

DeepConnect: A Visual Analytics System for Bridging Interdisciplinary Research Collaborations

Yingchaojie Feng, Zekai Shao, Yiqun Sun, Yixuan Tang, and Anthony K. H. Tung

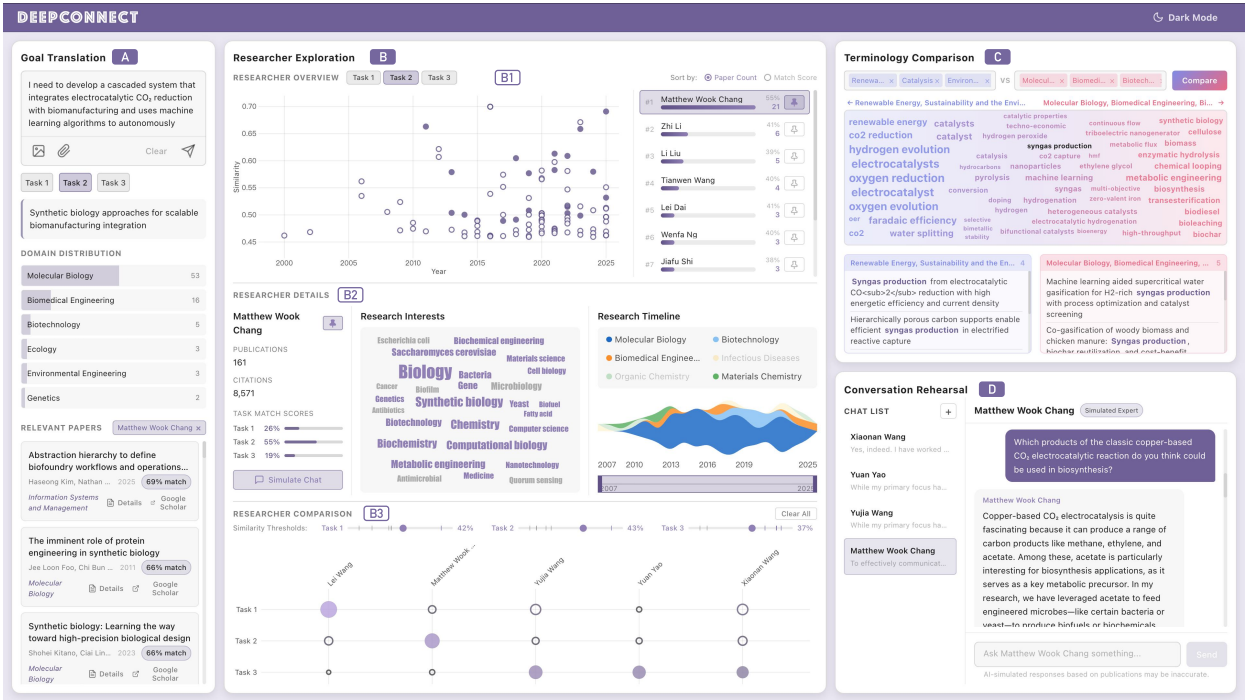


Fig. 1. DeepConnect is a visual analytics system designed to bridge interdisciplinary research collaboration. The interface features multiple coordinated views: (A) The Goal Translation view decomposes overarching research goals into actionable domain-specific tasks; (B) The Researcher Exploration view enables multi-dimensional evaluation of candidate expertise and complementarity; (C) The Terminology Comparison view bridges semantic gaps by visually mapping terminology gaps and overlaps; and (D) The Conversation Rehearsal environment supports publication-grounded communication rehearsal prior to formal engagement.

Abstract—Interdisciplinary research collaboration is crucial for scientific innovation, but it remains difficult to initiate in practice. Existing collaborator discovery approaches are often constrained by disciplinary boundaries and static researcher profiles that do not reflect the specific context of a new collaboration goal. As a result, researchers struggle to translate open-ended collaboration goals into domain-specific tasks, evaluate candidate researchers' fit and complementarity, and establish common ground before initial contact. To address these challenges, we present DeepConnect, an LLM-augmented visual analytics system for interdisciplinary collaborator discovery. DeepConnect translates collaboration ideas into domain-specific tasks, retrieves relevant papers to ground cross-domain exploration, and provides coordinated visualizations for exploring and comparing candidate researchers. It further reveals terminology gaps and overlaps across domains and supports publication-grounded conversation rehearsal to help users prepare for outreach. We evaluate DeepConnect through two case studies, a user study, and a component-level evaluation, showing its value for complementary team formation, idea refinement, and pre-contact preparation. The online demo is available at <https://deepconnect.sg>.

Index Terms—Visual analytics, interdisciplinary collaboration, collaborator discovery, semantic alignment

1 INTRODUCTION

Interdisciplinary collaboration is widely recognized as a crucial driver of scientific innovation [15, 40, 45, 49], particularly for complex prob-

lems that transcend traditional boundaries [40, 45]. Integrating diverse methodologies and knowledge bases enables researchers to address complex problems intractable within a single domain [56, 66]. Despite these benefits, initiating such partnerships remains practically difficult [10, 57]. A promising conceptual goal does not spontaneously translate into an actionable collaboration. Instead, researchers must navigate the intricate process of identifying suitable cross-domain candidates, assessing their complementary expertise, and establishing semantic common ground prior to formal engagement.

Current collaborator discovery practices inadequately support this complex transition [2, 27, 81]. While serendipitous encounters and collegial introductions foster initial trust, they rely on constrained

- Yingchaojie Feng, Yixuan Tang, and Anthony K. H. Tung are with the National University of Singapore. E-mail: [feng.y, dstyx, dstunga]@nus.edu.sg.
- Zekai Shao is with Fudan University. E-mail: zkshao23@m.fudan.edu.cn.
- Yiqun Sun is with Magellan Technology Research Institute. E-mail: duke.sun@mtri.co.jp.

Manuscript received xx xxx. 202x; accepted xx xxx. 202x. Date of Publication xx xxx. 202x; date of current version xx xxx. 202x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.202x.xxxxxxx

personal networks and lack scalability [5, 18, 41]. Academic search engines provide broader access but struggle with cross-domain exploration [1, 2, 81]. Their reliance on exact keyword matching forces users to guess unfamiliar domain terminology. Furthermore, because search results are predominantly publication-centric, users must manually synthesize isolated papers to reconstruct a candidate’s evolving trajectory and evaluate complementary fit. Although automated recommendation methods [21, 69, 77, 82] reduce search effort, they frequently prioritize historically similar or proximate researchers, failing to account for the specific context of novel collaborative objectives [37, 38].

Drawing on these limitations and our formative study, we identify three core challenges in interdisciplinary collaborator discovery. **(C1) Extracting Methodological Needs from Domain Contexts.** Researchers often articulate collaborative goals through the narratives and terminology of their own disciplines. This obscures the underlying methodological requirements, making it difficult to search for applicable methods and expertise from other domains [36, 66]. **(C2) Synthesizing Expertise from Fragmented Literature.** Assessing a candidate requires a comprehensive understanding of their research trajectory. Since standard search engines retrieve isolated publications, users struggle to manually aggregate these records to evaluate collaborative fit [34]. **(C3) Bridging Semantic Gaps in Communication.** Even when promising candidates are identified, differing academic paradigms complicate initial interactions. Researchers need support to overcome these communication barriers and assess project feasibility before initiating formal outreach [7].

DeepConnect is an LLM-augmented visual analytics system for interdisciplinary collaborator discovery. Starting from a researcher’s initial collaboration goal, the system translates it into structured tasks to facilitate broad exploration and cross-domain comparison. It then retrieves relevant literature from OpenAlex to ground these tasks in existing research, revealing how the goal connects to different academic fields. Through coordinated interactive visualizations, users can navigate candidate pools, trace research trajectories, and comparatively assess task alignment and complementary expertise. To bridge semantic divides prior to formal engagement, the system visualizes terminology differences and overlaps, while supporting publication-grounded conversation rehearsals. These capabilities help researchers evaluate collaboration feasibility and prepare for targeted initial outreach.

We evaluate DeepConnect through two case studies, a user study with 12 researchers, and a component-level evaluation. The results demonstrate that DeepConnect frames interdisciplinary collaborator discovery as a structured, evidence-based exploratory process. The case studies highlight its effectiveness in assembling complementary teams and iteratively refining collaboration ideas. Furthermore, the user study indicates that the system increases transparency in early-stage exploration, supports the contextual comparison of candidate researchers, and boosts users’ confidence in cross-domain communication. Taken together, these findings confirm that DeepConnect empowers researchers to transition from abstract collaborative intent to grounded candidate selection and informed initial contact.

The primary contributions of this paper are summarized as follows:

- We characterize the problem space of interdisciplinary collaborator discovery through a formative study, distilling these empirical challenges into actionable design requirements.
- We present DeepConnect, an LLM-augmented visual analytics system for translating goals, exploring researchers, and preparing targeted outreach across domains.
- We evaluate DeepConnect through two case studies, a user study, and a component-level evaluation, providing empirical evidence of its effectiveness in complementary team formation, cross-domain idea refinement, and communication preparation.

2 RELATED WORK

2.1 Scientific Collaborator Recommendation

Scientific collaborator recommendation [26] has transitioned from structural link prediction to semantic team formation. Early methods primarily leveraged topological features in co-authorship networks [21, 77].

Techniques such as random walks, heterogeneous meta-paths (e.g., PathPredict [58], ArnetMiner [63]), and random forest classifiers [48] were frequently used to predict future collaborations based on historical network proximity. However, these structural approaches are susceptible to homophily bias [3, 12]. They tend to recommend candidates from a researcher’s existing social circle, which can limit interdisciplinary discovery. To incorporate research content, subsequent studies framed academic matchmaking as a multi-constraint optimization problem [4, 6], aiming to balance complementary expertise with network communication costs. While these methods utilize semantic profiles [25], they typically rely on coarse-grained, static aggregations of a scholar’s publication history, which may not capture the specific requirements of a new project.

Although these methods are useful for identifying relevant scholars, they remain less suited to early-stage interdisciplinary exploration, where users may start with abstract goals rather than clear query terms or well-defined expertise needs. DeepConnect targets this missing step by translating collaboration intent into actionable, domain-specific tasks and grounding candidate retrieval in task-relevant literature.

2.2 Visual Analytics of Academic Data

Visual analytics is widely used to explore complex academic data [16, 20, 23], supporting tasks such as literature analysis, bibliometric evaluation, and trend discovery [19, 31, 78]. Existing tools are generally categorized into macro-level science mapping [39] and micro-level scholar profiling [11, 53, 74]. For example, CiteSpace [8] and VOSviewer [67] visualize broad disciplinary trends and citation networks. However, they often lack the granularity required to evaluate individual collaborators.

At the micro-level, ScholarPlot [35] provides visual summaries of academic achievements. To support deeper egocentric analysis, tools like Influence Flowers [54] map bi-directional citation influence, while GeneticFlow [34, 75] and GeneticPrism [62] track research evolution alongside citation impact. Furthermore, interactive systems such as VISPubCompAS [71] and ACSeeker [70] enable scholarly exploration through user-guided faceted filtering. Recently, VizCV [28] incorporated AI-assisted analysis to model career trajectories based on topic evolution, impact metrics, and collaboration dynamics.

These systems offer valuable but mainly retrospective views of scholarly activity, focusing on what fields or scholars have done rather than whether a candidate fits a new collaboration goal. DeepConnect addresses this need through task-oriented views that connect user-specific tasks with temporal publication evidence and candidate comparison.

2.3 Visual Analytics for Knowledge Alignment

Visual analytics has been extensively employed to bridge semantic gaps and align domain-specific terminology across disciplines [42, 79]. Early text-based tools focused on revealing lexical differences across document facets or tracking temporal word frequency trends [9, 29]. To support deeper semantic analysis, ConceptVector [47] leverages word embeddings to interactively construct concepts, while some other systems [30, 52, 73] map syntactic changes, model behaviors, or large language model (LLM) representations to human-understandable visual spaces. Beyond navigating terminological disparities, establishing cross-domain understanding requires broader visual sensemaking [50, 64, 72]. SenseMap [43] and SensePath [44] capture this cognitive process by recording analytic provenance to help users curate fragmented information. Ultimately, this individual sensemaking must transition into collaborative grounding, where visualizations act as shared “boundary objects” [33]. Foundational research demonstrates that shared visual interfaces and interaction minutiae are critical for establishing situational awareness, supporting consensus building, and reducing communication costs in cooperative environments [17, 22, 24].

While such work supports semantic analysis, sensemaking, and collaborative grounding, it largely focuses on knowledge artifacts or on interactions after collaboration begins. Before initial contact, researchers must interpret unfamiliar terminology and prepare for communication. DeepConnect targets this pre-collaboration stage with domain terminology comparison and publication-grounded rehearsal.

3 FORMATIVE STUDY

We conducted a formative study to investigate the practical challenges of initiating interdisciplinary collaboration and to inform the design of DeepConnect. Our goal was to understand how researchers articulate their requirements, search across domains, evaluate potential collaborators, and prepare for initial communication. Through analyzing archival collaboration requests and interviewing domain experts, we translated these observed barriers into concrete design requirements.

3.1 Methodology and Participants

Our study followed a two-phase qualitative approach, combining archival data analysis with expert interviews to systematically investigate this problem space.

First, we conducted a **thematic analysis** on an anonymized corpus of interdisciplinary collaboration proposals collected from a university-wide initiative. This corpus included requests from principal investigators across various STEM disciplines, including biological sciences, chemistry, physics, and food science, who were seeking external technical expertise. Analyzing these archival documents revealed how researchers naturally articulate their interdisciplinary needs.

Second, we conducted **semi-structured interviews** with six domain experts recruited through purposive sampling from our university. They included principal investigators and senior postdoctoral researchers from diverse STEM backgrounds, all with interdisciplinary collaboration experience or interest. Each interview lasted approximately 45–60 minutes and was conducted in person or online. The interviews examined collaborator search strategies, evaluation difficulties, and communication barriers. These expert interviews were covered by our university’s Departmental Ethics Review Committee approval (no. 001325), and all interviewees provided informed consent. Because the interviews involved potentially sensitive early-stage collaboration ideas, we documented them through detailed written notes rather than audio recordings. Two authors analyzed the proposals and interview notes using open coding and synthesized key findings.

3.2 Current Practice and Pain Points

Synthesizing the archival data and expert interviews revealed a clear mismatch between the highly specific requirements of interdisciplinary collaboration and the limited support offered by existing exploratory workflows. We identified recurring bottlenecks spanning the entire collaboration process, from defining initial goals to preparing for contact.

Coupling of Domain Context and Methodological Needs. Our analysis indicated that interdisciplinary collaboration goals often intertwine domain context with implicit methodological requirements. Researchers rarely articulate their needs using standardized technical terms recognizable across disciplines. Instead, they rely on detailed, domain-specific narratives grounded in their immediate scientific problems. For example, a biologist seeking a spatial segmentation method might describe the morphological complexity of cellular structures, while a materials scientist requiring predictive modeling might emphasize chemical synthesis conditions. Experts noted that this domain-specific framing obscures the underlying computational tasks from outsiders. Consequently, potential collaborators often fail to recognize the relevance of their methods, even when they possess the exact expertise required.

Bottlenecks in Current Discovery Channels. When actively seeking potential collaborators, experts reported substantial friction across their primary discovery channels.

Limited Support from In-Person Academic Networking. Although academic events serve as a common starting point for interdisciplinary discovery, experts noted that such interactions are inherently unscaleable and subject to high uncertainty. Informal conversations facilitate the exchange of broad interests but frequently fail to expose the task-level alignment necessary to determine project viability. Relying on physical networking restricts the search space to localized, serendipitous encounters, lacking the systematic reach required to comprehensively identify suitable collaborators for specific objectives.

Friction in Academic Search Platforms. Traditional academic search engines impose a substantial cognitive burden. Experts

described a persistent terminology mismatch: querying with familiar domain-specific terms predominantly returns immediate peers, whereas attempting to adopt unfamiliar terminology often yields an overwhelming number of irrelevant results. Furthermore, because these platforms are inherently document-centric, evaluating a candidate necessitates manual review of extensive publication records. This tedious process hinders the ability to form a holistic view of an individual’s expertise, track its evolution over time, or compare multiple candidates against a unified collaboration goal.

The Exploratory Limits of Text-Based LLMs. Several experts reported utilizing LLMs to circumvent the limitations of traditional search. While these systems can mitigate terminology mismatches, text-based conversational interfaces exhibit clear limitations for exploratory discovery. Their linear responses are difficult to inspect systematically and fail to provide a macro-level overview of the researcher landscape, complicating multidimensional comparisons. Experts emphasized that discovering interdisciplinary partners is a high-stakes process demanding publication-grounded evidence and interactive exploration, which purely text-based interfaces cannot adequately fulfill.

The Risks of Cross-Disciplinary Miscommunication. While identifying a suitable collaborator is a critical step, it does not resolve the inherent difficulties of subsequent communication. Experts emphasized that differing academic paradigms frequently cause conceptual misalignment. Specifically, domain-specific terminologies and habitual expressions act as significant barriers to mutual understanding, resulting in highly inefficient dialogue. Furthermore, evaluating the feasibility of a proposed idea during initial discussions is inherently difficult. Due to a lack of cross-domain familiarity, researchers often underestimate the methodological complexity of a task within the target discipline, or conversely, overlook potentially fruitful collaborative opportunities. Therefore, researchers highlighted a critical need to validate ideas and establish shared context prior to initial contact.

3.3 Design Requirements

To address these challenges, we distilled four design requirements (R1–R4) for supporting interdisciplinary collaboration initiation.

R1: Translate Collaboration Goals into Methodological Tasks.

To isolate technical needs from domain-specific contexts, the system must translate broad collaboration goals into actionable tasks. It should let researchers express intentions naturally and map them to methodological requirements recognizable across disciplines.

R2: Enable Comparative Evaluation of Researcher Fit. To alleviate the cognitive burden of manual candidate evaluation, the system must provide multidimensional visual evidence of a researcher’s expertise. It should summarize an individual’s academic trajectory, explicitly assess their relevance to the extracted tasks, and support visual comparisons of complementary strengths among candidates.

R3: Reveal Terminology Gaps and Overlaps Across Domains.

To mitigate friction caused by unfamiliar vocabularies, the system should visualize terminological differences and alignments between disciplines. By highlighting shared concepts and clarifying domain-specific expressions, the system can help users build shared context before communication.

R4: Enable Expertise-Grounded Conversation Rehearsal. To minimize unproductive discussions, the system must offer a safe, interactive environment to prepare for initial contact. It should enable users to systematically test the methodological feasibility of their ideas against a candidate’s domain expertise, allowing researchers to iteratively refine their communication strategies.

4 SYSTEM OVERVIEW

DeepConnect is a visual analytics system designed to facilitate interdisciplinary collaborator discovery using scholarly data from OpenAlex (Fig. 2D). The workflow is organized around the requirements identified in the formative study. It begins with **Goal Translation** (Fig. 2A), where we use an LLM to translate initial goals into actionable domain-specific tasks and retrieve cross-domain literature (R1). These tasks guide **Researcher Exploration** (Fig. 2B), which combines temporal scatterplots, detailed profiles, and comparative matrices to examine

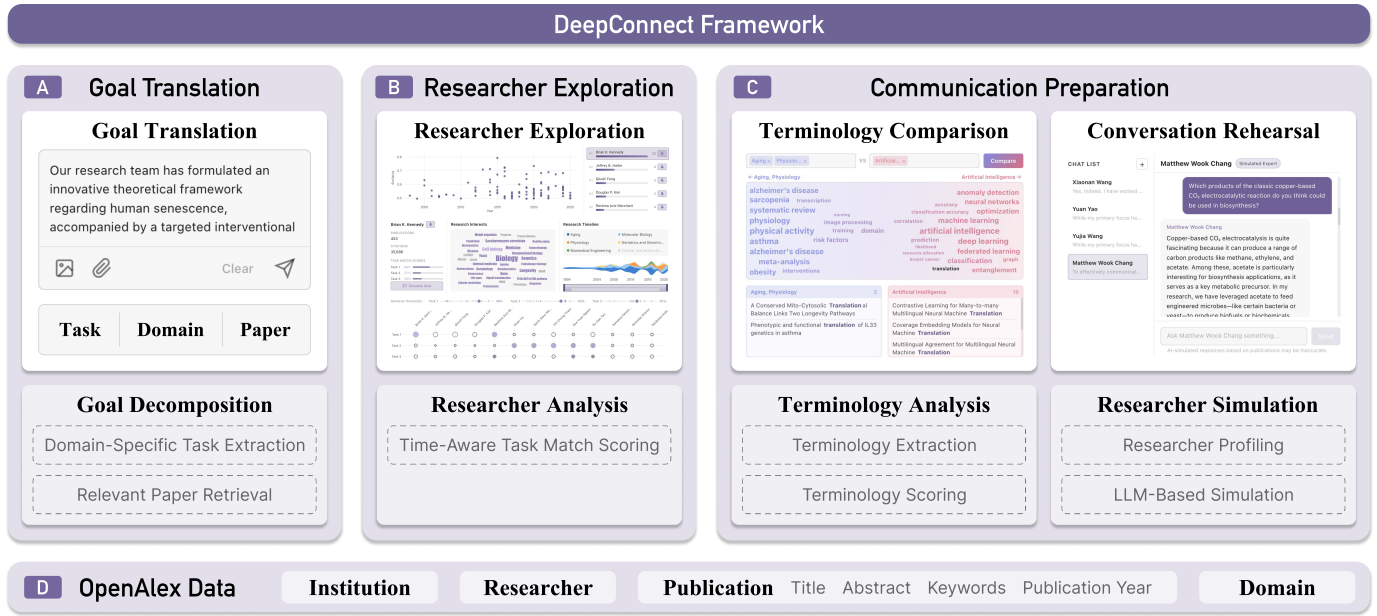


Fig. 2. The visual analytics framework of DeepConnect. The system comprises four coordinated modules: (A) Goal Translation decomposes collaborative objectives into domain-specific Tasks and retrieves Relevant Papers; (B) Researcher Exploration supports the macroscopic overview, detailed profiling, and comparative assessment of candidate scholars; (C) Communication Preparation reveals terminological overlaps and enables publication-grounded Conversation Rehearsals to bridge cognitive gaps; and (D) the foundational OpenAlex Data that grounds the aforementioned three modules by connecting institutions, researchers, publications, and domains.

evolving expertise and task fit (R2). **Communication Preparation** (Fig. 2C) visualizes terminology gaps and overlaps to establish common ground (R3) and provides publication-grounded conversation rehearsal for feasibility exploration and communication planning (R4).

4.1 Data Collection and Processing

We use OpenAlex¹, an open-access catalog of scholarly entities, as the primary data source for our system. OpenAlex represents scholarly data through interconnected entities, including *Institution*, *Author* (researcher), *Work* (publication), and *Domain*. Within this structure, authors are linked to their affiliated institutions, while works are linked to their contributing authors and associated domains. Each Work provides publication-level metadata, including its title, abstract, keywords, and publication year. These data provide the basis for semantic analysis, terminology comparison, and temporal trajectory tracking.

To ground our system in a practical application scenario, we constructed a representative local-university subset containing affiliated researchers and their associated publications. Facilitating collaborations within a single institution lowers administrative and spatial barriers, making the identified interdisciplinary opportunities more actionable and realistic for researchers. We use this curated dataset throughout our case studies and user study to evaluate the system’s practical utility.

4.2 Goal Decomposition

We design the Goal Decomposition module to transform open-ended collaboration goals into actionable domain-specific tasks and retrieve relevant papers to ground subsequent exploration (R1). To accommodate different stages of research ideation, DeepConnect accepts multimodal input, including textual descriptions, conceptual images, and project proposal documents.

Domain-Specific Task Extraction. We use GPT-5.1 to derive up to five semantically diverse tasks that cover the collaboration goal and support systematic cross-domain search.

Relevant Paper Retrieval. To guide candidate search, we ground the extracted tasks in relevant literature. We use the all-MiniLM-L6-v2 embedding model to encode both the extracted tasks and the OpenAlex

paper metadata (paper title, abstract, and keywords) into a shared semantic space [14, 59–61]. The system then computes the cosine similarity between the task embeddings and paper embeddings to rank the literature.

4.3 Researcher Analysis

We identify candidate researchers by extracting authors from the literature retrieved for each task t . Sourcing candidates directly from these relevant publications ensures that each scholar is associated with verifiable scholarly metadata, including specific papers, publication years, and research domains, which provides a reliable foundation for subsequent analysis (R2).

We use **Time-Aware Task Match Scoring** to estimate the alignment between a candidate researcher r and a task t based on both semantic relevance and publication recency. As research interests evolve over time, recent publications indicate current focus, while earlier works reflect foundational expertise. For each paper $p \in P_r$, we calculate a relevance score $Sim(t, p)$ using the cosine similarity between the task and paper embeddings. This score is then adjusted by a temporal weight $w(p)$ based on the publication year Y_p :

$$w(p) = \alpha + (1 - \alpha) \exp(-\lambda(Y_{\max} - Y_p)) \quad (1)$$

Here, $Y_{\max}$ represents the latest publication year in the dataset, λ determines the decay rate, and $\alpha \in [0, 1]$ establishes a lower bound for the weight of older publications (set to 0.5 in our implementation).

Let $Top_k(P_r, t)$ represent the set of k top-ranked papers in P_r based on the weighted relevance score $w(p) \cdot Sim(t, p)$, where $k = \min(x, |P_r|)$ and x defaults to 5. The final match score is defined as:

$$Score(r, t) = \frac{1}{k} \sum_{p \in Top_k(P_r, t)} w(p) \cdot Sim(t, p) \quad (2)$$

We aggregate only the top-matched papers to avoid diluting task-specific relevance with unrelated historical topics, especially for senior researchers with diverse publication records. This averaging approach supports robust ranking, filtering, and evaluation of candidates by capturing both sustained and current expertise.

¹<https://openalex.org/>

4.4 Terminology Analysis

To prepare users for target-domain concepts, we analyze domain terminology to reveal unfamiliar terms and shared vocabulary across fields (R3). This process consists of three steps.

Domain-Specific Terminology Extraction. We first construct a text corpus for domain terminology comparison. For each domain, the system retrieves the top 1000 most relevant papers using the embedding model, a threshold chosen to balance literature coverage with real-time performance. Grounded in this academic context, we deploy a lightweight LLM (GPT-4.1-mini) to extract conceptual terminology from the titles and abstracts, ensuring high-speed processing across thousands of documents. The extraction is designed to retain representative technical terms that reflect the conceptual focus of the domain. To keep the resulting terminology set analytically useful, the extraction excludes variable names, proper nouns, and generic academic verbs.

Terminology Scoring. After extraction, the system estimates the importance of each term with a domain-level TF-IDF formulation [51, 55]. For a term w in the domain D_A , the term frequency $TF(w, D_A)$ measures how frequently w is observed in the retrieved literature of D_A . The inverse document frequency is computed over the global set of retrieved papers P_{all} to downweight terms that broadly appear across many papers. The resulting importance score $S(w, D_A)$ is defined as

$$S(w, D_A) = TF(w, D_A) \cdot \log \left(\frac{|P_{all}|}{|\{p \in P_{all} : w \in p\}|} \right) \quad (3)$$

Here, $|\{p \in P_{all} : w \in p\}|$ denotes the number of papers in which the term w appears.

To quantify terminology overlap across domains, the system further computes a relative frequency ratio for each extracted term. Given the user domain D_B and a comparison domain D_A , let $Freq(w, D)$ denote the number of retrieved papers in domain D that contain term w . The relative ratio $R(w)$ is defined as

$$R(w) = \frac{Freq(w, D_A)}{Freq(w, D_A) + Freq(w, D_B)} \quad (4)$$

This metric normalizes the term distribution within the range $[0, 1]$. A value close to 1 indicates that the term is used primarily in D_A , while a value close to 0 indicates that it is used primarily in D_B . A value near 0.5 indicates shared terminology that appears with similar frequency in both domains. This metric provides the quantitative basis for the horizontal placement of terms in the Terminology Comparison view.

4.5 Publication-Grounded Researcher Simulation

To support feasibility exploration before initial contact, we provide an interactive, publication-grounded simulation for rehearsing questions and anticipating misunderstandings (R4).

Researcher Profiling. Upon selecting a candidate researcher r , the system aggregates their key metadata to construct a bounded grounding context $\mathcal{C}$. This context incorporates the researcher’s associated domains ($Domains_r$) and the abstracts of their complete publication record (P_r) to capture their academic trajectory and expertise.

$$\mathcal{C} = \{Name_r, Domains_r, Abstracts(P_r)\} \quad (5)$$

LLM-Based Simulation. By conditioning an LLM (GPT-5.1) on this localized context, the generation probability of any response R given a user query Q is explicitly modeled as $P(R | Q, \mathcal{C})$. To mitigate generative hallucinations, we impose strict semantic boundaries on the simulation [76]. The model is configured to ground its outputs exclusively in the provided publication records, reflect the candidate researcher’s established academic perspective, and acknowledge topics falling outside their documented expertise [80]. This evidence-based architecture ensures a reliable preparation environment for communication across domains.

5 VISUAL ANALYTICS SYSTEM

DeepConnect guides users from open-ended collaboration goals to actionable partnerships through an interactive visual analytics workflow. The exploratory process begins by translating user inputs into domain-specific tasks to establish a foundational literature base. With a structured task representation, users systematically evaluate and compare candidate researchers through coordinated visual evidence. Once promising candidates are identified, the system facilitates communication preparation by visualizing terminology gaps and supporting publication-grounded conversation rehearsal before initial contact.

5.1 Collaboration Goal Translation

The **Goal Translation** (Figure 1A) view is designed to structure users’ initial collaboration goals and connect them to verifiable literature evidence. Users first articulate their objectives in the *Research Goal* panel. The system processes this input and generates tabs for each extracted domain-specific task. These tabs serve as global contextual filters. Selecting a *Task* updates subsequent visual components to reflect the active analytical focus.

To provide a macroscopic perspective on the interdisciplinary nature of the active task, the *Domain Distribution* panel uses a horizontal bar chart to summarize the domain distribution of the top 100 retrieved publications. Interactive filtering is naturally supported: clicking a specific domain bar updates the adjacent *Relevant Papers* panel to display only literature from the selected field.

The *Relevant Papers* panel displays the retrieved seed publications, detailing their task relevance scores, publication years, and domain labels. Users can expand individual items to read full abstracts or navigate to the external Google Scholar website. When a candidate is selected later, this list updates to show only their task-matched publications.

5.2 Contextual Researcher Exploration

The **Researcher Exploration** (Figure 1B) view coordinates three vertically arranged panels to support both macroscopic discovery and individual-level profiling.

The *Researcher Overview* (Figure 1B1) panel supports initial candidate assessment. It features a scatterplot mapping publication years to the x-axis and embedding similarity scores to the y-axis, alongside a ranked list of candidate researchers sortable by publication volume or task match scores (Section 4.3). This panel tightly couples cross-view interactions. Within the scatterplot, brushing a specific region to isolate papers dynamically recalculates the candidate rankings, and clicking a data point links directly to the underlying literature in the *Relevant Papers* panel. Conversely, within the ranked list, hovering over a candidate researcher highlights their publications in the scatterplot to reveal temporal work patterns, and selecting a candidate updates the subsequent *Relevant Papers* panels for deeper inspection.

The *Researcher Details* (Figure 1B2) panel offers a microscopic profile of the selected candidate researcher. Alongside basic academic metadata and aggregate task match scores, a *Research Interests* word cloud maps keyword frequency to text size to reveal the individual’s overarching focus. Adjacent to this, a *Research Timeline* streamgraph encodes publication volume into stream width to visualize their evolutionary trajectory across disciplines. To guide analytical focus, keywords and streams highly relevant to the active task are highlighted, whereas irrelevant ones are in gray text or semi-transparency. If a candidate appears suitable, users can initiate a simulated dialogue in the *Conversation Rehearsal* view.

The *Researcher Comparison* (Figure 1B3) panel features an interactive dot matrix to help users balance expertise across disciplines. Rows represent the extracted tasks, while columns display candidate researchers pinned during the exploration phase. Dot size and opacity visually encode a candidate researcher’s match score for the corresponding task. By interactively adjusting independent score threshold sliders, dots meeting the criteria remain solid, whereas those falling below are rendered as hollow circles. This visual encoding makes task-level expertise gaps visible and supports users in comparing whether selected candidates collectively cover the task requirements.

5.3 Terminology Comparison

To bridge the semantic gap, the **Terminology Comparison** (Figure 1C) view maps terminology gaps and overlaps across domains. Based on the relative frequency ratio calculated in Section 4.4, a word cloud distributes the extracted terms along a horizontal axis. Terms positioned near the lateral edges represent domain-specific terminology that users should approach with caution or learn in advance to prevent misunderstandings. Conversely, centrally located terms indicate shared terminology that helps establish common ground. To reinforce this visual mapping, a color gradient transitions from blue on the left to red on the right, with shared terms blending into purple to symbolize the conceptual convergence of the compared domains.

To support contextual traceability, the system links these terms directly to the underlying literature. When a specific term is selected, the bottom panel displays sentences from the literature containing that term. This enables users to examine how unfamiliar domain terminology is practically applied in the target discipline’s literature.

5.4 Conversation Rehearsal

The **Conversation Rehearsal** (Figure 1D) view adopts a familiar messaging layout, featuring a *Chat List* on the left for managing multiple sessions and a main dialogue window on the right. To establish analytical trust and mitigate over-reliance on the LLM, the system marks the candidate researcher as a simulated expert and explicitly disclaims that all responses are derived from their past publications (Section 4.5). Reinforcing this factual grounding, the session initiates with a self-introduction generated directly from these academic records. This environment supports communication rehearsal, conceptual clarification, and feasibility exploration before contact.

6 CASE STUDY

To evaluate the effectiveness and utility of *DeepConnect* in real-world exploratory scenarios, we present two case studies. These cases involve *E1* and *E2*, corresponding to user-study participants *P11* and *P12*, respectively. Both used the system to address their actual interdisciplinary challenges. The first case focuses on task-driven team assembly for an AI-driven biohybrid system, while the second explores semantic alignment and iterative refinement in a sensor-vision fusion project. Together, they show how the system connects broad planning with domain-specific grounding.

6.1 Task-Driven Complementary Team Assembly

E1 is an environmental engineering researcher specializing in electrocatalytic CO₂ reduction. To overcome the efficiency bottlenecks of single chemical conversion, *E1* aimed to develop a cascaded system integrating upstream electrocatalysis with downstream biomanufacturing, regulated autonomously by machine learning algorithms. This objective required an interdisciplinary team spanning materials science, synthetic biology, and artificial intelligence. *E1* used *DeepConnect* to break down the project goal and identify researchers whose expertise could jointly support these needs.

Goal Translation and Task Grounding. *E1* first input this high-level goal into the Goal Translation view (Figure 1A). The system decomposed the request into three domain-specific tasks: (1) electrocatalytic reaction engineering for optimized CO₂ reduction processes, (2) synthetic biology approaches for scalable biomanufacturing, and (3) autonomous control systems using machine learning for real-time reactor optimization. *E1* used these tasks to systematically explore potential collaborators.

Contextual Profiling and Collaborator Identification. Focusing on the first task, *E1* found that the retrieved literature in the Domain Distribution was concentrated in renewable energy and sustainability. Although already familiar with reactor design, *E1* needed a materials scientist with strong expertise in catalyst development. In the ranked list of the Researcher Overview view, *E1* identified Prof. Lei Wang, who exhibited the highest matching score and publication count. To examine this candidate more closely, *E1* turned to the Researcher Details view (Figure 3). The Research Interests word cloud emphasized materials-related keywords such as *Electrochemistry*, *Electrocatalyst*,

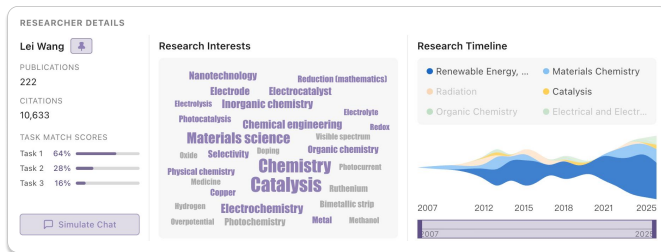


Fig. 3. **Case 1:** The Researcher Details panel illustrates Prof. Lei Wang’s expertise for the electrocatalytic reaction engineering task. The Research Interests and Research Timeline confirm sustained, task-relevant materials science output for the task context.

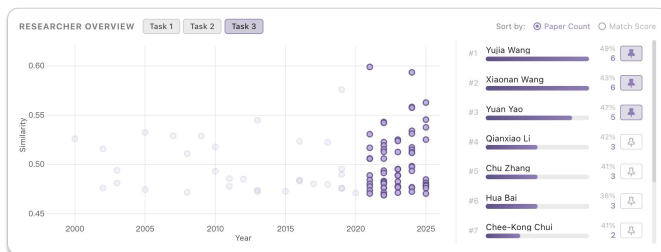


Fig. 4. **Case 1:** Temporal filtering within the Researcher Overview for the artificial intelligence Task. Brushing the scatterplot to isolate recent publications dynamically updates candidate rankings, enabling the user to identify active researchers such as Prof. Xiaonan Wang.

and *Carbon monoxide*. The Research Timeline streamgraph further showed sustained recent output in renewable energy. Taken together, these visual cues indicated a strong and active research focus aligned with the task, leading *E1* to pin Prof. Lei Wang.

Visual Semantic Alignment and Conversation Rehearsal. Proceeding to the second task, *E1* noted from the Domain Distribution a focus on molecular biology and biotechnology. Since this field fell outside their expertise, *E1* compared candidates using the scatterplot in the Researcher Overview (Figure 1B1). Prof. Matthew Wook Chang emerged as a strong candidate due to his high match score and recent publication volume, with the Research Timeline (Figure 1B2) verifying his sustained engagement in molecular biology. To prepare cross-disciplinary communication, *E1* inspected the Terminology Comparison (Figure 1C) for conceptual alignments and gaps. Unfamiliar biological concepts (e.g., *metabolic flux*, *enzymatic hydrolysis*) indicated areas requiring preliminary study. Meanwhile, domain-specific electrochemistry terms (e.g., *Faradaic efficiency*) prompted *E1* to consider how to explain them in language accessible to a biologist. Shared concepts like *Syngas production* established a practical conversational entry point. Guided by this semantic preparation, *E1* used the Conversation Rehearsal (Figure 1D) to discuss biological carbon sources and electrochemical metrics in accessible terms.

Interactive Visual Filtering and Qualitative Assessment. For the third task, the target literature centered on control systems and artificial intelligence. Observing that the initially top-ranked candidate primarily held older publications, *E1* brushed the scatterplot to filter for recent academic works (Figure 4). This dynamic update revealed three researchers with comparable relevance scores and publication volumes. *E1* pinned these candidates to conduct comparative conversation rehearsals. During discussions on parameter optimization, Prof. Xiaonan Wang offered the most relevant insights, mentioning specific experience in applying machine learning to complex chemical reactions like hydrothermal processes. This indicated robust dual expertise in algorithms and chemical engineering, prompting *E1* to select her for the AI-control task.

Visual Comparison and Team Assembly. Finally, *E1* examined the Researcher Comparison (Figure 1B3). After adjusting the similar-

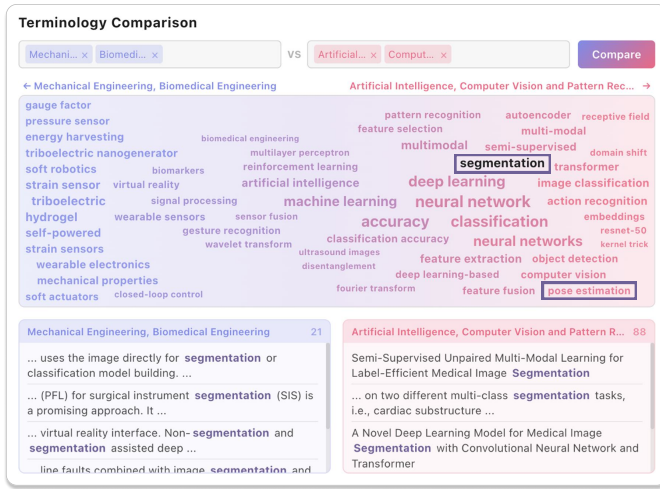


Fig. 5. **Case 2:** Semantic alignment within the Terminology Comparison for grasp intent recognition. Highlighting terminological gaps across fields related to mechanical engineering and artificial intelligence reveals some complementary concepts such as “segmentation” and “pose estimation”, prompting a novel sensor-vision fusion approach.

ity threshold sliders, the dot matrix confirmed that the three selected candidates (Prof. Lei Wang, Prof. Chang, and Prof. Xiaonan Wang) provided complementary and comprehensive coverage across the three tasks. Moreover, the matrix revealed that Prof. Xiaonan Wang exhibited higher relevance for the first two tasks compared to her peers. This visual pattern indicated a strong cross-disciplinary background, which could effectively lower communication barriers during the trilateral collaboration. Supported by this visual and conversational evidence, *E1* finalized the team assembly and proceeded to draft formal invitations.

6.2 Bridging Cognitive Gaps for Sensor-Vision Fusion

E2, a researcher in mechanical and biomedical engineering, developed a soft robotic exoskeleton glove equipped with high-density strain sensors and an electromyography (sEMG) module. A key bottleneck in this system was that physical sensor-based grasp intent recognition remained highly susceptible to environmental noise and user variance. To address this issue through artificial intelligence, *E2* utilized *DeepConnect* to discover cross-disciplinary expertise.

Initial Exploration and Conversation Rehearsal. *E2* inputted her collaborative objective into the Goal Translation: “I developed a soft robotic exoskeleton. I want to apply advanced deep learning algorithms to process my sEMG and soft strain sensor data to improve the accuracy of grasp intent recognition.” Following the system’s task extraction, *E2* reviewed the recommended candidates and selected Prof. Angela Yao due to her extensive publication record in deep learning and complex sequence processing. Seeking algorithmic guidance, *E2* initiated a Conversation Rehearsal to determine which models could enhance intent classification for noisy, user-dependent data. The simulated expert provided actionable insights, suggesting models capable of temporal aggregation (e.g., C2F-TCN) to capture long-range dependencies, alongside feature disentanglement techniques to isolate motion intent from user-specific noise.

Visual Sensemaking and Cognitive Shift. Acknowledging feature extraction and disentanglement as promising technical directions, *E2* sought to bridge her knowledge gap regarding these computational concepts. To understand how physical sensor data aligned with these algorithms, *E2* examined the Terminology Comparison (Figure 5). While her domain on the left was anchored by familiar terms like *strain sensors* and *wearable sensors*, *E2* noticed that the expert’s domain on the right predominantly featured computer vision terminology such as *segmentation* and *pose estimation*. This observation prompted a critical realization. She deduced that temporal disentanglement algorithms might underperform if applied exclusively to one-dimensional physical

signals. Since grasping inherently involves spatial interaction with the environment, *E2* recognized that visual information would provide spatial context missing from her current physical sensors.

Iterative Refinement. Building on this insight, *E2* hypothesized that integrating a micro-camera to capture spatial priors could significantly simplify the noise processing challenge. To validate this architectural shift, she returned to the Conversation Rehearsal. *E2* proposed a new framework that combined the expert’s feature extraction methods with visual inputs, suggesting the use of lightweight *segmentation* or *pose estimation* models to capture macroscopic environmental features for subsequent *feature fusion* with the microscopic strain data. The simulated expert confirmed that visual priors effectively compensate for the spatial blind spots of physical sensors. Through this iterative process, *DeepConnect* facilitated a crucial pivot from a unimodal algorithmic query to a comprehensive vision-motor fusion strategy.

7 USER STUDY

We conducted a user study to assess *DeepConnect*’s effectiveness and usability in interdisciplinary matchmaking. The study examined how researchers used the visual analytics workflow for cross-domain exploration and collected qualitative feedback on the system design. The user study was covered by the same ethics approval (no. 001325).

7.1 Participants and Apparatus

We recruited 12 researchers (*P1* – *P12*) from a local university through email invitations and targeted outreach, including one professor, four postdoctoral researchers, five doctoral candidates, and two master’s students. All were actively engaged in academic research, with research experience ranging from approximately 1 to 10 years. Their domains included computer and information sciences (AI, HCI, and clinical informatics), earth and environmental sciences (climate change, carbon cycling, hydrology, and environmental engineering), physical and life sciences (materials science, chemical biology, biomedical engineering, and bioenergy), and behavioral sciences (psychology). All had experience or strong interest in interdisciplinary research.

Each one-on-one study session lasted approximately 60 minutes. To accommodate individual preferences, sessions occurred either in person using a provided laptop or remotely via video conferencing with screen control capabilities. All participants provided informed consent covering participation, audio recording, and screen capture. Upon completion, each individual received a compensation of SGD 20.

7.2 Study Procedure

The evaluation comprised four structured phases.

Introduction and Background Survey (10 minutes). The session began with the project background and research motivation behind *DeepConnect*. Participants then completed a brief questionnaire on academic domains, research experience, and search strategies.

System Tutorial (10 minutes). An instructor guided participants through the *DeepConnect* interface using a predefined usage scenario. This tutorial explained the visual encodings, interactions, and data processing logic across the core modules. Participants then explored sample data to learn the workflow before the task.

Open-Ended Exploration (20 minutes). During the core task phase, participants formulated a realistic interdisciplinary idea aligned with their current interests. Using this objective, they used *DeepConnect* to discover, evaluate, and simulate communication with potential collaborators. Participants followed a think-aloud protocol, verbalizing their strategies and reasoning. The instructor observed silently and intervened only when clarification was needed.

Post-Study Feedback and Interview (20 minutes). Following the exploration, participants completed a 5-point Likert questionnaire (Figure 6) assessing the system’s effectiveness and usability. Finally, a semi-structured interview gathered rating rationales and design input.

7.3 Qualitative Feedback and Observations

The ratings and semi-structured interviews provided insights into the system’s utility and participants’ analytical processes. We organize the feedback by workflow stage.

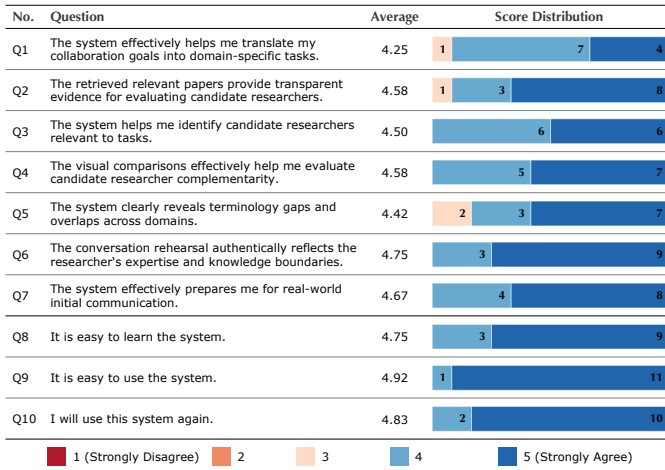


Fig. 6. Quantitative user-study results from the 5-point Likert-scale questionnaire. The score distributions indicate consistently positive participant feedback on the system's effectiveness in supporting intent translation, expertise evaluation, cognitive alignment, and overall usability.

7.3.1 Goal Translation and Evidence Transparency

Participants highly rated the system's ability to **translate collaboration goals into domain-specific tasks** (Q1: 4.25/5). P2 noted that the generated tasks bridged familiar domain objectives with target discipline requirements. P6 also confirmed the accuracy of the extracted tasks, stating, "The system actually reminded me of potential research directions that I had overlooked." However, P12 noted that some sub-tasks emphasized different aspects than expected. When the analytical focus of the extracted tasks required adjustments, P7 and P12 found the system highly flexible. They could easily guide the model to output more aligned task structures by refining their input descriptions.

Regarding the **exploration transparency** (Q2: 4.58/5), participants confirmed that the *Relevant Papers* established a robust foundation for exploration. P3 emphasized that visualizing these relevant papers offered crucial interpretability by justifying the high rankings of specific researchers. Furthermore, P7 observed that displaying publication years alongside these papers enabled users to quickly distinguish a candidate's recent active focus from past endeavors.

7.3.2 Researcher Evaluation

Participants confirmed that the *Researcher Exploration* effectively **identified researchers relevant to the decomposed tasks** (Q3: 4.50/5). An interesting pattern emerged regarding the profiles of recommended candidates. P1 observed that the system successfully surfaced boundary-spanning researchers with established dual expertise at the intersection of the targeted domains. For instance, P6, a medical researcher seeking an artificial intelligence collaborator, noticed that the system prioritized artificial intelligence scholars with medical crossover experience over purely theoretical computer scientists. P6 stated, "This pragmatic recommendation is extremely helpful for actual research, as it lowers the communication barrier and facilitates much more aligned collaborations." P3 suggested impact-factor filters and author-position indicators for assessing academic influence.

Participants confirmed that the visual comparisons were highly effective for **evaluating candidate complementarity** (Q4: 4.58/5). P1 commended the *Researcher Comparison*, emphasizing its substantial utility in constructing well-balanced collaborative groups. P2 contrasted this visual approach with traditional conversational interfaces, noting that standard text-based models demand repetitive prompting to extract and compare candidate profiles. The interactive matrix made complementary strengths immediately visible. P10 further suggested automatically recommending expert combinations for team assembly.

7.3.3 Communication Preparation

Participants confirmed that the *Terminology Comparison* effectively supported cross-disciplinary preparation by **revealing conceptual differences across domains** (Q5: 4.42/5). P6 noted that the visual encoding exposed ontological variations, contrasting the broad definition of artificial intelligence in public health with its specific sub-fields in computer science. To enhance this feature, P11 suggested brief LLM-generated definitions for displayed terms.

Participants widely utilized the *Conversation Rehearsal* to verify **candidates' knowledge boundaries** (Q6: 4.75/5). Users frequently employed a trust-building strategy by initially simulating chats with familiar researchers. After verifying these publication-grounded responses, participants proceeded to interact with unfamiliar candidates. P8 appreciated that the simulated researchers explicitly acknowledged their research boundaries while delivering comprehensive answers within their domain. Highlighting the efficiency of this feature, P12 stated, "Using simulated chats is much more efficient and less cognitively demanding than reading through a researcher's entire publication history." Furthermore, P7 observed that engaging with these digital personas helped differentiate candidates who shared nearly identical task match scores in the visual chart.

Participants confirmed that the *Terminology Comparison* and *Conversation Rehearsal* successfully **prepared them for real-world communication** (Q7: 4.67/5). Probing the simulation with controversial questions regarding data sharing, P6 found that the system accurately reflected a candidate's conservative stance, grounded in privacy concerns from their prior publications. P6 deemed this capability highly valuable for navigating sensitive topics during formal networking. P7 reported that simulations offered tailored advice to avoid outdated methods.

7.3.4 System Usability and Adoption

DeepConnect received high ratings for its **learnability** and **usability** (Q8: 4.75/5, Q9: 4.92/5). P3 found the interface intuitive to navigate. P5 noted that the visual encodings, specifically the *Research Timeline* and *Researcher Comparison*, proved accessible following the initial tutorial. Highlighting the rich interactive design, P1 observed that the seamless coordination across multiple visual modules significantly reduced context switching. This fluid transition helped users focus on academic evaluation.

Participants expressed a strong **reuse intention** (Q10: 4.83/5). Multiple users advocated for deploying DeepConnect on public academic platforms. P3 described the design as "incredibly impressive" and confirmed the intention to adopt the tool for future projects. Furthermore, P8 recognized the broader application potential of the system and inquired about future release plans, underscoring the practical utility of the visual analytics workflow in real-world collaborative scenarios.

7.4 Suggestions and Future Directions

Participants proposed several directions for future system iterations to further support interdisciplinary collaboration.

Broadening Collaboration Context. Real-world academic partnerships involve constraints beyond technical matching. P1 suggested that visualizing collaboration motives, such as rapid algorithmic development versus long-term domain impact, would help identify partners with aligned goals. Additionally, P3 noted that incorporating institutional affiliations could assist users in navigating policy restrictions and geographical preferences. P8 proposed mapping academic kinship networks through shared co-authors to reveal introduction pathways.

Enabling Proactive Coordination in Group Rehearsals. Evaluating team dynamics prior to formal contact remains challenging. To address this, P9 and P10 suggested expanding the *Conversation Rehearsal* to support multi-party discussions with several selected researchers simultaneously [46]. Within this simulated group environment, P10 proposed that the AI assume a proactive leadership role. The system could automatically decompose the user's Research Goal into a structured framework and actively drive the dialogue among the simulated researchers [13]. Additionally, the AI could strategically

Table 1. Measures used in the expert-user component-level evaluation.

Measure	Definition
Goal Translation	
Faithfulness (Faith)	Preserves the user’s intended collaboration goal.
Actionability (Act)	Yields concrete tasks for targeted search and evaluation.
Coverage (Cov)	Covers the key needs implied by the goal.
Task Match Scoring	
Task Fit	Aligns ranked candidates with the target task.
Expertise Recency (Recency)	Prioritizes recent task-relevant expertise.
Terminology Comparison	
Representativeness (Rep)	Captures representative domain concepts.
Comparison Clarity (Clar)	Makes terminology gaps and overlaps easy to compare.

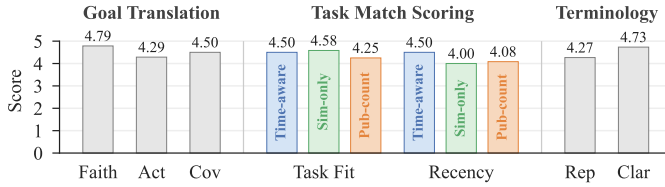


Fig. 7. Mean expert-user ratings for the evaluated goal translation, task match scoring, and terminology comparison outputs.

prompt the user for input at critical junctures to advance the conversation. This proactive coordination would alleviate the user’s burden of orchestrating complex cross-disciplinary discussions from scratch.

Modeling Academic Schools of Thought and Theoretical Conflicts. *P6* identified the need to model theoretical alignments and academic schools of thought. Scientific communities often contain distinct theoretical camps with conflicting foundational assumptions. Future systems could address this by mining semantic content and co-authorship networks to cluster scholars based on their theoretical inclinations. Visually encoding these distinct schools of thought would enable users to navigate theoretical conflicts or intentionally foster diverse intellectual debates before formal engagement.

8 COMPONENT-LEVEL EVALUATION

We further evaluate the effectiveness of three key components in DeepConnect: Goal Translation (Section 4.2), Task Match Scoring (Section 4.3), and Terminology Comparison (Section 4.4). Since no benchmark directly fits this evaluation setting, six expert users from the user study (*P2*, *P5*, *P6*, *P7*, *P9*, and *P11*) participated in a separate follow-up component-level evaluation, rating multiple outputs generated from their own open-ended system use.

Evaluation Design. We use 5-point Likert scales, with measures defined in Table 1, to assess whether component outputs appropriately capture users’ goals (R1), support task-relevant researcher discovery (R2), and reveal distinctive terminology patterns across domains (R3). For task match scoring, we compare Time-aware Task Match Scoring with two baselines: Similarity-only and Publication-count, which rank candidates by semantic similarity and task-relevant publication counts, respectively. The three ranking methods are anonymized and presented to users in randomized order.

Findings. Figure 7 summarizes the rating results across the three evaluated outputs. Goal translation received positive ratings for faithfulness (4.79), actionability (4.29), and coverage (4.50). *P7* noted that for an early-stage, open-ended goal, the generated tasks faithfully and comprehensively captured the intent, but could further specify concrete next steps. For task match scoring, Time-aware maintained comparable task fit to Similarity-only (4.50 vs. 4.58) while improving recency (4.50 vs. 4.00). Terminology comparison received ratings of 4.27 for representativeness and 4.73 for clarity, suggesting its value for supporting users’ understanding of terminology across domains.

9 DISCUSSION

We discuss design lessons for LLM-augmented visual analytics, DeepConnect’s potential as a bidirectional collaborative ecosystem, workflow generalizability, and limitations of our cross-sectional evaluation.

9.1 Design Lessons for LLM-Augmented Visual Analytics

DeepConnect demonstrates how visualization and LLM-based interaction jointly support interdisciplinary collaborator discovery. At the macro level, visualizations organize task, paper, and temporal evidence for side-by-side candidate comparison, helping users narrow the search space more efficiently than textual dialogue over fragmented evidence. At the micro level, LLM-based interaction enables flexible probing of feasibility, knowledge boundaries, and communication strategies for a project idea. Together, they provide a grounded, adaptive workflow for exploratory assessment.

These strengths suggest tighter coupling between visual and LLM-based interactions. For example, LLMs could explain unfamiliar concepts in terminology views, while visualizations could expose evidence behind LLM-based dialogue, reveal agent behavior, and diagnose retrieval failures or hallucination risks [32, 65, 68]. Future systems should make visual and language-based interactions mutually informative for transparent, coherent analysis.

9.2 Transitioning to a Collaborative Ecosystem

Currently, the system uses academic publications as proxies for human expertise. A broader platform could support direct exchange of methodological and physical resources. Future versions could also allow researchers to publish collaboration intents and available capabilities, turning the current egocentric workflow into bidirectional matchmaking. By integrating user-uploaded assets and domain-specific repositories, the system could map shared scientific assets, from datasets and models to specialized equipment and compounds. Together, these extensions would move the tool from personnel discovery toward a broader ecosystem for allocating intellectual and tangible scientific resources.

9.3 System Generalizability

Although the current implementation uses OpenAlex, the workflow is data- and domain-agnostic. Goal translation, semantic retrieval, and visual cognitive alignment could adapt to other structured and unstructured data, including preprint repositories, patent databases, and grant records. Such sources would capture early-stage research trends before peer review, giving users a more immediate view of emerging interdisciplinary opportunities. They could also reveal emerging expertise before it appears in mature citation or publication histories.

9.4 Longitudinal Evaluation Limitations

A primary limitation is the cross-sectional evaluation. Although the user study suggests that *DeepConnect* can reduce immediate cognitive load during exploration and communication preparation, interdisciplinary partnerships develop over time. Our framework focuses on the pre-collaboration phase and omits long-term team viability and relational dynamics. Future work should conduct longitudinal, in-the-wild deployments across institutions, connecting early exploration behavior with subsequent collaboration milestones. Tracking follow-up contacts and joint outputs would clarify how early exploration shapes sustained collaboration over time.

10 CONCLUSION

Initiating interdisciplinary collaboration requires goal translation, candidate evaluation, and semantic common ground. We present DeepConnect, an LLM-augmented visual analytics system that turns broad goals into actionable tasks, supports evidence-based *Researcher Exploration* with visual evidence of trajectories, fit, and complementarity, and bridges cognitive gaps through *Terminology Comparison* and publication-grounded *Conversation Rehearsal*. Across two case studies, a 12-researcher user study, and a component-level evaluation, the findings suggest reduced cognitive load, context-aware assessment, and more communicable early collaboration ideas.

ACKNOWLEDGMENTS

This research is supported by the National Research Foundation, Singapore and Infocomm Media Development Authority under its Trust Tech Funding Initiative. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of National Research Foundation, Singapore and Infocomm Media Development Authority.

REFERENCES

- [1] K. Balog, Y. Fang, M. De Rijke, P. Serdyukov, L. Si, et al. Expertise retrieval. *Found. Trends Inf. Retr.*, 6(2–3):127–256, 2012. doi: 10.1561/1500000024 2
- [2] F. Beck. PUREsuggest: Citation-based literature search and visual exploration with keyword-controlled rankings. *IEEE Trans. Vis. Comput. Graph.*, 31(1):316–326, 2025. doi: 10.1109/TVCG.2024.3456199 1, 2
- [3] G. Békés and G. I. Ottaviano. Cultural homophily and collaboration in superstar teams. *Manag. Sci.*, 71(10):8149–8168, 2025. doi: 10.1287/mnsc.2022.01799 2
- [4] N. Berktaş and H. Yaman. A branch-and-bound algorithm for team formation on social networks. *INFORMS J. Comput.*, 33(3):1162–1176, 2021. doi: 10.1287/ijoc.2020.1000 2
- [5] R. S. Burt. Structural holes and good ideas. *Am. J. Sociol.*, 110(2):349–399, 2004. doi: 10.1086/421787 2
- [6] M. Büyükbayraktar and A. Robbett. Team formation with complementary skills. *J. Econ. Manag. Strategy*, 28(4):713–733, 2019. doi: 10.1111/jems.12296 2
- [7] M. Caccamo, D. Pittino, and F. Tell. Boundary objects, knowledge integration, and innovation management: A systematic review of the literature. *Technovation*, 122, art. no. 102645, 2023. doi: 10.1016/j.technovation.2022.102645 2
- [8] C. Chen. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.*, 57(3):359–377, 2006. doi: 10.1002/asi.20317 2
- [9] C. Collins, F. B. Viegas, and M. Wattenberg. Parallel tag clouds to explore and analyze faceted text corpora. In *Proc. IEEE VAST*, pp. 91–98. IEEE, Piscataway, 2009. doi: 10.1109/VAST.2009.5333443 2
- [10] J. N. Cummings and S. Kiesler. Collaborative research across disciplinary and organizational boundaries. *Soc. Stud. Sci.*, 35(5):703–722, 2005. doi: 10.1177/0306312705055535 1
- [11] T. Dang, H. N. Nguyen, and V. Pham. WordStream: Interactive visualization for topic evolution. In J. Johansson, F. Sadlo, and G. E. Marai, eds., *EuroVis 2019 - Short Papers*, pp. 103–107. The Eurographics Association, 2019. doi: 10.2312/evs.20191178 2
- [12] F. Fabbri, F. Bonchi, L. Boratto, and C. Castillo. The effect of homophily on disparate visibility of minorities in people recommender systems. In *Proc. ICWSM*, vol. 14, pp. 165–175. AAAI Press, Palo Alto, 2020. doi: 10.1609/icwsml.v14i1.7288 2
- [13] Y. Feng, H. Luo, Z. Lin, Y. Sun, P. Wei, L. B. Hsieh et al. OrchMAS: Orchestrated reasoning with multi collaborative heterogeneous scientific expert structured agents. *arXiv preprint arXiv:2603.03005*, 2026. doi: 10.48550/arXiv.2603.03005 8
- [14] Y. Feng, Y. Sun, Y. Sun, M. Zhu, Q. Huang, A. K. H. Tung et al. Don't reinvent the wheel: Efficient instruction-following text embedding based on guided space transformation. In *Proc. ACL*, pp. 24511–24525. Association for Computational Linguistics, Vienna, Austria, 2025. doi: 10.18653/v1/2025.acl-long.1196 4
- [15] S. Fortunato, C. T. Bergstrom, K. Börner, J. A. Evans, D. Helbing, S. Milojević et al. Science of science. *Science*, 359(6379):eaao0185, 2018. doi: 10.1126/science.aao0185 1
- [16] T.-L. Fung, J.-K. Chou, and K.-L. Ma. A design study of personal bibliographic data visualization. In *Proc. IEEE PacificVis*, pp. 244–248. IEEE, Piscataway, 2016. doi: 10.1109/PACIFICVIS.2016.7465279 2
- [17] D. Gergle, R. E. Kraut, and S. R. Fussell. Language efficiency and visual technology: Minimizing collaborative effort with visual information. *J. Lang. Soc. Psychol.*, 23(4):491–517, 2004. doi: 10.1177/0261927X04269589 2
- [18] M. S. Granovetter. The strength of weak ties. *Am. J. Sociol.*, 78(6):1360–1380, 1973. doi: 10.1086/225469 2
- [19] Z. Guo, J. Tao, S. Chen, N. V. Chawla, and C. Wang. SD²: Slicing and dicing scholarly data for interactive evaluation of academic performance. *IEEE Trans. Vis. Comput. Graph.*, 29(8):3569–3585, 2023. doi: 10.1109/TVCG.2022.3163727 2
- [20] H. Hao, Y. Cui, Z. Wang, and Y.-S. Kim. Thirty-two years of IEEE VIS: Authors, fields of study and citations. *IEEE Trans. Vis. Comput. Graph.*, 29(1):1016–1025, 2023. doi: 10.1109/TVCG.2022.3209422 2
- [21] D. Hassan and M. A. Hasan. Supervised link prediction in co-authorship networks based on author node-based features. *arXiv preprint arXiv:2505.21673*, 2025. doi: 10.48550/arXiv.2505.21673 2
- [22] J. Heer and M. Agrawala. Design considerations for collaborative visual analytics. *Inf. Vis.*, 7(1):49–62, 2008. doi: 10.1057/palgrave.ivs.9500167 2
- [23] F. Heimerl, Q. Han, S. Koch, and T. Ertl. CiteRivers: Visual analytics of citation patterns. *IEEE Trans. Vis. Comput. Graph.*, 22(1):190–199, 2016. doi: 10.1109/TVCG.2015.2467621 2
- [24] L. Homaeian, J. R. Wallace, and S. D. Scott. Joint action storyboards: A framework for visualizing communication grounding costs. *Proc. ACM Hum.-Comput. Interact.*, 5(CSCW1), art. no. 28, 27 pp., 2021. doi: 10.1145/3449102 2
- [25] L. Huang, X. Ni, X. Chen, and Y. Zhang. Measuring interdisciplinary interactions using citation analysis and semantic analysis. In *Proceedings of the 1st Workshop on AI + Informetrics (AII 2021)*, co-located with *iConference 2021*, vol. 2871 of *CEUR Workshop Proceedings*, pp. 140–152. CEUR-WS.org, 2021. 2
- [26] O. Husain, N. Salim, R. A. Alias, S. Abdelsalam, and A. Hassan. Expert finding systems: A systematic review. *Appl. Sci.*, 9(20), art. no. 4250, 2019. doi: 10.3390/app9204250 2
- [27] S. Latif and F. Beck. VIS author profiles: Interactive descriptions of publication records combining text and visualization. *IEEE Trans. Vis. Comput. Graph.*, 25(1):152–161, 2019. doi: 10.1109/TVCG.2018.2865022 1
- [28] V. Lazárik, M. Agus, B. Kozlíková, and P.-P. Vázquez. VizCV: AI-assisted visualization of researchers' publications tracks. *arXiv preprint arXiv:2505.08691*, 2025. doi: 10.48550/arXiv.2505.08691 2
- [29] B. Lee, N. H. Riche, A. K. Karlson, and S. Carpendale. SparkClouds: Visualizing trends in tag clouds. *IEEE Trans. Vis. Comput. Graph.*, 16(6):1182–1189, 2010. doi: 10.1109/TVCG.2010.194 2
- [30] H. Li, Z. Wen, Q. Jiang, C. Li, Y. Wu, Y. Yang et al. ConceptViz: A visual analytics approach for exploring concepts in large language models. *IEEE Trans. Vis. Comput. Graph.*, 32(1):57–67, 2026. doi: 10.1109/TVCG.2025.3634806 2
- [31] Z. Li, C. Zhang, S. Jia, and J. Zhang. GaleX: Exploring the evolution and intersection of disciplines. *IEEE Trans. Vis. Comput. Graph.*, 26(1):1182–1192, 2020. doi: 10.1109/TVCG.2019.2934667 2
- [32] J. Lu, B. Pan, J. Chen, Y. Feng, J. Hu, Y. Peng et al. AgentLens: Visual analysis for agent behaviors in LLM-based autonomous systems. *IEEE Trans. Vis. Comput. Graph.*, 31(8):4182–4197, 2025. doi: 10.1109/TVCG.2024.3394053 9
- [33] L. F. Luna-Reyes, L. J. Black, W. Ran, D. L. Andersen, H. Jarman, G. P. Richardson et al. Modeling and simulation as boundary objects to facilitate interdisciplinary research. *Syst. Res. Behav. Sci.*, 36(4):494–513, 2019. doi: 10.1002/sres.2564 2
- [34] Y. Luo, L. Shi, M. Xu, Y. Ji, F. Xiao, C. Hu et al. Impact-oriented contextual scholar profiling using self-citation graphs. In *Proc. ACM SIGKDD*, pp. 4572–4583. ACM, New York, 2023. doi: 10.1145/3580305.3599845 2
- [35] D. Majeti, E. Akleman, M. E. Ahmed, A. M. Petersen, B. Uzzi, and I. Pavlidis. Scholar Plot: Design and evaluation of an information interface for faculty research performance. *Front. Res. Metr. Anal.*, 4, art. no. 6, 2020. doi: 10.3389/frma.2019.00006 2
- [36] Y. Mao, D. Wang, M. Muller, K. R. Varshney, I. Baldini, C. Dugan et al. How data scientists work together with domain experts in scientific collaborations: To find the right answer or to ask the right question? *Proc. ACM Hum.-Comput. Interact.*, 3(GROUP), art. no. 237, 23 pp., 2019. doi: 10.1145/3361118 2
- [37] S. M. McNee, J. Riedl, and J. A. Konstan. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *Extended Abstracts of Proc. CHI*, pp. 1097–1101. ACM, New York, 2006. doi: 10.1145/1125451.1125659 2
- [38] M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annu. Rev. Sociol.*, 27(1):415–444, 2001. doi: 10.1146/annurev.soc.27.1.415 2
- [39] A. Narechania, A. Karduni, R. Wesslen, and E. Wall. VITALITY: Promoting serendipitous discovery of academic literature with transformers & visual analytics. *IEEE Trans. Vis. Comput. Graph.*, 28(1):486–496, 2022. doi: 10.1109/TVCG.2021.3114820 2
- [40] National Academy of Sciences, National Academy of Engineering, and

- Institute of Medicine. *Facilitating interdisciplinary research*. National Academies Press, Washington, DC, 2005. doi: 10.17226/11153 1
- [41] M. E. Newman. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. U.S.A.*, 98(2):404–409, 2001. doi: 10.1073/pnas.98.2.404 2
- [42] D. Newman-Griffis, V. Sivaraman, A. Perer, E. Fosler-Lussier, and H. Hochheiser. TextEssence: A tool for interactive analysis of semantic shifts between corpora. In *Proc. NAACL-HLT Demos*, pp. 106–115. Association for Computational Linguistics, Online, 2021. doi: 10.18653/v1/2021.naacl-demos.13 2
- [43] P. H. Nguyen, K. Xu, A. Bardill, B. Salman, K. Herd, and B. W. Wong. Sensemap: Supporting browser-based online sensemaking through analytic provenance. In *Proc. IEEE VAST*, pp. 91–100. IEEE, Piscataway, 2016. doi: 10.1109/VAST.2016.7883515 2
- [44] P. H. Nguyen, K. Xu, A. Wheat, B. W. Wong, S. Attfield, and B. Fields. SensePath: Understanding the sensemaking process through analytic provenance. *IEEE Trans. Vis. Comput. Graph.*, 22(1):41–50, 2016. doi: 10.1109/TVCG.2015.2467611 2
- [45] K. Okamura. Interdisciplinarity revisited: evidence for research impact and dynamism. *Palgrave Commun.*, 5(1), art. no. 141, 2019. doi: 10.1057/s41599-019-0352-4 1
- [46] B. Pan, J. Lu, K. Wang, L. Zheng, Z. Wen, Y. Feng et al. AgentCoord: Visually exploring coordination strategy for LLM-based multi-agent collaboration. *Comput. Graph.*, 132, art. no. 104338, 2025. doi: 10.1016/j.cag.2025.104338 8
- [47] D. Park, S. Kim, J. Lee, J. Choo, N. Diakopoulos, and N. Elmqvist. ConceptVector: Text visual analytics via interactive lexicon building using word embedding. *IEEE Trans. Vis. Comput. Graph.*, 24(1):361–370, 2018. doi: 10.1109/TVCG.2017.2744478 2
- [48] M. Pavlov and R. Ichise. Finding experts by link prediction in co-authorship networks. *FEWS*, 290:42–55, 2007. 2
- [49] A. Porter and I. Rafols. Is science becoming more interdisciplinary? measuring and mapping six research fields over time. *Scientometrics*, 81(3):719–745, 2009. doi: 10.1007/s11192-008-2197-2 1
- [50] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. A. Keim. Knowledge generation model for visual analytics. *IEEE Trans. Vis. Comput. Graph.*, 20(12):1604–1613, 2014. doi: 10.1109/TVCG.2014.2346481 2
- [51] G. Salton and C. Buckley. Term-weighting approaches in automatic text retrieval. *Inf. Process. Manag.*, 24(5):513–523, 1988. doi: 10.1016/0306-4573(88)90021-0 5
- [52] C. Schätzle, M. Hund, F. Dennig, M. Butt, and D. Keim. HistoBankVis: Detecting language change via data visualization. In *Proceedings of the NoDaLiDa 2017 workshop on processing historical language*, pp. 32–39. Linköping University Electronic Press, Linköping, 2017. 2
- [53] L. Shi, C. Wang, Z. Wen, H. Qu, C. Lin, and Q. Liao. 1.5D egocentric dynamic network visualization. *IEEE Trans. Vis. Comput. Graph.*, 21(5):624–637, 2015. doi: 10.1109/TVCG.2014.2383380 2
- [54] M. Shin, A. Soen, B. T. Readshaw, S. M. Blackburn, M. Whitelaw, and L. Xie. Influence flowers of academic entities. In *Proc. IEEE VAST*, pp. 1–10. IEEE, Piscataway, 2019. doi: 10.1109/VAST47406.2019.8986934 2
- [55] K. Sparck Jones. A statistical interpretation of term specificity and its application in retrieval. *J. Doc.*, 28(1):11–21, 1972. doi: 10.1108/eb026526 5
- [56] A. Stirling. A general framework for analysing diversity in science, technology and society. *J. R. Soc. Interface*, 4(15):707–719, 2007. doi: 10.1098/rsif.2007.0213 1
- [57] D. Stokols, K. L. Hall, B. K. Taylor, and R. P. Moser. The science of team science: overview of the field and introduction to the supplement. *Am. J. Prev. Med.*, 35(2):S77–S89, 2008. doi: 10.1016/j.amepre.2008.05.002 1
- [58] Y. Sun, R. Barber, M. Gupta, C. C. Aggarwal, and J. Han. Co-author relationship prediction in heterogeneous bibliographic networks. In *Proc. ASONAM*, pp. 121–128. IEEE, Piscataway, 2011. doi: 10.1109/ASONAM.2011.112 2
- [59] Y. Sun, Q. Huang, Y. Tang, A. K. H. Tung, and J. Yu. A general framework for producing interpretable semantic text embeddings. *arXiv preprint arXiv:2410.03435*, 2024. doi: 10.48550/arXiv.2410.03435 4
- [60] Y. Sun, Q. Huang, A. K. H. Tung, and J. Yu. Position: Text embeddings should capture implicit semantics, not just surface meaning. In *Proc. ICML PMLR*, 2026. To appear. doi: 10.48550/arXiv.2506.08354 4
- [61] Y. Sun, Q. Huang, Z. Xu, Y. Sun, Y. Tang, and A. K. H. Tung. One swallow does not make a summer: Understanding semantic structures in embedding spaces. *arXiv preprint arXiv:2512.00852*, 2025. doi: 10.48550/arXiv.2512.00852 4
- [62] Y. Sun, Z. Liu, Y. Luo, L. Xia, and L. Shi. GeneticPrism: Multifaceted visualization of citation-based scholarly research evolution. *IEEE Trans. Vis. Comput. Graph.*, 31(10):9024–9038, 2025. doi: 10.1109/TVCG.2025.3589485 2
- [63] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. ArnetMiner: Extraction and mining of academic social networks. In *Proc. ACM SIGKDD*, pp. 990–998. ACM, New York, 2008. doi: 10.1145/1401890.1402008 2
- [64] S.-O. Tergan. The use of digital concept maps as cognitive tools for managing knowledge and knowledge resources. In *AAAI Spring Symposium: Reasoning with Mental and External Diagrams: Computational Modeling and Spatial Assistance*, pp. 73–76. AAAI Press, Palo Alto, 2005. 2
- [65] H. Tian, Y. Feng, Z. Wen, H. Li, M. Zhu, and W. Chen. RAGEExplorer: A visual analytics system for the comparative diagnosis of RAG systems. *IEEE Trans. Vis. Comput. Graph.*, 32(6):4807–4817, 2026. doi: 10.1109/TVCG.2026.3694443 9
- [66] B. Uzzi, S. Mukherjee, M. Stringer, and B. Jones. Atypical combinations and scientific impact. *Science*, 342(6157):468–472, 2013. doi: 10.1126/science.1240474 1, 2
- [67] N. Van Eck and L. Waltman. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2):523–538, 2010. doi: 10.1007/s11192-009-0146-3 2
- [68] K. Wang, B. Pan, Y. Feng, Y. Wu, J. Chen, M. Zhu et al. XGraphRAG: Interactive visual analysis for graph-based retrieval-augmented generation. In *Proc. IEEE PacificVis*, pp. 1–11. IEEE, Piscataway, 2025. doi: 10.1109/PacificVis64226.2025.00005 9
- [69] W. Wang, J. Liu, T. Tang, S. Tuarob, F. Xia, Z. Gong et al. Attributed collaboration network embedding for academic relationship mining. *ACM Trans. Web*, 15(1), art. no. 4, 20 pp., 2020. doi: 10.1145/3409736 2
- [70] Y. Wang, T.-Q. Peng, H. Lu, H. Wang, X. Xie, H. Qu et al. Seek for success: A visualization approach for understanding the dynamics of academic careers. *IEEE Trans. Vis. Comput. Graph.*, 28(1):475–485, 2022. doi: 10.1109/TVCG.2021.3114790 2
- [71] Y. Wang, M. Yu, G. Shan, H.-W. Shen, and Z. Lu. VISPubCompAS: a comparative analytical system for visualization publication data. *J. Vis.*, 22(5):941–953, 2019. doi: 10.1007/s12650-019-00585-2 2
- [72] L. Weng, X. Wang, J. Lu, Y. Feng, Y. Liu, H. Feng et al. InsightLens: Augmenting LLM-powered data analysis with interactive insight management and navigation. *IEEE Trans. Vis. Comput. Graph.*, 31(6):3719–3732, 2025. doi: 10.1109/TVCG.2025.3567131 2
- [73] S. Wu, L. Shi, L. Xia, C. Wu, Z. Liu, Y. Feng et al. Visualization of machine learning models through their spatial and temporal listeners. *arXiv preprint arXiv:2603.27527*, 2026. doi: 10.48550/arXiv.2603.27527 2
- [74] Y. Wu, N. Pitipornvivat, J. Zhao, S. Yang, G. Huang, and H. Qu. egoSlider: Visual analysis of egocentric network evolution. *IEEE Trans. Vis. Comput. Graph.*, 22(1):260–269, 2016. doi: 10.1109/TVCG.2015.2468151 2
- [75] F. Xiao and L. Shi. GeneticFlow: Exploring scholar impact with interactive visualization. In *Proc. IEEE VIS*, pp. 66–70. IEEE, Piscataway, 2023. doi: 10.1109/VIS54172.2023.00022 2
- [76] Z. Xu, S. Jain, and M. Kankanhalli. Hallucination is inevitable: An innate limitation of large language models. *arXiv preprint arXiv:2401.11817*, 2024. doi: 10.48550/arXiv.2401.11817 5
- [77] Q. Yu, C. Long, Y. Lv, H. Shao, P. He, and Z. Duan. Predicting co-author relationship in medical co-authorship networks. *PLOS ONE*, 9(7):e101214, 2014. doi: 10.1371/journal.pone.0101214 2
- [78] W. Zeng, A. Dong, X. Chen, and Z. Cheng. VIStory: interactive storyboard for exploring visual information in scientific publications. *J. Vis.*, 24(1):69–84, 2021. doi: 10.1007/s12650-020-00688-1 2
- [79] J. Zhang, H. Lv, and Z. Niu. AnchorTextVis: A visual analytics approach for fast comparison of text embeddings. *IEEE Comput. Graph. Appl.*, 45(6):29–43, 2025. doi: 10.1109/MCG.2025.3598262 2
- [80] X. Zhang, F. Zhao, Y. Liu, P. Chen, Y. Wang, X. Wang et al. TreeQA: Enhanced LLM-RAG with logic tree reasoning for reliable and interpretable multi-hop question answering. *Knowl.-Based Syst.*, 330, art. no. 114526, 2025. doi: 10.1016/j.knsys.2025.114526 5
- [81] Z. Zhang, B. G. Patra, A. Yaseen, J. Zhu, R. Sabharwal, K. Roberts et al. Scholarly recommendation systems: a literature survey. *Knowl. Inf. Syst.*, 65(11):4433–4478, 2023. doi: 10.1007/s10115-023-01901-x 1, 2
- [82] X. Zhou, L. Ding, Z. Li, and R. Wan. Collaborator recommendation in heterogeneous bibliographic networks using random walks. *Inf. Retr. J.*, 20(4):317–337, 2017. doi: 10.1007/s10791-017-9300-3 2