

LITFORAGER: Exploring Multimodal Literature Foraging Strategies in Immersive Sensemaking

Haoyang Yang , Elliott H. Faa , Weijian Liu , Shunan Guo , Duen Horng Chau , and Yalong Yang 

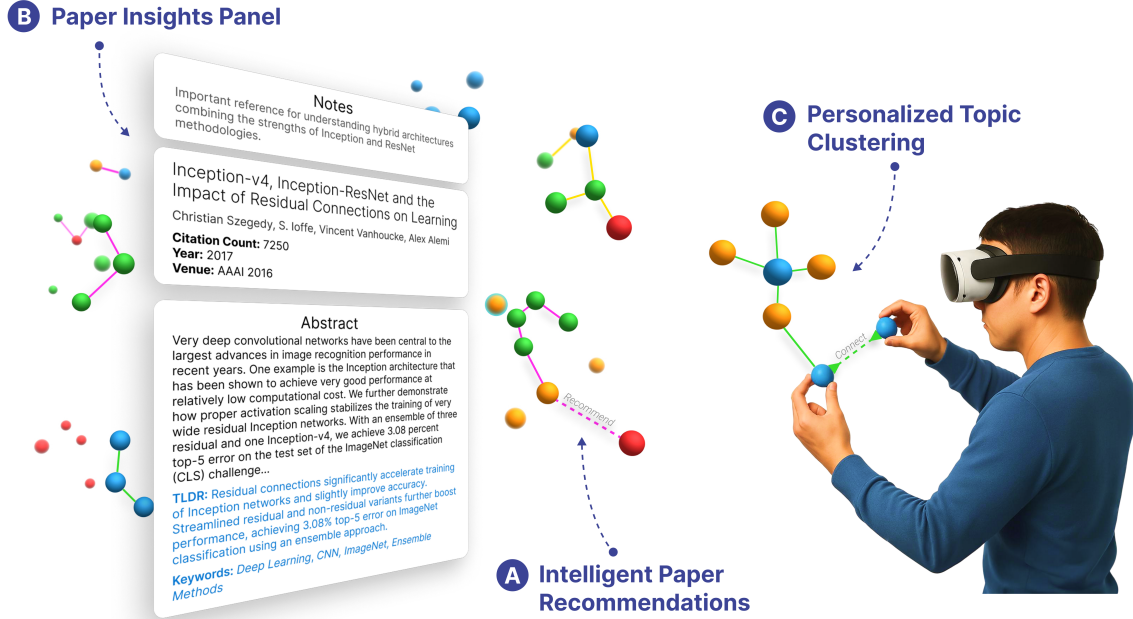


Fig. 1: LITFORAGER supports multimodal interactions to enable researchers to forage and visualize paper nodes within an immersive literature network graph in VR. **(A) Intelligent Paper Recommendations** help forage relevant literature via paper connections, such as citations, co-authorship, and thematic similarities. **(B) Paper Insights Panel** provides paper summaries and conceptual highlights while supporting user-authored annotations to capture reflections and emerging ideas. **(C) Personalized Topic Clustering** automatically groups papers by theme, with gesture controls allowing flexible, user-driven graph customization.

Abstract—Exploring and comprehending relevant academic literature is a vital yet challenging task for researchers, especially given the rapid expansion in research publications. This task fundamentally involves sensemaking—interpreting complex, scattered information sources to build understanding. While emerging immersive analytics tools have shown cognitive benefits like enhanced spatial memory and reduced mental load, they predominantly focus on *information synthesis* (e.g., organizing known documents). In contrast, the equally important *information foraging* phase—discovering and gathering relevant literature—remains underexplored within immersive environments, hindering a complete sensemaking workflow. To bridge this gap, we introduce LITFORAGER, an interactive literature exploration tool designed to facilitate information foraging of research literature within an immersive sensemaking workflow using network-based visualizations and multimodal interactions. Developed with WebXR and informed by a formative study with researchers, LITFORAGER supports exploration guidance, spatial organization, and seamless transition through a 3D literature network. An observational user study with 15 researchers demonstrated LITFORAGER’s effectiveness in supporting fluid foraging strategies and spatial sensemaking through its multimodal interface.

Index Terms—Immersive analytics, immersive sensemaking, information foraging, literature exploration, multimodal interaction

1 INTRODUCTION

Literature review is a crucial part of virtually every research endeavor, yet staying on top of the influx of new papers has become increasingly

challenging [31]. Each year, millions of new research articles are published [17, 24, 28], making it difficult for researchers to discover and digest the latest relevant information.

Literature review is fundamentally a sensemaking activity that involves both information foraging and synthesis. Pirolli and Card’s sensemaking model [47, 48] formalizes this process as two interlinked loops: an *information foraging* loop of searching, filtering, and collecting pertinent data sources, followed by an *information synthesis* loop of analyzing and synthesizing insights from those sources. Effective sensemaking is an iterative process between the two phases, but most prior research focused on only one phase rather than integrating both, leaving the full iterative workflow underexplored [16]. This gap is particularly evident in the emerging field of immersive analytics. Despite

- Haoyang Yang, Elliott H. Faa, Weijian Liu, Duen Horng Chau, and Yalong Yang are with Georgia Tech. E-mail: {alexanderyang, efaa3, wliu430, polo, yalong.yang}@gatech.edu
- Shunan Guo is with Adobe Research. E-mail: sguo@adobe.com

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org.
Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxxx

the recognized cognitive benefits of immersive environments—such as enhanced spatial memory and reduced cognitive load—current systems are mainly designed to support synthesis tasks, such as helping users organize pre-existing document collections or visualize data in 3D [13, 34, 35]. Support for the upstream task of information foraging within the immersive space remains limited, reducing the potential of immersive technologies to facilitate the complete, end-to-end sensemaking workflow.

To address this gap and explore the potential of immersive sensemaking with integrated foraging support, we introduce LITFORAGER, an immersive system explicitly designed to streamline information foraging during literature exploration. To shape the design of LITFORAGER, we conducted a formative study with expert interviews to identify researchers’ goals, needs, and challenges during literature review. Insights from this study highlighted the need for effective **exploration guidance**, flexible **spatial organization**, and **seamless transition** between foraging and synthesis, which became the core design goals for LITFORAGER.

To provide **exploration guidance**, LITFORAGER presents literature as an interactive 3D network graph, visualizing relationships like citations, authorship, and thematic similarities. Users can dynamically expand this network using integrated recommendations from Semantic Scholar [1, 38], and rapidly comprehend content with the aid of LLM-powered summaries and keyword extraction. LITFORAGER supports flexible **spatial organization**—users can freely position papers within the 3D space or apply automatic clustering by topic similarity. This allows them to leverage spatial memory and externalize their mental model. To enable **seamless transition** between foraging and synthesizing in the iterative workflow, the system integrates multimodal input, combining intuitive hand gestures for direct manipulation with voice commands for efficient control over exploration, annotation, and workspace management. Implemented with Babylon.js [12] and WebXR [21] and guided by user requirements identified in our formative study, LITFORAGER provides an embodied, web-based environment tailored for effective literature foraging. We evaluated LITFORAGER through an observational user study with 15 researchers who engaged in immersive literature discovery tasks to assess the system’s usability and interaction design.

In summary, our work has the following primary contributions:

- **A formative study with six academic researchers** that identifies critical user needs, task breakdown, and pain points faced during literature discovery and sensemaking. The findings informed the key requirements and design considerations necessary to effectively support immersive literature exploration.
- **LITFORAGER, a fully developed, open-source¹ immersive system explicitly designed to support literature foraging in the sensemaking workflow.** LITFORAGER integrates literature network visualization, spatial organization, and multimodal interactions, allowing users to explore and comprehend complex relationships within the literature and enhance cognitive engagement and memory.
- **An observational study with 15 academic researchers** that provides insights into how LITFORAGER can support literature foraging and sensemaking workflow. Our findings suggest that spatial organization features promote externalized thinking, while multimodal interaction enables fluid exploration.

2 RELATED WORK

Immersive Sensemaking. Early cognitive models of sensemaking, such as the framework by Pirolli and Card, describe iterative *foraging* and *synthesis* loops through which analysts gather information and develop hypotheses [47, 48, 54]. *Foraging* involves searching and filtering information into schemas, while *synthesis* entails refining a mental model based on these schemas. Building on this foundation, later research has highlighted the importance of externalizing thought processes to amplify cognition. Notably, the Space to Think system [2]

demonstrated that using a large display allowed researchers to spatially arrange information and offload memory, significantly improving sensemaking performance.

Leveraging space as an extension of cognition is central to the emergence of immersive analytics. Immersive systems provide an expansive 3D canvas and engage the user’s body through movement and gesture, enabling spatial cognition and embodied interaction for complex analysis [62, 63]. Drawing inspiration from Space to Think, Lisle et al. introduced Immersive Space to Think (IST), a VR-based sensemaking environment for document analysis [35], and later compared the cognitive effects of traditional 2D and immersive displays [37], showing how immersive environments leverage spatial memory and externalization to reduce context-switching costs and enhance sensemaking. Tahmid et al. [60, 61] extended IST by integrating gaze-driven multimodal interactions and recommendation cues to support dynamic information discovery and relevance prediction. However, both IST and its extensions rely on pre-selected document corpora and offer limited support for integrating newly discovered materials into the workflow, restricting their adaptability for real-world, iterative foraging tasks across evolving datasets.

Our tool addresses these limitations by envisioning a complete sensemaking workflow with built-in discovery and recommendation capabilities. In contrast to prior systems that required manually pre-loaded document sets, LITFORAGER enables users to fluidly discover new papers and insert them into their analysis environment on the fly, bringing the traditionally separate discovery phase into the same immersive workflow.

Literature Discovery. Literature discovery is the process of exploring academic literature to uncover insights, patterns, or connections that are not explicitly stated [58, 59]. It differs from the standard literature search in that it is a more exploratory process rather than a straightforward goal-oriented query. In a standard search, a researcher typically knows what to look for: specific keywords, authors, or papers. Literature discovery, on the other hand, can be more beneficial when the researcher does not have a precise target in mind; instead, the goal is to explore a research area and uncover unexpected yet relevant connections.

Many interactive systems support literature discovery by helping researchers navigate relationships between works. Kang et al. [25] focus on author networks, allowing users to form a “committee” of key authors and expand recommendations via co-authorship and citation links. PaperQuest [49] visualizes citation networks and applies relevance algorithms to guide exploration. Palani et al. [45] extract and compile related work sections to highlight key prior studies and recurring themes. PUREsuggest [3] combines citation-based recommendations with interactive visual exploration, allowing users to guide the process using seed papers and keywords. Building on these approaches, our system supports literature discovery by enabling researchers to explore, connect, and uncover insights through immersive foraging.

Multimodal and Embodied Interaction in VR. Prior research in multimodal interaction shows that combining input modalities enhances usability across workflows. For example, Saktheeswaran et al. [55] explored how touch and speech support network analysis, while Data@Hand [26] enabled personal data exploration on smartphones using the same modalities. Reinders et al. [51] paired conversational agents with tactile displays to improve accessibility for blind users. Luo et al. [39] demonstrated how bimanual gestures and spatial manipulation in AR support intuitive document arrangement. Collectively, these studies highlight the versatility of multimodal interfaces in improving user experience.

In immersive environments, speech input enables natural command execution and querying without typing. MEinVR [68] combined voice and controller input to navigate 3D molecular structures more intuitively. Multimodal interactions are also central in commercial systems: the Microsoft HoloLens 2 blends eye gaze, hand gestures, and voice for seamless interaction [41], reducing reliance on abstract UI elements like WIMP menus.

Immersive analytics also benefits from embodied interaction and spatial layout, which offload cognitive effort through physical action

¹LITFORAGER’s source code is publicly available at <https://github.com/AlexanderHYang/LitForager>.

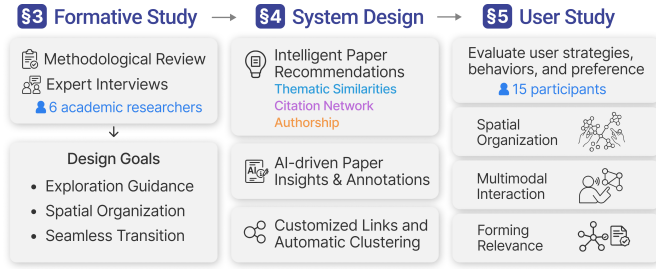


Fig. 2: Overview of our research workflow. We conducted a formative study (§ 3) to identify design goals for immersive literature foraging. Based on these goals, we developed a system (§ 4) with interactive foraging features. We then evaluated the system (§ 5) through a user study with 15 participants, focusing on spatial organization, multimodal interaction, and relevance formation.

and spatial memory [11, 19, 22, 66, 69]. By allowing analysts to move and organize content in 3D space, immersive systems act as external memory aids [35, 57], minimizing the need for internal recall or auxiliary note-taking [20, 27, 43, 67]. Lisle et al. [36] and Luo et al. [40] demonstrated that both VR and AR environments enable such memory externalization by allowing documents to be anchored relative to virtual objects or physical structures. Similarly, Andrews et al. [2] found that users working with large displays leveraged spatial positioning to track document relationships and maintain context over time. Such embodied interaction reduces memory load and mirrors real-world behaviors—like grouping items or glancing around to recall a location—making complex tools more intuitive and learnable.

3 FORMATIVE STUDY

In this section, we present a formative study based on semi-structured expert interviews that investigates how academic researchers conduct literature reviews. We examine their primary tasks, objectives, and pain points in the process, which informed the design goals of LITFORAGER.

3.1 Procedures

Semi-Structured Expert Interviews. We reviewed influential educational resources and guidelines on literature reviews, including those from university libraries and educational institutions [7, 14, 46, 50] to ground our understanding of discovery behaviors, strategies for evaluating relevance, and synthesis practices. Booth et al.’s systematic framework [5] further guided the segmentation of the literature review process into distinct stages (searching, organizing, and synthesizing), which shaped the structure of our study procedure. Together, these sources influenced the design of our open-ended interview questions, including how we framed prompts around identifying starting points, expanding search strategies, assessing importance, and integrating insights.

We conducted semi-structured interviews with 6 academic researchers (R1–6; 2 female, 4 male; ages 20–25), including five PhD students and one master’s student, all with extensive literature review experience. Prior to the interviews, participants completed a brief survey detailing their research experience, literature review frequency, and objectives, providing additional context for discussion. The interview questions were organized according to distinct stages of the literature review process, with greater emphasis placed on the phases of literature searching and discovery, reflecting our focus on information-foraging behaviors. Specifically, we asked participants about their strategies for identifying starting points, expanding their search, organizing literature, evaluating relevance, and synthesizing insights when conducting a literature review. We iteratively refined the interview protocol through two pilot sessions. Each interview lasted approximately 45–60 minutes and was conducted in person.

Data Collection and Analysis. All interviews were recorded (with consent) and transcribed for analysis. We collected additional notes on any specific tools or techniques participants mentioned (e.g., using

reference managers, annotation practices, or ad-hoc physical organization of papers). Using Braun and Clarke’s reflective approach [6, 8], we performed a thematic analysis on the transcripts to extract recurrent themes regarding user tasks, motivations, pain points, and strategies. Three authors independently coded one of the transcripts to develop an initial schema. The thematic coding was refined through discussion and then applied to re-code all six transcripts. Consistency and saturation were reached by the fifth and sixth interviews, as no substantially new themes emerged.

3.2 Findings & Insights

We present formative study findings that reveal a multi-dimensional view of researchers’ practices and challenges in the literature review process.

3.2.1 Context, Data Needs, and Practices

From the expert interviews, we identified both distinctive strategies and common themes across three interrelated dimensions (Fig. 3): **Context** (When and Why researchers conduct literature reviews), **Data** (What information researchers seek), and **Methodology** (How they source, evaluate, and organize the information) :

Context. Researchers reported engaging in literature reviews at various points throughout their projects. During the early stages of a research project (initiation and ideation), researchers conduct literature reviews primarily to *identify research gap*, which helps define the project’s scope and refine research questions, “... to do something really novel, I have to find papers that I can build on” (R2). Later, in the mid-project phase (development and problem-solving), researchers revisit the literature review primarily for research *framework alignment*, positioning their findings within broader scholarly conversations and updating evolving theoretical frameworks. This alignment ensures their work remains relevant to current trends and clearly demonstrates the novelty of their contributions. In the paper writing phase, the literature review process shifts towards *idea synthesis*. As researchers gather more insights from additional readings, they organically develop subsections within their related work, structuring emerging sub-concepts to frame their findings within a broader scholarly conversation, “As I read more, I finalize these subconcepts into the subsections of my related work” (R4).

Data. The interview data reveal that researchers look for a variety of information types during the literature review process. Many researchers indicated that they build a comprehensive understanding of their field by starting with *seminal papers* or *survey papers*, which encapsulate foundational concepts and frameworks, “If I’m completely new to a field, I start with literature surveys or seminal papers” (R3). There is also a significant focus on emerging trends and identifying gaps: many researchers sift through *highly-cited publications* such as conference proceedings or journal articles, “I lean more on quantitative metrics like citations, but also on the reputation of the venue” (R1), to capture current developments and highlight areas where further inquiry is needed. Additionally, *interdisciplinary insights* also play a vital role. Researchers noted that sometimes, by exploring literature from other fields, they uncover novel perspectives, “I notice that certain techniques used elsewhere might be applicable—even if the fields seem disconnected” (R6) and “I came across a recommended study about policy and governance aspects of the field that wasn’t originally on my radar” (R3). This diverse set of data ensures that their work is well-grounded in established knowledge while simultaneously aligning it with state-of-the-art developments in emerging research.

Methodology. In tandem with the information they seek, researchers have mentioned various methodologies for sourcing, evaluating, and organizing the literature. They usually start out with a few *seed papers*, often recommendations from peers and guidance from experts or advisors, “Often, I begin my literature research with a few ‘seed’ papers that my peers or my advisor have given me or that I happened to catch during a conference talk” (R3). When sourcing relevant studies, they leverage targeted searches using specific *keywords*, often complemented by various types of *literature network* to trace connections

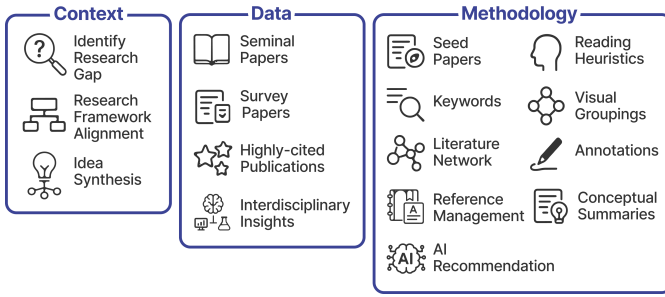


Fig. 3: Three interrelated dimensions of literature review practices identified from expert interviews.

across the research landscape, such as citation network, reference network, or author network. Additionally, **AI recommendation** systems further enhance the discovery process by efficiently identifying relevant studies, particularly interdisciplinary works that span diverse academic fields, “*I also use tools like Semantic Scholar’s AI-powered recommendation system to suggest new papers based on the one I’m reading*” (R2). Once potential sources are identified, researchers often use **reading heuristics** by quickly scanning abstracts, introductions, and conclusions to determine relevance, “*I typically start by skimming the abstract, introduction, and conclusion*” (R3), followed by a more detailed evaluation of the research methodology. To synthesize the gathered insights, some researchers mentioned creating **visual groupings**, “*I think a mental map is good, especially a tree or network structure ...*” (R5), **annotations**, and **conceptual summaries** that help identify common themes and discrepancies across studies. Finally, effective organization is achieved through the use of **reference management** systems, such as Zotero or Mendeley.

3.2.2 Pain Points in the Literature Review Process

Information Overload. A recurring concern is the sheer scale of available literature. Many mentioned the risk of missing critical studies amid the vast number of publications. This overwhelming volume not only complicates the discovery process but also increases the likelihood of redundancy or overlooked work, “*It’s hard to be completely comprehensive ... Not finding literature in a particular niche doesn’t necessarily mean there isn’t any*” (R1). Consequently, researchers highlighted the need for effective exploration guidance that can navigate large datasets and surface key papers, ensuring that important contributions are not lost in the noise. Evaluating the credibility and relevance of each source is another labor-intensive aspect of the literature review process. Researchers reported the challenge of having to read through a large number of papers, emphasizing the necessity for reading guidance to streamline the evaluation process.

Interdisciplinary Blind Spots. Many researchers mentioned the concern that, despite their best efforts, critical literature might be inadvertently overlooked. This challenge is particularly acute in interdisciplinary research areas, where seminal studies and innovative methodologies may be scattered across diverse fields and publication venues. Researchers often rely on traditional search strategies that focus on their primary discipline, which can lead to a narrow perspective that misses valuable insights from adjacent or seemingly unrelated domains, “*I wish recognition systems could combine multiple references or keywords simultaneously to yield more interdisciplinary results*” (R3).

Difficulty in Synthesis and Organization. Despite the availability of digital tools, synthesizing information from a wide range of sources remains a significant challenge. Researchers noted that while reference managers help with storage, they fall short in supporting synthesis. For instance, tools like Zotero typically organize literature as a linear list, which limits their ability to support visualizations or map thematic grouping. While importing literature into dedicated visual mapping tools could be a solution, it adds workflow friction. Participants expressed a need for tools that allow them to visually map connections and structure their ideas, while also integrating the foraging aspect

of exploring a large corpus of literature, “*I envision a system where I could offload some cognitive effort by interacting with a dynamic graph that expands into new nodes as I explore ideas further*” (R6). Researchers require a seamless, iterative workflow that enables them to transition effortlessly from discovering and storing relevant studies to synthesizing and constructing a coherent narrative.

3.3 Design Goals

Drawing from the insights of our formative study and grounded in established HCI and sensemaking principles, we distilled three core design goals (G1-G3) to guide the development of LITFORAGER. These goals directly address the identified researcher pain points and aim to integrate effective foraging support within an immersive sensemaking workflow.

G1. Intelligent exploration guidance via recommendations and insights. The challenge of navigating information overload and evaluating relevance highlights the need for intelligent exploration guidance. Aligned with Information Foraging Theory [47], our design focuses on surfacing relevant and high-quality literature through targeted recommendations and AI-driven insights, helping researchers efficiently navigate large datasets and emerging trends while reducing reliance on exhaustive manual search.

G2. Flexible spatial organization of the literature network. Organizing and synthesizing diverse sources using traditional foraging tools can be cognitively demanding. Leveraging the affordances of immersive environments [2, 22, 23, 35], LITFORAGER introduces a 3D literature network that enables researchers to externalize their understanding by spatially arranging papers. Supporting both automatic clustering and customizable layouts, this design aims to help offload cognitive effort and facilitates intuitive exploration of thematic and methodological relationships.

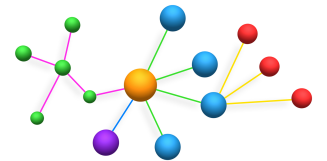
G3. Seamless transition between foraging and synthesizing. Synthesizing information from a vast corpus of literature imposes significant cognitive load, especially when traditional tools separate the processes of finding and organizing information. To address this disconnect, we aim to create a seamless transition between foraging and synthesizing by integrating robust exploration with fluid multimodal interaction. By reducing context switching and interface complexity [48], our goal is to help researchers maintain context and focus throughout the iterative sensemaking process.

4 LITFORAGER

Building on the insights from our formative study (§ 3) and guided by our core design goals (§ 3.3), we developed LITFORAGER, an interactive prototype designed to make literature foraging easier within an immersive sensemaking workflow. This section details the core components and features of LITFORAGER, organized by the literature network visualization, interaction model, and key features enabling intelligent immersive foraging and sensemaking.

4.1 Literature Network Visualization

The core representation within LITFORAGER is an interactive 3D force-directed graph displaying academic papers as nodes and their relationships using distinct edge colors to convey meaning: white for *thematic similarities*, magenta for *citations & references*, yellow for *author-centric*, and green for *user-created links*.



Web-based tools like Connected Papers [10], ResearchRabbit [53], and Argo Scholar [33] use network visualization to reveal structures, clusters, and pathways within research areas. Building on this proven approach and research suggesting immersive environments enhance complex network exploration via improved spatial understanding and interaction [29, 30], LITFORAGER employs immersive 3D network visualization. To populate the visualization, LITFORAGER retrieves metadata for requested papers (e.g., authors, abstracts, publication year,

and citation relationships) from Semantic Scholar’s scholarly corpus via their public API [1, 38]. The visualization, implemented using Babylon.js [12] and WebXR [21], is accessible via web browsers on standard VR headsets.

Initially, the network’s spatial arrangement is determined by a force-directed layout, which automatically clusters related papers in close proximity. To simulate the layout of nodes in 3D space, LITFORAGER employs d3-force-3d [64], a three-dimensional extension of D3’s physical force-directed layout engine [42]. This provides an immediate visual structure that helps users grasp the overall landscape (G2). Rendering this network in VR aims to capitalize on distinct advantages over conventional 2D screens. The expansive 3D space allows users to leverage their natural spatial reasoning and memory [35, 57, 67]. Users can physically navigate around and within the graph, examining connections from multiple perspectives. This embodied navigation transforms the virtual space into an external cognitive aid [2, 20, 27, 52, 65], potentially making complex relationships easier to comprehend compared to flat representations. This spatial layout forms the foundation for all subsequent foraging and sensemaking activities within LITFORAGER.

4.2 Multimodal Interaction Model

To ensure users can control the system easily and support transitioning between finding and analyzing literature, LITFORAGER uses a combination of interaction methods: direct hand gestures, voice commands, and a WIMP-style hand-menu interface (G3). Offering multiple input modes is crucial in complex VR environments because it provides flexibility; users can choose the most efficient or comfortable method for a given task, reducing interaction friction and mental load [26, 41, 55].

Voice Commands. For actions like requesting summaries (“*Summarize paper*”) or triggering recommendations (“*Recommend papers by thematic similarities*”), voice input offers a fast, hands-free alternative. This modality is particularly beneficial when users’ hands are occupied with manipulation tasks or when navigating complex menu structures would be cumbersome [9]. Speech allows for efficient execution of commands and queries using natural language, reducing the need for users to divert visual attention or exert physical effort, thus supporting fluid interaction (G3).

Embodied Gesture Interaction. Users interact directly with paper nodes using hand tracking. A simple pinch gesture selects a paper, which can then be moved freely in 3D space (6DoF) and pinned. This leverages the principles of direct manipulation [56], making spatial organization tasks feel intuitive and mapping directly to users’ actions. Such embodied interactions are particularly powerful in VR, allowing users to leverage their innate spatial reasoning and proprioception, effectively turning the virtual space into an extension of their cognition [15, 27, 32]. Gestures like bringing two papers together to link them (bimanual interaction) (Fig. 4A) or the steeping gesture for clustering (Fig. 4B) further enhance this embodied control, aligning with findings that physical actions can offload cognitive burden in immersive analytics [11, 23, 66].

Hand Menu. A standard menu attached to the user’s non-dominant hand provides reliable, visible access to all features using familiar point-and-click interaction with holographic buttons. This WIMP-style interface acts as a dependable fallback, enhances the discoverability of system functions [43], and ensures usability for users who may prefer explicit control, are new to voice/gesture commands, or encounter limitations with other modalities. Such menus remain a cornerstone of VR interaction design for providing structured access to functionality [32].

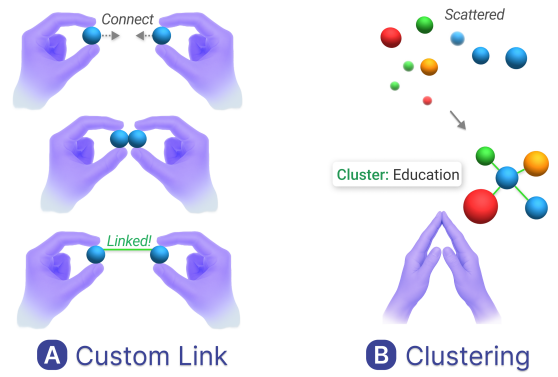


Fig. 4: Two bimanual gestures for organizing papers in LITFORAGER. (A) Bringing two papers together with a bimanual pinch gesture creates a **Custom Link**. (B) A bimanual steeping gesture triggers automatic **Clustering** of nearby papers.

4.3 Intelligent Foraging

LITFORAGER includes several features designed specifically to help with finding, understanding, and organizing literature.

4.3.1 Intelligent Network Expansion via Recommendations

LITFORAGER offers multiple ways to expand the literature network, helping users discover relevant papers as they forage (G1). This is primarily achieved through an intelligent paper recommendation system. Users can request recommendations via the hand-menu or corresponding voice commands (e.g., “*Recommend papers by citations*”). The system offers three key recommendation modes—*thematic similarities*, *citations & references*, and *author-centric*, reflecting common literature exploration strategies and corresponding to the literature network’s visualization types (§ 4.1). These modes respectively capture content-based, structural, historical, and community-driven discovery, providing a comprehensive yet manageable toolkit for navigating the literature landscape.

Thematic Similarities. Users can request recommendations based on thematic similarities to one or more selected papers, enabling the addition of topically relevant yet potentially novel works to the current network view. A powerful application of this is creating thematic “bridges”: by selecting papers from seemingly unrelated clusters and requesting recommendations, users can uncover non-obvious connections between research areas, surface novel trends, and receive guided suggestions for exploration.

Citations & References. Users can expand the network by exploring citation links. When a single paper is selected, it can retrieve its “forward citations” (papers citing the selected work) or “backward references” (papers cited by the selected work). This common strategy allows researchers to trace the lineage and impact of ideas over time, providing insights into a field’s development and trajectory.

Author-Centric. Users can explore an author’s papers by selecting their name on a paper node to request other publications by that author. This helps researchers identify key experts within specific areas and discover relevant work through established authors.

4.3.2 AI-Enabled Thematic Clustering and Custom Links

To support flexible organization within literature networks, LITFORAGER provides both automated and manual structuring tools for users to personalize their literature network (G2). These features help reduce cognitive load while enabling researchers to map connections, identify patterns, and externalize their evolving understanding of the research landscape.

LLM-enabled Topic Clustering. LITFORAGER uses Google’s Gemini 2.5 Flash [18] to automatically group papers into thematic clusters based on title and abstract analysis. These clusters are spatially arranged and labeled with their core themes, effectively categorizing the network

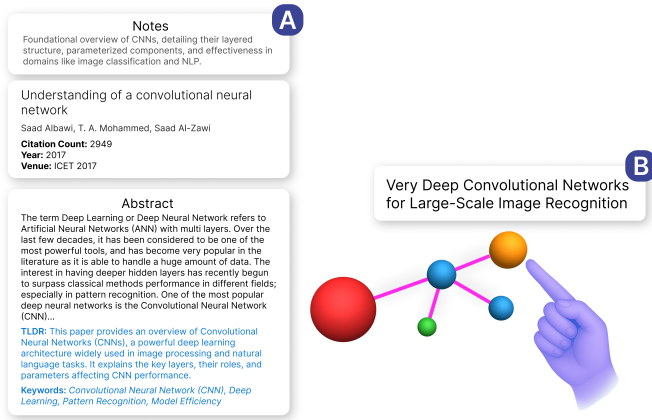


Fig. 5: User's point of view of the paper assessment interface. (A) **Paper Insights Panel** with detailed metadata (e.g., authors, venue), AI insights including TLDR and keywords, and user-authored annotations. (B) Hovering over a paper node reveals its title above the node for quick identification without opening the full panel.

into key domains. This visualization acts as a powerful guide for exploration and enables efficient spatial organization and thematic externalization (G1, G2). The clustering feature is readily accessible via a hand-menu button, the voice command (“Cluster papers”), or the bimanual steeping gesture (Fig. 4B).

Customized Node Links. Beyond automatically generated connections, users can manually create custom links between nodes. By physically bringing two nodes close together (one held in each hand), users establish a direct link. This allows researchers to define unique relationships that may not be captured by citation data or recommendation algorithms, enabling them to flexibly structure the network according to their personal conceptual models and spatially externalize their understanding (G2).

4.3.3 Intelligent Paper Insights and Annotation

To help users navigate information overload [31] and quickly assess relevance, LITFORAGER integrates AI-powered analysis based on each paper’s metadata via Google’s Gemini 2.5 Flash API [18]. To support synthesis, it also offers lightweight annotation tools.

Paper Insights Panel. Hovering over a node reveals its title for quick scanning (Fig. 5B), while long pressing opens a detailed panel with the title, authors, and abstract of the paper (Fig. 5A). The panel remains pinned near the node and oriented toward the user to preserve readability and spatial context. These interactions provide quick access to essential information, reducing cognitive load and supporting fluid navigation (G1, G3).

AI-Driven Content Analysis. From the panel, users can request a quick summary (“TLDR”) or keywords via menu or voice commands (Fig. 5A). Powered by LLM, these features help distill key information, reducing the effort and time needed to assess a paper’s relevance (G3).

Speech-Based Annotations. Users can dictate notes onto a paper via a hand-menu button or the voice command (“Start/stop annotating”). The system transcribes the speech and attaches the notes to the panel (Fig. 5A). This allows researchers to capture immediate thoughts without breaking their exploration flow, smoothly bridging foraging and synthesis (G3), which aligns with sensemaking models emphasizing iterative reflection.

5 USER STUDY

We conducted an observational study to understand how researchers utilize immersive tools for literature exploration and synthesis, focusing on their strategies, behaviors, and preferences. Rather than measuring task performance, our qualitative emphasis reflects LITFORAGER’s

aim to support open-ended sensemaking and research ideation. The study scenario was informed by insights from our formative expert interviews and grounded in common document foraging tasks described in literature review best practices. This alignment ensured that the scenario realistically exercised the system’s core features and provided comprehensive coverage of relevant user workflows. Accordingly, we formulated the following research questions:

RQ1: *How do users choose between different input modalities (gesture, speech, or hand-based menu) when managing the exploration and organization process?*

We explore whether modality preferences are task-specific, user-dependent, or influenced by system affordances, and how input choice affects user flow, agency, and control in immersive interaction.

RQ2: *How do users externalize their understanding by organizing, annotating, or clustering papers within the 3D immersive space?*

We seek to understand the diversity of spatial structuring and annotation strategies users employ, and how these externalization practices contribute to managing complexity and potentially offloading cognitive load during immersive literature foraging.

RQ3: *How do users develop and refine their sense of relevance during immersive literature exploration to guide subsequent foraging actions?*

We aim to understand the process by which users determine paper relevance, identify the factors influencing these judgments (e.g., metadata, network cues, task goals), and analyze how their relevance criteria evolve throughout the exploration session.

Participants and System Setup. We recruited 15 participants (P1–P15; 8 female, 7 male; ages 20–27), including 9 PhD students, 5 master’s students, and 1 undergraduate student, with backgrounds in information visualization, HCI, XR, machine learning, and education. Most had 1–4 years of research experience and regularly engaged in literature reviews, with 10 identifying as experts, and five as occasional. All but one had used reference managers, and several had tried academic visualization or AI-based tools. Most had little or no prior experience with immersive technologies.

The experiment was conducted within a 10×10 ft lab area, ensuring sufficient space for participants to move freely. A Meta Quest 3 headset was used. Participants interacted with the LITFORAGER using the headset’s built-in hand tracking and voice commands facilitated by an external microphone. The experimenters observed the session by monitoring the participant’s VR view cast to a laptop and their physical movements.

Task Scenario and Goals. The study used a task scenario simulating the initial phase of a literature review. Participants were asked to construct a *Related Work* graph for a research topic relevant to their background. This task was designed to mirror the early stages of scholarly inquiry, where researchers identify, evaluate, and organize literature to understand a research landscape. Each participant received a topic aligned with their academic background. To initiate exploration, three seed papers representing distinct sub-topics were preloaded into LITFORAGER. These papers served as the starting point. Participants had full autonomy in deciding which papers to investigate further or expand upon. Participants were instructed to explore the literature using the system’s recommendation features, assess the relevance of discovered papers, and arrange them spatially. We intentionally avoided prescribing a specific structure or output, encouraging participants to arrange papers in a personally meaningful way that externalized their reasoning and sensemaking approach.

Experimental Procedure. Each study session lasted approximately 90 minutes and consisted of four parts: pre-study briefing, system tutorial, main exploration task, and post-task debriefing. The session began with a 5-minute pre-study phase for consent, a brief welcome, and a short demographic questionnaire.

Next, participants proceeded to a 30-minute interactive tutorial. This extended duration was deemed necessary to ensure that participants achieved proficiency with the system’s multiple interaction modalities (e.g., hand tracking, voice commands) and core literature exploration

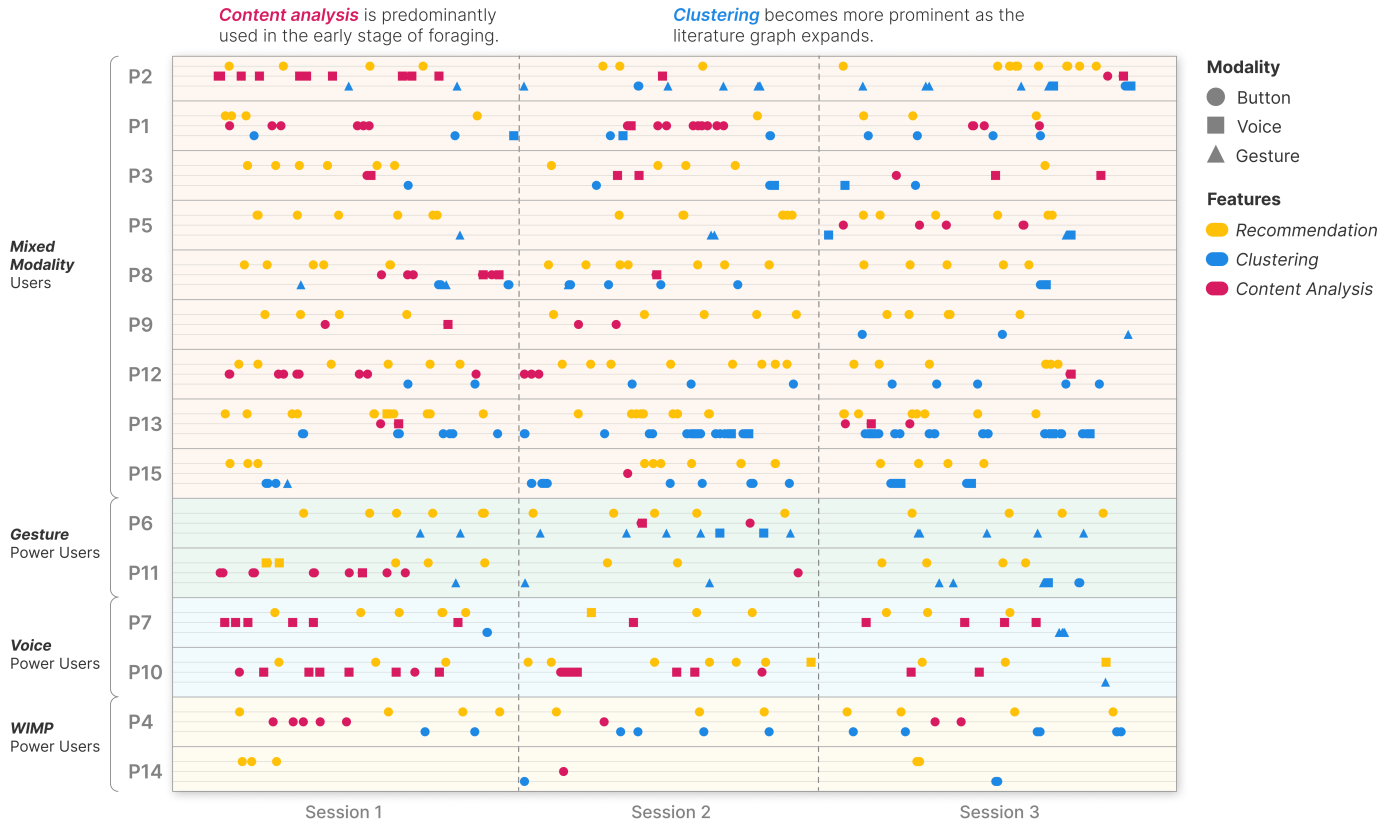


Fig. 6: Interaction logs of all participants across three sessions, showing input modalities (button, voice, gesture) and feature usage (recommendation, clustering, content analysis). Users exhibited diverse modality preferences, including mixed-modality users and modality-specific (gesture, voice, WIMP) power users.

features. Wearing the VR headset, they familiarized themselves with these functions under the real-time guidance of the researcher observing via the cast view. They were encouraged to ask questions throughout the walkthrough and were given time to practice interactions at their own pace afterwards. The main exploration task lasted 40 minutes, divided into three 10-minute sessions with short breaks in between. This structure helped mitigate potential VR fatigue while enabling us to observe how participants’ strategies evolved over time. At the end of each session, we conducted a brief 1-minute interview in which participants summarized their goals, progress, and plans for the next phase. These periodic reflections served as a lightweight alternative to traditional think-aloud protocols, capturing participants’ reasoning without disrupting task focus or interfering with voice-based interactions. Each summary was followed by a 4-minute rest break, during which participants could relax or ask questions.

Following the main task, a 15-minute semi-structured interview session was conducted to gather rich qualitative feedback on their experience and probe specific observed behaviors.

Data Collection and Analysis. We collected behavioral data through interaction logs, brief in-task reflections, and a post-task semi-structured interview.

Interaction Logs. The system automatically recorded timestamped user interactions, capturing exploration sequences (e.g., paper selections, recommendation requests), spatial manipulations (e.g., moving, clustering papers), and input modality usage (gesture, speech, WIMP menu). These detailed behavioral traces enable the reconstruction and analysis of users’ navigation paths and interaction patterns throughout the session.

Self-Reflections and Interviews. During short breaks between sessions, participants provided brief verbal reflections, articulating their goals, progress, and next steps. These reflections offered insight into participants’ evolving strategies in real time. After the main task, we conducted a semi-structured interview to further explore their experi-

ences, including their relevance judgment criteria, spatial structuring rationale, and input modality preferences.

6 RESULTS & DISCUSSION

Our analysis aims to identify distinct information foraging patterns in the immersive environment. We incorporated both quantitative and qualitative data in our analysis: interaction logs and Likert-scale responses were examined to identify behavioral trends, while interviews and observational notes were analyzed using thematic analysis. In this section, we discuss how immersive affordances affect iterative foraging and synthesis behaviors.

6.1 Input Modality Preferences and Tradeoffs (RQ1)

To explore RQ1, we analyzed all 15 participants’ interaction logs within our immersive literature exploration system during the user study. This analysis aims not only to identify modality usage patterns, but also to understand the underlying strategies, tradeoffs, and implications for usability, expressiveness, and cognitive effort in immersive environments. We structured our analysis along two key dimensions: (1) Input Modality—**gesture**, **speech**, or **hand menu** (WIMP-style input); and (2) System Features—*Recommendation*, *Clustering* (system-generated topic clusters and customized node links), and *Content Analysis* (AI-based summarization, keyword extraction, and user-created annotations). Visualizing modality and feature usage across sessions revealed several key patterns (Fig. 6).

WIMP Interface as the Ubiquitous Default. Across all foraging categories, **hand menu** emerged as the most frequently used modality. This was especially pronounced for *Recommendation*, where every participant relied almost exclusively on button input over speech. This strong reliance on WIMP-style menus reflects users’ comfort with explicit, visible UI controls—particularly for discrete, function-oriented tasks like querying for related papers.

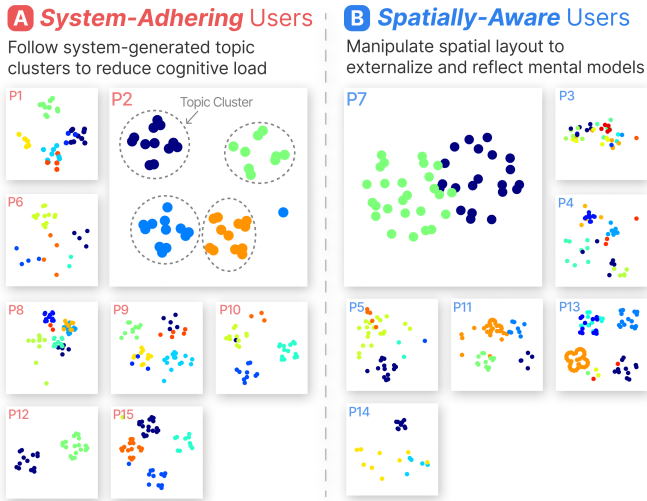


Fig. 7: Final spatial layouts of participants' literature networks (bird's-eye view), illustrating two distinct spatial organization behaviors. Colors indicate system-generated clusters. (A) **System-Adhering Users** (left, e.g., P2) largely preserved system-defined clusters, while (B) **Spatially-Aware Users** (right, e.g., P7) actively reorganized papers to reflect their own mental models.

Mixed Modality Use as a Marker of Adaptivity and Expertise. Most participants (9/15) exhibited active mixed modality usage—combining at least two input modalities over the course of the sessions. Typically, users began with **hand menu** interactions, gradually incorporating gestures or voice as they gained comfort and familiarity. Among these mixed modality users, P2 stood out as particularly adaptive and fluent. P2 not only explored all three modalities but also demonstrated a high degree of contextual sensitivity in modality choice—using **gesture** for *Clustering*, **voice** for *Content Analysis*, and **hand menu** for other precision or unfamiliar actions. We also observed several *power users* who demonstrated proficiency in specific modality-feature pairings. For clustering tasks, P6 and P11 consistently leveraged the bimanual steeping gesture to invoke clustering. For content analysis, P7 and P10 became strong voice users, frequently invoking content analysis through speech. This flexible, situational use of input demonstrates a desirable interaction pattern for immersive systems, where users fluidly combine multiple modalities to best suit their preferences, context, and task.

Foraging Strategies Blended Recommendation-Driven and Spatial Exploration. Our analysis revealed an important evolution in participants' foraging strategies. Initially, users relied predominantly on system guidance via *Recommendation* and *Content Analysis*. As their spatial understanding of the literature collection grew, their activity shifted towards self-directed exploration using *Clustering* and manipulating the layout. This demonstrates a natural blending of strategies, leveraging both system guidance and the emerging spatial context. This underscores the need for immersive systems to seamlessly integrate recommendation features within the spatial exploration environment.

Summary. Our findings show participants initially relied on hand menus for precise tasks like selecting recommendations, reflecting a preference for familiar, structured interactions that reduce errors [43]. As users became more comfortable, many started mixing input methods, using gestures and voice for actions where those modalities felt more natural or expressive. This reflects how users often combine different inputs to leverage their unique strengths [4, 44]. Importantly, even with novel VR inputs available, visible menus remained essential support, especially for complex tasks or new users, highlighting the need for clear guidance in 3D environments [32]. Therefore, future immersive systems should offer a balanced mix of input options, tailored to the specific task.

6.2 Spatial Externalization Strategies (RQ2)

To address RQ2, we analyzed participants' spatial organization behaviors throughout their literature exploration sessions. All participants demonstrated meaningful and sometimes complex use of spatial manipulation to externalize their evolving understanding of the literature network in immersive space. By examining both the participants' final configurations and temporal progression of layouts, we identified key strategies used to support sensemaking.

Clustering and Spatial Grouping Behavior. A consistent pattern observed across almost all participants was the use of clustering to organize papers spatially. Most participants' final layouts exhibited at least one or more clearly defined, visually distinct clusters (Fig. 7), often aligned with the system-generated thematic clusters (color-coded). This suggests that automated clustering served as a helpful starting point for organization and, in many cases, as a foundation for participants' externalization strategies.

We observed a spectrum of spatial behaviors that reflected how closely participants' mental models aligned with the system-generated structure. We distinguish between **System-Adhering** and **Spatially-Aware** users. **System-Adhering** users (e.g., P2) relied heavily on the automatic clustering with minimal manual reorganization (Fig. 7A); 8 out of 15 participants followed this approach. This suggests that the LLM-based clustering often captured meaningful thematic boundaries sufficient for users' foraging needs.

In contrast, **Spatially-Aware** users (e.g., P5, P7, P11, P13) employed more nuanced and active spatial strategies (Fig. 7B). These participants merged, split, or repositioned clusters based on personal interpretations of topic relevance and research goals. For instance, P7 described their process as “creating a spider web [of papers],” deliberately emphasizing central topics and constructing “meta-clusters” to indicate relationships between system-defined themes. Some also used spatial properties to encode importance or intent (e.g., placing key papers in direct view and moving less relevant ones behind them). These behaviors underscore how embodied spatial interaction in VR enabled personalized and expressive externalization beyond static system outputs.

Evolution of Spatial Organization Over Time. Beyond analyzing final layouts, we also analyzed the temporal evolution of participants' spatial layouts to understand their sensemaking process (Fig. 8 shows P2's progression). Initially, layouts tended to be sparse, reflecting early exploration based on seed papers and system recommendations. However, as participants gathered more literature, they began actively structuring their space—repositioning items, creating thematic clusters, and refining the overall organization. Final layouts were typically more structured and clustered, visually representing the user's developed understanding. This progression demonstrates that spatial externalization was not fixed but emerged dynamically through the user's iterative cycle of foraging, organizing, and understanding. Clustering stood out as an intuitive method for managing the growing complexity, helping users reduce clutter while maintaining important thematic connections.

Summary. Overall, these findings demonstrate that spatial externalization was a core strategy in participants' immersive literature foraging. Participants actively engaged in clustering, iterative reorganization, and personalized layout construction to manage complexity and externalize their understanding within the 3D environment. This use of space to aid thinking is a known cognitive strategy [20, 27] and matches behaviors seen in similar VR [35] and large-screen [2] sensemaking tools. The combination of automated clustering and manual spatial manipulation enabled users to create meaningful, interpretable representations of the literature network.

6.3 Forming and Adjusting Relevance (RQ3)

This section examines how participants judged paper relevance, what shaped those judgments, and how their strategies evolved over time. Our thematic analysis [6, 8] was collaboratively conducted by two authors to identify recurring patterns. As summarized in Table 1, users engaged in a layered sensemaking process that combined familiar meta-data heuristics with network structure and system recommendations.

Table 1: Summary of relevance-judgment themes for RQ3 (N=15 participants).

Theme	n/15	Representative Quote
Thematic similarities as primary channel	15	"I relied mostly on thematic similarities because I wanted papers with similar context." (P1)
Citation / reference trails for bridging	10	"I relied heavily on citation and reference recommendations, then used thematic to broaden." (P10)
Pivoting among recommendation modes	9	"At first I relied on references, then switched to thematic when I needed new ideas." (P2)
Metadata scanning (title, keywords, summary)	7	"Due to time limits, I mostly relied on titles; if one seemed off-topic, I excluded it." (P10)
Author-based recs seldom used	5	"I didn't use author recs at all — I'm not familiar with authors in this field." (P5)
Ad-hoc → patterned workflows	5	"I started kind of sporadically, but then developed a pattern as I went along." (P1)
Graph-based heuristics (centrality / clusters)	4	"I selected papers based on how closely they were connected to others in the network." (P3)
Quantity vs. precision shifts	4	"Initially I was picky, but later I realized I'd just add anything that might be relevant." (P2)
Citation maturity checks	3	"A lot of recent papers with zero citations felt odd; I preferred older, foundational work." (P8)
Personal interest override	3	"I decided relevance mostly by personal interest, especially around XR in education." (P7)
Stable approach (little or no change)	1	"My core method didn't really change through the sessions." (P7)

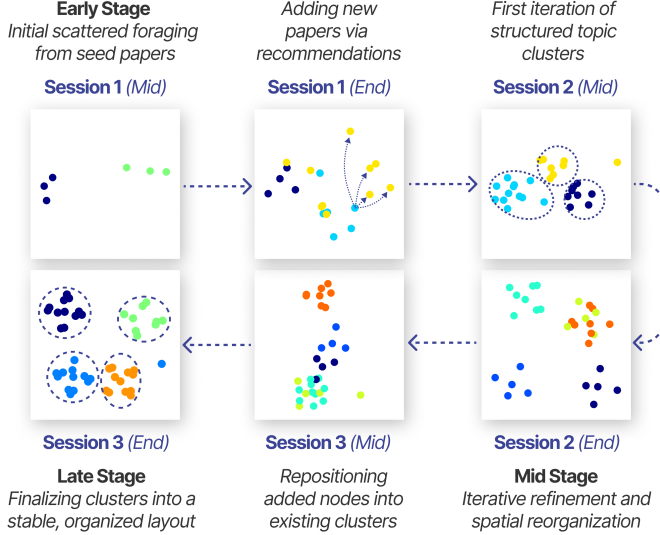


Fig. 8: Temporal evolution of Participant 2's spatial layout, showing six snapshots of the literature network from a bird's-eye view. *Mid* and *End* indicate the middle and end of each session, respectively.

How Relevance Was Judged. Most participants began with fast filtering based on visible metadata—especially paper titles. Seven out of 15 participants made keep-or-discard decisions primarily from title scans, often skipping abstracts. A few also referenced prompt-related keywords or short summaries. A subset (3/15) checked citation counts to assess paper maturity, preferring well-cited works. Three participants admitted that personal interest influenced their choices, sometimes retaining papers only loosely related to their search goals. Together, these strategies reflect a common filtering process: participants typically began with a quick assessment based on easily visible cues (like titles or keywords), then applied more careful filtering using abstracts or citation counts when needed, and occasionally explored papers out of personal curiosity. Across participants, these heuristics prioritized efficiency and ease of use, while still allowing flexibility to follow individual interests and goals during immersive literature foraging.

Factors Influencing Relevance Judgments. Recommendation tools strongly shaped exploration patterns. All participants used recommendations by thematic similarities for topical consistency and interdisciplinary insights. Two-thirds of participants (10/15) also used citation or reference views to discover foundational papers or bridge clusters. Author-based recommendations were rarely used (5/15), mostly due to unfamiliarity with domain experts. Spatial layout also influenced judgments. Four out of 15 participants referenced cluster centrality or connectivity, using network position as a proxy for importance when metadata was insufficient.

How Strategies Evolved Over Time. Most participants developed dynamic adaptive strategies throughout the study. Five participants began with ad-hoc browsing before settling into structured workflows.

The majority (9/15) reported switching recommendation modes—for example, moving from citation-based to thematic recommendations—when their progress slowed or when trying to seek different types of connections. Some (4/15) adjusted their level of selectivity over time, shifting between quantity-first (gather broadly) and precision-first (filter narrowly) approaches. Two participants even developed new goals mid-task, such as connecting previously separate clusters of papers. Notably, only one participant maintained a consistent strategy throughout the session. This adaptability underscores the inherently serendipitous nature of literature exploration—not merely a process of retrieving known items, but an evolving search for connections, gaps, and unexpected insights within a complex and expanding information space.

Summary. Across sessions, participants combined filtering heuristics with recommendation tools and network cues to refine their relevance judgments. Participants also exhibited distinct foraging styles—such as *metadata scanners* (filtering by titles or keywords), *network navigators* (navigating the graph structure), and *adaptive pivots* (frequently switching strategies). These styles suggest opportunities for future personalization in immersive systems. More broadly, these patterns highlight how immersive spatial tools actively shape how users assess and construct relevance during open-ended, exploratory tasks like literature review.

7 CONCLUSION & FUTURE WORK

We present LITFORAGER, an immersive analytics system designed to facilitate the integrated workflow of literature exploration. Through a formative study with academic researchers, we identified essential user needs and established key design goals: exploration guidance, spatial organization, and seamless transitions between foraging and synthesis. LITFORAGER achieves these goals by combining interactive 3D literature network visualization, flexible spatial organization, and multimodal interactions. Our user study with 15 researchers revealed distinct patterns in modality preference, spatial organization, and relevance formation. This work highlights the potential of immersive technologies to enhance literature foraging and contributes to immersive analytics by addressing previously underexplored aspects of the sensemaking workflow.

Future work will explore enhancing LITFORAGER by integrating personalized recommendation algorithms and expanding multimodal interaction capabilities. Additionally, longitudinal studies will investigate how sustained use impacts researchers' productivity and cognitive workflows in real-world scenarios.

ACKNOWLEDGMENTS

This work was supported in part by NSF grant IIS-2441310.

REFERENCES

- [1] B. W. Ammar, D. Groeneveld, C. Bhagavatula, I. Beltagy, M. Crawford, D. Downey, J. Dunkelberger, A. Elgohary, S. Feldman, V. A. Ha, R. M. Kinney, S. Kohlmeier, K. Lo, T. C. Murray, H.-H. Ooi, M. E. Peters, J. L. Power, S. Skjonsberg, L. L. Wang, C. Wilhelm, Z. Yuan, M. van Zuylen, and O. Etzioni. Construction of the literature graph in semantic scholar. In

- North American Chapter of the Association for Computational Linguistics, 2018. 2, 5
- [2] C. Andrews, A. Endert, B. Yost, and C. North. Information visualization on large, high-resolution displays: Issues, challenges, and opportunities. *Information Visualization*, 10(4):341–355, 2011. 2, 3, 4, 5, 8
 - [3] F. Beck. PUREsuggest: Citation-Based Literature Search and Visual Exploration with Keyword-Controlled Rankings. *IEEE Transactions on Visualization and Computer Graphics*, 31(1):316–326, Jan. 2025. doi: 10.1109/TVCG.2024.3456199 2
 - [4] R. A. Bolt. “put-that-there” voice and gesture at the graphics interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, pp. 262–270, 1980. 8
 - [5] A. Booth, A. Sutton, and D. Papaioannou. *Systematic approaches to a successful literature review*. Sage, Los Angeles, second edition ed., 2016. 3
 - [6] V. Braun and V. Clarke. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4):589–597, Aug. 2019. doi: 10.1080/2159676X.2019.1628806 3, 8
 - [7] Brown University Library. Organizing and creating information. <https://libguides.brown.edu/organize/litreview>. Accessed: 2025-03-10. 3
 - [8] D. Byrne. A worked example of Braun and Clarke’s approach to reflexive thematic analysis. *Quality & Quantity*, 56(3):1391–1412, June 2022. doi: 10.1007/s11135-021-01182-y 3, 8
 - [9] P. R. Cohen and S. L. Oviatt. The role of voice input for human-machine communication. *proceedings of the National Academy of Sciences*, 92(22):9921–9927, 1995. 5
 - [10] Connected Papers team. Connected papers. <https://www.connectedpapers.com/>, 2020. Accessed: 2025-04-01. 4
 - [11] M. Cordeil, A. Cunningham, T. Dwyer, B. H. Thomas, and K. Marriott. ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST ’17, pp. 71–83. Association for Computing Machinery, New York, NY, USA, Oct. 2017. doi: 10.1145/3126594.3126613 3, 5
 - [12] David Catuhe and Babylon.js contributors. Babylon.js. <https://www.babylonjs.com>, 2013. Accessed: 2025-04-05. 2, 5
 - [13] K. Davidson, L. Lisle, K. Whitley, D. A. Bowman, and C. North. Exploring the Evolution of Sensemaking Strategies in Immersive Space to Think. *IEEE Transactions on Visualization and Computer Graphics*, 29(12):5294–5307, Dec. 2023. doi: 10.1109/TVCG.2022.3207357 2
 - [14] Department of Psychology, UC San Diego. Writing a literature review. <https://psychology.ucsd.edu/undergraduate-program/undergraduate-resources/academic-writing-resources/writing-research-papers/writing-lit-review.html>. Accessed: 2025-03-10. 3
 - [15] P. Dourish. *Where the action is: the foundations of embodied interaction*. MIT press, 2001. 5
 - [16] A. Endert, W. Ribarsky, C. Turkay, W. Wong, I. Nabney, I. D. Blanco, and F. Rossi. The State of the Art in Integrating Machine Learning into Visual Analytics. *Computer Graphics Forum*, 36(8):458–486, Dec. 2017. doi: 10.1111/cgf.13092 1
 - [17] A. Ghasemi, P. Mirmiran, K. Kashfi, and Z. Bahadoran. Scientific Publishing in Biomedicine: A Brief History of Scientific Journals. *International Journal of Endocrinology and Metabolism*, 21(1):e131812, Dec. 2022. doi: 10.5812/ijem-131812 1
 - [18] Google DeepMind. Gemini 2.5 Flash. <https://deepmind.google/technologies/gemini/>, 2024. Accessed: 2025-06-16. 5, 6
 - [19] H. H. Huang, H. Pfister, and Y. Yang. Is embodied interaction beneficial? a study on navigating network visualizations. *Information Visualization*, 22(3):169–185, 2023. doi: 10.1177/14738716231157082 3
 - [20] E. Hutchins. *Cognition in the Wild*. MIT press, 1995. 3, 5, 8
 - [21] Immersive Web Working Group. WebXR. <https://www.w3.org/TR/webxr/>, 2024. Accessed: 2025-04-05. 2, 5
 - [22] S. In, E. Krokos, K. Whitley, C. North, and Y. Yang. Evaluating Navigation and Comparison Performance of Computational Notebooks on Desktop and in Virtual Reality. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI ’24, pp. 1–15. Association for Computing Machinery, New York, NY, USA, May 2024. doi: 10.1145/3613904.3642932 3, 4
 - [23] S. In, T. Lin, C. North, H. Pfister, and Y. Yang. This is the Table I Want! Interactive Data Transformation on Desktop and in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 30(8):5635–5650, Aug. 2024. doi: 10.1109/TVCG.2023.3299602 4, 5
 - [24] R. Johnson, A. Watkinson, and M. Mabe. The stm report. *An overview of scientific and scholarly publishing. 5th edition October*, 94, 2018. 1
 - [25] H. B. Kang, N. Soliman, M. Latzke, J. C. Chang, and J. Bragg. ComLittee: Literature Discovery with Personal Elected Author Committees. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–20. ACM, Hamburg Germany, Apr. 2023. doi: 10.1145/3544548.3581371 2
 - [26] Y.-H. Kim, B. Lee, A. Srinivasan, and E. K. Choe. Data@Hand: Fostering Visual Exploration of Personal Data on Smartphones Leveraging Speech and Touch Interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI ’21, pp. 1–17. Association for Computing Machinery, New York, NY, USA, May 2021. doi: 10.1145/3411764.3445421 2, 5
 - [27] D. Kirsh. The intelligent use of space. *Artificial intelligence*, 73(1-2):31–68, 1995. 3, 5, 8
 - [28] P. Knoth, D. Herrmannova, M. Cancellieri, L. Anastasiou, N. Pontika, S. Pearce, B. Gyawali, and D. Pride. CORE: A Global Aggregation Service for Open Access Papers. *Scientific Data*, 10(1):366, June 2023. doi: 10.1038/s41597-023-02208-w 1
 - [29] J. Kotlarek, O.-H. Kwon, K.-L. Ma, P. Eades, A. Kerren, K. Klein, and F. Schreiber. A study of mental maps in immersive network visualization. In *2020 IEEE pacific visualization symposium (PacificVis)*, pp. 1–10. IEEE, 2020. 4
 - [30] O.-H. Kwon, C. Muelder, K. Lee, and K.-L. Ma. A study of layout, rendering, and interaction methods for immersive graph visualization. *IEEE transactions on visualization and computer graphics*, 22(7):1802–1815, 2016. 4
 - [31] E. Landhuis. Scientific literature: Information overload. *Nature*, 535(7612):457–458, July 2016. doi: 10.1038/nj7612-457a 1, 6
 - [32] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017. 5, 8
 - [33] K. Li, H. Yang, E. Montoya, A. Upadhayay, Z. Zhou, J. Saad-Falcon, and D. H. Chau. Visual exploration of literature with argo scholar. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pp. 4912–4916, 2022. 4
 - [34] L. Lisle, X. Chen, J. Edward Gitre, C. North, and D. A. Bowman. Evaluating the Benefits of the Immersive Space to Think. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 331–337, Mar. 2020. doi: 10.1109/VRW50115.2020.00073 2
 - [35] L. Lisle, K. Davidson, E. J. Gitre, C. North, and D. A. Bowman. Sensemaking Strategies with Immersive Space to Think. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 529–537, Mar. 2021. doi: 10.1109/VR50410.2021.00077 2, 3, 4, 5, 8
 - [36] L. Lisle, K. Davidson, E. J. K. Gitre, C. North, and D. A. Bowman. Different realities: a comparison of augmented and virtual reality for the sensemaking process. *Frontiers in Virtual Reality*, 4, Aug. 2023. Publisher: Frontiers. doi: 10.3389/frvir.2023.1177855 3
 - [37] L. Lisle, K. Davidson, L. Pavanatto, I. A. Tahmid, C. North, and D. A. Bowman. Spaces to Think: A Comparison of Small, Large, and Immersive Displays for the Sensemaking Process. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1084–1093, Oct. 2023. ISSN: 2473-0726. doi: 10.1109/ISMAR59233.2023.00125 2
 - [38] K. Lo, L. L. Wang, M. Neumann, R. M. Kinney, and D. S. Weld. S2orc: The semantic scholar open research corpus. In *Annual Meeting of the Association for Computational Linguistics*, 2020. 2, 5
 - [39] W. Luo, M. O. Ellenberg, M. Satkowski, and R. Dachsel. Documents in Your Hands: Exploring Interaction Techniques for Spatial Arrangement of Augmented Reality Documents. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI ’25, pp. 1–22. Association for Computing Machinery, New York, NY, USA, Apr. 2025. doi: 10.1145/3706598.3713518 2
 - [40] W. Luo, A. Lehmann, H. Widengren, and R. Dachsel. Where Should We Put It? Layout and Placement Strategies of Documents in Augmented Reality for Collaborative Sensemaking. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI ’22, pp. 1–16. Association for Computing Machinery, New York, NY, USA, Apr. 2022. doi: 10.1145/3491102.3501946 3
 - [41] Microsoft Mixed Reality. Instinctual interactions. <https://learn.microsoft.com/en-us/windows/mixed-reality/design/interaction-fundamentals>, 2022. Accessed: 2025-03-10.

- [42] Mike Bostock, Jason Davies, Jeffrey Heer, Vadim Ogievetsky, and community. D3.js. <https://www.babylonjs.comhttps://d3js.org>, 2011. Accessed: 2025-06-16. 5
- [43] A. Norman Donald. *The design of everyday things*. MIT Press, 2013. 3, 5, 8
- [44] S. Oviatt. Ten myths of multimodal interaction. *Communications of the ACM*, 42(11):74–81, 1999. 8
- [45] S. Palani, A. Naik, D. Downey, A. X. Zhang, J. Bragg, and J. C. Chang. Relatedly: Scaffolding Literature Reviews with Existing Related Work Sections. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–20. ACM, Hamburg Germany, Apr. 2023. doi: 10.1145/3544548.3580841 2
- [46] G. Paré and S. Kitsiou. Chapter 9 Methods for Literature Reviews. In *Handbook of eHealth Evaluation: An Evidence-based Approach [Internet]*. University of Victoria, Feb. 2017. 3
- [47] P. Pirolli and S. Card. Information foraging. *Psychological review*, 106(4):643, 1999. 1, 2, 4
- [48] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, vol. 5, pp. 2–4. McLean, VA, USA, 2005. 1, 2, 4
- [49] A. Ponsard, F. Escalona, and T. Munzner. PaperQuest: A Visualization Tool to Support Literature Review. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 2264–2271. ACM, San Jose California USA, May 2016. doi: 10.1145/2851581.2892334 2
- [50] Purdue Online Writing Lab, Purdue University. Writing a literature review. https://owl.purdue.edu/owl/research_and_citation/conducting_research/writing_a_literature_review.html. Accessed: 2025-03-10. 3
- [51] S. Reinders, M. Butler, I. Zukerman, B. Lee, L. Qu, and K. Marriott. When Refreshable Tactile Displays Meet Conversational Agents: Investigating Accessible Data Presentation and Analysis with Touch and Speech. *IEEE Transactions on Visualization and Computer Graphics*, 31(1):864–874, Jan. 2025. arXiv:2408.04806 [cs]. doi: 10.1109/TVCG.2024.3456358 2
- [52] G. Reiske, S. In, and Y. Yang. Multi-focus querying of the human genome information on desktop and in virtual reality: an evaluation. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1123–1131. IEEE, 2023. doi: 10.1109/ISMAR59233.2023.00129 5
- [53] ResearchRabbit team. Researchrabbit. <https://www.researchrabbit.ai/>, 2021. Accessed: 2025-04-01. 4
- [54] D. Russell, M. Stefik, P. Pirolli, and S. Card. The cost structure of sense making. Jan. 1993. 2
- [55] A. Saktheeswaran, A. Srinivasan, and J. Stasko. Touch? Speech? or Touch and Speech? Investigating Multimodal Interaction for Visual Network Exploration and Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 26(6):2168–2179, June 2020. doi: 10.1109/TVCG.2020.2970512 2, 5
- [56] B. Shneiderman. Direct manipulation: A step beyond programming languages. *Computer*, 16(08):57–69, 1983. 5
- [57] R. Skarbez, N. F. Polys, J. T. Ogle, C. North, and D. A. Bowman. Immersive Analytics: Theory and Research Agenda. *Frontiers in Robotics and AI*, 6:82, Sept. 2019. doi: 10.3389/frobt.2019.00082 3, 5
- [58] N. R. Smalheiser. Literature-based discovery: Beyond the ABCs. *Journal of the American Society for Information Science and Technology*, 63(2):218–224, 2012. doi: 10.1002/asi.21599 2
- [59] N. R. Smalheiser. Rediscovering Don Swanson: the Past, Present and Future of Literature-Based Discovery. *Journal of data and information science (Warsaw, Poland)*, 2(4):43–64, Dec. 2017. doi: 10.1515/jdis-2017-0019 2
- [60] I. A. Tahmid, L. Lisle, K. Davidson, K. Whitley, C. North, and D. A. Bowman. Evaluating the Feasibility of Predicting Information Relevance During Sensemaking with Eye Gaze Data. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 713–722. IEEE, Sydney, Australia, Oct. 2023. doi: 10.1109/ISMAR59233.2023.00086 2
- [61] I. A. Tahmid, C. North, K. Davidson, K. Whitley, and D. Bowman. Enhancing Immersive Sensemaking with Gaze-Driven Recommendation Cues. In *Proceedings of the 30th International Conference on Intelligent User Interfaces, IUI '25*, pp. 641–659. Association for Computing Machinery, New York, NY, USA, Mar. 2025. doi: 10.1145/3708359.3712103 2
- [62] W. Tong, H. Li, M. Xia, W. Kam-Kwai, T.-C. Pong, H. Qu, and Y. Yang. Exploring spatial hybrid user interface for visual sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, 2025. doi: 10.1109/TVCG.2025.3538771 2
- [63] W. Tong, M. Xia, K. K. Wong, D. A. Bowman, T.-C. Pong, H. Qu, and Y. Yang. Towards an understanding of distributed asymmetric collaborative visualization on problem-solving. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 387–397. IEEE, 2023. doi: 10.1109/VR55154.2023.00054 2
- [64] vasturiano. d3-force-3d. <https://github.com/vasturiano/d3-force-3d>, 2025. Accessed: 2025-06-16. 5
- [65] Y. Yang, M. Cordeil, J. Beyer, T. Dwyer, K. Marriott, and H. Pfister. Embodied navigation in immersive abstract data visualization: Is overview+detail or zooming better for 3d scatterplots? *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1214–1224, 2020. doi: 10.1109/TVCG.2020.3030427 5
- [66] Y. Yang, T. Dwyer, K. Marriott, B. Jenny, and S. Goodwin. Tilt Map: Interactive Transitions Between Choropleth Map, Prism Map and Bar Chart in Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(12):4507–4519, Dec. 2021. doi: 10.1109/TVCG.2020.3004137 3, 5
- [67] S. R. Yousif, M. D. Rosenberg, and F. C. Keil. Using space to remember: Short-term spatial structure spontaneously improves working memory. *Cognition*, 214:104748, Sept. 2021. doi: 10.1016/j.cognition.2021.104748 3, 5
- [68] Z. Yuan, Y. Liu, and L. Yu. MEinVR: Multimodal Interaction Paradigms in Immersive Exploration. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 85–90, Oct. 2022. doi: 10.1109/ISMAR-Adjunct57072.2022.00026 2
- [69] Q. Zhu, T. Lu, S. Guo, X. Ma, and Y. Yang. Compositingvis: Exploring interactions for creating composite visualizations in immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 2024. doi: 10.1109/TVCG.2024.3456210 3