information networks, but it remains at abstract path templates instead of evolution of research behaviors over real conceptual knowledge. While KPN extends beyond path template search to capture authentic evolution of millions of research career behaviors across 19 disciplines, thereby enable applications beyond EvoPath's original scope (cross-disciplinary mapping, personalized learning path navigation, research planning). Yang et al. (2025) identify six IMRaD-based functional interdisciplinarity patterns within single bioinformatics papers, while our KPN shifts to a macro cross-disciplinary career perspective, thus exposing systemic structural divergence rather than intra-paper functional heterogeneity. Xu and Yang (2024) investigate organic growth knowledge production within a cMOOC by combining LDA topic clustering with refined CIE behavior coding to uncover intra-topic behavioral sequences and inter-topic interaction mechanisms, whereas our KPN framework extends beyond a single course to large-scale, multi-disciplinary career trajectories, quantifying macro-level concept precedence divergence that complements those micro interaction insights.

The study of prerequisite relation (PR) (Tang et al., 2023) has employed a range of methods, including statistical, machine or deep learning, and LLM-based approaches. Early statistical methods identify PR through hypothesis testing and Fisher's exact test (Bart & Read, 1984; Gasparetti et al., 2015). Various unsupervised methods leverage textual structure (Wang et al., 2016) (chapter or curricular hierarchy), hyperlink-based reference distance (Liang et al., 2015), or embedding-based link prediction (Li et al., 2020) to infers PR. Supervised methods range from video reference distance (Pan et al., 2017) and strict partial order and active learning (Liang, Ye, Wang, et al., 2018; Liang, Ye, Zhao, et al., 2018) to directional feature engineering over course co-occurrence/order. An RNN probabilistic model Chen, Lan, et al. (2018) predicts behaviors from clickstreams while learning PR-matrix R, a supervised approach (Xiao et al., 2022) constructs features from course descriptions and Wikipedia to classify PR. Recent unsupervised graph autoencoders such as R-VGAE (Li et al., 2020) and CD-VGAE (Li, Yan, & Radev, 2021) model PR learning as a link prediction, leveraging concept-resource graph embedding $\hat{X}$ to predict PR-matrix $\hat{A} = \hat{X}^{\top} R \hat{X}$, and adapt to across domains (Li, Yan, Li, et al., 2021) to improve scalability. Supervised ConLearn (Sun et al., 2022) leverage BERT embedding initialization combined with Gated GNN and self-attention for relation prediction, TCPL (Tang et al., 2023) uses MOOC video captions for continual BERT pre-training, then fuses textual representation and resource–concept graph structural embedding to predict PR. The CGPrompt (Yang, Yang, et al., 2024) directly employs zero-shot LLM prompting for concept graph recovery. EBCPL (Zhang et al., 2024) introduces an evidence-based framework that first explicitly extracts evidence sentences from educational documents and then applies a BiLSTM relation extraction model to infer concept PR. GKROM (Zhang et al., 2025) jointly optimizes concept–concept, document–concept, and document–document relations through a global multi-objective framework with two auxiliary tasks to enrich node embeddings and enhance concept prerequisite prediction.

While these research has significantly advanced the understanding of knowledge evolution and identification of prerequisite relations, several critical limitations persist. First, existing supervised and unsupervised CPN methods fail to recover authentic large-scale knowledge advancement patterns rooted in real temporal and sequential scientific dynamics, largely because they rely on curated sources (textbooks, MOOCs, expert annotations) that reflect idealized rather than genuinely dynamic and innovative knowledge evolution. This reliance limits their ability to capture the dynamic and innovative pathways through which scientific concepts are adopted and integrated in real-world research. Second, these approaches predominantly focus on domain-specific contexts, lacking the breadth to reveal universal patterns or differences across disciplines. Third, although recent works have explored macro-level knowledge evolution, they rarely investigate fine-grained, micro-level progression or the specific precedence relationships that shape the hierarchical structure of scientific knowledge. Furthermore, the manual annotation of prerequisite relations is labor-intensive, subjective, and difficult to scale for large and evolving knowledge networks. Consequently, these methods cannot capture the nuanced and meaningful patterns of knowledge advancement, nor can they detail the specific evolution pathways or precedence networks between disciplines. Addressing these gaps is crucial for a deeper understanding of the dynamics of scientific knowledge evolution.

In contrast, our study explores shifts in scientific research topics along career paths and uncovers patterns of knowledge progression. We addresses these gaps by leveraging large-scale publication trajectories from nearly five million scientists across 19 disciplines, enabling the automatic extraction of KPNs through LLM-assisted annotation. This approach not only overcomes the limitations of data representativeness and scalability but also bridges the gap between theoretical models and real scientific discovery, providing a more realistic and comprehensive view of knowledge progression.

3. Methods

In this section, we introduce the **CoCiTCD** for cross-domain KPN construction: an integration of Co-Citing (**CoCi**) networks and Temporal Community Detection (**TCD**) used to extract research communities, identify representative concepts, and derive temporal community concept pairs that encode innovative knowledge transitions. We then refine the resulting transition matrices to highlight statistically concept precedences, enabling robust identification of temporal relationship across 19 disciplines. Section 3.1 assembles the cross-discipline concept and author–work data, focusing on hierarchical concepts across 19 disciplines and sub-concepts in four selected fields, refines the prerequisite relation label space into five types (Table 1), and constructs the SciConPreq dataset by

Table 1
Definition of concept prerequisite relations.

Type	Related	Preq	Notation	Description
−2	√		$A \backsim B$	A and B are related but not prerequisite
−1	√	√	$A \leftarrow B$	B is a prerequisite of A
0			$A \perp B$	A and B are not related (independent)
1	√	√	$A \rightarrow B$	A is a prerequisite of B
2	√	√	$A \leftrightarrow B$	A and B are contextual mutual prerequisites

aggregating five LLMs' labels via majority voting for all cross-field concept pairs (Table 2) and for each of the four discipline-specific concept pairs (Table 3); Section 3.2 builds per-author co-citing (bibliographic coupling) networks and augments them with adaptive semantic similarity edges (concept sequence extraction and refinement details in Section Appendix B); Section 3.3 performs community detection and selects minimally sufficient representative concepts via Work Coverage Rate (WCR) and its survival statistics (Table 4); Section 3.4 orders communities by Average Publication Year (APY) and filters community pairs to a mid-range APY difference (APYD) window to derive Temporal Concept Pair (TCP); Section 3.5 aggregates all TCPs into the raw Knowledge Precedence Matrix (KPM), enforces directionality by retaining the larger of each symmetric pair (Algorithm 1), and optimize the matrix using cumulative value mass (CVM) and cumulative frequency mass (CFM) thresholds to form optimized KPMs (OKPMs) for specific disciplines (Eqs. (10) and (11), Algorithm 2), yielding the optimized precedence matrix (Fig. 4) to construct the KPNs. Together these sections define a reproducible transformation from publication and citation data to an optimized directed cross-disciplinary KPN.

3.1. Dataset

We utilize entities such as Authors, Works, and Concepts[1] from the OpenAlex dataset (Priem et al., 2022) updated up to the end of 2022, which is publicly accessible.[2] Specifically, we select authors with more than 10 works for CCN construction (KPN construction), yielding 4,969,403 candidates who collectively cover 80,388,510 works. For subsequent semantic network refinement, we further narrow our focus to a subset of 1,312,048 authors who have published more than 50 papers. This subset serves as the foundation for constructing concept sequences and learning concept embeddings. The minimum threshold of 50 papers ensures sufficient sequence length and high-quality embeddings, as detailed in Appendix B. For each author, we construct a CCN based on all their published works within the dataset. Each work of an author is tagged[3] with multiple Wikipedia concepts representing its main topics. The OpeAlex dataset comprises 65,068 concepts in total, organized across six hierarchical levels ranging from level 0 to level 5. The 19 level-0 concepts represent 19 disciplines, which further branch into 284 level-1 concepts. The method for hierarchical concept classification is detailed in SciConNav (Xiang et al., 2025). Ultimately, our work involves a total of 303 cross-discipline concepts, including 19 level-0 concepts and 284 level-1 concepts. In addition, to facilitate cross-domain validation, we select 250 top sub-discipline concepts from each of the disciplines of Computer Science, Engineering, Mathematics, and Physics, ranked by associated works count.

Existing work models prerequisite relations of concept pair (A, B) as a ternary function $\mathrm{Preq}(A, B) \in \{-1, 0, 1\}$ (Bai et al., 2025), where 0 aggregates all remaining cases without distinction. We refine this undifferentiated 0 into a triad shown in Table 1: Type 0 now denotes independence ($A \perp B$); the new Type −2 ($A \sim B$) for related but non-prerequisite pairs (motivated by Talukdar and Cohen (2012) and annotator feedback); and to ensure the completeness of the relation space and to resolve ties when aggregated model votes for −1 ($B \rightarrow A$) and 1 ($A \rightarrow B$) are equal, we add Type 2 ($A \leftrightarrow B$) for contextual mutual prerequisites (CMP), as the concept understanding is context-sensitive (Hsu et al., 1998). Note that type 2 relation uses notation $A \leftrightarrow B$, which is not a simultaneous bidirectional, but a conditional prerequisite $A \rightarrow B$ or $B \rightarrow A$ depending on context. Anti-symmetry is ensured via the unidirectional inference result instantiated contextual CPN, where only one direction is permitted per context. More detailed description of type 2 relation, together with their interpretation, are presented in Appendix C.

Previous research has primarily established ground-truth prerequisite relations using manual expert labeling, and several public datasets have been compiled to support research in this field.[4] Existing domain-specific datasets lack cross-field prerequisite labels for our selected concepts across 19 disciplines, and given that exhaustive manual annotation is impractical (large relation space, dispersed expertise, bias risks, and high cost). Prior supervised or unsupervised pipelines require abundant supporting text (courses, textbooks, resource graphs, comprehensive Wikipedia coverage) and produce inferred instead of ground-truth relations; with only concept names and gaps in Wikipedia coverage, these corpus-dependent methods are inapplicable. Recent work on LLM-based prerequisite annotation has progressed through AI-assisted concept-pair ranking with selective expert verification (Aytekin et al., 2024) and task-specific fine-tuning for prerequisite detection (Aytekin & Saygín, 2025) to direct prompt-based prerequisite recovery (Yang, Yang, et al., 2024), which reports GPT-4 as the best zero-shot model. Therefore, we employ a structured prompt to annotate prerequisite relations. The task requires five LLMs to output an integer from $\{-2, -1, 0, 1, 2\}$ for each of the $\binom{303}{2}$ concept

1 https://docs.openalex.org/api-entities/concepts
2 https://docs.openalex.org/download-all-data/download-to-your-machine
3 https://docs.google.com/document/d/1OgXSLriHO3Ekz0OYoaoP_h0sPcuvV4EqX7VgLLblKe4
4 https://github.com/xiangshb/Concept-Prerequisite-Datasets

Table 2
Response distribution of five LLMs in prerequisite relation annotation.

	GPT-4o	Claude-3-5-sonnet	ERNIE-3.5-8K	qwen-plus	Spark4.0-Ultra	Majority vote
−2	2781	8008	10 550	7389	8020	**6148 (13.44%)**
−1	1550	1472	1907	631	2807	**1609 (3.52%)**
0	39 931	35 796	31 744	36 998	33 356	**36849 (80.54%)**
1	1491	475	1552	718	1569	**1136 (2.48%)**
2		2		17	1	**11 (0.24‰)**

Table 3
Top 250 concepts (levels 0 to 4) ranked by works count in four disciplines.

Discipline	0	1	2	3	4
Computer science	1	29	175	37	8
Engineering	1	40	148	56	5
Mathematics	1	11	194	39	5
Physics	1	23	184	40	2

pairs, indicating the relation type defined in Table 1. The final annotation is derived by aggregating responses from five LLMs through majority voting (MV). The corresponding annotation prompt for 303 cross-discipline concepts is shown in Fig. A.1. Given the heterogeneity of concepts across disciplines, each concept is supplemented by a concise GPT-4 summary of its Wikipedia definition to reduce tokens and costs. In the current snapshot, 12 of the 303 concepts lack exact Wikipedia pages as shown in Table B.7, 11 were manually matched to related pages containing their definitions, while "Risk analysis (engineering)" received a GPT-4 definition.

LLM Response Statistics. Table 2 reports the distribution of LLM responses for pairs drawn from 303 cross-field concepts across the five relation types defined in Table 1: 94% of pairs are non-prerequisites (Type 0: 80.54%, Type -2: 13.44%), while the prerequisite relations (Types -1, 1, 2) constitute only 6%, indicating severe class imbalance. CMPs (Type 2) are particularly rare (20 single-model responses; 11 final MV labels). The 20 Type 2 responses in Table 2 each correspond to a distinct concept pair (Table C.8, rows 1–20) and all failed the final MV due to the single-model support. In contrast, the 11 final Type 2 assignments (Table C.8, rows 21–31) arose solely from tie patterns (two Type 1 vs two Type -1) without any explicit Type 2 response.

To further address the effectiveness and potential bias of our method, we performed cross-domain evaluations by assessing author-level precision for authors from each discipline across the four disciplines, providing a cross-domain validation of topic transition patterns. Addition to the 303 level 0 and level 1 concepts that cross 19 disciplines, we conducted additional analyses on sub-concepts from four representative disciplines: Computer Science (C.S.), Engineering (Eng.), Mathematics (Math.), and Physics (Phys.). For each discipline, we selected the top 250 concepts from levels 0 to 4, ranked by the number of associated works, as shown in Table 3. We then annotate $\binom{250}{2}$ pairwise prerequisite relations for each discipline using five recent LLMs (GPT-4.1, Claude 3.7 Sonnet, Gemini 2.5 Pro, Grok 3 Beta, DeepSeek-V3) as shown in Figure A.2. To support further research, we publicly release our LLM-annotated dataset, **Sci**entific **Con**cept **Pre**re**q**uisite (**SciConPreq**), a collection of concept prerequisite relations annotated by LLMs via majority voting, in our code repository[5].

3.2. Co-citing network construction of researchers

We select authors with more than 10 published works and create a co-citing (Zeng et al., 2019) network for each one of these authors, which is also called bibliographic coupling (Kessler, 1963; Weinberg, 1974) in bibliometrics. The nodes of the CCN are published works of each author, listed in chronological order based on publication year. Denote R_i as the set of references for work w_i. If two works w_i and w_j share common references, i.e., $R_i \cap R_j \neq \emptyset$, then a directed edge e_{ij} is formed from the earlier work to the later one. An illustration of constructing a CCN is shown in Fig. 1. The work w_1 and w_3 in Fig. 1(a) share at least one reference marked by the two red directed arrows, and the publication year of w_1 is earlier than w_3, thus resulting a directed edge $w_1 \rightarrow w_3$ shown in Fig. 1(c).

To address the risk of missing critical knowledge dependencies, we introduce a semantic refinement to the construction of CCNs. Beyond traditional co-citing links, we train a concept embedding model to quantify the semantic similarity between each pair of works based on their associated scientific concepts. For each author, we compute the embedding of each work as the average of its associated concept embeddings, and then calculate pairwise cosine similarities among all works. We enhance the original CCN by adding edges between works whose semantic similarity exceeds an adaptively determined quantile threshold, and by removing edges with weak semantic connections (similarity below 0.1). To ensure that the refined network is neither overly dense nor sparse, we set the average degree as a logarithmic function of network size, following empirical recommendations from the literature. This adaptive thresholding ensures that nearly all networks maintain an average degree within the optimal range for community detection. Full details of the semantic refinement procedure and parameter selection are provided in Appendix B. This approach enables our method to more robustly capture latent knowledge dependencies, thereby improving the assessment of prerequisite alignments.

[5] https://github.com/xiangshb/KPN-Mining/tree/main/llm_annotation

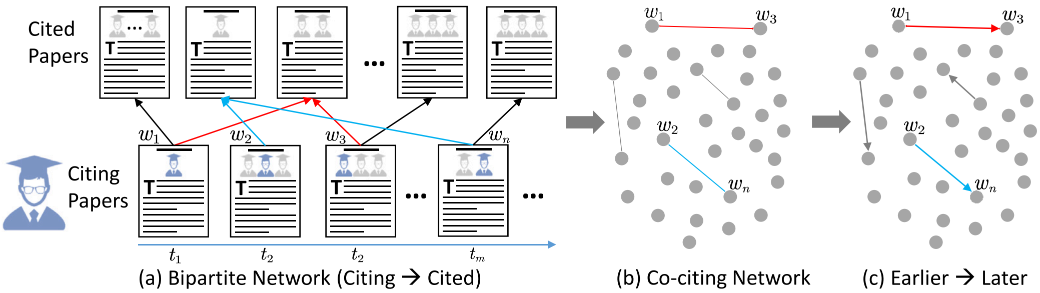


Fig. 1. Example of constructing the co-citing network.

Table 4
Survival probability of the top k concepts decreases as the WCR τ increases.

k	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
top 1	0.97	0.937	0.917	0.796	0.751	0.68	0.585	0.54	0.467	0.398	0.357	0.332
top 2	0.997	0.993	0.991	0.954	0.936	0.896	0.829	0.78	0.695	0.596	0.522	0.466
top 3	1.0	0.999	0.999	0.988	0.984	0.969	0.936	0.906	**0.84**	0.743	0.652	0.566
top 4	1.0	1.0	1.0	0.997	0.996	0.991	0.977	0.962	0.921	0.842	0.748	0.637
top 5	1.0	1.0	1.0	0.999	0.999	0.997	0.992	0.985	0.962	0.905	0.819	0.69

3.3. Community representative concepts as research field

3.3.1. Representative concepts of the community

Let $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ be the set of authors. For each author a_i, let $\mathcal{W}_i = \langle w_{i1}, \ldots, w_{im_i} \rangle$ denote the sequence of their m_i publications in the OpenAlex database, indexed by chronological (publication) order so that w_{i1} is the earliest and w_{ij} is the jth published work. Let $G_i = (\mathcal{W}_i, \mathcal{E}_i)$ be the corresponding CCN of author a_i. For a detected community $\pi \subseteq \mathcal{W}_i$ and each work $w \in \pi$, let $C(w)$ denote the set of concepts labeled to w. Define the set of all concepts in community π by $C(\pi) = \bigcup_{w\in\pi} C(w)$. For each concept $c \in C(\pi)$, its in-community frequency is $f_\pi(c) = |\{w \in \pi : c \in C(w)\}|$. Let $\mathrm{TOP}(\pi, k) \subseteq C(\pi)$ be the set of the "top k" concepts with the largest values of $f_\pi(c)$, where $1 \leqslant k \leqslant |C(\pi)|$. The collectively covered works (CCW) of top k concepts in community π is represented as

$$\mathrm{CCW}(\pi, k) = \left\{ w \mid w \in \pi, C(w) \cap \mathrm{TOP}(\pi, k) \neq \emptyset \right\}, \tag{1}$$

which are works in π that are labeled with at least one concept from $\mathrm{TOP}(\pi, k)$.

The WCR of top k concepts in community π is defined as the proportion of CCW of top k concepts that

$$\mathrm{WCR}(\pi, k) = |\mathrm{CCW}(\pi, k)| \,/\, |\pi| , \tag{2}$$

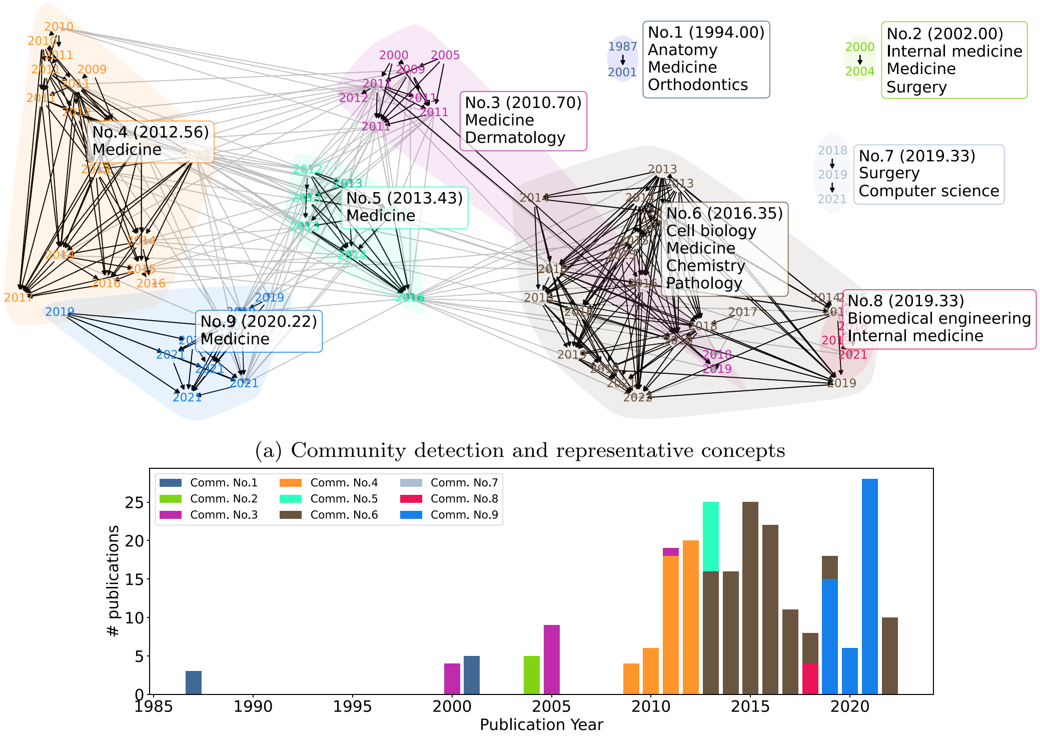
which can be used to measure the representativeness of top k concepts $\mathrm{TOP}(\pi, k)$ and determine the representative concepts. An example of calculating the WCR of top 3 concepts is shown in Fig. 3(a). This community contains 35 works, the top 3 most representative concepts collectively cover 31 works, yielding the $\mathrm{WCR}(\pi, 3) = 31/35 \approx 0.886$. Overall, we select the set of representative concepts (RC) of community π by selecting the minimum top k concepts $\mathrm{TOP}(\pi, k_{\min})$ that achieves a WCR τ of works in π, formally

$$\mathrm{RC}(\pi, \tau) = \mathrm{TOP}(\pi, k_{\min}), \quad k_{\min} = \min\{k \mid \mathrm{WCR}(\pi, k) \geqslant \tau\}. \tag{3}$$

Let X_k be the WCR achieved by the top k concepts in a randomly selected community. We further assess typicality of the top k concepts via the empirical survival function $S_k(\tau) = P(X_k \geqslant \tau)$, the proportion of communities whose top k concepts achieve a WCR of at least τ. As τ increases, the survival probability $S_k(\tau)$ for the top k concepts decreases monotonically. Table 4 reports $S_k(\tau)$ for $k = 1, \ldots, 5$ at thresholds $\tau \in \{0.4, 0.45, \ldots, 0.95\}$. For instance, $S_3(0.8) = 0.84$ indicates that the top 3 concepts survive in 84% of communities when covering at least 80% of the works. These top three concepts thus form a small yet highly representative sample. When the WCR increases to 0.85, the survival probability $S_3(0.85)$ decreases to 0.743. To assess the impact of the WCR threshold τ on the selected concepts $\mathrm{RC}(\pi, \tau)$, we perform a sensitivity analysis by varying τ from 0.6 (lenient) to 1.0 (stringent) in 0.05 increments ($\tau \in [0.6, 0.65, \ldots, 1.0]$), computing the resulting prerequisite relation AUC at each value (see Section 4.4.1 for details).

3.3.2. Co-citing network for a real-world author example

The CCN of a real author example is shown in Fig. 2, with 9 detected communities, where nodes in the same community are marked by a unique color. Each node (work) is labeled as its publication year, and each community is labeled as its APY, followed by several RCs around the center. These RCs characterize the community as a distinct research field. We sort the nine communities



(a) Community detection and representative concepts

(b) The number of publications of different communities by year

Fig. 2. The co-citing network of a real author.

by its APY in ascending order from top to bottom, which exhibit the shift dynamics of this author's research topics. The communities with different APY and RCs are shown in Fig. 2(a). For example, the first community π_1 is displayed as "No. 1 (1994.00)", where $\mathrm{APY}(\pi_1) = 1994.00$, and the RCs are "Anatomy", "Medicine" and "Orthodontics", respectively. Fig. 2(b) displays the number of publications per year for each research community (research field), where each color represents a specific community. We observe that research topics within each community tends to focus on specific time periods, comprising multiple works published in adjacent years. While the range of publication years across different communities shifts backward as the its APY increases, the distinct temporal gaps between different communities indicate that research topics evolve over time, revealing the shifts in focus across various periods shown in Fig. 2(b). For instance, the publications of community π_4 (No. 4) precede those of community π_5 (No. 5), followed by community π_6 (No. 6). The most recent works, located at the right end, include communities π_8 (No. 8) and π_9 (No. 9).

3.4. Temporal concept pairs

For any community (subset) $\pi \subseteq \mathcal{W}_i$, its average publication year (APY) is defined as

$$\mathrm{APY}(\pi) = \frac{1}{|\pi|} \sum_{w \in \pi} \mathrm{PY}(w), \tag{4}$$

where $\mathrm{PY}(w)$ is the publication year of work w. Let $\Pi_i = \langle \pi_{i1}, \dots, \pi_{ik_i} \rangle$ be the sequence of k_i communities detected from G_i ordered by ascending APY, so that $\mathrm{APY}(\pi_{i1}) \leqslant \mathrm{APY}(\pi_{i2}) \leqslant \cdots$, the indices thus reflect their temporal progression.

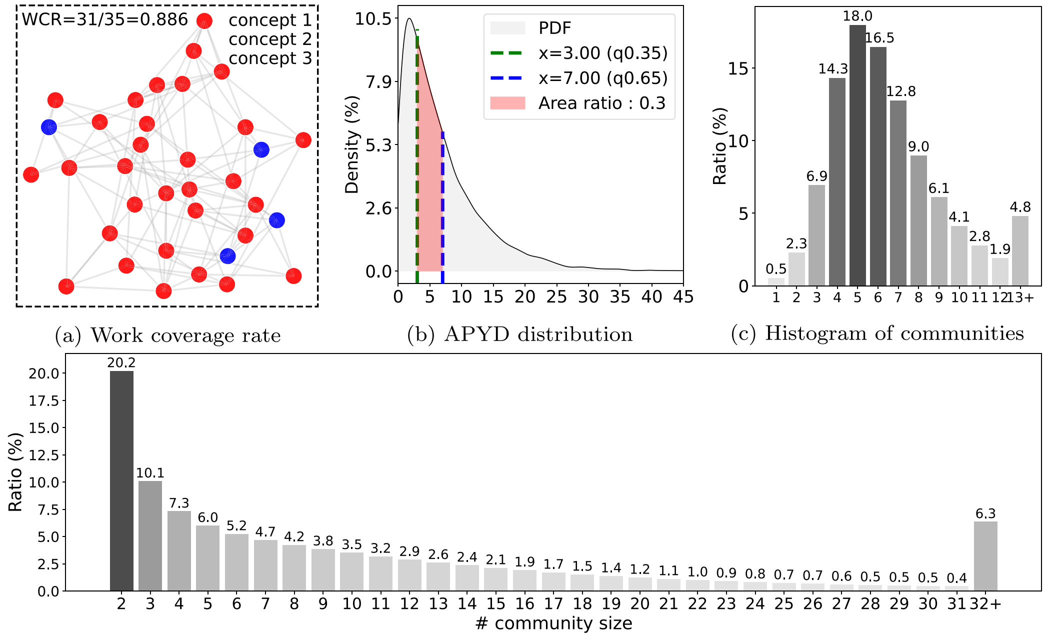
3.4.1. Construction of temporal concept pairs

Our objective is to reveal the precedence relationship between scientific concepts. For each author $a_i \in \mathcal{A}$, and any two distinct communities $r, s \in \Pi_i$, we define $r \prec s$ as a strict partial order on Π_i iff $\mathrm{APY}(r) < \mathrm{APY}(s)$, which captures the temporal progression among research topics that $\mathrm{RC}(r) \rightarrow \mathrm{RC}(s)$. We denote the community pairs (CP) of author a_i as

$$\mathrm{CP}(a_i) = \{(r, s) \in \Pi_i \times \Pi_i \mid r \neq s, r \prec s\}, \tag{5}$$

and then select the pair (r, s) whose average publication year difference (APYD) is moderate, neither too small nor too large, to adequately measure the natural shifts in the research topics. The APYD of community pair (r, s) is defined as:

$$\mathrm{APYD}(r, s) = \mathrm{APY}(s) - \mathrm{APY}(r). \tag{6}$$



(a) Work coverage rate (b) APYD distribution (c) Histogram of communities

(d) Histogram of community size

Fig. 3. Statistics of co-citing network communities.

To determine the suitable APYD interval, for each author a_i we compute $\mathrm{APYD}(r,s)$ for every ordered community pair $(r,s) \in \mathrm{CP}(a_i)$, and aggregating across all authors yields

$$\mathcal{D}_{\mathrm{APYD}} = \bigcup_{a_i \in \mathcal{A}} \{ \mathrm{APYD}(r,s) \mid (r,s) \in \mathrm{CP}(a_i) \}. \tag{7}$$

The empirical distribution of these differences is shown in Fig. 3(b). Let Q_x be the xth quantile of the APYD distribution. We select the pair (r,s) whose APYD falls within the interquartile range $\left[Q_{0.35}, Q_{0.65}\right]$ that encompasses 30% of the observed values, corresponding to $\mathrm{APYD}(r,s) \in [3,7]$ years.

For any author a_i and community pair $(r,s) \in \mathrm{CP}(a_i)$ with $\mathrm{APYD}(r,s) \in [3,7]$, we regard r as temporally preceding s, thus concept $c_i \in \mathrm{RC}(r)$ precedes concept $c_j \in \mathrm{RC}(s)$. The desired TCPs between a community pair (r,s) are obtained via Cartesian product, given by

$$\mathrm{TCP}(r,s) = \left\{ (c_i, c_j) \,\middle|\, c_i \in \mathrm{RC}(r), c_j \in \mathrm{RC}(s) \right\}. \tag{8}$$

In the CCN depicted in Fig. 2(a), research field No. 1 (π_1) encompasses concepts: Anatomy, Medicine, and Orthodontics, and research field No. 2 (π_2) comprises concepts: Internal Medicine, Medicine, and Surgery. Since $\mathrm{APYD}(s_1, s_2) = 6$, which falls within the range $[3,7]$, it satisfies the requirement for constructing the TCPs. Table B.5 lists the constructed TCPs of the CCN example in Fig. 2(a).

3.4.2. Statistical information of co-citing networks

Finally, we present the statistical information regarding the CCN G_i. We calculate the number of communities $|\Pi_i|$ of all authors, and the community size $|\pi|$ for each community. Our statistical analysis of the CCN G_i reveals key sociological insights into scientific knowledge production. As shown in Fig. 3(c), the number of communities per author $|\Pi_i|$ follows an approximately normal distribution centered at five (the most common value at 19.97%), with fewer authors participating in either few or many fields, and only 4.81% exceeding thirteen, demonstrating a long-tail effect that reflects broad interdisciplinarity and diverse research interests. Meanwhile, the distribution of community size $|\pi|$ in Fig. 3(d) follows a power-law pattern: most research directions are small, with 20.2% of author field consisting of just one or two works, and the proportion of fields steadily declines as the number of papers increases. While the majority of research fields contain only a few papers, there also exist a small proportion of large fields with high publication counts, as 6.35% of communities contain more than 32 papers. This indicates that scientific research is highly decentralized, characterized by the coexistence of numerous niche topics and a few dominant fields. Together, these patterns highlight the fragmented yet pluralistic structure of scientific research, where individual researchers are typically involved in several research fields, and most research fields are represented by only a small number of papers. This structure not only facilitates the integration and recombination of diverse knowledge, but also encourages cross-disciplinary exploration and supports the continual progression of scientific fields.

3.5. Optimization of knowledge precedence matrix

Let $C = \{c_1, c_2, \dots, c_N\}$ be the set of interested concepts (level 1 and level 0) across 19 disciplines. Chen, Lu, et al. (2018) define the prerequisite matrix as a binary matrix $\mathbf{M}$, where $m_{ij} = 1$ if concept c_i is a prerequisite for concept c_j, and 0 otherwise. In our work, we extend the binary limitation of the prerequisite matrix by defining m_{ij} as a continuous value representing the relative strength of concept c_i as a precedence for concept c_j. We construct a cross-domain knowledge precedence Matrix (KPM) $M_{N\times N}$ by transforming TCPs of around 5 million candidate authors. Each concept pair (c_i, c_j) contributes an increment of 1 to the matrix element m_{ij}, thus the ultimate m_{ij} represents the final frequency of pair (c_i, c_j), illustrating the transition pattern of the temporal knowledge flow among the group of scientists. We set the principal diagonal element $m_{ii} = 0$ to avoid self-circulation. For symmetrical elements m_{ij} and m_{ji}, we keep the larger one while setting the smaller to be 0, hence resulting in a directed graph without self-loops. The algorithm for constructing the KPM is shown in Algorithm 1.

The original KPM is complex with a large number of concepts, hence we focus on the sub-matrix concerning specific disciplines. Denote C_i as the set of concepts classified to discipline d_i, $N_i = |C_i|$ and $\mathcal{I}_i$ as the corresponding indices of concepts in C_i in $\mathbf{M}$. The sub-matrix from discipline $d_i \to d_j$ is extracted from $\mathbf{M}$ by selecting the corresponding indices $\mathcal{I}_i$ and $\mathcal{I}_j$ such that $\mathbf{M}^{sub} = \mathbf{M}\left[\mathcal{I}_i, \mathcal{I}_j\right]_{N_i\times N_j}$, meaning that concepts in C_i are precedence of concepts in C_j. Due to the large value discrepancies in $\mathbf{M}^{sub}$, we normalize each column to obtain $\mathbf{M}^{norm}$, ensuring the sum of each column $\sum_{k=1}^{N_i} m_{kj}^{norm} = 1$ for all j. We leverage the matrix elements of $\mathbf{M}^{sub}$ and its corresponding frequencies to determine the threshold of $\mathbf{M}^{sub}$ to select the corresponding values in $\mathbf{M}^{norm}$. Supposing that there are K unique values in $\mathbf{M}^{sub}$, let $x_1 > x_2 > \cdots > x_K$ be the sequence of ordered unique values in $\mathbf{M}^{sub}$, and let f_i be the frequency of x_i such that $\sum_{i=1}^{K} f_i = N_i \times N_j$. The threshold x_k $(k < K)$ is selected by retaining the appropriate top k unique elements. The optimized knowledge precedence matrix (OKPM) from $d_i \to d_j$ is obtained via

$$\mathrm{OKPM}(d_i, d_j) = \mathbf{M}^{norm}\left[\mathbf{M}^{sub} \geqslant x_k\right], \tag{9}$$

where the mask $\left[\mathbf{M}^{sub} \geqslant x_k\right]$ selects indices of elements in $\mathbf{M}^{sub}$ that are greater than or equal to x_k. These indices are used to extract corresponding values from $\mathbf{M}^{norm}$, which together form the final OKPM. In this way, $\mathrm{OKPM}(d_i, d_j)$ captures the significant values in $\mathbf{M}^{sub}$ as precedence relations, with the corresponding values in $\mathbf{M}^{norm}$ serving as their weights.

To determine the appropriate index $k < K$, we define the Cumulative Value Mass (CVM) and the Cumulative Frequency Mass (CFM) of the top k highest values respectively as

$$\mathrm{CVM}(k) = \frac{\sum_{i=1}^{k} x_i}{\sum_{i=1}^{K} x_i}, \quad \mathrm{CFM}(k) = \frac{\sum_{i=1}^{k} f_i}{N_i \times N_j}, \tag{10}$$

where the CVM indicates the importance of the top k sorted matrix values, and the CFM indicates the proportion of effective edges. We generally maintain a CVM $\geqslant 0.95$ and restricting the number of edges (CFM) to less than 20% of the total edges. The ultimate threshold index is $k = \min(k_1, k_2)$, where

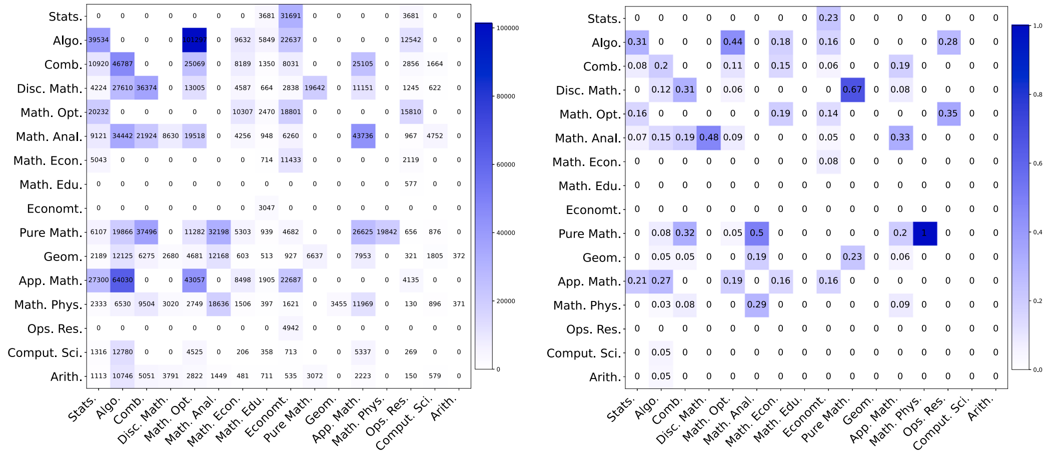
$$k_1 = \min\{k \mid \mathrm{CVM}(k) \geqslant 0.95\}, \quad k_2 = \max\{k \mid \mathrm{CFM}(k) \leqslant 0.2\}. \tag{11}$$

Based on extensive sensitivity analysis across multiple disciplines, we set CVM $\geqslant 0.95$, which ensures that we capture the most significant knowledge flows while filtering out noise. The CFM upper bound of 0.2 was chosen to maintain network sparsity and interpretability. This optimization process is outlined in Algorithm 2.

We demonstrate the matrix optimization for Mathematics as an illustrative example in Fig. 4. The original matrix represents knowledge flow among concepts in Mathematics shown in Fig. 4(a), which is complex and cluttered with numerous insignificant elements, obscuring the essential relationships. The refined network weight matrix is presented in Fig. 4(b), which serves as the foundation for constructing the precedence network among concepts of Mathematics. This optimization not only simplifies the matrix, but also accentuates the most significant precedence order, facilitating a clearer understanding of the knowledge structure within the discipline.

4. Experiments and ablation study

In this section, we first assess alignment with LLM prerequisite benchmarks. Second, we analyze four subfields, evaluating authors' precision both within their own subfield and across the other three. This cross-validation demonstrates the robustness of our method: authors achieve high precision within their own domains but considerably lower precision in others, supporting the method's validity and explaining the lower recall observed in cross-domain scenarios. Third, we conduct ablation and sensitivity analyses, including examining the sensitivity of WCR to investigate how varying WCR thresholds affect KPN construction and author-level precision. We also assess KPNs across different career stages. Our findings indicate that Early-career KPNs show lower recall but higher precision, indicating a focused, prerequisite-consistent exploration strategy, and display relatively higher alignment with prerequisite relationships compared to those in the Mid-career or Late-career. Finally, we analyze the ROC curve to evaluate the overall performance of the obtained KPM matrix.



(a) Original KPM (b) KPM optimization

Fig. 4. KPM construction and optimization in Mathematics (row concepts precede column concepts).

Table 5
Optimal hyper-parameter settings.

Param	WCR	AveDegree	LouvResolution	CVM	CFM
Setting	1.0	log	1.0	0.95	0.2

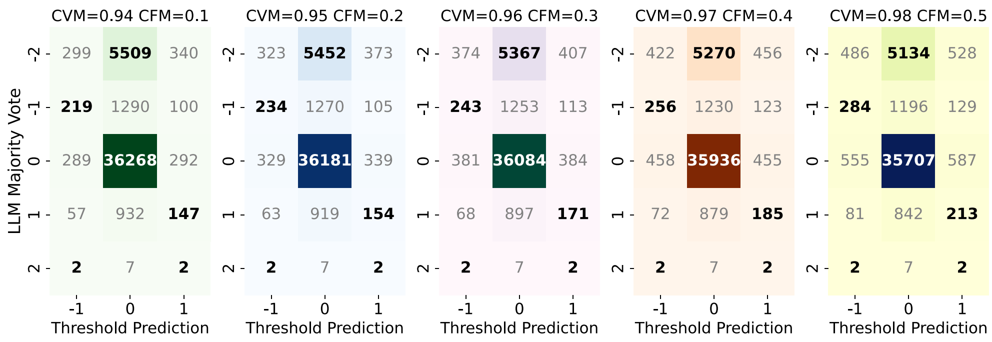
4.1. Parameter settings

We present the jointly tuned optimal hyperparameter settings for final KPN construction obtained via an anchor-based bidirectional scanning procedure as shown in Table 5: WCR = 1.0; size-adaptive average degree $\langle k\rangle = \log(10n)$ (natural log, n is the number of concepts); Louvain resolution = 1.0; CVM = 0.95; CFM = 0.2. These settings are selected for cross-discipline structural stability: WCR 1.0 maximizes transition signal recall (25.77% in Table 6); the logarithmic degree target prevents over-density in large fields and fragmentation in small ones; resolution 1.0 balances semantic over-merging (< 0.8) and unstable micro-communities (> 1.2); CVM 0.95 with CFM 0.2 preserves dominant temporal flow edges while filtering low-confidence links. This reproducible joint configuration is a global compromise rather than a concatenation of local optima.

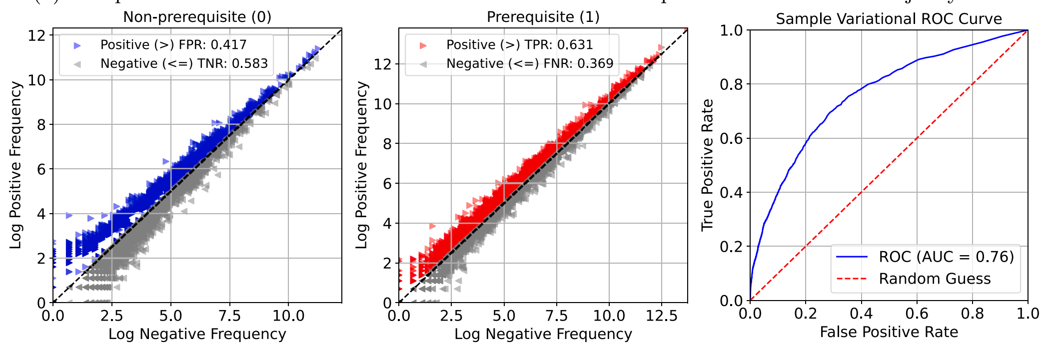
We applied an anchor-based bidirectional scanning procedure, and we justify the selection principles and reason of these parameter settings as follows: (1) **WCR Thresholds**: We first inspected the survival probabilitys of top k concepts (Table 4), and chosen a relative high baseline WCR = 0.8, as the retained top k concepts exhibit consistently high survival probability, indicating stable representativeness. We expanded the search symmetrically to the range 0.6–1.0 (Table 6), with WCR = 1.0 yields the highest prerequisite recall and adopted for subsequent sensitivity analyses. (2) **Average Degree** $\langle k\rangle$: we select from literature reports workable mean degrees in the interval 4–10 (see Appendix B), and we adopted $\langle k\rangle = 8$ as a central setting and probed sensitivity with $\langle k\rangle \in \{4, 6, 8, 10, 12\}$, plus the dynamic $\log(10n)$ formulation (Eq. B.1), to cover both fixed and scale-adaptive regimes (Eq. B.2). (3) **Louvain Resolution**: Starting from the commonly used default 1.0, we varied the resolution parameter over a balanced range (0.4–1.6, step 0.2), with $\langle k\rangle = 8$ fixed, to assess community granularity effects on prerequisite recall without overfitting to a single partition scale. (4) **KPM Optimization**: We scanned the narrow high-specificity band 0.90–0.95 as the CVM is intended to retain significant elements. CFM governs the proportion of maximum edges, we explored 0.10–0.50 to span from strict pruning to moderately permissive linkage density. Overall, we select CVM = 0.95 and CFM = 0.2 based on the sensitivity analysis (see Figure B.3).

4.2. Alignment with LLM prerequisite benchmarks

Our analysis underscores the dynamic and complex nature of scientific discovery, as evidenced by our evaluation of the alignment between KPNs and SciConPreq benchmarks through theoretical prerequisite relation recall. The alignment results of our KPN and LLM-derived CPN are shown in Fig. 5. The LLMs majority vote annotations assign an integer in $\{-2, 1, 0, 1, 2\}$ for each pair (A, B) of concepts as shown in Table 1. Our network-based method compare the corresponding positive concept pair (A, B) and negative concept pair (B, A) frequencies, resulting an integer in $\{-1, 0, 1\}$. Given the absence of prior research on evaluating knowledge temporal patterns, we compare the precedence relations derived from co-citing analysis and theoretical concept prerequisite relations annotated by LLMs to calculate the theoretical prerequisite relation recall of the resulting KPN. We hope this initial assessment will lay the groundwork for future research in this direction.



(a) Comparison of confusion matrices across different threshold predictions and LLM majority votes



(b) LLM non-prerequisites (c) LLM prerequisites (d) Sample variational ROC curve

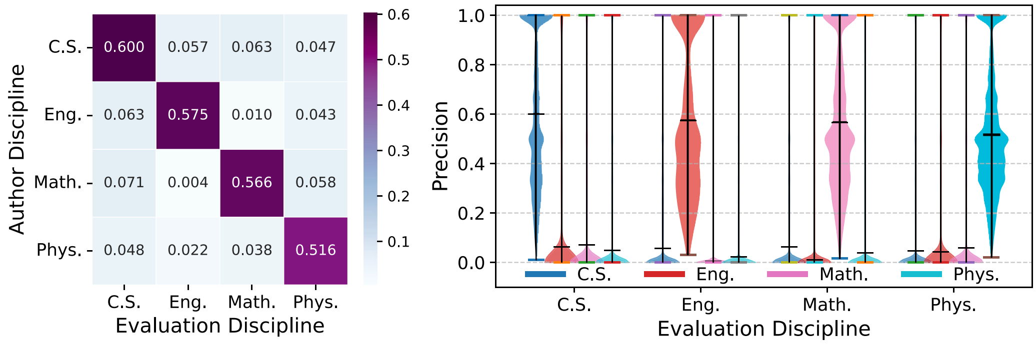
Fig. 5. Assessment of KPN alignment with LLM annotations.

Fig. 5(a) presents the confusion matrices under different CVM and CFM conditions. In the first matrix (CVM = 0.94, CFM = 0.1), our network-based approach with matrix optimization successfully identifies the majority of unrelated concept pairs (36,268) as 0. However, a substantial portion of prerequisite relations are unrelated in the macro-level knowledge progression patterns. Without CCN refinement, only 12.66% of LLM prerequisite relations are recovered. After optimization with the CCN through concept semantic similarity, the recall increases to just 25.77% in best case. This finding suggests that the evolution of scientific knowledge is highly dynamic and often diverging from conventional prerequisite structures, revealing a previously underappreciated complexity in scientific knowledge progression. While the extreme imbalance between prerequisite and non-prerequisite concept pairs in our dataset and the unsupervised nature of our approach partly explain the lower recall, the more fundamental reason may lie in the striking discrepancy between theoretical knowledge prerequisites and actual knowledge progression patterns in the scientific community. We later investigate in detail the low alignment through author-level precision in Section 4.5.

Figs. 5(b) and 5(c) illustrate the overall alignment without matrix optimization, where Fig. 5(b) displays all concept pairs without prerequisite relations, and Fig. 5(c) shows all concept pairs with prerequisite relations. Each triangular point represents a concept pair (A, B), the y-axis denotes the logarithm of the frequency of the concept pair with positive relation $(A \rightarrow B)$, and the x-axis denotes the logarithm of the frequency of concept pair with negative relation $(A \leftarrow B)$. The comparison results in an overall false positive rate (FPR) of 41.7% shown in Fig. 5(b) and an overall true positive rate (TPR) of 63.1% shown in Fig. 5(c).

To evaluate the model performance under different parameters, we attempt to construct a ROC-like curve by varying the CFM threshold. Unlike traditional methods that generate continuous prediction scores for each concept pair, our approach determines prerequisite ($-1/1$) and non-prerequisite (0) based on frequency patterns, making conventional threshold-based ROC analysis inapplicable. Instead, we utilize the CFM as a threshold parameter in Eq. (10) to control the fraction of significant edges, and further calculate the corresponding TPR and FPR, hence we called it the sample variational ROC curve. A higher CFM threshold selects the most frequent concept pairs, while a lower threshold gradually includes pairs with weaker precedence patterns. The resulting ROC curve is shown in Fig. 5(d) with an AUC of 0.76, indicating a moderately strong ability to distinguish prerequisite relations from incorrect relations. The curve illustrates the trade-off between true positive and false positive rates across different sample sizes, using frequency thresholds on varying edge subsets instead of prediction scores on a fixed dataset. This approach preserves the TPR–FPR trade-off principle used to assess discrimination.

Although some discrepancies exist with theoretical prerequisite relationships, the resulting KPNs effectively capture the precedence order of knowledge advancement and the hierarchical dependencies within scientific communities. These KPNs reflect actual scientific practices and provide significant value for research guidance, particularly for young scholars, by offering them meaningful insights.



(a) Mean precision heatmap.

(b) Violin plots of cross-domain precision evaluation.

Fig. 6. The analysis of cross-discipline precision validation. (a) Heatmap showing mean precision values for each author-evaluation discipline pair. (b) Violin plots showing the corresponding distribution of precision values across evaluation disciplines, with colors indicating author disciplines.

4.3. Validation of cross-disciplinary precision

In the above section, we obtain the results with low recall of LLM prerequisite relations from KPNs through alignment analysis. To under stand this discrepency and the low recall from KPN, we further analyze the author-level precision, which give our meta understanding of the individual perspective. In addition to the knowledge transition pattern across 19 disciplines, we further refine our KPN analysis within each discipline Computer Science, Engineering, Mathematics, and Physics.

For an author a_i whose primary discipline is d_i, we evaluate their precision in topic transitions that align with prerequisite relations within the sub-concepts of any discipline d_j (including their own). Specifically, the author-level precision of a_i in discipline d_j is defined as

$$\Pr(a_i, d_j) = \frac{TP}{TP + FP}, \tag{12}$$

where TP (true positives) is the number of TCPs of a_i in d_j that correctly match the LLM-annotated prerequisite relations, and FP (false positives) is the number of pairs that do not match. This metric enables us to assess how well the topic transitions of authors from any discipline adhere to the prerequisite structure within any given discipline.

Further, for each discipline d_i, we consider the set of all authors whose primary discipline is d_i. The collection of their precision scores for topic transitions within discipline d_j can be represented as:

$$\mathcal{P}(d_i, d_j) = \{\Pr(a, d_j) \mid a \in \mathcal{A}(d_i)\}, \tag{13}$$

where $\mathcal{A}(d_i)$ denotes the set of authors whose primary discipline is d_i, and $\Pr(a, d_j)$ is the precision of author a in discipline d_j. This formulation allows us to analyze the distribution of prerequisite aligned topic transition precision for authors from discipline d_i when evaluated on discipline d_j.

The results on the above four disciplines are shown in Fig. 6. Fig. 6(a) presents a heatmap where rows represent the authors' primary disciplines and columns represent the evaluated disciplines. For example, in the first column, the first row shows that Computer Science authors achieve an average precision of 0.6 for prerequisite relations among 250 C.S. concepts. In contrast, Engineering, Mathematics, and Physics authors have much lower average precisions on C.S. concepts: 0.063, 0.071, and 0.048, respectively. The distribution of these precision scores is further illustrated in Fig. 6(b), where the first four violin plots with different color correspond to the precision distributions of authors from each discipline on C.S. concepts. The x-axis in Fig. 6(b) indicates the evaluated discipline, and colors distinguish authors' primary disciplines. Notably, only C.S. authors exhibit a well-distributed precision on C.S. concepts, while authors from other disciplines show precision distributions concentrated at lower values. The average precision values and corresponding precision distributions of authors from other disciplines on all four evaluated disciplines can be found in Fig. 6(a) and Fig. 6(b), respectively.

The validation of cross-disciplinary precision demonstrates that authors achieve higher precision on prerequisite relations within their own discipline, confirming that our method aligns with real-world expertise boundaries. This also indicates that the TCPs identified by our approach exhibit strong domain specificity, with high average precision values reflecting the overall accuracy of our extraction process. Conversely, the consistently low precision of authors on out-of-domain concepts further verifies that our method is unbiased in extracting TCPs across disciplines. Together, these results enhance the reliability and credibility of our approach.

Table 6
Performance of prerequisite extraction at various WCR.

WCR	60%	65%	70%	75%	80%	85%	90%	95%	100%
Precision	22.21%	22.18%	21.16%	20.99%	20.38%	19.19%	18.67%	17.65%	16.29%
Recall	11.60%	12.07%	12.58%	13.12%	**14.20%**	15.65%	17.93%	21.11%	25.77%

Table 7
Performance of prerequisite relation extraction at different career stage.

Career	Early (1–10)	Middle (11–20)	Late (21+)
Precision	**21.91%**	19.60%	16.23%
Recall	16.73%	**19.77%**	18.61%

4.4. Ablation study and sensitivity analyses

In this section, we conduct a comprehensive series of sensitivity analyses to evaluate the robustness and reliability of our KPN construction. Specifically, we examine how varying the WCR, career stage, semantic similarity thresholds, and community detection resolution parameters affect the precision and recall of prerequisite relations. By systematically assessing the impact of these parameters, both at the cross-discipline and sub-discipline levels, and benchmarking against a citation-based knowledge network, we demonstrate that our approach maintains stable and competitive performance across a wide range of settings. These analyses provide important insights into the methodological stability of our work and offer practical guidance for parameter selection in future applications.

4.4.1. Sensitivity analyses of WCR

To ensure the completeness of TCPs, we conducted an additional sensitivity analysis to examine how different WCR thresholds influence our results. Initially, we selected concepts that collectively cover 80% of the works as representative concepts. However, without comparison, the effectiveness of our choice for KPN construction and its robustness in matching LLM-annotated concept pairs remains unclear. Therefore, we introduced a range of coverage thresholds from 60% to 100% in increments of 5% to systematically analyze how varying WCR values affect KPN construction and its alignment with LLM results. This approach allows us to evaluate the robustness of our method and to provide a comprehensive reference for future studies. We assess the KPNs with Precision and Recall of LLM prerequisite relations under different WCR, the results are shown in Table 6. Additionally, the performance of prerequisite relation extraction for sub-discipline concepts under different WCRs is presented in Table B.3. As shown in Table B.4, the proportion LLM-labeled positive pairs increases from 19% (Engineering) to 33.31% (Physics), which partially explains the low recall observed in Table B.3. For the Computer Science discipline, recall rises from 1.43% to 7% as the WCR increases from 60% to 100%, while precision varies from 46.75% to 40.93%. Notably, this precision remains consistently higher than the cross-domain performance reported in Table 6.

Table 6 shows the performance of prerequisite extraction under different WCR thresholds (CVM = 0.95, CFM=0.2). To further compare the accuracy level, we constructed a citation-based network and optimized it using the same parameters. The results show a precision of 26.6%, which is slightly higher than the 25.77% achieved by our method. However, the recall is 15.16%, which remains consistently lower than the values observed across different WCRs. As the WCR threshold increase from 60% to 100%, precision gradually decrease from 22.21% to 16.29%, while recall increase from 11.60% to 25.77%. At WCR = 100%, the method achieves the highest recall (25.77%) but at the cost of lower precision (16.29%), as more non-prerequisite transitions are included. Overall, precision remains within a relatively narrow range (16.29% to 22.21%) across all WCR thresholds, indicating that many scientists' topic transition patterns do not align with LLM-identified prerequisite relations. Nevertheless, under more relaxed WCR conditions, a higher proportion of concept transitions are identified as prerequisites. Based on a trade-off between precision and recall, we select WCR = 0.8 as the threshold for presenting our main results.

4.4.2. KPN analysis across career stages

In this section, we analyze the Knowledge Prerequisite Networks (KPNs) constructed for three distinct career stages: Early-career (years 1–10), Mid-career (years 11–20), and Late-career (years 21+). We assess variations in author-level precision and recall, as well as the alignment between the resulting KPNs and LLM-annotated concept pairs. Table 7 presents the evaluation results for KPNs at each career stage. As shown, Early-career researchers achieve the highest precision (21.91%), indicating that their topic transitions are more focused and closely follow LLM-annotated prerequisite relationships. However, their recall is 16.73%, suggesting that Early-career researchers engage with fewer topics and transitions overall. In the middle stage, recall increases to 19.77%, while precision decreases to 19.6%, implying that topic transitions become more frequent but less strictly aligned with prerequisite relations. In the Late-career, both recall (18.61%) and precision (16.23%) decline, indicating that topic transitions become less frequent and less likely to adhere to prerequisite patterns. Additionally, the performance of prerequisite relation extraction for sub-discipline concepts across different career stages is presented in Table B.4. Similar to the results shown in Table 6, recall is lower due to the increased number of ground truth pairs. However, the precision remains consistently higher than that observed in the cross-domain setting, as reported in Table 7.

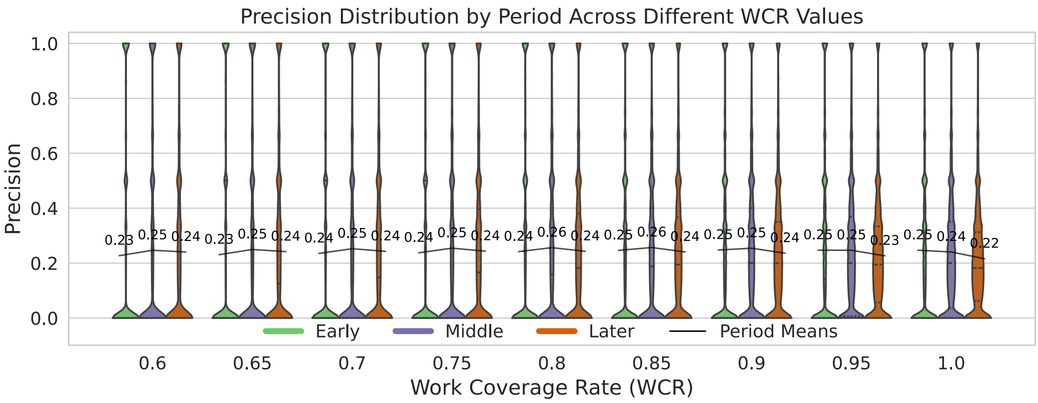


Fig. 7. Violin plots of precision distribution for Early, Middle, and Late career stages under different WCR thresholds.

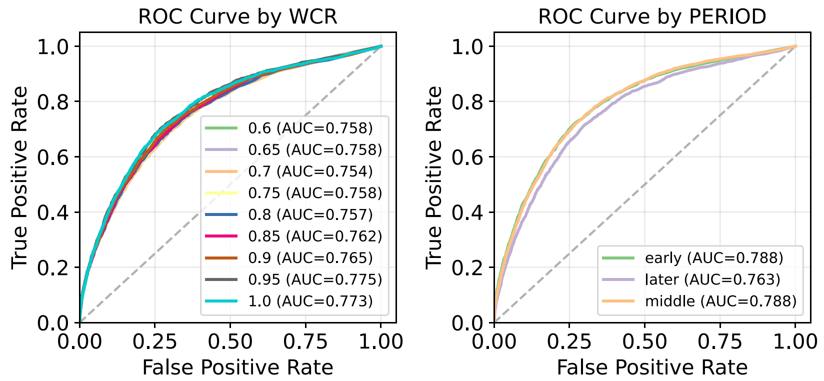
By comparing these metrics across career periods, we aim to reveal how knowledge transition with prerequisite relationships differ throughout an academic career. Table 7 demonstrates variations in topic transitions across different career stages, while Fig. 7 further examines the distribution of precision under varying WCR thresholds. The higher precision of 21.91% in Table 7 is observed in Early-career researchers, further validated by the violin plots in Fig. 7, on average, precision values for Early-career, Mid-career, and Late-career stages are relatively similar. However, as the WCR threshold becomes more relaxed, allowing more concept pairs to be included, a slight and indistinct pattern emerges: Early-career transitions exhibit higher mean precision when WCR is 1. The higher precision observed in the early career stage can be attributed to researchers' limited experience and more focused, systematic research trajectories, they typically concentrate on a narrower range of topics in Early-career within well-defined domains, following established academic pathways that closely align with conventional prerequisite structures, which naturally leads to stronger adherence to prerequisite relationships as they build foundational expertise systematically.

Consequently, when stricter criteria (e.g., WCR = 0.6) are applied, many of their transitions are excluded from the analysis. As the criteria become more relaxed (e.g., WCR = 1), more of their transitions are included, and since most of these transitions are aligned with prerequisite structures, their mean precision becomes slightly higher than that of Mid-career researchers when WCR = 1. The Early-career transition precision is higher than Mid-career and Late-career, further supporting the conclusion. Therefore, the observed fluctuation in precision across WCR thresholds actually supports the idea that Early-career researchers' transitions are fewer and align more with prerequisite relations, thus more sensitive to the selection criteria of WCR. This pattern implies that KPNs can serve as reliable guides for Early-career scholars by identifying topics that constitute essential milestones or precedence steps in their academic trajectories, thereby improving the efficiency of academic development. In the Late-career, both precision and recall decline, possibly reflecting either a stabilization of interests or more unconventional topic shifts. These findings highlight the dynamic nature of knowledge development and suggest that prerequisite relationships shift in complexity as careers advance, offering valuable insights for academic career planning and support.

4.4.3. ROC analysis across WCRs and career stages

Fig. 8 presents ROC curves evaluating the performance of KPN prerequisite relation extraction across 303 concepts from 19 disciplines. In Fig. 8(a), we compare ROC curves under different WCR thresholds. The AUC increases slightly from 0.758 at WCR = 0.6 to 0.773 at WCR = 1.0, indicating that while relaxing the WCR threshold includes more transitions (increasing recall but reducing precision), the overall model performance remains relatively stable but with a subtle improvement. This suggests that the method is robust to WCR threshold variation, likely because most topic transitions does not correspond to true prerequisite relations. Nevertheless, threshold selection is still important in practice. In Fig. 8(b), the ROC curves across career stages (Early, Middle, Late) show that AUC is highest in the Early-career and Mid-career (up to 0.788), with the Late-career slightly lower at 0.763. This indicates that, although individual-level differences may exist, the model's performance is consistent across career stages at the group aggregate level.

Fig. 9 presents ROC curve comparisons for four sub-disciplines under varying WCR thresholds (first row) and across different career stages (second row). In the first row, each subplot shows the ROC curves for a specific sub-discipline (e.g., Computer Science) evaluated under WCR thresholds ranging from 0.6 to 1.0. Compared to the cross-disciplinary analysis, the number of LLM-annotated prerequisite pairs within each sub-discipline is substantially higher. As a result, if researchers' topic transitions within these sub-disciplines are not representative or the transition effects are weak, the model may fail to effectively recall the annotated prerequisite relations. It is important to note that LLM annotations serve only as a reference—actual topic transitions by researchers are driven by their research needs rather than strict prerequisite logic. Thus, while theoretical prerequisite structures may be clear within a discipline, real-world topic transitions may not strictly follow them. This can lead to greater variability in ROC curves across different WCR thresholds at the sub-discipline level. Overall, all four disciplines exhibit variations in ROC performance with changing WCR, and the AUC generally increases as the WCR threshold becomes more relaxed, indicating that the method maintains robustness in recalling prerequisite relations at the sub-discipline level.



(a) ROC curve by WCR across disciplines(b) ROC curve by period across disciplines

Fig. 8. ROC curve comparison across disciplines: (a) by wcr, (b) by period.

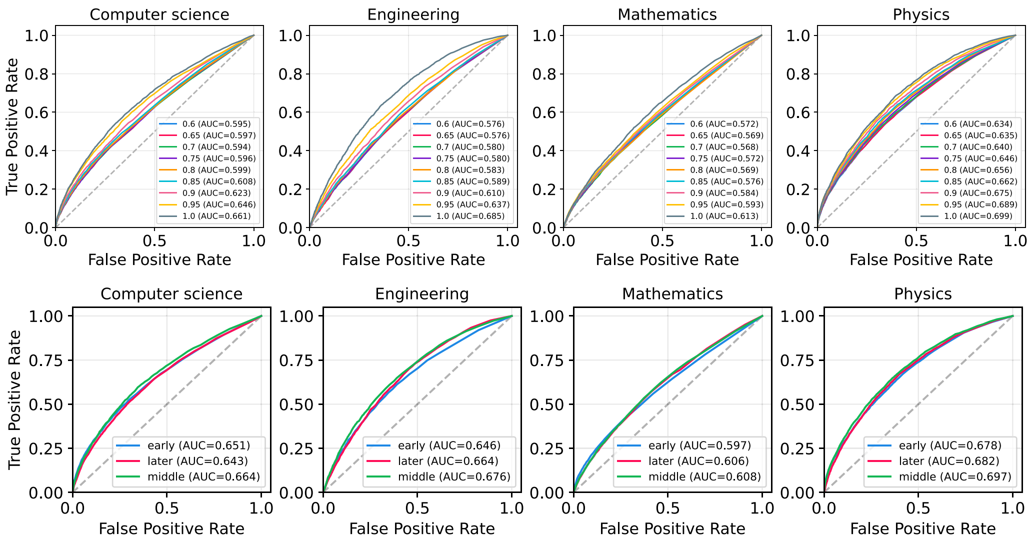


Fig. 9. Comparison of sub-discipline ROC curves: (top) by WRC, (bottom) by Period.

The second row displays the ROC curves for the four sub-disciplines across different career stages. For Computer Science, Early-career researchers achieve higher performance than late-career researchers, but both are outperformed by Mid-career researchers. For Engineering, Mathematics, and Physics, Early-career researchers show the lowest recall robustness, with late-career researchers performing better, and Mid-career researchers achieving the highest AUC. This suggests that topic transitions are most pronounced during the Mid-career stage across these disciplines.

In summary, Figs. 8 and 9 systematically analyze the robustness and variability of prerequisite extraction results under different WCR thresholds and career stages. Overall, the method demonstrates greater robustness at WCR = 1.0. In terms of career stages, Mid-career researchers exhibit the richest and most pronounced topic transitions, leading to higher recall of prerequisite relations. Importantly, our primary objective is not merely to recover prerequisite relations, but to uncover the underlying patterns in scientists' topic transitions. The use of LLM-based evaluation provides a useful lens for understanding the KPNs of researchers. By comparing these differences, we gain valuable insights into how scientists navigate and transition between research topics.

4.4.4. Complementary knowledge citation network

To further validate the precision and recall, we construct a Knowledge Citation Network (KCN) (Yang, Wu, & Lyu, 2024) based on the same dataset as KPN. For each author a_i, and each of their published work $w_a \in \mathcal{W}_i$, we enumerate each citation $w_a \to w_b$ where w_a (citing) references w_b (cited). For each citation we define the citation concept pair (CCP) set as

$$\mathrm{CCP}(w_a, w_b) = \left\{ (c_i, c_j) \,\middle|\, c_i \in C(w_b), c_j \in C(w_a) \right\}. \tag{14}$$

We then construct the KCN by aggregating the frequencies of these CCPs across all authors and all citation events. This process results in a directed, weighted network, where nodes represent concepts, and the weight of each directed edge reflects the number of times a specific concept pair appears in citation relationships. This author-centric approach enables us to systematically analyze how knowledge flows between different concepts through the citation activities of scientists.

The resulting KCN attains a prerequisite recall of 26.6% (precision 15.16%), slightly above (i) the KPN's best configuration (recall 25.77%, precision 16.29%; Table 6) under fully inclusive WCR = 1.0 with logarithmic average degree and resolution = 1.0, and (ii) the local structural optimum from the sensitivity analysis (recall 25.73%, precision 17.80%) at fixed average degree = 5 and resolution = 1.4 (Tables 8 for average-degree sensitivity and 9 for Louvain-resolution sensitivity). Compared to topic transition behavior in career trajectories, the KCN exploits explicit citation links as direct signals of knowledge inheritance, naturally attaining a slightly higher prerequisite recall (26.6%) than the local optimum from our KPM parameter sensitivity analysis (25.73%). This modest gain suggests that citation-based pairing approximates a practical empirical upper bound on recall for purely structure-driven approaches. However, this broader coverage incurs a precision cost: the KCN's precision (15.16%) is consistently lower than that of KPM across parameter settings, reflecting that many significant citation links encode methodological reuse or thematic associations rather than strict prerequisite relations. Overall, the uniformly low recall across KCN and KPM approaches indicates that canonical prerequisite ordering is only weakly instantiated in real-world scientific practice, reinforcing the suitability of our KPN method and underscoring the intrinsically complex and creative nature of scientific knowledge evolution.

4.4.5. Sensitivity of semantic thresholds and Louvain resolution

During the refinement of the CCN, our model constructs CCNs by combining co-citing relationships and semantic similarity between their publications. The co-citing relationships provide the base edges, and the additional edges based on the semantic similarity between papers are added. Instead of directly setting a semantic similarity threshold, we select the threshold that results in the desired average degree for the similarity matrix. This ensures that the density of the refinement is controlled by the target average degree, rather than by an arbitrary similarity cutoff. The final network thus combines the co-citing structure with a controlled number of similarity-based refinement edges. To evaluate the robustness of our KMP generation approach, we conducted a comprehensive sensitivity analysis examining how two key parameters affect the precision and recall of prerequisite detection: the average degree of the work similarity matrix and the Louvain resolution parameter for community detection.

The semantic similarity threshold is determined based on either fixed average degree constraints (4, 6, 8, 10, 12) or a dynamic logarithmic average degree (defined in Eq. B.1 and denoted as "log"). Community detection is then performed using the Louvain algorithm with varying resolution (from 0.4 to 2.0 with a step size of 0.2). We then construct the corresponding KPMs under these parameter combinations and evaluate their precision and recall to assess parameter sensitivity. Based on the CVM and CFM threshold settings with Eq. (11), Fig. 10 presents the parameter sensitivity analysis for community detection performance.

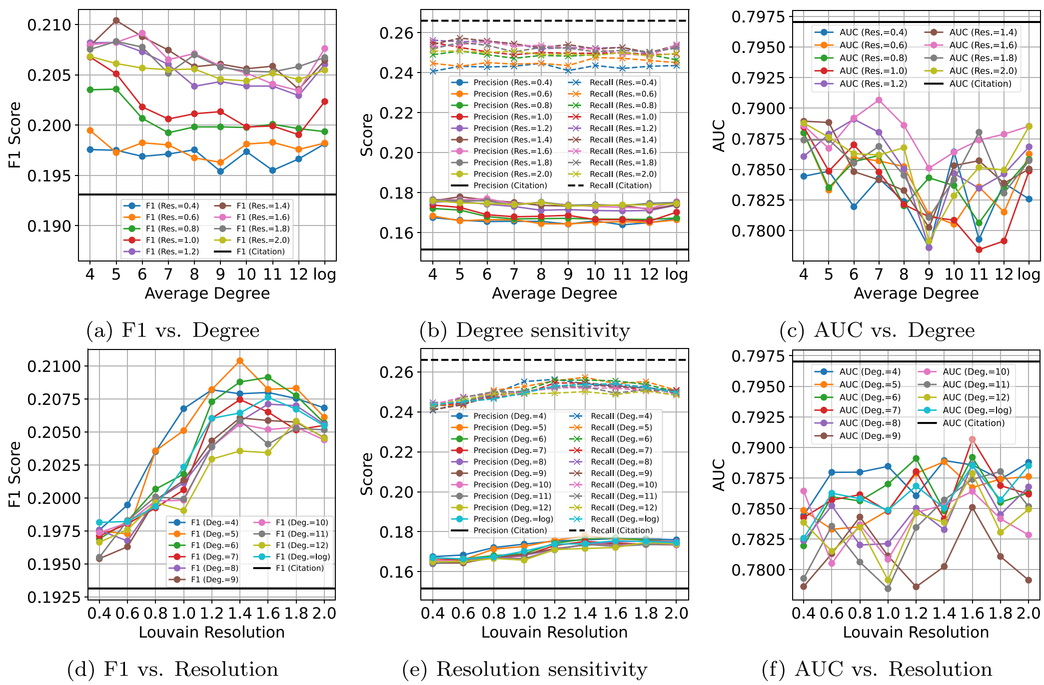
Figs. 10(a) and 10(d) show the F1 scores of the KPMs under various parameter combinations. Specifically, Fig. 10(a) plots F1 score versus average degree, with each curve representing a different resolution. The black horizontal line indicates the baseline F1 score of the KCN described in Section 4.4.4, demonstrating that our method consistently outperforms KCN across all parameter settings. The results indicate that the average degree has a limited effect on F1 score, while variation in the resolution parameter leads to more pronounced differences, suggesting that performance is more sensitive to resolution than to average degree. This observation is further confirmed in Fig. 10(d), where F1 score is plotted against the Louvain resolution parameter for different average degrees. Here, F1 score increases with resolution, reaching a peak around 1.4, and then gradually declines.

Figs. 10(b) and 10(e) provide a more detailed breakdown of precision and recall under the same parameter settings of Figs. 10(a) and 10(d). Solid lines indicate precision, while dashed lines represent recall. Both metrics show low sensitivity to average degree, as curves for different resolutions are nearly flat, but exhibit greater variation across different resolutions. In Fig. 10(e), recall slightly increases with resolution, peaking near 1.4, and then stabilizes. Precision curves in Fig. 10(e) also show a slight increase with resolution, attaining its highest level around 1.4 before gradually declining. These results indicate that resolution has a stronger influence on both precision and recall than average degree. Compared to the KCN, our approach achieves consistently higher precision, while KCN tends to have slightly higher recall but substantially lower precision. This suggests that our CVM and CFM threshold provides a more stable and balanced overall performance.

Figs. 10(c) and 10(f) assess overall performance using the area under the ROC curve (AUC), calculated across all parameter combinations rather than fixed CVM and CFM thresholds. KCN achieves relatively higher AUC values. In Fig. 10(c), our method attains the highest AUC at resolution 1.6 across all average degrees, a trend that is also evident in Fig. 10(f). However, there is no consistent pattern with respect to average degree, though the combination of average degree = 7 and resolution = 1.6 yields the highest AUC. In summary, these results demonstrate that the method is robust to changes in average degree but more sensitive to the resolution parameter, with optimal performance typically achieved at a resolution value around 1.4–1.6.

Based on the results in Figs. 10(b) and 10(e), the optimal combination of average degree (5) and resolution (1.4) yields a local optimum recall of 25.73% and a precision of 17.8%. Tables 8 and 9 summarize the top-performing parameter settings. Specifically, Table 8 shows that, with resolution fixed at 1.4, precision ranges from 16.58% to 17.38% and recall from 24.9% to 25.51% across different average degrees; the dynamic logarithmic threshold ("log") also performs competitively. Table 9 reports that, with average degree fixed at 5, precision varies from 16.45% to 17.5% and recall from 24.47% to 25.37% across different resolutions. Although KCN's recall (26.6%) is slightly higher, our method maintains stable performance across parameter settings, consistently achieving higher precision than the KCN baseline (15.16%) while attaining comparable recall. Higher resolution values (e.g., 1.4, 1.6) tend to improve precision, suggesting that finer community granularity may better capture prerequisite relationships. Our approach thus demonstrates robustness and outperforms the citation-only baseline in precision while maintaining competitive recall.

In summary, our sensitivity analysis demonstrates that the KPM method is robust to changes in average degree and more sensitive to Louvain resolution, with optimal performance typically achieved at a resolution around 1.4–1.6. Across all tested parameter



(a) F1 vs. Degree (b) Degree sensitivity (c) AUC vs. Degree

(d) F1 vs. Resolution (e) Resolution sensitivity (f) AUC vs. Resolution

Fig. 10. Parameter sensitivity analysis for community detection performance. (a) and (d) show F1 scores across degree and resolution parameters. Black horizontal lines indicate the baseline performance using KCN. (b) and (e) show precision and recall metrics across different degree and resolution parameters, with solid lines representing precision and dashed lines representing recall. (c) and (f) display AUC values across degree and resolution parameters respectively. Different colored lines represent various parameter combinations, demonstrating the robustness of across parameter space.

Table 8
Precision and recall of the KPMs with optimized threshold under different average degree settings (4 to 12, and log) with a fixed resolution parameter of 1.4.

Degree	4	5	6	7	8	9	10	11	12	log	KCN
Precision	17.59%	17.8%	17.63%	17.51%	17.36%	17.33%	17.37%	17.37%	17.16%	17.4%	15.16%
Recall	25.41%	25.73%	25.59%	25.44%	25.26%	25.41%	25.19%	25.26%	25.01%	25.37%	26.6%

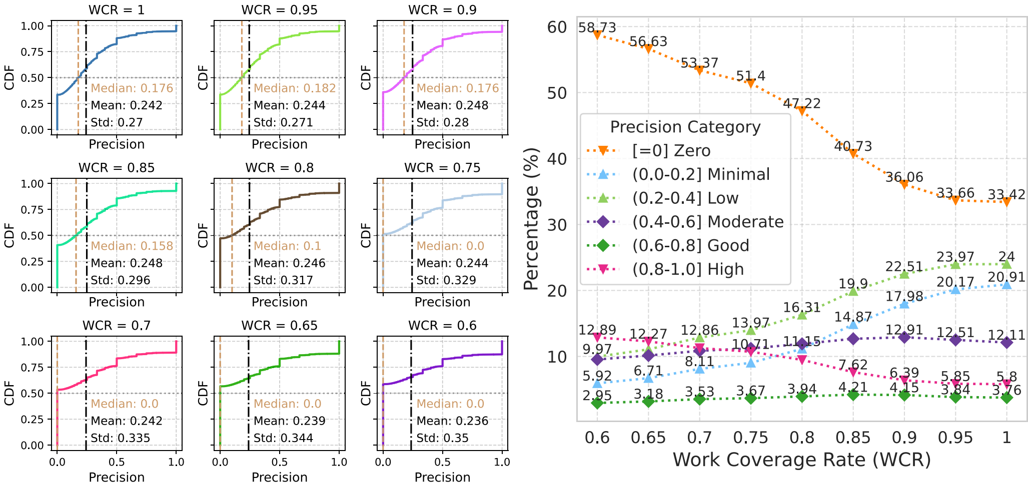
Table 9
Precision and recall of the KPMs with optimized threshold under different resolution parameter settings (0.4, 0.6 to 2.0) with a fixed average degree of 5.

Resolution	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0	KCN
Precision	16.63%	16.59%	17.13%	17.26%	17.57%	17.8%	17.64%	17.6%	17.49%	15.16%
Recall	24.32%	24.32%	25.08%	25.26%	25.55%	25.73%	25.41%	25.51%	25.08%	26.6%

settings, KPM consistently outperforms the citation-based KCN baseline in precision (up to 17.8% vs. 15.16%) and F1 score, while maintaining comparable recall (sensitivity local optimum 25.73% vs. 26.6%). These results highlight the effectiveness and stability of our approach in validating the low prerequisite alignment, confirming that KPM offers a more precise and reliable alternative to citation-only methods.

4.5. Investigating the low prerequisite recall in KPNs

To understand the fundamental causes behind the low prerequisite recall observed in our evaluation, we conduct a comprehensive investigation from four complementary perspectives. First, we analyze author-level precision to examine how individual researchers' topic transitions align with LLM-annotated prerequisites, revealing that approximately 80% of authors exhibit zero to low precision. Second, we demonstrate the inherent diversity in research trajectories through quantitative analysis of concept entropy, unique topic counts, and research concentration patterns, showing that scientists systematically broaden and diversify their research scope over time. Third, we contextualize our findings within recent literature on topic-switching patterns and knowledge advancement pathways. Fourth, we compare prerequisite recall based on citation-derived KCNs under the same evaluation criteria, finding that



(a) Precision CDF Subplots

(b) Precision categories trend by WCR

Fig. 11. Author-level precision in identifying KPN prerequisite relations across varying WCR thresholds. (a) Cumulative distribution functions (CDFs) of precision for all authors across 303 concepts from 19 disciplines, with each subplot corresponding to a different WCR threshold. (b) Proportion trends of authors grouped into six precision levels (Zero, Minimal, Low, Moderate, Good, High) under varying WCR thresholds.

while citation-based KCNs achieve a slightly higher maximum recall (26.6%) than our method (25.77%), our approach consistently attains higher precision and F1 scores. This further demonstrates that the low prerequisite alignment is not only evident in individual research careers, but also manifests in citation behaviors, reinforcing the notion that scientific progress often follows individualized and innovative pathways that differ from the structured learning sequences typically found in educational settings.

4.5.1. Author-level precision analysis

To investigate the discrepancy between KPNs and LLM-annotated prerequisites, we analyzed author-level precision, defined as the proportion of topic transitions that align with the LLM-annotated CPN. This analysis also helps explain the source of overall low recall from the perspective of individual researchers. To facilitate a more intuitive interpretation of precision values, we categorized the results into "Zero Precision" (0) and other five meaningful intervals: "Minimal Precision" (0.0–0.2), "Low Precision" (0.2–0.4), "Moderate Precision" (0.4–0.6), "Good Precision" (0.6–0.8), and "High Precision" (0.8–1.0). This classification allows us to assess author level alignment under different WCRs.

Fig. 11 provides two complementary perspectives on author-level precision across varying WCR thresholds. In Fig. 11(a), the cumulative distribution functions (CDFs) show that a substantial proportion of authors have zero precision, indicated by abrupt jumps at Precision = 0. This suggests that many authors' topic transition patterns do not correspond to LLM-identified prerequisite relations. Both the mean and median precision, marked by black and light brown lines respectively, remain low across all WCR values.

Fig. 11(b) further groups authors into six precision levels (Zero, Minimal, Low, Moderate, Good, High) and tracks how the proportion in each category changes as WCR increases from 0.6 to 1. While the proportion of zero-precision authors decreases with higher WCR, the share of authors in the minimal and low precision categories increases, and only minor gains are seen in the higher precision levels. Overall, these results highlight persistently low precision, which largely accounts for the limited alignment between author topic transitions and LLM prerequisite relations, regardless of the WCR threshold.

To further investigate the reasons behind the low prerequisite recall observed for author topic transitions, we analyzed the distribution of TCPs across different author precision levels, as shown in Fig. 12. The results show that authors with zero precision typically have only 1 to 3 concept pairs, indicating limited topic transitions. In contrast, minimal and low precision authors often have 3 to 10 concept pairs, and sometimes even more than 20. For moderate and good precision authors, concept pairs are generally concentrated in the range of 2 to 5, whereas high precision authors typically have only 1 to 3 pairs. Building on Fig. 11(b) that approximately 80% of authors fall into the zero, minimal, or low precision categories, and minimal or low precision authors tend to have more TCPs, thus we examined how the number of concept pairs per author relates to precision. A correlation analysis suggests only a weak negative relationship between the number of concept pairs and precision (correlation coefficient: −0.0339), suggesting that simply having more topic transitions does not fully explain the low precision.

4.5.2. Demonstration of diversity in research trajectory

The low alignment rate between KPN transitions and LLM-annotated prerequisites reflects the diverse and non-linear nature of scientific research. The non-linear nature means scientists often do not follow prerequisite pattern or standard trajectory in acquiring knowledge or selecting topics, as indicated by the ratio of Zero and Minimal precision curve in Fig. 11(b), the ratio

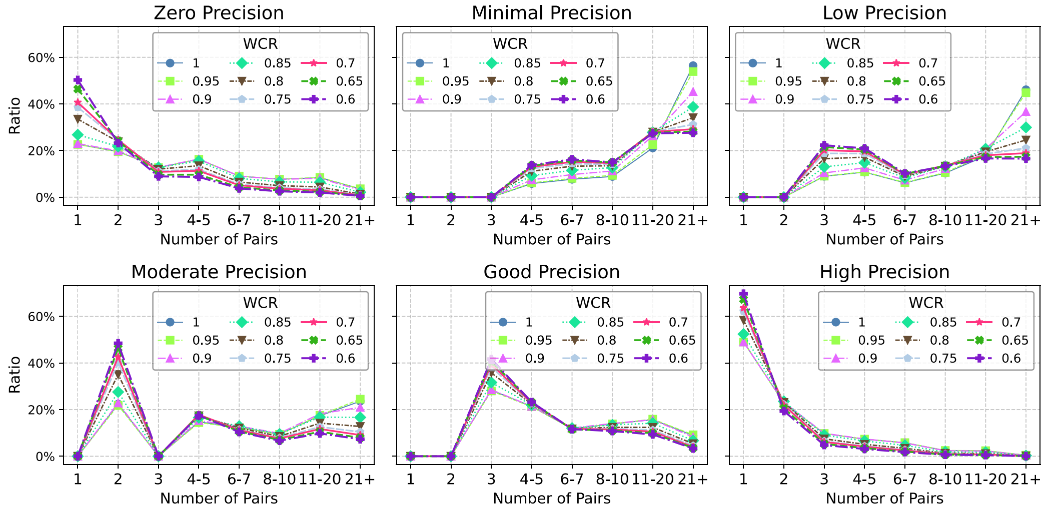


Fig. 12. Temporal concept pairs distribution of authors by precision level.

statistically support that most authors follow non-prerequisite switch patterns. The diversity of scientific research is reflected in two main aspects. First, individual scientists differ in both the range of research topics and the pace at which their research topics expand, which results in highly personalized research paths. Second, for an individual scientist, the diversity of scientific paths exhibits in the broaden range of research topics over time, later topics often differ from earlier ones, demonstrating an increasing trend in concept entropy. To assess the diversity of a researcher's scholarly path as their publication count increases, we employ three metrics including Unique Concepts, Shannon Entropy, and Herfindahl–Hirschman Index (HHI).

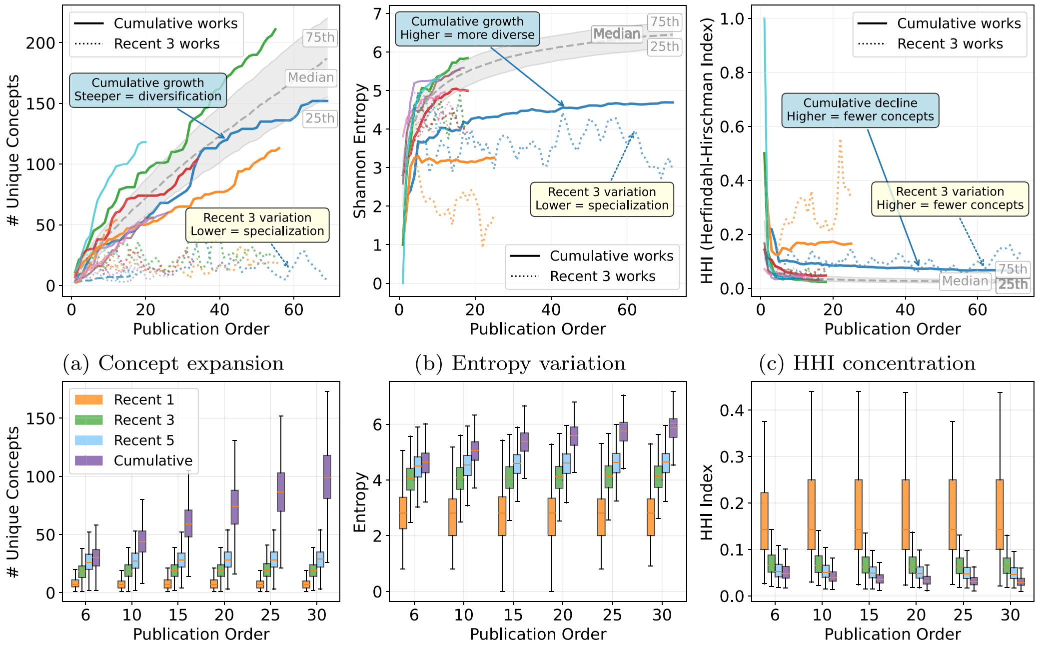
1. **Unique Concepts**: The cumulative number of distinct concepts covered by the researcher's publications, plotted as a function of the number of papers.
2. **Shannon Entropy**: Measures the distributional evenness of concepts; higher entropy indicates greater diversity

$$H = -\sum_{i=1}^{N} p_i \log p_i, \tag{15}$$

   where p_i is the probability of the ith concept.
3. **Herfindahl–Hirschman Index (HHI)**: Quantifies concentration, higher values indicate greater focus on fewer concepts

$$HHI = \sum_{i=1}^{N} p_i^2. \tag{16}$$

We then conduct sequential analysis to track how unique concepts, entropy, and HHI evolve as more papers are published. We also calculate the diversity metrics within a fixed window (e.g., the most recent 3 or 5 papers) to observe changes in research focus over time. Fig. 13 presents a comprehensive analysis of the diversity and evolution of research topics among scientists at both individual and collective levels. Figs. 13(a), 13(b), and 13(c) show how unique concept counts (breadth), Shannon entropy (evenness), and HHI (concentration) evolve with publication order and jointly trace the co-evolution of breadth and focus. Each plot displays eight randomly selected authors in distinct colors with two trajectories: a cumulative curve up to the kth publication (solid) and a sliding 3-paper window (dashed). The gray band is the cross-author interquartile range (25th–75th percentiles) and the gray dashed line is the median. In Fig. 13(a) cumulative unique concepts rise at author-specific rates indicating heterogeneous long-run diversification while 3-paper window counts remain comparatively flat showing a stable narrower active concept set. Fig. 13(b) shows a rapid early rise in entropy followed by a slowdown indicating diminishing marginal gains in evenness while lower window entropy indicates a tighter short-term focus. Fig. 13(c) shows an early sharp decline in HHI then a plateau while higher window HHI indicates sustained short-term specialization. Figs. 13(d), 13(e), and 13(f) further compare cumulative and recent (1/3/5 paper window) metrics at publication indices 6, 10, 15, 20, 25, and 30 for authors with at least 30 works. Cumulative concept counts expand steadily while the recent counts stay relatively stable and incremental novel concepts accumulate across windows to drive long-run breadth. Cumulative entropy increases over careers while short-window entropy shows little systematic change. Cumulative HHI declines then stabilizes while window HHI remains higher with the single-paper window highest. Taken together, the metrics show a composite diversity pattern: gradual long-run diversification alongside persistent short-run topical concentration, embedded within cross-author diversity in research directions, publication volume, and rates of expansion.



(a) Concept expansion (b) Entropy variation (c) HHI concentration

(d) Concepts distribution (e) Entropy distribution (f) HHI distribution

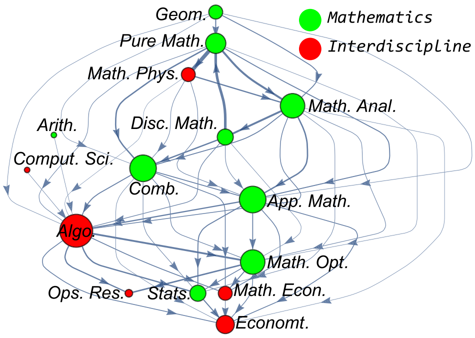
Fig. 13. Topical diversity and concentration over publication sequences. (a–c) For eight randomly selected authors, cumulative (solid) and 3-paper window (dashed) trajectories of unique concepts (breadth), Shannon entropy (evenness), and HHI (concentration). Shaded regions show the cross-author interquartile range; the dashed interior line is the median. (d–f) Distributions (cumulative and recent-1/3/5 paper window) at six publication milestones (6, 10, 15, 20, 25, 30) for authors with at least 30 papers.

4.5.3. Complementary perspectives from related studies

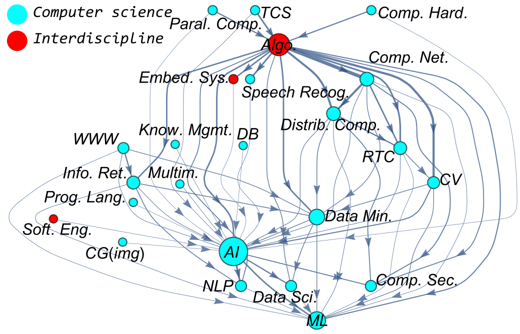
Beyond our statistical analysis from the perspective of author-level precision and diverse research topics, recent studies on topic-switching patterns further support our findings. Yang's (Yang, 2025) statistical framework reveals that frequent topic transitions drive scientific innovation and disruption. As innovation often involves breaking conventions and integrating diverse fields, scientists who switch research topics are less likely to follow traditional prerequisite chains in their knowledge transitions. Thus, these findings provide support for our empirical finding of low alignment between real knowledge transitions and LLM-annotated prerequisites. The methodological convergence of Yang's topic-switching patterns and our low prerequisite alignment (25.77%) suggests that the observed low precision reflects deliberate exploration strategies that prioritize innovation-seeking over conventional learning sequences.

Additionally, the analysis of ICA Fellows by Barnett and Park (Barnett & Park, 2023; Park & Barnett, 2024) revealed that elite scholars exhibit sparse collaboration networks and that the alignment between co-ethnicity and collaboration or citation networks is generally weak. Their studies further highlight the fragmented and diverse structure of scholarly collaboration, rather than a standardized network. The increasingly diverse and global nature of communication research, as well as the diversity of citation flows are consistent with our findings regarding research topic diversity. Our low alignment between observed knowledge transitions and theoretical prerequisite relations may be partially attributed to the influence of social, cultural, and geographical factors on scientists' research trajectories.

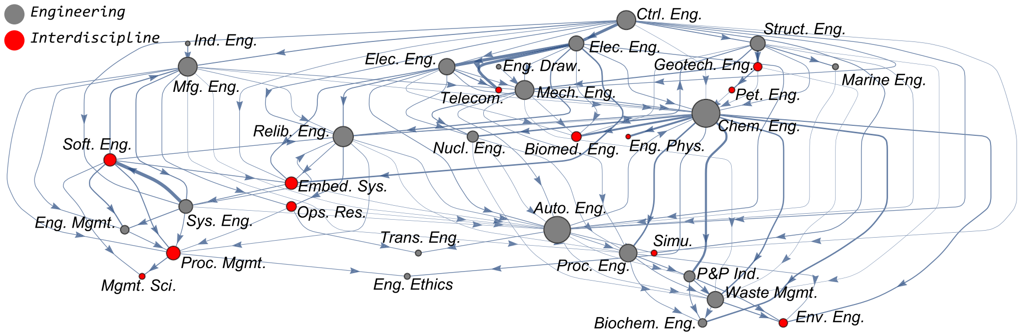
Taken together, our multi-perspective analyses reveal that the low alignment between KPN transitions and LLM-annotated prerequisite relations fundamentally reflects the authentic, diverse, and non-linear nature of scientific exploration. First, author-level precision analysis shows that nearly 80% of authors exhibit zero to low precision, indicating that most researchers do not follow prerequisite-driven topic sequences. This suggests that while some researchers adopt exploitation strategies by adhering to established paths, most topic selections are guided by individualized considerations or deliberate exploration strategies that prioritize innovation, rather than by formal learning orders. Second, our quantitative diversity analysis demonstrates that scientists systematically broaden their research scope over time, with increasing concept entropy and topic variety, further highlighting the prevalence of personalized and evolving research trajectories. Third, our findings are corroborated by recent studies: frequent topic switching is linked to scientific innovation and disruption (Yang, 2025), while scholarly collaboration and citation networks are themselves fragmented and shaped by social, cultural, and geographical factors rather than intellectual dependencies (Barnett & Park, 2023; Park & Barnett, 2024). Fourth, our comparative analysis with citation-based KCNs shows that, even when using citation relations as proxies for knowledge transition, the observed prerequisite recall remains low, indicating citation-based KCNs also capture only a limited portion of the prerequisite structure identified by LLM annotations. This further supports the notion that knowledge transfer in science often occurs through invisible channels such as self-study, mentorship communication,



(a) KPN within Mathematics concepts



(b) KPN within Computer science concepts



(c) KPN within Engineering concepts

Fig. 14. KPN within Mathematics, Computer science, and Engineering.

interdisciplinary collaboration, and funding-driven shifts in research, all of which frequently disrupt linear prerequisite patterns. Collectively, these results indicate that low prerequisite alignment is a robust signature of the complex, innovation-seeking, and individualized diverse pathways that characterize real-world scientific advancement.

5. Real-world knowledge precedence network

The precedence relationships hold strong educational significance. We leverage the OKPM for concepts in Mathematics, Computer Science, and Engineering to build KPNs. This network allows us to analyze precedence relationships both within each discipline and across these different fields. Each non-zero element in the matrix represents a directed edge in the network. The concepts represented by the row indices are potential prerequisites for the concepts represented by the column indices. For clarity, Table B.6 provides both the full names and abbreviations of all the concepts used in Fig. 4.

5.1. The KPN within discipline

To demonstrate in-domain precedence relations within each discipline, we focus on Mathematics, Computer Science and Engineering. We construct domain-specific KPNs based on the optimized OKPM(d_i, d_i). Fig. 14 presents the resulting KPNs for: Mathematics (Fig. 14(a)), Computer Science (Fig. 14(b)) and Engineering (Fig. 14(c)). These visualizations highlight the significance of precedence relationships within each field, offering insights into the structure of knowledge in these disciplines. We now focus on describing the key information and significant implications derived from the KPN of Mathematics and Computer Science.

In the KPN within the discipline of Mathematics in Fig. 14(a), we observe that Geometry and Pure mathematics have played a crucial role as precedents, fostering the development and learning of various sub-disciplines such as Mathematical physics, Mathematical analysis, Discrete mathematics, Arithmetic, Computational science, Combinatorics, and Algorithm. These foundational concepts also contribute to the learning of advancing Applied mathematics, Mathematical optimization, Operations Research, Statistics, Mathematical Economics, and Econometrics. Some interdisciplinary concepts, such as Mathematical physics, Algorithm, Operations Research, Mathematical Economics, and Econometrics, indicate that Mathematics, as a fundamental discipline, not only fosters the growth of its sub-disciplines but also contributes to the development of other disciplines related to Mathematics.

In the KPN of Computer Science in Fig. 14(b), concepts at the top are Theoretical Computer Science, Parallel Computing, and Computer Hardware. These concepts promote the application of Algorithms and further serve as prerequisites for Speech Recognition, Data Mining, AI (Artificial Intelligence), NLP (Natural Language Processing), Data Science, and Machine Learning. AI plays a crucial role as a descendent of sub-disciplines like Soft Engineering, Computer Graphics, Information Retrieval, Distributed

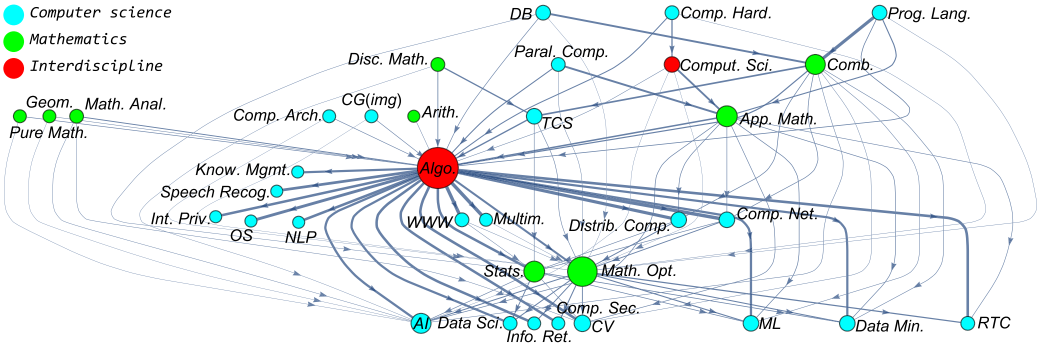


Fig. 15. KPN between Mathematics and Computer science.

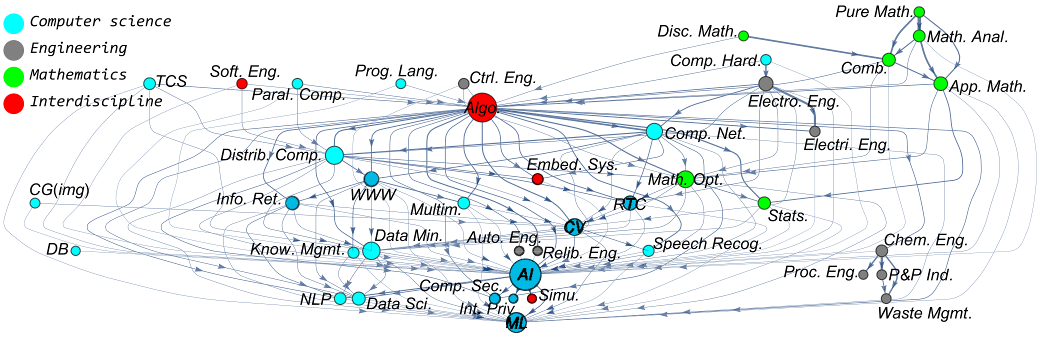


Fig. 16. KPN between Mathematics, Computer science and Engineering.

Computing, Real-time Computing, and Computer Vision. Additionally, AI also plays a notable antecedent to NLP, Data Science, Machine Learning, and Computer Security, highlighting the comprehensive and essential role of Algorithms and AI in Computer Science. We also display the KPN with Engineering in Fig. 14(c), more KPN examples within the disciplines of Physics, Economics, and Chemistry can be seen in Fig. D.5, and disciplines of Biology, Medicine, and Geology in Fig. D.6.

5.2. The KPN across multiple disciplines

In addition to the KPN within Mathematics, Computer Science and Engineering, respectively. We reveal the mutual precedence relationships between two fields to construct a cross-domain KPN. All edges from two OKPMs including OKPM(d_i, d_j) from $d_i \rightarrow d_j$ and OKPM(d_j, d_i) from $d_j \rightarrow d_i$ are combined to create the final KPN. We construct the KPN between Mathematics and Computer Science as depicted in Fig. 15. The concept of Algorithm serves as a critical interdisciplinary bridge between two fields. Preceding concepts in Computer Science include Database, Computer Architecture, Parallel Computing, and Computational Science, while in Mathematics, they include Pure Mathematics, Mathematical Analysis, Discrete Mathematics, and Combinatorics. The Algorithm commonly precedes advanced fields such as Statistics, Mathematical Optimization, AI, Data Science, Information Retrieval, Computer Vision, and Machine Learning. Descendants like Statistics and Mathematical Optimization, further drive the development of Data Science, Computer Vision, Machine Learning, and Data Mining. This KPN highlights the integration of Mathematics as a foundational discipline with Computer Science, fostering advancements in Computational Science and Algorithms.

The KPN among Mathematics, Computer Science, and Engineering is illustrated in Fig. 16, further underscores the interdisciplinary importance and bridging role of the Algorithm. In this network, Mathematics concepts like: Pure Mathematics, Mathematical Analysis, Combinatorics, and Applied Mathematics occupy the highest positions. Applied fields such as AI, NLP, Data Science, and Machine Learning appear at the bottom, indicating their reliance on foundational concepts. Mathematical Optimization and Statistics facilitate the development of downstream applied sciences. Engineering concepts, although interacting less with mathematics than computer science, still show significant interplay. For example, Control Engineering temporally precedes Algorithm, and Computer Hardware is essential for Electronic Engineering. Automotive and Reliable Engineering contribute to AI, while Chemical Engineering engages more within Engineering, interacting with Process Engineering, the Pulp and Paper Industry, and Waste Management. More KPN examples across diverse disciplines can be seen in Fig. D.7 and Fig. D.8.

Overall, the knowledge precedence network derived from scientists' research trajectories effectively captures the real-world patterns of how researchers transition between topics in different fields. This provides valuable insights into cross-disciplinary research shift patterns, despite the technical limitations in recall rate.

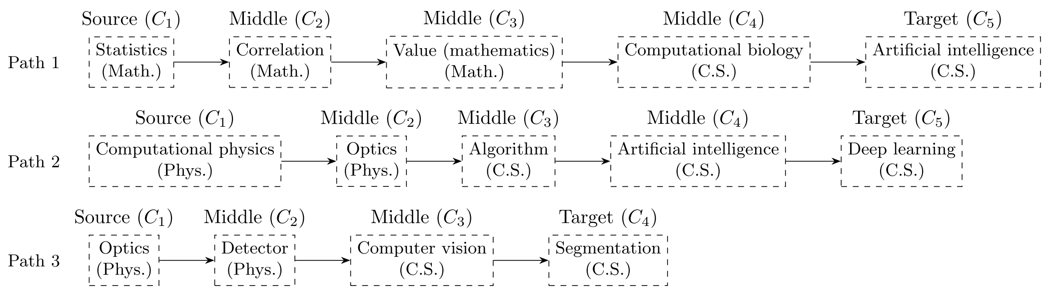


Fig. 17. Exploratory learning path navigation extracted from the personalized subgraph in Figure C.4. Three exemplary paths demonstrate how concepts from different disciplines form coherent learning sequences. Each path shows the progression from source concepts to target concepts, with discipline annotations provided for clarity. The C_i labels indicate the sequential position of concepts within each learning path.

6. Applications and implications of KPNs

This section explores the practical applications and theoretical implications of KPNs, demonstrating the potential to transform academic research and learning. We first present personalized learning path navigation that guide researchers from familiar concepts to target domains through empirically-derived pathways, comparing our approach with existing concept recommendation systems. We then examine the broader theoretical and practical implications of KPNs for scientific progression, interdisciplinary collaboration, and educational planning. Finally, we discuss our key finding that only 25.77% of scholarly transitions align with traditional prerequisites, revealing novel insights into how knowledge actually evolves in real-world research contexts.

6.1. Personalized learning path navigation

Our KPN enables personalized subgraph extraction based on researchers' existing expertise and target interests, guiding both topic learning and research strategy, as illustrated in Figure C.4. These subgraphs demonstrate KPN's capacity to navigate customized interdisciplinary learning pathways that connect a researcher's current knowledge with desired domains. Users can explore the generated subgraph and select learning paths aligned with their interests. To enhance clarity and reduce visual complexity, we extract representative learning paths from these personalized subgraphs, as shown in Fig. 17. For instance, when a researcher proficient in Statistics and Correlation seeks to enter Artificial Intelligence, our method extracts a tailored knowledge subgraph from the KPN for navigation as shown in Figure C.4a, which presents various possible learning trajectories derived from millions of scholars' behavioral patterns, with one exemplary trajectory highlighted as “Path 1” in Fig. 17. Similarly, “Path 2” represents the learning trajectory for a physicist familiar with Computational physics and Optics who pursues Deep learning, extracted from the personalized subgraph shown in Figure C.4b. “Path 3” illustrates another scenario where a researcher with expertise in Optics aims to learn Segmentation, derived from the subgraph in Figure C.4c.

Unlike ConceptGCN (Alatrash et al., 2024), which learns learner and concept embeddings via a GCN over a heterogeneous KG and recommends the top-5 concepts based on highest learner-concept similarity to remediate a marked “Do Not Understand” (DNU) concept. Our KPN differs fundamentally by constructing a purely concept-only network from millions of observed scholarly transition sequences. It returns personalized bridging subgraphs that surface multi-step concept-only trajectories from a user's current expertise toward target domains or research topics, enabling active exploration of concept pathways rather than passive proximity-based suggestion. Furthermore, compared to LLM-based path generation methods that rely on user-specific historical data (Abrar et al., 2025; Cheng et al., 2025), such as Yang and Liang (2025), who integrate GPT-4 with multimodal student behaviors, emotions, and physiological signals for extracurricular path recommendations; Chen et al. (2024) use GPT-4 with Chain-of-Thought prompting to generate personalized content from teacher-uploaded materials tailored to students' interests and goals; Liu, Qian, and Zhao (2025) employ NLP and GCN with minimum spanning trees to create adaptive English learning paths based on exercise difficulties and performance data; or Ng and Fung (2024), who leverage LLMs like Llama-2-70B and GPT-4 with prompt engineering for dynamic, multi-turn educational paths. Our KPN operates on broader scientific concepts without requiring individualized user histories, knowledge points, or exercise outcomes. These prior systems predominantly target instructional or formative learning scenarios centered on fine-grained learner activities, exercises, or knowledge states. In contrast, our work addresses research-oriented knowledge exploration and scientific discovery, shifting the learning unit from student-facing micro-tasks to higher-order scientific concepts and their epistemic relations. Accordingly, KPN operates on broad scientific concepts derived from large-scale scholarly transition sequences, emphasizing evolutionary, behavior-grounded interdisciplinary precedence relations rather than prescriptive educational sequences aimed at remediation or stepwise skill mastery. Independent of individualized learner histories, enumerated knowledge points, or exercise outcomes, it provides a reusable concept-only infrastructure for interactive exploratory navigation and flexible research or learning strategies. Users can proactively extract, compare, and adapt multi-step concept trajectories for research planning. Overall, KPN enables navigation of real-world innovation patterns that extend beyond textbook advances, surpassing the linear, history-dependent paths produced by learner-centered LLM systems.

Limitations: The current navigation pathways derived from KPNs are constrained by a restricted high-level concept set, yielding coarse subgraphs that provide only macro-level, directional guidance rather than fine-grained instructional sequencing. The low prerequisite alignment shows that KPNs predominantly capture empirically observed, research-driven transition behaviors rather than validated, fine-grained pedagogical progressions or textbook prerequisite logic, thereby limiting their direct use in precise curriculum design or mastery scaffolding. Nonetheless, this behavioral orientation is precisely what gives KPNs value for exploratory, interdisciplinary navigation, as broader and finer-grained concept layers are incorporated, future versions may support more nuanced and pedagogically actionable pathways.

6.2. Theoretical and practical implications of KPNs

The constructed KPNs reveal the hierarchical chain of knowledge dependencies, illustrating scientific propagation and advancement patterns, as shown by the real-world KPNs in Section 5 for Mathematics, Computer Science, and Engineering. We verify that scientific progress is driven by the development of new theories, technologies, and discoveries that build upon existing foundations. This structured progression highlights how basic theories and methods lay the groundwork for theoretical and practical implications as following:

1. **Theoretical foundation**: Our results illustrate how fundamental knowledge, particularly in Mathematics, forms the basis for scientific exploration and innovation, such as how foundational theories (e.g., Pure mathematics and Computer hardware) underpin the development of advanced algorithms and optimization methods.
2. **Interdisciplinary collaboration**: The KPNs reveal a clear transition from foundational knowledge to advanced applications, emphasizing the bridging role of fields like Algorithms, Statistics, and Optimization in enabling cross-domain integration, knowledge transfer, and innovation in fields like AI and data science.
3. **Technology advancement**: KPNs highlight the technological advancements of cutting-edge fields like AI, Data Science, Information Retrieval, and Computer Vision, demonstrating how they systematically evolve from foundational Mathematics and intermediate interdisciplinary fields such as Algorithms, Statistics, and Optimization.
4. **Educational implications**: The hierarchical structure of KPNs provides valuable insights for education and research planning, enabling personalized learning navigation, guiding research strategies, fostering interdisciplinary exploration, and advancing scientific discovery.
5. **Novel insights into scientific progression**: The LLM prerequisite benchmark evaluation reveals a striking divergence between theoretical prerequisites and real-world knowledge progression. Only 25.77% of transitions align with theory in the best case, while most follow alternative or reverse patterns, challenging conventional views of scientific progression and highlighting the complexity and flexibility of knowledge evolution.

6.3. Limitations of KPNs

The low alignment rate indicates that researchers often follow knowledge trajectories that deviate from textbook-defined prerequisite sequences. This highlights the KPN's strength in capturing the diversity and complexity of real-world scientific progression. However, such deviation also limits the usefulness of KPNs in contexts where a clear, stepwise acquisition of foundational knowledge is required, such as in formal education or curriculum design, where high alignment with established prerequisite structures is essential.

In these cases, conventional CPNs which are explicitly constructed to represent established prerequisite relationships, are more appropriate for supporting prerequisite-based learning and curriculum planning. By contrast, KPNs are particularly advantageous for understanding real-world research practices, interdisciplinary knowledge transfer, and innovation pathways. For instance, when analyzing how breakthroughs in artificial intelligence emerge from the convergence of computer science, mathematics, and engineering, KPNs reveal the innovative precedence pathways through which researchers traverse knowledge domains—often bypassing standard prerequisites. Similarly, when mapping scientific workforce mobility or the evolution of emerging fields, KPNs capture atypical or cross-disciplinary transitions that conventional CPNs may overlook.

The impact of low prerequisite alignment on the usefulness or acceptability of the KPN framework is therefore context-dependent. In exploratory research or when mapping the innovative pathways of scientific discovery, low recall is expected and even desirable, as it reflects the broad range of knowledge trajectories that researchers actually follow. Conversely, in contexts where comprehensive coverage of foundational knowledge is critical, such as curriculum design or formal training, the CPN is more suitable for ensuring that all prerequisite relationships are represented. Thus, low recall of KPN is a strength for exploratory analyses, but a limitation for applications requiring completeness and structure; in such cases, CPNs are the preferred framework.

In summary, the constructed KPNs provide a novel framework for analyzing scientific knowledge progression patterns while offering insights for interdisciplinary collaboration, technological development, and scientific research planning. By quantitatively capturing dynamic knowledge dependencies, this approach highlights the gap between theoretical knowledge structures and real-world scientific practices.

7. Discussion

This study investigates the collective behavior of scientists, focusing on the evolving patterns of research topics. Analyzing precedence relationships across 19 scientific domains using the OpenAlex dataset, we examine 4,969,403 authors with over 10 publications, focusing on level-0 and level-1 concepts. The identified KPNs underscore the foundational role of basic disciplines like Mathematics, which serve as foundations for advanced fields and as cornerstones for scientific innovation. These dependencies highlight the dynamic and interconnected nature of scientific progress and knowledge evolution.

Our research introduces a network-based method to mine precedence relations between concepts by analyzing scientists' career paths. This method reveals inherent temporal relationships in fields like Mathematics, Computer Science, and Engineering, without explicit feature extraction or parameter learning. Through constructing CCNs and identifying bibliographic couplings, we uncover the structure of scholarly connections based on shared references. The Louvain algorithm is employed for community detection, identifying cohesive clusters that represent distinct research fields. Key concepts for each community are determined by analyzing the most frequent concepts covering 80% of the works, ensuring a robust depiction of the main topics. TCPs are established through temporal community pairs, revealing precedence and succession patterns in scientific domains. However, several limitations remain: (i) the coarse concept granularity collapses fine-grained educational micro-progressions; (ii) the fixed APYD window [3,7] may misalign with heterogeneous field maturation rates; and (iii) while improving interpretability, the thresholding strategy can prune low-frequency yet pedagogically meaningful links.

The analysis spans the careers of millions of scholars across various fields, revealing patterns in knowledge inheritance and progression. This pattern aids in understanding the hierarchical nature of scientific knowledge evolution, in which basic concepts serve as foundations and intermediate concepts bridge basic theories and advanced applications, paving the way for advancements in fields like AI and Data Science. Understanding these relationships provides strategic value for helping trace and forecast scientific topic trends, shaping educational policies, enhancing self-directed learning, fostering interdisciplinary collaboration and innovation, and guiding academic research and topic selection. While this analysis provides strategic value for educational, research, and policy decisions, there are limitations to consider. Since the precedence relations are rooted in constantly evolving research topics, it is imperative to engage researchers and knowledge from diverse disciplines, while temporal dynamics of scientific development will further refine and validate these KPNs. Further, the observed precedence encodes transition frequency rather than prerequisite necessity, hence the resulting navigation paths should be treated as exploratory aids rather than validated mastery sequences. Future integration with curated CPN layers may allow dual-channel guidance: KPN for exploratory cross-domain expansion; CPN for prerequisite-constrained mastery sequencing.

This alignment evaluation compares two fundamentally different data sources: patterns derived from scientists' research career statistics (positive and negative concept pair frequency), and prerequisite relations obtained from LLMs combined with Wikipedia definitions. Notably, the low alignment between these sources reveals a significant finding: scientists' actual research transitions largely deviate from theoretical precedence relationships. This suggests that real-world research behavior is primarily driven by factors beyond pure knowledge prerequisites, reflecting more complex and practical transition patterns. While traditional prerequisite relationships represent theoretical knowledge dependencies, our behavior-based approach uncovers how scientists actually navigate between research fields in practice. This novel perspective complements prerequisite-centric methods by revealing practice-driven dynamics at a coarser conceptual granularity, offering unique insights into the real-world patterns of knowledge field transitions that often operate independently of theoretical prerequisites.

Overall, this study maps the complex landscape and precedence relations of scientific knowledge, emphasizing the dynamic and interconnected nature of scientific progress. The analysis of TCPs provides a powerful tool for identifying knowledge dependencies, research trends, and the knowledge evolution network, championing interdisciplinary research as a catalyst for innovation and scientific advancement. Future work may augment Wikipedia representations with heterogeneous KGs, and scholarly metadata, which involve fusing multi-source embeddings and assessing universality via cross-source alignment, semantic similarity, coverage overlap, and temporal stability analyses to ensure robustness of the findings.

CRediT authorship contribution statement

Shibing Xiang: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Bing Liu:** Writing – review & editing. **Xin Jiang:** Writing – review & editing. **Zhengan Huang:** Writing – review & editing. **Yifang Ma:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Conceptualization.

Funding

This work was supported by the National Natural Science Foundation of China [NSFC62006109, NSFC12031005]; the 13th Five-year Plan for Education Science Funding of Guangdong Province, China [2020GXJK457]; the Stable Support Plan Program of Shenzhen Natural Science Fund, China under [20220814165010001]; and the Major Key Project of PCL, China [PCL2023A09].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Any potential conflicts of interest have been disclosed.

Acknowledgments

The computation in this study was supported by the Center for Computational Science and Engineering of SUSTech and by Pengcheng Laboratory, China. All authors have seen and approved the final version of the manuscript being submitted.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ipm.2025.104424.

Data availability

The OpenAlex dataset is publicly accessible for use by researchers and the wider community, our code is publicly available at: https://github.com/xiangshb/KPN-Mining.

References

Abrar, M., Aboraya, W., Khaliq, R. A., Subramanian, K. P., Al Husaini, Y., & Al Hussaini, M. (2025). AI-powered learning pathways: Personalized learning and dynamic assessments. *International Journal of Advanced Computer Science & Applications*, *16*(1).

Adorni, G., Alzetta, C., Koceva, F., Passalacqua, S., & Torre, I. (2019). Towards the identification of propaedeutic relations in textbooks. In *International conference on artificial intelligence in education* (pp. 1–13). Springer.

Alatrash, R., Chatti, M. A., Ain, Q. U., Fang, Y., Joarder, S., & Siepmann, C. (2024). ConceptGCN: Knowledge concept recommendation in MOOCs based on knowledge graph convolutional networks and SBERT. *Computers and Education: Artificial Intelligence*, *6*, Article 100193.

Alzetta, C., Miaschi, A., Adorni, G., Dell'Orletta, F., Koceva, F., Passalacqua, S., & Torre, I. (2019). Prerequisite or not prerequisite? That's the problem! an NLP-based approach for concept prerequisite learning. In *Proceedings of the 6th Italian conference on computational linguistics* (pp. 1–8).

Alzetta, C., Torre, I., & Koceva, F. (2024). Annotation protocol for textbook enrichment with prerequisite knowledge graph. *Technology, Knowledge and Learning*, *29*(1), 197–228.

Aytekin, M. C., & Saygín, Y. (2025). Discovering prerequisite relations using large language models. *Interactive Learning Environments*, *33*(2), 1670–1688.

Aytekin, M. C., Saygın, Y., et al. (2024). ACE: AI-assisted construction of educational knowledge graphs with prerequisite relations. *Journal of Educational Data Mining*, *16*(2), 85–114.

Bai, Y., Liu, Z., Guo, T., Hou, M., & Xiao, K. (2025). Prerequisite relation learning: A survey and outlook. *ACM Computing Surveys*, *57*(11), 1–28.

Barnett, G. A., & Park, H. W. (2023). Co-authorship among the fellows of the international communication association. *Scientometrics*, *128*(6), 3401–3418.

Bart, W. M., & Read, S. A. (1984). A statistical test for prerequisite relations. *Educational and Psychological Measurement*, *44*(2), 223–227.

Bawden, D., & Robinson, L. (2020). Information overload: An overview. In D. P. Redlawsk (Ed.), *The oxford encyclopedia of political decision making*. Oxford University Press.

Bornmann, L., Haunschild, R., & Mutz, R. (2021). Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications*, *8*(1), 1–15.

Chen, W., Lan, A. S., Cao, D., Brinton, C., & Chiang, M. (2018). Behavioral analysis at scale: Learning course prerequisite structures from learner clickstreams. In *Proceedings of the 11th international conference on educational data mining* (pp. 66–75). International Educational Data Mining Society.

Chen, E., Lee, J.-E., Lin, J., & Koedinger, K. (2024). Gptutor: Great personalized tutor with large language models for personalized learning content generation. In *Proceedings of the eleventh ACM conference on learning scale* (pp. 539–541).

Chen, P., Lu, Y., Zheng, V. W., & Pian, Y. (2018). Prerequisite-driven deep knowledge tracing. In *2018 IEEE international conference on data mining* (pp. 39–48). IEEE.

Cheng, X., Zhang, Z., Wang, J., Fang, L., He, C., Guan, Q., Pan, S., & Luo, W. (2025). GraphRAG-induced dual knowledge structure graphs for personalized learning path recommendation. arXiv preprint arXiv:2506.22303.

Doignon, J.-P., & Falmagne, J.-C. (2016). Knowledge spaces and learning spaces. In *New handbook of mathematical psychology* (pp. 274–321).

Dominici, G., Barbiero, P., Zarlenga, M. E., Termine, A., Gjoreski, M., Marra, G., & Langheinrich, M. (2025). Causal concept graph models: Beyond causal opacity in deep learning. In *Proceedings of the 13th international conference on learning representations*. ICLR.

Dowling, C. E., Koch, U., & Quante, K. A. (1996). A new interface for querying experts on prerequisite relationships. In *Proceedings sixth Australian conference on computer-human interaction* (pp. 320–321). IEEE.

Gan, W., Sun, Y., & Sun, Y. (2022). Knowledge structure enhanced graph representation learning model for attentive knowledge tracing. *International Journal of Intelligent Systems*, *37*(3), 2012–2045.

Gasparetti, F., Limongelli, C., & Sciarrone, F. (2015). Exploiting wikipedia for discovering prerequisite relationships among learning objects. In *2015 International conference on information technology based higher education and training* (pp. 1–6). IEEE.

He, Y., Wang, H., Pan, Y., Zhou, Y., & Sun, G. (2022). Exercise recommendation method based on knowledge tracing and concept prerequisite relations. *CCF Transactions on Pervasive Computing and Interaction*, *4*(4), 452–464.

Hsu, W.-L., Chen, Y.-S., Wang, Y.-K., et al. (1998). A context sensitive model for concept understanding. *Proceedings of ITALLC*, *98*, 161–169.

Huang, S., Lu, W., Cheng, Q., Luo, Z., & Huang, Y. (2024). Evolutions of semantic consistency in research topic via contextualized word embedding. *Information Processing & Management*, *61*(6), Article 103859.

Jia, C., Shen, Y., Tang, Y., Sun, L., & Lu, W. (2021). Heterogeneous graph neural networks for concept prerequisite relation learning in educational data. In *Proceedings of the 2021 conference of the North American chapter of the association for computational linguistics: human language technologies* (pp. 2036–2047).

Jian, Y., Luo, X., Li, Z., Zhang, M., Zhang, Y., Xiao, K., & Hou, X. (2025). APKGC: Noise-enhanced multi-modal knowledge graph completion with attention penalty. In *Proceedings of the AAAI conference on artificial intelligence*: *vol. 39*, (14), (pp. 15005–15013).

Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, *14*(1), 10–25.

Kirillovich, A., Falileeva, M., Nevzorova, O., Lipachev, E., Dyupina, A., & Shakirova, L. (2021). Prerequisite relationships of the ontomathedu educational mathematical ontology. In *Applied computer sciences in engineering: 8th workshop on engineering applications* (pp. 517–524). Springer.

Koppen, M., & Doignon, J.-P. (1990). How to build a knowledge space by querying an expert. *Journal of Mathematical Psychology*, *34*(3), 311–331.

Kumar, N. A., Feng, W., Lee, J., McNichols, H., Ghosh, A., & Lan, A. (2023). A conceptual model for end-to-end causal discovery in knowledge tracing. In *Proceedings of the 16th international conference on educational data mining* (pp. 413–418).

Li, Z., Chen, L., Jian, Y., Wang, H., Zhao, Y., Zhang, M., Xiao, K., Zhang, Y., Deng, H., & Hou, X. (2025). Aggregation or separation? Adaptive embedding message passing for knowledge graph completion. *Information Sciences*, *691*, Article 121639.

Li, I., Fabbri, A., Hingmire, S., & Radev, D. (2020). R-vgae: Relational-variational graph autoencoder for unsupervised prerequisite chain learning. In D. Scott, N. Bel, & C. Zong (Eds.), *Proceedings of the 28th international conference on computational linguistics* (pp. 1147–1157).

Li, I., Fabbri, A. R., Tung, R. R., & Radev, D. R. (2019). What should i learn first: Introducing lecturebank for nlp education and prerequisite chain learning. In *Proceedings of the AAAI conference on artificial intelligence*: *vol. 33*, (01), (pp. 6674–6681).

Li, Z., Huang, W., Gong, X., Luo, X., Xiao, K., Deng, H., Zhang, M., & Zhang, Y. (2025). Decoupled semantic graph neural network for knowledge graph embedding. *Neurocomputing*, *611*, Article 128614.

Li, I., Yan, V., Li, T., Qu, R., & Radev, D. R. (2021). Unsupervised cross-domain prerequisite chain learning using variational graph autoencoders. In *Annual meeting of the association for computational linguistics* (pp. 1005–1011). Association for Computational Linguistics.

Li, I., Yan, V., & Radev, D. (2021). Efficient variational graph autoencoders for unsupervised cross-domain prerequisite chains. In *ENLSP workshop, neurIPS 2021* (pp. 1–7).

Liang, C., Wu, Z., Huang, W., & Giles, C. L. (2015). Measuring prerequisite relations among concepts. In *Proceedings of the 2015 conference on empirical methods in natural language processing* (pp. 1668–1674).

Liang, C., Ye, J., Wang, S., Pursel, B., & Giles, C. L. (2018). Investigating active learning for concept prerequisite learning. In *Proceedings of the AAAI conference on artificial intelligence*: *vol. 32*, (1).

Liang, C., Ye, J., Wu, Z., Pursel, B., & Giles, C. (2017). Recovering concept prerequisite relations from university course dependencies. In *Proceedings of the AAAI conference on artificial intelligence*: *vol. 31*, (1).

Liang, C., Ye, J., Zhao, H., Pursel, B., & Giles, C. L. (2018). Active learning of strict partial orders: A case study on concept prerequisite relations. arXiv preprint arXiv:1801.06481.

Liu, S., Cheng, H., Wang, Y., He, Y., Fan, C., & Liu, Z. (2025). EvoPath: Evolutionary meta-path discovery with large language models for complex heterogeneous information networks. *Information Processing & Management*, *62*(1), Article 103920.

Liu, D., Qian, Y., & Zhao, J. (2025). A personalized path generation model for english learning based on natural language processing. *Journal of Combinatorial Mathematics and Combinatorial Computing*, *127*, 1093–1116.

Lu, W., Zhou, Y., Yu, J., & Jia, C. (2019). Concept extraction and prerequisite relation learning from educational data. In *Proceedings of the AAAI conference on artificial intelligence*: *vol. 33*, (01), (pp. 9678–9685).

Ma, J., Li, K., Zhang, F., Wang, Y., Luo, X., Li, C., & Qiao, Y. (2024). TaReT: Temporal knowledge graph reasoning based on topology-aware dynamic relation graph and temporal fusion. *Information Processing & Management*, *61*(6), Article 103848.

Manrique, R., Sosa, J., Marino, O., Nunes, B. P., & Cardozo, N. (2018). Investigating learning resources precedence relations via concept prerequisite learning. In *2018 IEEE/WIC/aCM international conference on web intelligence* (pp. 198–205). IEEE.

Nafa, F., Babour, A., & Melton, A. (2022). Prerequisite relations among knowledge units: A case study of computer science domain. *Computer Modeling in Engineering & Sciences (CMES)*, *133*(3), 639–652.

Ng, C., & Fung, Y. (2024). Educational personalized learning path planning with large language models. arXiv preprint arXiv:2407.11773.

Omerovic, S., Milutinovic, V., & Tomazic, S. (2001). Concepts, ontologies, and knowledge representation.

Pan, L., Li, C., Li, J., & Tang, J. (2017). Prerequisite relation learning for concepts in moocs. In *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 1: long papers)* (pp. 1447–1456).

Park, H. W., & Barnett, G. A. (2024). ICA fellows' networking patterns in terms of collaboration, citation, and bibliographic coupling and the relevance of co-ethnicity. *Scientometrics*, *129*(9), 5433–5452.

Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv preprint arXiv:2205.01833.

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review*, *14*(2), 249–255.

Scheines, R., Silver, E., & Goldin, I. M. (2014). Discovering prerequisite relationships among knowledge components. In *Proceedings of the 7th international conference on educational data mining* (pp. 355–356).

Shao, Z., Zhao, R., Yuan, S., Ding, M., & Wang, Y. (2022). Tracing the evolution of AI in the past decade and forecasting the emerging trends. *Expert Systems with Applications*, *209*, Article 118221.

Stavrinides, P., & Zuev, K. M. (2023). Course-prerequisite networks for analyzing and understanding academic curricula. *Applied Network Science*, *8*(1), 19.

Sun, J., He, Y., Xu, Y., Sun, J., & Sun, G. (2024). A learning-path based supervised method for concept prerequisite relations extraction in educational data. In *Proceedings of the 33rd ACM international conference on information and knowledge management* (pp. 2168–2177).

Sun, H., Li, Y., & Zhang, Y. (2022). Conlearn: Contextual-knowledge-aware concept prerequisite relation learning with graph neural network. In *Proceedings of the 2022 SIAM international conference on data mining* (pp. 118–126). SIAM.

Talukdar, P., & Cohen, W. (2012). Crowdsourced comprehension: predicting prerequisite structure in wikipedia. In *Proceedings of the seventh workshop on building educational applications using NLP* (pp. 307–315).

Tang, X., Liu, K., Xu, H., Xiao, W., & Tan, Z. (2023). Continual pre-training of language models for concept prerequisite learning with graph neural networks. *Mathematics*, *11*(12), 2780.

Wang, X., Chen, L., Lyu, D., Ban, T., Guan, Y., & Chen, Q. (2022). Research concept link prediction via graph convolutional network. In *2022 8th International conference on big data and information analytics* (pp. 220–225). IEEE.

Wang, B., Li, J., Chen, Z., Xu, B., & Xie, X. (2023). A new type of polytomous surmise system. *Journal of Mathematical Psychology*, *116*, Article 102803.

Wang, S., Ororbia, A., Wu, Z., Williams, K., Liang, C., Pursel, B., & Giles, C. L. (2016). Using prerequisites to extract concept maps fromtextbooks. In *Proceedings of the 25th acm international on conference on information and knowledge management* (pp. 317–326).

Weinberg, B. H. (1974). Bibliographic coupling: A review. *Information Storage and Retrieval*, *10*(5–6), 189–196.

Xiang, S., Jiang, X., Liu, B., Huang, Y., Tian, C., & Ma, Y. (2025). SciConNav: Knowledge navigation through contextual learning of extensive scientific research trajectories. *Journal of the Association for Information Science and Technology*, *76*(10), 1308–1339.

Xiao, K., Bai, Y., & Wang, S. (2021). Mining precedence relations among lecture videos in MOOCS via concept prerequisite learning. *Mathematical Problems in Engineering*, *2021*(1), Article 7655462.

Xiao, K., Bai, Y., & Wang, Z. (2022). Extracting prerequisite relations among concepts from the course descriptions. *International Journal of Software Engineering and Knowledge Engineering*, *32*(04), 503–523.

Xu, Y., & Yang, Y. (2024). Research on the patterns of the organic growth knowledge production based on a dual perspective of content and behaviour. *British Journal of Educational Technology*, *55*(3), 1145–1166.

Yang, A. J. (2025). Unraveling topic switching and innovation in science. *Information Processing & Management*, *62*(4), Article 104171.

Yang, X., Ding, L., Wang, W., & Yang, J. (2025). Identification of interdisciplinary research patterns based on the functional structures of IMRaD. *Information Processing & Management*, *62*(3), Article 104063.

Yang, J., & Hu, J. (2025). Scientific knowledge role transition prediction from a knowledge hierarchical structure perspective. *Journal of Informetrics*, *19*(1), Article 101612.

Yang, Q., & Liang, C. (2025). A second-classroom personalized learning path recommendation system based on large language model technology. *Applied Sciences*, *15*(14), 7655.

Yang, J., Wu, L., & Lyu, L. (2024). Research on scientific knowledge evolution patterns based on ego-centered fine-granularity citation network. *Information Processing & Management*, *61*(4), Article 103766.

Yang, R., Yang, B., Ouyang, S., She, T., Feng, A., Jiang, Y., Lecue, F., Lu, J., & Li, I. (2024). Leveraging large language models for concept graph recovery and question answering in NLP education. arXiv preprint arXiv:2402.14293.

Zeng, A., Shen, Z., Zhou, J., Fan, Y., Di, Z., Wang, Y., Stanley, H. E., & Havlin, S. (2019). Increasing trend of scientists to switch between topics. *Nature Communications*, *10*(1), 3439.

Zhang, M., Wang, J., Xiao, K., Wang, S., Zhang, Y., Chen, H., & Li, Z. (2025). Learning concept prerequisite relation via global knowledge relation optimization. In *Proceedings of the AAAI conference on artificial intelligence*: *vol. 39*, (2), (pp. 1638–1646).

Zhang, J., Zhang, X., Du, X., & Lu, T. (2024). EBCPL: A novel evidence-based method for concept prerequisite relation learning. In *International conference on database systems for advanced applications* (pp. 390–405). Springer.