Garden of Papers: Finding, Reading, and Organizing Research Papers in a Visual, Integrated, and Flexible Workspace

Donghyeok Ma
Department of Industrial Design
KAIST
Daejeon, Republic of Korea
donghyeok.ma@kaist.ac.kr

Hanbee Jang
Department of Industrial Design
KAIST
Daejeon, Republic of Korea
hanbee.jang@kaist.ac.kr

Joon Hyub Lee
DRB-KAIST
SketchTheFuture Research Center
KAIST
Daejeon, Republic of Korea
joonhyub.lee@kaist.ac.kr

Seok-Hyung Bae
Department of Industrial Design
KAIST
Daejeon, Republic of Korea
seokhyung.bae@kaist.ac.kr

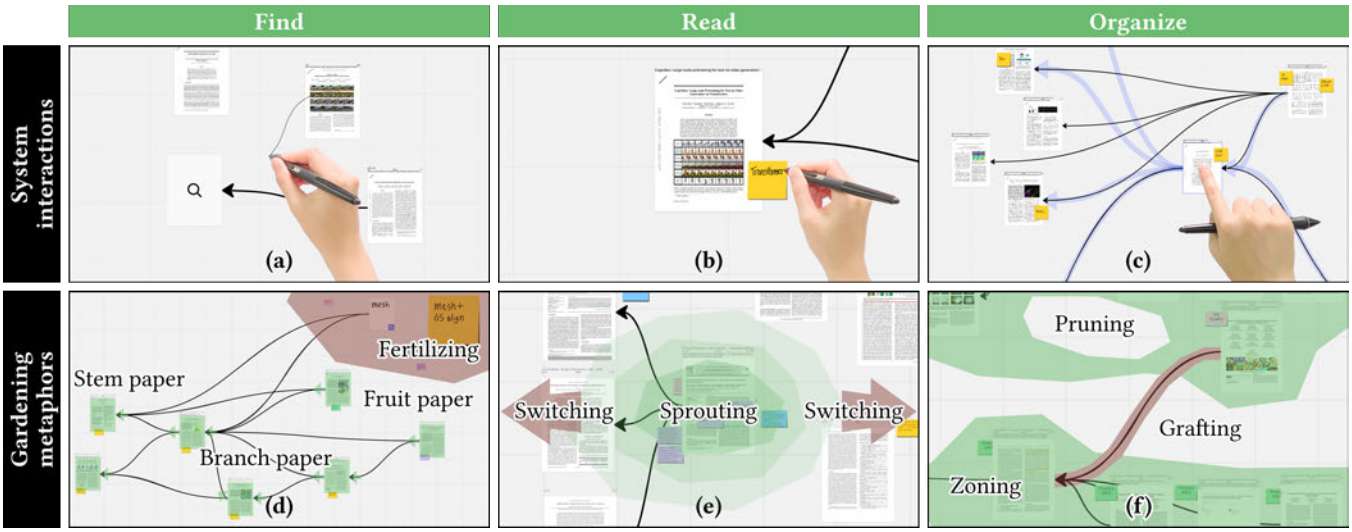

Figure 1: In our novel interactive system, *Garden of Papers*, the researcher can use sketch-based pen and multi-touch gestures on a 2D canvas to (a) find new papers using citation relationships, (b) read papers and jot down thoughts directly within the same workspace, and (c) organize papers based on these citation relationships and thoughts. (d) Beyond simply finding new papers that cite or are cited by *seed papers*, the researcher can intuitively identify foundational papers (*stem papers*), intermediate papers (*branch papers*), and converging papers (*fruit papers*). The researcher can also keep incomplete queries and return to them later (*fertilizing*). (e) Beyond simply reading papers in traditional PDF viewers, the researcher can immediately bring additional materials to the workspace and peruse them side by side (*sprouting*), while seamlessly transitioning between finding, reading, and organizing activities (*switching*). (f) Beyond simply storing papers in folders, the researcher can arrange them into meaningful clusters (*zoning*), create new connections between clusters (*grafting*), and remove unnecessary ones (*pruning*), cultivating the research landscape as if tending a garden.

Abstract

Writing a research paper requires significant time and effort to find, read, and organize many related papers, which are complex and intensive knowledge tasks. We present *Garden of Papers*, a novel sketch-based interactive system that allows users to perform these tasks quickly and easily on a 2D canvas. Users can intuitively create paper nodes and citation links, naturally accumulate thoughts onto them, and flexibly expand and restructure them, progressively shaping their personalized node-link diagrams of related papers. We first conducted a short-term lab study (4 participants, 1 week) using a proof-of-concept system. Based on the findings from this study, we refined the system and conducted a long-term in-the-wild study (12 participants, 4 weeks). From these studies, we quantitatively and qualitatively demonstrate that our system's visual, integrated, and flexible workspace facilitates 9 usage patterns that enhance real-world literature reviews. We describe these patterns using gardening metaphors, emphasizing how literature review tools grounded in visual thinking and spatial memory can help researchers not only make sense of, but also actively cultivate their evolving research landscapes.

UIST '25, Busan, Republic of Korea

ACM ISBN 979-8-4007-2037-6/25/09
https://doi.org/10.1145/3746059.3747637

CCS Concepts

• **Human-centered computing → Interaction techniques**.

Keywords

Literature review, node-link diagram

ACM Reference Format:

Donghyeok Ma, Hanbee Jang, Joon Hyub Lee, and Seok-Hyung Bae. 2025. Garden of Papers: Finding, Reading, and Organizing Research Papers in a Visual, Integrated, and Flexible Workspace. In *The 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25), September 28–October 01, 2025, Busan, Republic of Korea.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3746059.3747637

1 INTRODUCTION

Researchers write Related Work sections to explain that their papers are based on an existing body of knowledge, differ from it, and contribute to it. To do so, they investigate hundreds of papers over several weeks or months, cite dozens of related papers, and describe their relationships in writing.

A common method of investigating related papers is to first find *seed papers* that are considered highly relevant and then conduct an exhaustive survey of the papers that the seed papers cite and the papers that cite the seed papers [5, 6]. By doing so, researchers can collect a comprehensive list of old and new papers on the topic.

However, each paper usually cites dozens of others. Also, it may be cited by hundreds of others. Therefore, as the above process is repeated, the number of papers to collect can increase exponentially. At this point, it is easy for researchers to feel anxious about missing out on important papers, overwhelmed by the growing pile of papers to read and analyze, and frustrated by the lack of a visual representation to help them see the big picture.

We thus present *Garden of Papers*, a novel interactive system designed to assist researchers in conducting literature reviews. This system is the first to integrate graph-based paper finding, reading, and organizing within a unified workspace, with all features translated into elegant and intuitive pen and multi-touch gestures.

Specifically, the system allows users to find papers by simply sketching their citation relationships with other papers on a 2D canvas. By repeating this intuitive process, users can construct node-link diagrams composed of paper nodes and citation links, leveraging their visual thinking and spatial memory. As users read the collected papers directly from these diagrams, they can freely annotate their thoughts and flexibly expand or restructure the diagrams as needed, progressively cultivating a personalized landscape of related papers (Figure 1a-c).

We conducted 2 user studies that complement each other and found 9 usage patterns facilitated by our system during literature review processes (Figure 1d-f). We describe these patterns using gardening metaphors to reflect the organic and evolving nature of literature reviews.

This paper makes the following contributions:

• We designed *Garden of Papers*, a novel interactive system that supports literature reviews by enabling researchers to construct and navigate node-link diagrams of papers and citations through sketch-based interactions.

• We implemented a proof-of-concept system and conducted a short-term lab study (4 participants, 1 week) with researchers who had previously published papers, resulting in 7 findings (F1-F7) on how our system could potentially benefit literature reviews.

• We upgraded the system and conducted a long-term in-the-wild study (12 participants, 4 weeks), resulting in 2 additional findings (F8, F9), along with quantitative and qualitative evidence supporting all findings (F1-F9) on how our system indeed benefits real-world literature reviews.

• We summarized meaningful graph topologies of research papers and the corresponding activities that produce them through gardening metaphors, suggesting a conceptual framework for future literature review tools.

2 RELATED WORK

In this section, we analyze previous work on visualization and interaction techniques relevant to literature reviews. We also discuss how our proposed system differs from existing tools designed to support literature reviews.

2.1 Node-Link Diagrams for Papers & Citations

In Related Work sections, researchers explain previous work relevant to their own work and position their contributions within the overall research landscape of the field. This requires finding, reading, and classifying numerous papers, as well as understanding the citation relationships among them. Various commercial tools have been developed to support this process of reviewing and managing related work. Many of these tools provide list-based interfaces [7, 12], which are suitable for presenting search results and their classifications but less so for showing the complex graphs of citation relationships among a multitude of papers within a large-scale research landscape [25].

Previous work aimed to better visualize citation relationships among papers using node-link diagrams [5, 6, 9, 11, 19, 21, 30, 39, 41, 43, 46, 51]. On a 2D canvas, papers are represented as nodes and their citations as links, helping researchers easily identify relationships between papers. However, even when represented as a node-link diagram, an increasing number of papers can make it difficult for researchers to revisit and process the information they have found [56].

People are known to remember an item's location better when they place it in a location of their own choosing [42, 54]. Our proposed system also employs a node-link diagram, but unlike previous work that automatically determined the placement of paper nodes and the citation links between them [6, 9, 11, 19, 21, 30, 39, 41, 43, 46, 51], it lets researchers manually position paper nodes and determine their citation links. This approach helps researchers easily recall the position of each paper and the local structures of their research landscape [5, 49].

People are also known to recognize images more easily than text [31, 52]. Therefore, unlike previous work that represented nodes using abstract symbols and text labels [5, 6, 9, 11, 19, 21, 30, 39, 41, 43, 46, 49, 51], our system represents nodes as thumbnails that display actual pages of papers [2, 31], along with any annotations left on those pages [53], helping researchers more easily recognize and

remember specific papers and their citation relationships within a complex, overall research landscape.

2.2 Interactions for Node-Link Diagrams

To find and organize papers with a node-link diagram, user interactions for effectively expressing papers and their citation links are needed. Xia et al. [55] proposed a system that interprets quickly drawn sketches and transforms them into interactive objects, enabling users to continue their work without disrupting their flow of thought. Subramonyam et al. [50] proposed a system that converts portions of text highlighted by users into a node-link diagram. Similarly, our system, when researchers quickly sketch papers and their citation relationships, automatically retrieves the corresponding papers from academic databases and transforms them into interactive paper nodes and citation links, enabling researchers to organically generate new ideas and further develop them.

To search for and add new information, many systems rely on a search box that users can type text into, while some systems make use of node-link diagrams. For example, when users create and label a node as an "artist" and create a link indicating that the node has been "influenced by" another known artist node, the system searches for people who match those conditions [18, 26, 37]. In our system, researchers can sketch a blank sheet and its relationships with existing papers, along with keywords, to formulate diverse search conditions. By combining keyword-based and graph-based queries, researchers can efficiently find desired papers and integrate them to the existing node-link diagram with minimal effort.

As new papers are added, new citation relationships emerge, causing the node-link diagrams of papers and citations to rapidly expand in size and complexity. We thus designed a coherent set of effective interactions for selecting and moving nodes, forming clusters, and creating and deleting links, by referring to various pen and multi-touch interaction techniques [20, 55]. We also designed interactions for reading and annotating papers that harmonize with the node-link diagram interactions. In addition, we implemented mouse and keyboard interactions to be compatible with the pen and multi-touch interactions originally designed for tablets, enabling researchers to use them on desktops as well.

2.3 Tools for Supporting Literature Reviews

Literature review is a nonlinear process in which the tasks of finding, reading, and organizing papers frequently alternate [13]. Tools that support this process should enable such frequent task switching without disturbing the continuous flow of thought [58]. Most of the aforementioned work proposed systems specialized for individual tasks, but some attempted to integrate multiple tasks. For example, while researchers are reading a paper, such systems may display a list of papers cited in a selected portion of the paper [22], provide brief summaries of cited papers in the margins [40], present a list of new papers retrieved through keyword-based search [36] and large language models [23], or recommend new papers that match the content the researchers have previously read and their known research interests [4].

To the best of our knowledge, our system is the first to integrate all three tasks of finding, reading, and organizing papers within a single, unified 2D workspace. In this workspace, researchers can use node-link diagrams to search for and collect papers, construct an overall research landscape, read papers actively while recording ideas and notes, and develop the node-link diagrams further, which can lead to a virtuous cycle.

We previously introduced a subset of the interaction techniques for such a system [27, 28], but did not validate its usefulness in actual research practices. In this paper, we present a long-term in-the-wild study involving researchers from diverse fields, reporting both quantitative and qualitative findings that demonstrate our system's ability to effectively support literature reviews in actual research practices.

3 PROOF OF CONCEPT

In this section, we derive design goals for a novel interactive system supporting literature reviews from analysis of related work and introduce interactions that meet these goals.

3.1 Design Goals

We aim to address the cognitive load researchers face when finding, reading, and organizing a large number of papers. To this end, we set 3 key design goals that align with the natural flow of thought:

- **Visual.** The system should visualize all information in alignment with researchers' intentions on a 2D canvas; the system should represent citation relationships between papers through a node-link diagram; the system should support finding papers using a node-link diagram.
- **Integrated.** The system should support the complete workflow of finding, reading, and organizing papers within a single workspace, enabling a continuous flow of researchers' thinking.
- **Flexible.** The system should allow researchers to switch seamlessly at any time between finding, reading, and organizing papers; the system should allow researchers to freely utilize every region of a 2D canvas for different purposes; the system should allow researchers to freely utilize papers and sticky notes for different purposes.

3.2 Interactions

Building on our design goals, we introduce novel node-link diagram techniques that allow researchers to find, read, and organize papers directly on a 2D canvas. We implemented a proof-of-concept system based on pen and multi-touch inputs, as these modalities best demonstrate our sketch-based interactions.

- **Find.** Researchers can create a blank sheet on the canvas and specify search conditions to discover and collect related papers (Figure 2) that match various keywords and citation relationships (Figure 3).
- **Read.** Researchers can flip through the full-text pages of collected papers directly on the canvas (Figure 4a) and jot down thoughts that come to mind while reading (Figure 4b).
- **Organize.** Researchers can selectively display specific citation links among paper nodes (Figure 4c), select multiple nodes and links to manipulate simultaneously (Figure 4d), and remove unnecessary entities from diagrams (Figure 4e).

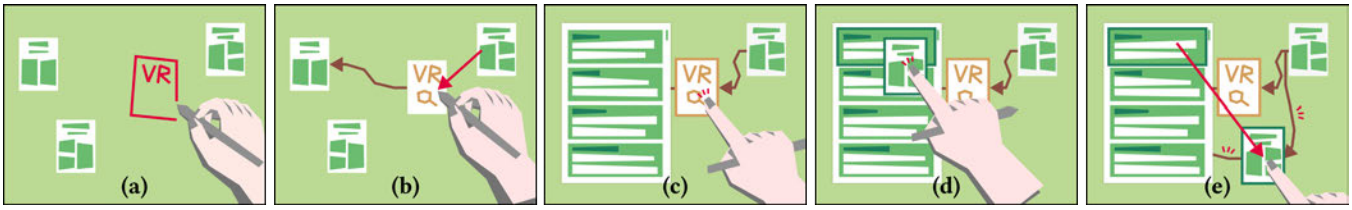

Figure 2: The user can sketch a blank sheet on a 2D canvas by (a) drawing a box, writing one or more keywords, and (b) drawing one or more citation relationships. (c-d) Then, the recognized keywords and citation relationships from the sketch are used as search conditions to find papers in an external database [15]. (e) The user can collect papers at desired locations on the canvas. Citation links to and from the collected papers are automatically generated.

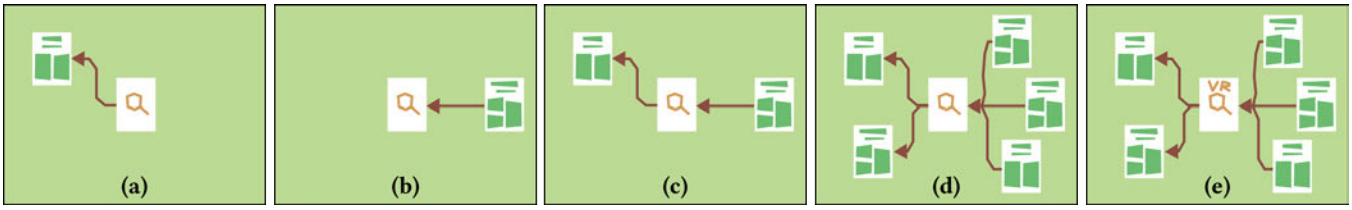

Figure 3: By drawing various node-link diagrams, the user can query (a) all papers citing a paper, (b) all papers cited by a paper, (c) all papers citing a paper and cited by another paper, (d) all papers citing multiple papers and cited by other multiple papers, and (e) all papers containing specific keywords, citing multiple papers, and cited by other multiple papers.

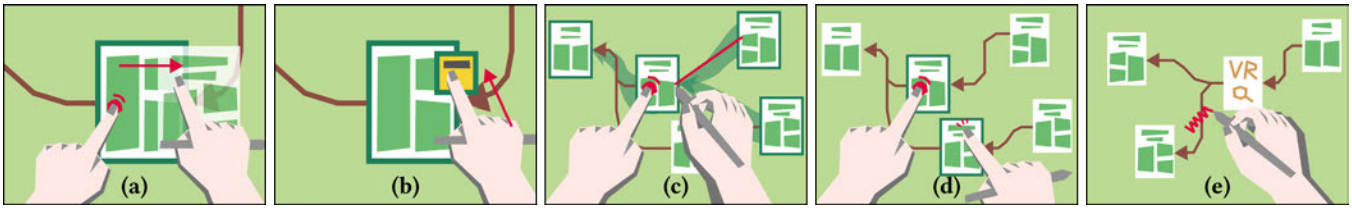

Figure 4: The user can (a) navigate through pages of a paper by holding it down with a finger and flicking it with another finger, (b) create a sticky note by sketching a small square and attach it to a page by dragging it, (c) view all existing citation relationships between a paper and the other papers by long-pressing it and curate preferred relationships by drawing over them, (d) select multiple papers by long-pressing a paper with a finger and tapping other papers with another finger, and (e) erase unnecessary information by scribbling over entities.

4 SHORT-TERM LAB STUDY

To understand how our proof-of-concept system can support literature reviews in a controlled setting and to identify its limitations for real-world use, we conducted a short-term lab study. Over the course of 1 week, we asked experienced researchers from various fields to use our system for literature reviews, and conducted pre- and post-interviews with them to compare our system against their usual workflows.

• **Participants.** We recruited 4 researchers (2 males and 2 females, ages 26-29) from various fields (Table 1). All participants had previously written and published research papers, with an average of 5.0 years of research experience and 3.5 publications as first authors or co-authors.

• **Apparatus.** We implemented our system using Unity and built a tablet application supporting pen and multi-touch inputs. We used Selvy Pen SDK [44] to recognize handwriting, and Serp API [45], a third-party API for Google Scholar [15], and Semantic Scholar API [24] to retrieve paper metadata, including the title, authors, publication venue, PDF URL, and citation relationships for each paper. We deployed our system on a 14.4" Microsoft Surface Laptop Studio, a convertible device that we primarily used in tablet mode, paired with a Microsoft Slim Pen.

Table 1: Participant demographics for the lab study. G: gender, A: age, Y: years of research experience, F: number of first-author papers, C: number of co-author papers.

	G	A	Y	F	C	Major	Research area
P1	M	26	5	5	0	Electrical engineering	Audio deep learning
P2	M	26	4	1	3	Mechanical engineering	Vision deep learning
P3	F	29	6	0	1	Biological science	Optogenetics
P4	F	27	5	1	3	Biomolecular engineering	RNA engineering

• **Procedure.** On Day 1, we visited the participants individually at their workplaces to conduct pre-interviews about their usual literature review workflows. We then provided 30-minute tutorial sessions, during which we demonstrated each feature of the system step by step and guided the participants through hands-on practice until they became fully familiar with it. We also provided printed manuals describing each feature.

Table 2: System usage log per participant during the lab study.

	Time spent (hh:mm)	Paper nodes (#)			Citation links (#)				Queries (#)				All citation views (#)	Sticky notes (#)
		Added	Deleted	Total	Generated	Added	Deleted	Total	Keyword	Graph	Both	Total		
P1	03:43	63	25	38	58	15	34	39	16	60	13	89	22	37
P2	03:29	85	34	51	92	4	37	59	17	54	42	113	3	29
P3	03:29	65	25	30	41	3	19	25	22	48	42	112	40	12
P4	03:36	38	5	33	4	26	5	25	56	20	16	92	69	29
Total	14:16	251	89	162	195	48	95	148	111	182	113	406	134	147
Avg.	03:34	62.8	22.3	40.5	48.8	12.0	23.8	37.0	27.8	45.5	28.3	101.5	33.5	26.8

From Day 2 to Day 5, we visited the participants at their workplaces every day and observed them conducting literature reviews on topics of personal interest using our system. Each supervised session lasted approximately 1 hour. We also responded to their questions in person whenever they encountered difficulties. After each session, we encouraged them to review the collected papers on their own to deepen their understanding and familiarity with those papers.

On Day 6, we conducted post-interviews in which we asked the participants to review their usage logs and node-link diagrams arranged by date, and to reflect on the perceived advantages and disadvantages of our system compared to their usual workflows.

• **Measurements.** We recorded the total usage time of our system, the number of paper nodes, citation links, and sticky notes created and deleted over time, and the frequency and duration of all invoked system features. We also captured screen recordings of the participants' interactions for qualitative analysis.

4.1 Pre-Interview

According to the participants, they typically begin literature review processes by conducting keyword searches in databases like Google Scholar [15], or by exploring recent papers on publisher websites such as Nature [48] and Science [3] or prepublication repositories such as arXiv [10] and bioRxiv [8]. They prioritize papers based on the title, venue, publication year, citation count, abstract, figures, and introduction. Additionally, they explore papers cited by these papers, as well as papers that cite them.

As the number of papers grows, the participants utilize tools such as Mendeley [12], Notion [34], Google Sheets [16], and Google Slides [17] to group related papers for easy retrieval later. They organize papers into dedicated folders within Mendeley, and create separate Notion documents to summarize key points, take notes, and collect notable figures (P1, P2). Additionally, they record metadata and assign custom keywords to each paper in Google Sheets, enabling easy retrieval through keyword filtering (P3). Finally, they use Google Slides to consolidate notes and notable figures from multiple papers into coherent narrative blocks (P4).

4.2 Results

Over the 1-week period, the 4 participants used our system for a total of 14 hours and 16 minutes (Table 2). On average, each participant spent 3 hours and 34 minutes using the system, creating 4 canvases containing 162 paper nodes, 148 citation links, and 147 sticky notes. Each canvas contained 40.5 paper nodes and 37.0 citation links from 101.5 queries using keyword-only (27.8 (27.4%)), graph-only (45.5 (44.8%)), and both (28.3 (27.9%)) search conditions, while also viewing all citation relationships 33.5 times (Table 2). Each canvas also contained 26.8 sticky notes (Table 2).

Notably, most participants (P1, P2, P3) primarily collected papers using graph-based queries, which resulted in higher numbers of automatically generated citation links (Figure 2e) than manually added ones. In contrast, P4 mainly relied on keyword-based queries, which resulted in a higher number of manually added citation links than automatically generated ones.

4.3 Post-Interview

Through the analysis of the results and follow-up interviews, we found 7 usage patterns in which researchers benefited from using our system's features for finding and organizing research papers (F1-F7). Here, we compare constructing and expanding node-link diagrams from a few key papers to planting and nurturing trees in a garden, extending the commonly used term *seed papers*.

4.3.1 ***Finding Papers.*** We found that the participants could identify 3 types of papers essential to conducting literature reviews (F1-F3), as well as employ a strategy for deferred discovery (F4).

• **Stem paper (F1).** We found that the participants could identify papers serving as the knowledge foundation for many subsequent papers, either by querying for papers commonly cited by multiple existing papers (Figure 5a) (P1, P2) or by examining citation relationships among all collected papers (Figure 4c) (P3).

• **Fruit paper (F2).** We found that the participants could identify papers that integrate knowledge from prior papers, whether close ones (self-pollination) or distant ones (cross-pollination), either by querying for papers that cite multiple existing papers (Figure 5b) (P1) or by examining citation relationships among all collected papers (Figure 4c) (P4).

• **Branch paper (F3).** We found that the participants could identify intermediate papers connecting stem papers to fruit papers, illustrating how knowledge evolved over time or how alternative approaches were applied to similar problems, either by querying for papers that cite existing papers and are cited by other existing papers (Figure 5c) (P1, P2) or by examining citation relationships among all collected papers (Figure 4c) (P3, P4).

• **Fertilizing (F4).** We found that the participants could accumulate spontaneous and abstract thoughts for future investigations by jotting down relevant keywords on sticky notes (Figure 5d) (P4) or blank sheets (P3, P4), which they used as placeholders for undeveloped directions of inquiry.

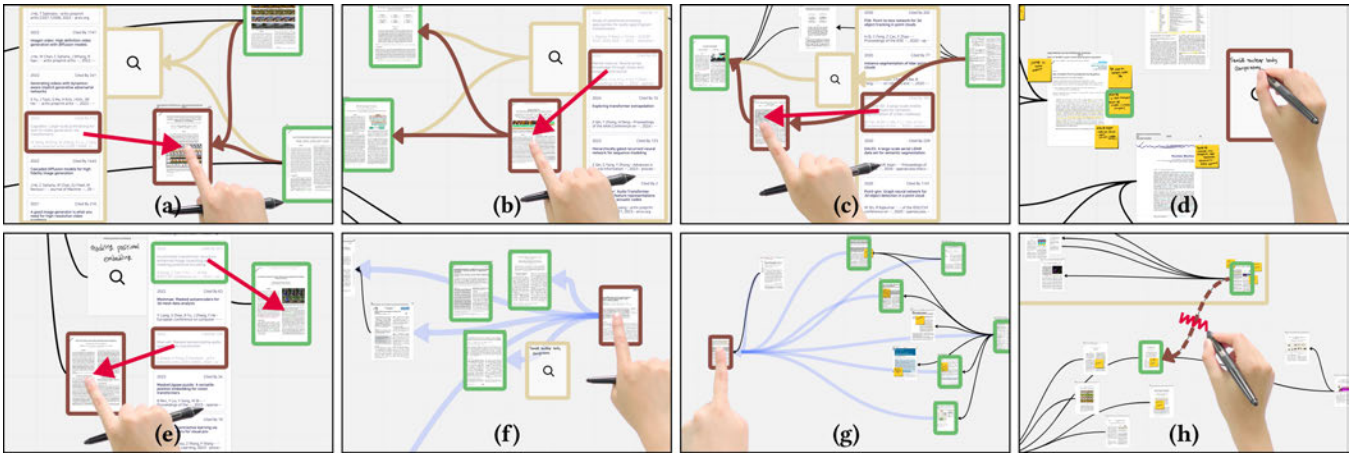

Figure 5: Reproduced snapshots of the lab study participants using our proof-of-concept system for literature reviews. (a) P2, specializing in computer vision, collected an important paper (brown) by searching for papers (beige) cited by 2 video-generating AI papers (green). (b) P1, specializing in audio and speech processing, collected a cross-pollinating paper (brown) by searching for papers (beige) citing both a speech deep learning paper (green, lower) and a vision deep learning paper (green, upper). (c) P2 collected a paper (brown) similar to his own research by searching for papers (beige) that cite a paper (green, left) on a 3D point dataset he had used and also cited by a survey paper (green, right) on 3D point networks. (d) P4, specializing in RNA engineering, referred to a handwritten sticky note (green) containing keywords about proteins associated with Alzheimer's disease, then transcribed those keywords onto a blank sheet (brown) to initiate a search for related papers. (e) P1 collected a vision-related paper (brown) on the left side and a language-related paper (green) on the right side of the search result panel. (f) P4 collected 3 papers (green) that belonged to a group and 1 paper (brown) that belonged to another group from the search results using the same sheet (beige) and examined citation relationships among them (light blue) to determine which was most relevant to her research and should be read first. (g) P3, specializing in optogenetics, newly learned about calcium ion reactions in cancer cells from a paper (brown) that belonged to a group and explored the highlighted citation links (light blue) to see whether previously collected papers (green) that belonged to another group mentioned these reactions. (h) P2 removed an unnecessary citation link (brown) to disconnect a diffusion model paper (green, upper) from another generative model paper (green, lower), and constructed a visually divided research landscape (beige) focused on diffusion models for autonomous driving.

4.3.2 ***Organizing Papers.*** We found 3 additional usage patterns (F5-F7) of the participants that reflect their strategies for semantically and functionally organizing collected papers, ultimately working toward coherent storytelling across their entire research landscapes.

• **Zoning (F5).** We found that the participants could designate semantic or functional regions on the canvas by clustering papers with similar topics (P1) (Figure 5e), separating unread from read papers (P2), and arranging papers in their preferred reading order (P1, P2, P4).

• **Grafting (F6).** We found that the participants could create new relationships between disparate zones, representing the convergence of seemingly unrelated fields and the synthesis of new insights, by examining potentially overlooked citation relationships among collected papers (Figure 5f, g) (P3, P4) or by reorganizing papers in innovative ways (P1, P2).

• **Pruning (F7).** We found that the participants could curate essential relationships between papers necessary for coherent storytelling, either by examining citation relationships among collected papers and showing necessary ones (Figure 4c) (P1, P3, P4) or hiding unnecessary ones (Figure 5h) (P1, P2, P3).

4.3.3 ***Limitations.*** All participants expressed a strong interest in using our system for their personal research projects even after their trial had ended. However, the inherent limitations of our proof-of-concept system, designed specifically for a lab study, prevented them from doing so.

• **Lack of support for active reading.** We focused on testing whether the participants could easily learn our interactions and use them to find and organize research papers. Therefore, by design, our lab study excluded the active reading of papers in our system and required the participants to study their collected papers elsewhere, on their own. However, active reading is indispensable in real-world literature reviews, as it is during this activity that papers are organized and many new ones are discovered.

• **Controlled environment and equipment.** The participants used our system in a controlled environment with provided equipment. In reality, however, researchers find, read, and organize papers alongside various research activities, such as programming, experimenting, attending seminars, and writing papers, on devices that may not support pen or multi-touch input. Moreover, since these activities often span multiple devices, cloud-based data synchronization is essential.

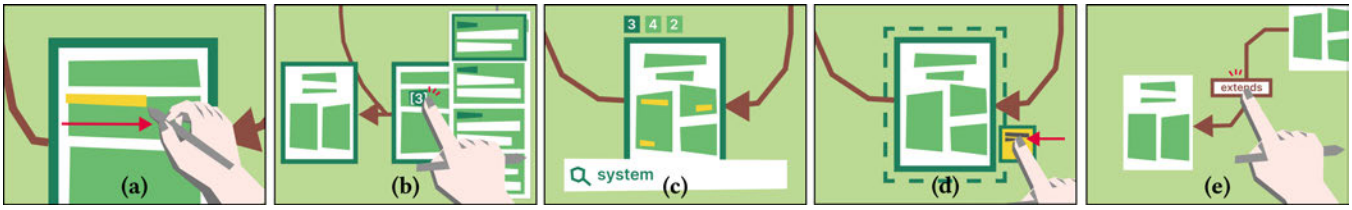

Figure 6: The user can (a) select or highlight text in a paper by dragging, and (b) find a cited paper by double-tapping the citation number within the text. (c) By pressing Ctrl + F, the user can open a search box and enter one or more keywords. The number of matches per page is displayed at the top of the paper, and tapping a result navigates to the corresponding page. (d) When the user attaches a sticky note near a paper, the note and paper become grouped and can be moved together. (e) Double-tapping a citation link allows the user to create a text label, which can then be dragged and positioned along the path of the link.

5 GARDEN OF PAPERS

Building on the potential demonstrated in the short-term lab study, we upgraded our interactive system into a fully-fledged workspace suitable for conducting literature reviews in real-world research environments, by adding features for studying the content of collected papers and also making the system cross-platform. We call this workspace *Garden of Papers*, inspired by the lab study participants' gardening-like behaviors of planting and nurturing ideas.

• **Active reading.** We implemented key features commonly found in commercial PDF viewers [1], including text selection and highlighting (Figure 6a), citation hyperlinks (Figure 6b), and text search (Figure 6c). To give researchers versatile options for capturing their thoughts, we provided features to attach sticky notes to a specific page of a paper (Figure 4b) or directly to the paper itself (Figure 6d), and to create and attach text labels to citation links (Figure 6e). Additionally, we implemented contextual menu interfaces to freely adjust font size, text alignment, note size, and note color (Figure 7).

• **Cross-platform.** In implementing cross-platform compatibility for both desktop and tablet environments, we ensured that all system interactions could be performed with mouse and keyboard inputs, in addition to pen and touch inputs (Figure 8). We also developed a cloud database system to seamlessly synchronize data across devices, enabling researchers to utilize the system flexibly in diverse research settings.

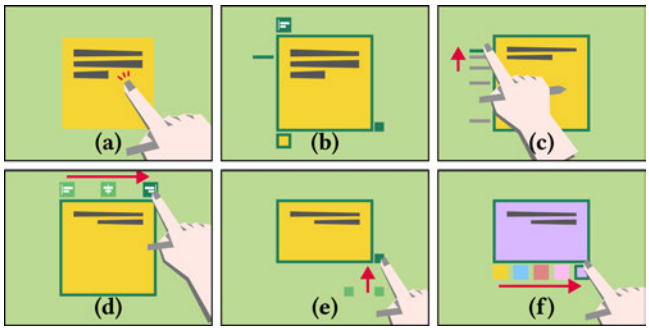

Figure 7: (a) When the user selects a sticky note, (b) icons for adjusting the note's properties appear. Dragging across (c) left icons changes the font size, (d) top icons changes the text alignment, (e) right icons changes the note size, and (f) bottom icons changes the note color.

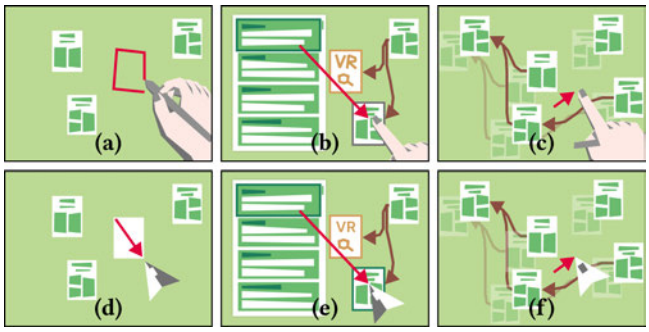

Figure 8: In our system, interactions for creation, manipulation, and navigation are performed via (a) pen and (b, c) touch inputs on the tablet, and via (d) right-click, (e) left-click, and (f) middle-click mouse inputs on the desktop, respectively.

6 LONG-TERM IN-THE-WILD STUDY

We conducted a long-term in-the-wild study in which researchers from various fields used our system on their own devices to find, read, and organize papers necessary for their personal research. The goal was to verify whether our system could help experienced researchers conduct actual literature reviews in real-world settings, beyond controlled environments.

To enable a more meaningful evaluation of the holistic user experience, we chose to examine our system in its entirety and without arbitrary restrictions, rather than individual features in isolation and out of context. This choice was in consideration of the nature of our system, which is complex and highly integrated, with many features often used together in a continuous workflow.

Moreover, we chose not to include a common baseline condition in this study. Pre-interviews during our lab study revealed that researchers use different applications for each task, with combinations varying by field and personal preference. As no single combination could reasonably represent all participants' workflows, imposing a uniform baseline would misrepresent their actual practices and thus hinder an accurate assessment of how well our system fits into and enhances them.

• **Participants.** We recruited 12 researchers (6 males and 6 females, ages 25-32) from diverse fields (Table 3). All participants had previously written and published research papers, with an average of 4.3 years of research experience and 4.4 publications as first authors or co-authors. All of them owned a desktop and a tablet.

Table 3: Participant demographics for the in-the-wild study. G: gender, A: age, Y: years of research experience, F: number of first-author papers, C: number of co-author papers.

	G	A	Y	F	C	Major	Research area
P1	F	28	6	4	0	Industrial design	HCI
P2	F	27	3	1	1	Industrial design	HCI
P3	M	30	5	3	0	Biological science	Synthetic biology
P4	F	25	5	5	0	Biological science	Synthetic biology
P5	M	28	2	1	4	Mechanical engineering	Robotics
P6	M	25	6	0	1	Mechanical engineering	Robotics
P7	M	32	5	2	1	Electrical engineering	Computer vision
P8	F	25	4	2	0	Electrical engineering	Computer vision
P9	M	28	3	1	2	Materials science	Catalysis
P10	M	28	6	3	8	Materials science	Catalysis
P11	F	28	3	4	4	Computer science	NLP
P12	F	29	5	1	3	Computer science	NLP

• **Apparatus.** We implemented our system using Unity and built multiple platform-specific clients: a desktop version with mouse and keyboard inputs, and a tablet version with pen and multi-touch inputs. We used Serp API [45], a third-party API for Google Scholar [15], and Semantic Scholar API [24], to retrieve paper metadata, including the title, authors, publication venue, PDF URL, and citation relationships for each paper. We used MongoDB Atlas [33] for a cloud database and Google Cloud Platform [14] for a cloud server in implementing cross-platform data synchronization. We asked the participants to install our system on their personal desktops and tablets.

• **Procedure.** On Day 1, we visited the participants individually at their workplaces to introduce the overall purpose and procedure of the study. We then provided 30-minute tutorial sessions, during which we demonstrated each feature of both the desktop and tablet applications step by step and guided the participants through hands-on practice until they became fully familiar with it. We also provided online manuals with video clips describing each feature, which the participants could refer to at any time.

From Day 2 to Day 29, the participants used the system freely as part of their ongoing literature reviews. They were free to choose how and when to use the system on their preferred devices, though we encouraged them to use the system instead of their existing tools and methods. Each week, we held regular 30-minute in-person sessions and responded to any questions they had.

On Day 30, we asked the participants to complete a 5-point Likert scale survey assessing their satisfaction with our system. We then conducted post-interviews in which we asked the participants to review their usage logs and node-link diagrams arranged by date, demarcate clusters on their canvases manually, and reflect on the perceived advantages and disadvantages of our system compared to their usual workflows.

• **Measurements.** We recorded the total usage time of our system, along with the number of paper nodes, citation links, and sticky notes created, deleted, and manipulated over time. In addition, we counted the number of the participant-defined clusters and the number of nodes and links within and across the clusters. Specifically, we counted stem papers that had at least two incoming citation links, fruit papers that had at least two outgoing links, and branch papers that had both at least one incoming and one outgoing link, all within the same cluster. We also recorded the frequency and duration of all invoked system features and classified them by task and device.

6.1 Results

Over the 4-week period, the 12 participants used our system for a total of 83 hours and 31 minutes, creating 21 canvases (Table 4) containing 88 clusters (Table 5), 503 paper nodes, 344 citation links, and 498 sticky notes (Table 4), by invoking system features 39,520 times (Figure 9).

On average, the participants spent 3 hours and 59 minutes on each canvas, finding papers for 28 minutes (11.9%), reading papers for 2 hours and 37 minutes (66.0%), and organizing papers for 53 minutes (22.1%) (Table 4). Each canvas contained 4.2 clusters (Table 5), 24.0 paper nodes, 16.4 citation links, and 23.7 sticky notes (Table 4). Specifically, out of the 24.0 paper nodes were 1.2 stem, 1.9 fruit, and 2.7 branch paper nodes (Table 5, left). Also, out of the 16.4 citation links were 10.1 within-cluster and 5.3 across-cluster citation links (Table 5, right).

On average, the participants switched between finding, reading, and organizing tasks 304.6 times per canvas, spending 54.8 seconds on each task (Table 6, Figure 9). Out of the 12 participants, 10 participants who used both desktop and tablet devices invoked 69.5% of features on desktop and 30.5% on tablet (Table 6).

6.2 Survey

In the 5-point Likert scale survey assessing the participants' satisfaction with our system (Figure 10), all survey items scored above the neutral point (Q1-Q15, average: 4.4/5.0). Specifically, the system's features (Q1-Q9: 4.4/5.0), usability (Q10-Q11: 3.8/5.0) and usefulness (Q12-Q15: 4.7/5.0) all received scores above neutral.

7 DISCUSSION

Based on the findings from both the short-term lab study and the long-term in-the-wild study, this section discusses how our system can benefit literature reviews, using gardening metaphors (Figure 11) to illustrate how the participants could make sense of and actively cultivate their evolving research landscapes.

We conducted a quantitative analysis of usage logs and a qualitative analysis of survey responses and follow-up interview transcripts. The analyses confirmed the previous 7 findings on usage patterns from the lab study (F1-F7), and revealed 2 additional findings on usage patterns arising from the added features for active reading (F8) and cross-platform support (F9).

7.1 Finding Papers

During the long-term in-the-wild study, the 4 previously identified usage patterns for finding papers were confirmed: identifying stem (F1), fruit (F2), and branch (F3) papers, and fertilizing (F4).

7.1.1 ***Stem paper (F1).*** On average, each canvas contained 1.2 stem papers that each had 2.7 incoming citation links (Table 5). The participants agreed that the system was helpful for identifying foundational papers (Q1: 4.5/5.0) (Figure 10).

Table 4: System usage log per canvas during the in-the-wild study.

	Time spent (hh:mm (%))				Paper nodes (#)					Citation links (#)					Sticky notes (#)					
	Finding	Reading	Organizing	Total	Collected by sheet	Collected by in-text	Deleted	Studied	Total	Generated	Added	Deleted	Labeled	Total	Created	Deleted	Attached to canvas	Attached to paper	Attached to page	Total
P1-A	00:26 (19.6)	01:06 (49.2)	00:42 (31.2)	02:14	14	78	35	2	57	81	1	27	0	55	68	7	46	7	8	61
P1-B	00:13 (2.8)	06:44 (88.0)	00:42 (9.2)	07:39	15	13	10	7	18	13	1	4	7	10	89	13	12	9	55	76
P2-A	00:30 (17.5)	01:54 (66.2)	00:28 (16.2)	02:53	55	0	37	10	18	0	4	1	1	3	45	28	3	0	14	17
P2-B	00:47 (20.0)	01:59 (50.5)	01:09 (29.5)	03:55	118	0	79	13	39	24	39	31	5	32	33	9	1	4	19	24
P3-A	00:44 (12.7)	02:46 (47.7)	02:18 (39.6)	05:48	74	0	36	13	38	0	51	12	0	39	50	15	14	0	21	35
P4-A	00:53 (17.0)	02:25 (45.8)	01:58 (37.2)	05:16	81	0	44	6	37	16	26	22	0	20	30	21	2	5	2	9
P5-A	00:45 (7.0)	09:28 (88.1)	00:32 (4.9)	10:45	31	8	26	7	13	16	14	20	0	10	16	14	0	0	2	2
P5-B	00:16 (16.3)	01:06 (66.0)	00:18 (17.7)	01:39	17	4	7	2	14	4	10	5	0	9	7	3	2	0	2	4
P5-C	00:20 (40.4)	00:05 (9.6)	00:25 (50.0)	00:50	27	0	8	0	19	1	15	10	0	6	0	0	0	0	0	0
P6-A	00:24 (8.8)	02:34 (56.2)	01:36 (35.0)	04:34	55	1	34	5	22	4	22	17	7	9	81	40	8	10	23	41
P7-A	00:33 (16.9)	01:57 (59.8)	00:46 (23.3)	03:16	66	0	43	9	23	6	14	6	0	14	48	22	6	6	14	26
P7-B	00:13 (19.5)	00:35 (50.6)	00:21 (29.9)	01:09	28	0	17	3	11	1	12	1	5	12	24	10	2	2	10	14
P7-C	00:03 (8.7)	00:23 (58.0)	00:13 (33.3)	00:40	16	0	9	1	7	0	18	3	0	15	8	3	1	3	1	5
P8-A	00:14 (6.4)	02:50 (78.8)	00:32 (14.8)	03:36	23	0	16	1	7	3	12	3	0	9	20	19	0	0	1	1
P8-B	00:32 (62.0)	00:07 (12.8)	00:13 (25.2)	00:52	16	0	9	1	7	3	8	6	0	5	4	2	2	0	0	2
P9-A	00:44 (11.3)	04:23 (67.4)	01:23 (21.3)	06:30	61	1	31	20	31	27	40	66	0	1	63	11	14	19	19	52
P10-A	00:29 (7.4)	04:33 (68.5)	01:36 (24.1)	06:39	47	0	23	13	24	0	8	3	0	5	100	64	2	12	22	36
P10-B	00:19 (7.4)	03:12 (74.0)	00:48 (18.6)	04:20	46	0	22	12	24	7	15	7	0	15	35	10	0	22	3	25
P11-A	00:25 (11.6)	02:24 (66.5)	00:47 (21.9)	03:36	40	23	41	16	22	30	7	26	0	11	24	11	4	5	4	13
P12-A	00:46 (12.9)	03:50 (65.0)	01:18 (22.1)	05:54	54	0	8	11	46	7	60	25	0	42	64	24	12	3	25	40
P12-B	00:16 (18.1)	00:47 (54.2)	00:24 (27.7)	01:27	31	0	5	7	26	21	13	12	0	22	22	7	4	1	10	15
Total	09:55 (11.9)	55:07 (66.0)	18:28 (22.1)	83:31	915	128	540	159	503	264	387	307	25	344	831	333	135	108	255	498
Avg.	00:28 (11.9)	02:37 (66.0)	00:53 (22.1)	03:59	43.6	6.1	25.7	7.6	24.0	12.6	18.4	14.6	1.2	16.4	39.6	15.9	6.4	5.1	12.1	23.7

Table 5: Graph topology and graph layout per canvas during the in-the-wild study. †: the average number of links per node, ‡: the average number of nodes or links per cluster, ‖: the number of nodes or clusters connected by across-cluster links.

	Graph topology											Graph layout							
	Stem papers (#)			Fruit papers (#)			Branch papers (#)					Clusters (#)	Within-cluster (#)				Across-cluster (#)		
	Nodes	In-links	In-l/N†	Nodes	Out-links	Out-l/N†	Nodes	In-links	In-l/N†	Out-links	Out-l/N†		Nodes	Links	N/C‡	L/C‡	Links	Nodes‖	Clusters‖
P1-A	0	0	0.0	2	55	27.5	0	0	0.0	0	0.0	5	57	30	11.4	6.0	22	24	5
P1-B	0	0	0.0	1	6	6.0	1	3	3.0	1	1.0	7	17	6	2.4	0.9	4	5	3
P2-A	0	0	0.0	1	2	2.0	0	0	0.0	0	0.0	3	16	3	5.3	1.0	0	0	0
P2-B	0	0	0.0	2	10	5.0	7	18	2.6	10	1.4	8	38	13	4.8	1.6	18	20	8
P3-A	5	18	3.6	10	25	2.5	6	18	3.0	10	1.7	8	38	17	4.8	2.1	23	18	7
P4-A	4	11	2.8	1	5	5.0	3	5	1.7	5	1.7	4	31	15	7.8	3.8	3	4	2
P5-A	2	4	2.0	2	6	3.0	4	5	1.3	4	1.0	4	11	7	2.8	1.8	3	4	3
P5-B	1	3	3.0	2	5	2.5	2	2	1.0	2	1.0	4	13	7	3.3	1.8	2	3	3
P5-C	0	0	0.0	0	0	0.0	3	5	1.7	3	1.0	3	10	4	3.3	1.3	2	3	3
P6-A	2	4	2.0	1	2	2.0	2	3	1.5	5	2.5	8	21	7	2.6	0.9	0	0	0
P7-A	1	3	3.0	3	7	2.3	4	6	1.5	6	1.5	4	23	14	5.8	3.5	0	0	0
P7-B	2	4	2.0	2	6	3.0	3	6	2.0	6	2.0	3	11	10	3.7	3.3	2	3	2
P7-C	3	10	3.3	1	5	5.0	3	5	1.7	10	3.3	2	7	5	3.5	2.5	10	7	2
P8-A	1	2	2.0	2	5	2.5	4	7	1.8	4	1.0	2	7	6	3.5	3.0	3	5	2
P8-B	1	2	2.0	1	2	2.0	0	0	0.0	0	0.0	3	6	3	2.0	1.0	1	2	2
P9-A	0	0	0.0	0	0	0.0	0	0	0.0	0	0.0	4	31	1	7.8	0.3	0	0	0
P10-A	0	0	0.0	1	3	3.0	0	0	0.0	0	0.0	3	21	3	7.0	1.0	0	0	0
P10-B	1	2	2.0	2	5	2.5	2	7	3.5	3	1.5	2	21	11	10.5	5.5	1	2	2
P11-A	0	0	0.0	3	7	2.3	0	0	0.0	0	0.0	3	21	7	7.0	2.3	3	5	2
P12-A	2	4	2.0	2	6	3.0	7	9	1.3	11	1.6	4	31	17	7.8	4.3	3	4	2
P12-B	0	0	0.0	0	0	0.0	5	12	2.4	16	3.2	4	24	11	6.0	2.8	11	13	4
Total	25	67	29.7	39	162	81.2	56	111	29.8	96	25.4	88	445	197	112.8	50.5	111	122	52
Avg.	1.2	3.2	2.7	1.9	7.7	4.2	2.7	5.3	2.0	4.6	1.7	4.2	22.5	10.1	5.4	2.4	5.3	5.8	2.5

For example, P8, who is researching robot vision, wanted to explore related work starting from a recent paper recommended to her in a lab seminar. She began by collecting the references cited in that paper, then recursively gathered the references cited in those papers to trace further back through the literature, eventually arriving at foundational papers commonly cited by many of her accumulated papers (Figure 12a). She likened this experience to simultaneous localization and mapping (SLAM), a common technique in her field, where one becomes better aware of one's surroundings as well as one's place in it, at the same time (P8-A).

P5, who is researching reinforcement learning for quadrupedal robots, sought to extend his work to robots with different leg configurations, where his previous techniques were no longer effective. Unsure of which specific keywords to search for, he instead discovered several highly cited papers on more generalizable reinforcement learning techniques through YouTube. He used these papers

Table 6: Number of task switches and average task duration, task duration ratio by device used, and device usage ratio. F: finding, R: reading, O: organizing, D: desktop, T: tablet.

	Task switching		Desktop			Tablet			Device usage	
	Count	Avg. time (s)	F (%)	R (%)	O (%)	F (%)	R (%)	O (%)	D (%)	T (%)
P1-A	265	30.4	19.6	49.2	31.2	-	-	-	-	-
P1-B	448	61.5	2.8	88.0	9.2	-	-	-	-	-
P2-A	250	41.5	21.1	65.5	13.4	0.8	69.9	29.4	82.4	17.6
P2-B	398	35.4	20.7	47.9	31.4	15.8	65.2	19.0	85.2	14.8
P3-A	345	60.5	13.2	47.8	38.9	11.6	47.5	40.9	67.5	32.5
P4-A	562	33.7	19.4	44.8	35.8	14.1	46.9	38.9	54.2	45.8
P5-A	202	191.7	7.0	88.1	4.9	-	-	-	-	-
P5-B	126	47.3	16.3	66.0	17.7	-	-	-	-	-
P5-C	55	54.0	40.4	9.6	50.0	-	-	-	-	-
P6-A	629	26.1	9.7	53.9	36.5	0.7	78.6	20.7	90.6	9.4
P7-A	209	56.3	18.1	54.2	27.7	16.5	61.8	21.7	25.5	74.5
P7-B	107	38.7	19.5	50.6	29.9	-	-	-	-	-
P7-C	44	54.1	31.4	15.6	53.0	2.3	69.9	27.8	21.9	78.1
P8-A	158	82.0	5.2	85.0	9.8	13.8	39.0	47.2	86.4	13.6
P8-B	41	76.3	62.0	12.8	25.2	-	-	-	-	-
P9-A	521	44.9	10.8	68.6	20.5	13.2	62.2	24.6	81.4	18.6
P10-A	873	27.4	8.1	70.2	21.7	5.0	62.2	32.8	78.6	21.4
P10-B	383	40.7	8.1	70.9	20.9	5.3	82.9	11.8	74.8	25.2
P11-A	325	39.8	14.2	67.6	18.2	2.2	62.5	35.3	78.2	21.8
P12-A	339	62.6	16.8	56.6	26.7	0.0	92.9	7.1	76.8	23.2
P12-B	117	44.8	18.1	54.2	27.7	-	-	-	-	-
Total	6397	1149.9	-	-	-	-	-	-	-	-
Avg.	304.6	54.8	18.2	55.6	26.2	7.8	64.7	27.5	69.5	30.5

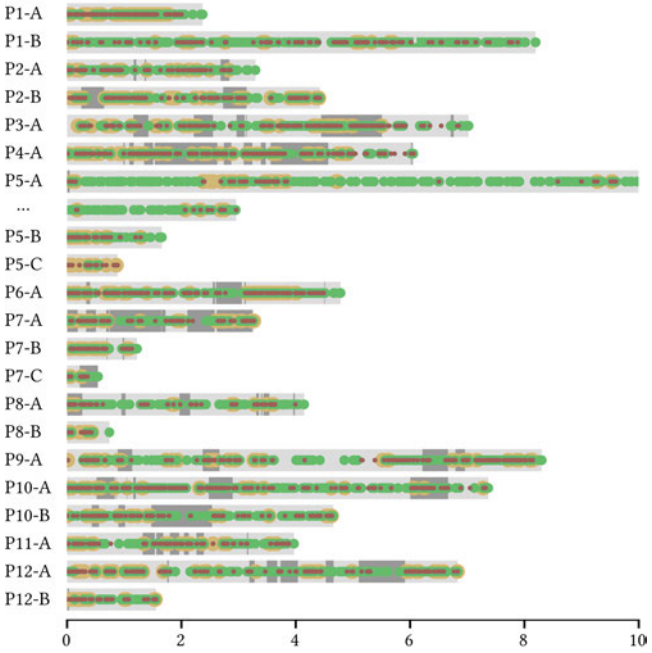

Figure 9: Task-switching patterns in each canvas over usage time in hours. Among the 12 participants, 10 participants used both desktop and tablet devices, and 2 participants used desktop devices only (P1, P5). Brown: find, green: read, beige: organize, dark grey: tablet, light grey: desktop.

as entry points to collect subsequent works that illustrated how the field evolved over time. He noted that foundational papers were especially helpful when venturing into new research areas (P5-B).

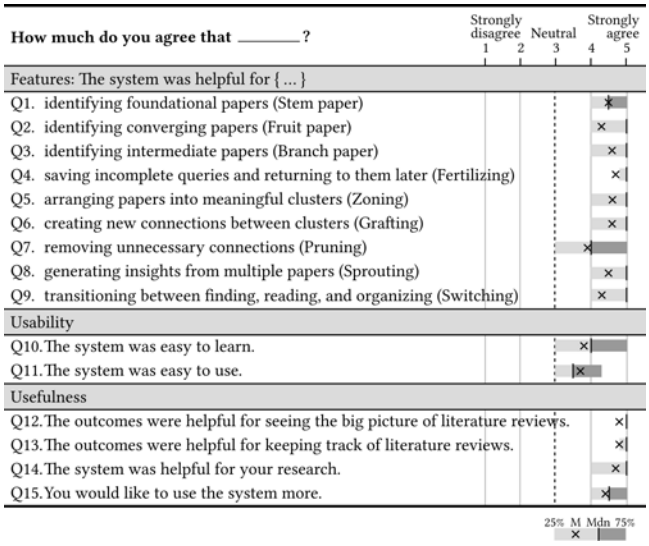

Figure 10: 5-point Likert scale survey results on our system's features, usability, and usefulness. Dashed line: neutral (higher is better).

7.1.2 ***Fruit paper (F2).*** On average, each canvas contained 1.9 fruit papers that each had 4.2 outgoing citation links (Table 5). The participants agreed that the system was helpful for identifying converging papers (Q2: 4.3/5.0) (Figure 10).

For example, the aforementioned P5 also wanted to collect papers that commonly cited both the ones he had gathered on robot manipulation and another paper he had collected on a reinforcement learning technique not specific to robots. To do so, he needed to exhaustively retrieve all citing papers of each and identify their intersections. However, he explained that "*seminal machine learning papers typically have very high citation counts,*" and that his usual workflow involved searching in a web browser and opening dozens of papers in multiple tabs, which quickly became overwhelming. "*It was practically impossible to go through every citing paper, so I wouldn't even attempt it,*" he noted, describing how difficult it was to identify such papers amid the clutter. In contrast, using our system, he used outgoing citation links to multiple papers he wanted as his search criteria. He said that "*the search results were instantly and dramatically narrowed down, giving me the confidence that I could review them one by one*" (P5-A).

P7, who is researching vision AI, searched for papers that cited recent ones he already knew, and found a highly up-to-date paper that was published just a few days earlier (Figure 12b). He noted that his field was evolving so rapidly that he had developed a habit of regularly checking blogs that introduce new papers. Notably, through our system, he discovered a brand-new paper not yet covered by those blogs, and remarked with excitement that "*I found fresh, hot-off-the-press information even before blogs*" (P7-A).

7.1.3 ***Branch paper (F3).*** On average, each canvas contained 2.7 branch papers that each had 2.0 incoming citation links and 1.7 outgoing citation links (Table 5). The participants agreed that the system was helpful for identifying intermediate papers (Q3: 4.6/5.0) (Figure 10).

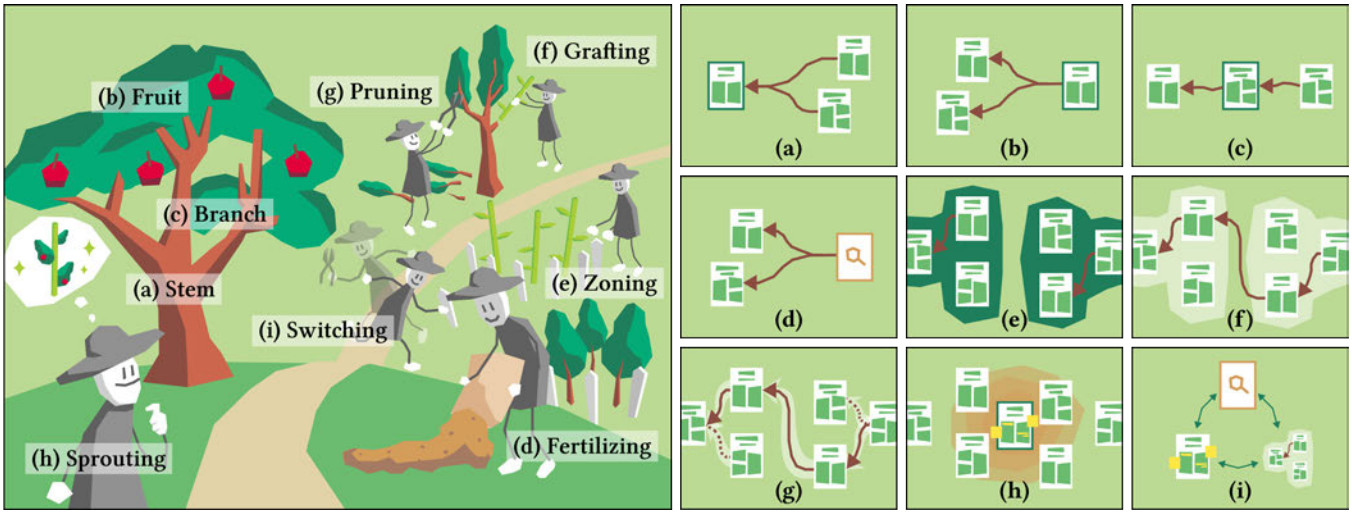

Figure 11: We confirmed 9 findings on usage patterns (F1-F9) through which researchers conduct literature reviews using our system in real-world research environments. We describe these patterns using gardening metaphors likening the process to planting and nurturing trees. (a) Stem paper (F1) provides shared knowledge foundational to different research directions. (b) Fruit paper (F2) integrates multiple strands of knowledge to achieve advanced or interdisciplinary outcomes. (c) Branch paper (F3) shows how knowledge has developed over time. (d) Fertilizing (F4) refers to spreading placeholders at relevant positions even when search ideas are still unclear, expecting future exploration. (e) Zoning (F5) refers to arranging related papers close together and unrelated papers far apart to create a meaningful spatial structure. (f) Grafting (F6) refers to connecting different groups of papers to discover new insights. (g) Pruning (F7) refers to removing unnecessary information and retaining only what is relevant. (h) Sprouting (F8) refers to reading papers alongside existing ones to generate fresh ideas. (i) Switching (F9) refers to transitioning between finding, reading, and organizing papers as needed, without interruption.

For example, the aforementioned P7 also began constructing a diagram out of papers he had read in the past, in order to identify papers that can serve as performance baselines against which to compare his own work. He was specifically looking for papers that were grounded in foundational work and had sufficiently influenced subsequent work. His usual workflow would have involved printing out papers individually, reading them carefully, and storing them sequentially in a physical binder. In contrast, our system allowed him to visually grasp, at a glance, each paper's position within academic lineage, making it easier for him to spot well-positioned papers that met his needs (P7-A, P7-B, P7-C).

Note that this pattern was particularly common among AI researchers (P5, P7, P8), who each created an average of 2.3 canvases with 7.7 branch papers, over 1.5 times more canvases and 3.1 times more branch papers than the others (Table 5). They used these canvases to identify baselines and datasets commonly used by papers tackling similar problems to theirs but with different approaches.

7.1.4 ***Fertilizing (F4).*** On average, the participants collected 43.6 paper nodes originating from blank sheets where they specified keywords or citation links as search conditions, per canvas. Also, they attached 6.4 sticky notes directly to canvas for recording investigation ideas (Table 4). The participants agreed that the system was helpful for saving incomplete queries and returning to them later (Q4: 4.7/5.0) (Figure 10).

For example, the aforementioned P7 also noticed that recent papers positioned at the far right edge of the canvas sparked his curiosity as he thought to himself, "*there must have been more recent advancements in this area.*" He acted on this impulse by creating a blank sheet and adding outgoing citation links from it to the recent papers, and kept it as a placeholder for his curiosity about the latest research. He periodically returned to it and refreshed the search results throughout his investigation (P7-A).

P1 and P10 both reported frequently coming up with ideas for follow-up investigations while reading papers. In their usual workflows using conventional PDF viewers, they would open a paper at a time and jot down their ideas within the document using annotation tools. However, P1 lamented that "*after some time passes, I just remember that I wrote something down, but often forget which paper I wrote it in,*" since annotations remain out of sight unless the exact PDF file is reopened. In contrast, they found our system more effective, as it allows them to capture higher-level ideas by placing sticky notes directly on the canvas instead of tying them to a specific page of a paper. P10 noted that "*I could always revisit my notes and eventually act on the ideas*" (Figure 12c) (P1-B, P10-B).

7.2 Organizing Papers

During the long-term in-the-wild study, the 3 previously identified usage patterns for organizing papers were confirmed: zoning (F5), grafting (F6), and pruning (F7).

7.2.1 ***Zoning (F5).*** On average, each canvas contained 4.2 user-defined clusters, with each cluster containing 5.4 paper nodes and 2.4 citation links (Table 5). The participants agreed that the system was helpful for arranging papers into meaningful clusters (Q5: 4.6/5.0) (Figure 10).

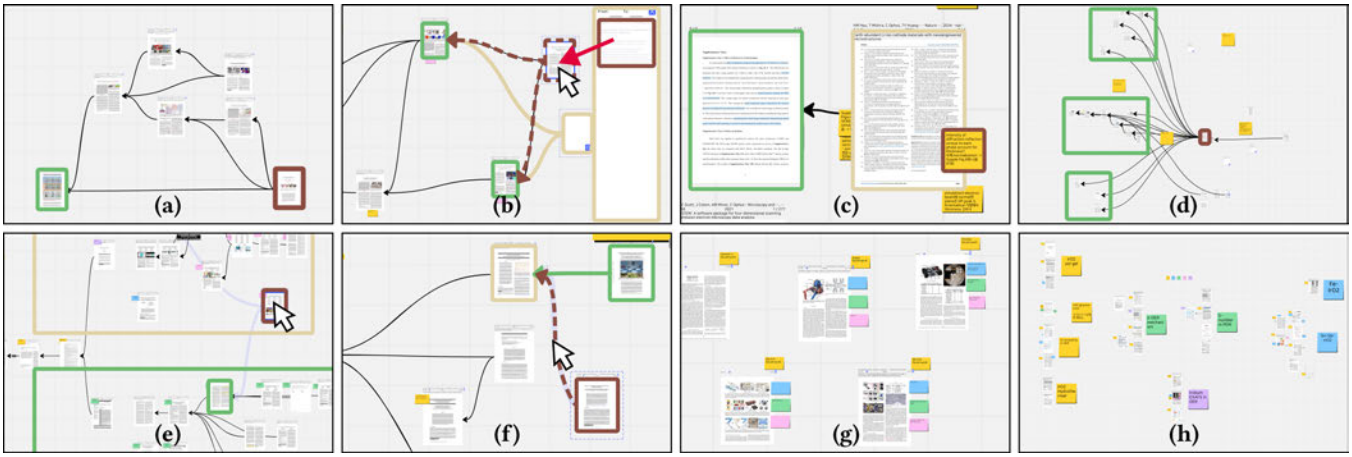

Figure 12: Reproduced snapshots of the in-the-wild study participants using our system for literature reviews. (a) P8, specializing in computer vision, mapped out a research landscape on LiDAR, starting from a recent paper (brown) recommended in a seminar, and identified a foundational paper (green) commonly cited within the citation graph. (b) P7, specializing in computer vision, collected a recent state-of-the-art paper (brown) by searching for papers (beige) citing 2 recent Gaussian Splatting papers (green). (c) P10, specializing in catalysis, read a paper (beige) on Li-ion cathode materials, noticed a referenced supplementary material, left a sticky note (brown) reminding himself to review it, and then collected the paper (green) later, to the left of the paper. (d) P12, specializing in NLP, quickly formed 3 distinct clusters of relevant papers (green) just as she collected them starting from a paper (brown) on measuring social bias in LLMs. (e) P2, specializing in HCI, examined all citation relationships between a paper (brown) and the other collected papers and discovered that it addressed both financial literacy (green) and AI services (beige) fields. (f) P5, specializing in robotics, collected a paper (green) on quadrupedal robot control and a paper (brown) on bipedal robot control citing the same paper (beige), but intentionally removed an automatically generated citation link (brown) to emphasize the relationship (green) most relevant to his research. (g) P6, specializing in robotics, attached colored sticky notes next to robot-design papers, comparing PC modules (blue), IMU sensors (green), and motors (pink). (h) P9, specializing in catalysis, visually organized papers by placing material properties of Iridium in the center, catalytic reactions with other materials to the right, experimental methods to the left, and experimental equipment at the bottom.

For example, P10 shared that he usually relied on tagging to organize papers into groups. However, he found the tagging process cognitively burdensome, noting that "*I added tags to categorize papers, but I gave them different names every time, so in the end, they just caused more confusion.*" In contrast, he appreciated how our system allowed papers to be grouped based on spatial arrangement, noting that "*I could loosely group them first and then refine the organization later.*" This usage pattern aligns with previous work [29, 47], which emphasizes supporting implicit spatial relationships to reduce users' cognitive load (P10-A, P10-B).

P12 also found spatial organization helpful when selecting which papers to read. She explained that she usually dumped all unread papers into a single "*unread*" folder, where only the paper titles were visible. Because the titles alone provided little context, it was difficult for her to distinguish between papers or decide which papers to read first. She noted, "*since I hadn't read the papers yet, I couldn't decide which folder they belonged to, other than the unread folder.*" In contrast, with our system, she was able to immediately judge newly collected papers from their appearances and spatially arrange them by topic. "*As I worked, they naturally became grouped by theme,*" she noted with satisfaction (P12-A, P12-B) (Figure 12d).

7.2.2 ***Grafting (F6).*** On average, each canvas contained 5.3 citation links that crossed boundaries of user-defined clusters, connecting 5.8 paper nodes across 2.5 clusters (Table 5). The participants agreed that the system was helpful for creating new connections between clusters (Q6: 4.6/5.0) (Figure 10).

For example, P3, who is researching AI-based protein design, complained that his Markdown text editor with a graph view for interlinked notes [35] automatically rearranged the graph layout each time a new relationship was added, often causing large shifts in the entire diagram and making it difficult for him to maintain a meaningful spatial organization. In contrast, our system allowed him to deliberately arrange design case papers and AI model papers in separate areas and examine citation links between them. He noted that this allowed him to clearly see which AI methods were applied to which design cases (P3-A).

P2, who is researching AI services for financial literacy, previously attempted to organize papers using a free-form diagramming tool [32]. However, she found the process of manually gathering metadata and constructing the diagrams herself too tedious, and eventually gave up. In contrast, with our system, all collected papers and citations were automatically loaded as interactive objects. Notably, she could examine every possible citation link between any collected papers with no effort required on her part (Figure 12e). This allowed her to quickly trace connections between conceptual foundations and practical applications, and she enthusiastically shared, "*I found the exact paper I was looking for*" (P2-B).

7.2.3 ***Pruning (F7).*** On average, each canvas contained 16.4 citation links, accounting for 52.9% of the total 31.0 cumulatively created citation links, including 12.6 that were automatically generated and 18.4 that were manually added (Table 4). The participants agreed that the system was helpful for removing unnecessary connections (Q7: 3.9/5.0) (Figure 10).

For example, P5 used our system to find robot manipulation papers that employed a reinforcement learning technique proposed in a well-known foundational paper. These papers were collected using graph-based queries, and the system automatically displayed citation links from them to the foundational paper. However, because his research focused specifically on quadrupedal robots, he intentionally removed the citation link to a paper on bipedal robots and put it aside "*just in case it's needed*" (P5-C) (Figure 12f).

Meanwhile, P1, P2, P7, and P8 preferred to actively keep a larger portion of all possible citation links visible between the collected papers, keeping 65.4% of the cumulatively created links on average, which is approximately 1.2 times more than the proportion retained by the others. They explained that their goal was to understand the broad academic landscape, and showing all citation links helped them gain that overview.

7.3 Reading Papers

During the long-term in-the-wild study, 2 additional usage patterns for reading papers with our system emerged: sprouting (F8) and switching (F9).

7.3.1 ***Sprouting (F8).*** On average, each canvas contained 7.6 paper nodes that had either sticky notes attached or text highlighted. 5.1 sticky notes were attached around paper nodes, and 12.1 sticky notes were attached to specific pages (Table 4). Additionally, each canvas contained 1.2 citation links that had text labels attached (Table 4). The participants agreed that the system was helpful for generating insights from multiple papers (Q8: 4.5/5.0) (Figure 10).

For example, P6, who is researching legged robots, collected 5 papers presenting widely known robot design methods. He used sticky notes in different colors to write down what kind of PC modules, IMU sensors, and motors were used in each paper (Figure 12g). By viewing multiple papers side by side and comparing their features, he was able to choose and purchase the specific components needed for his ongoing robot design project (P6-A).

P1 used to annotate new ideas directly on PDFs and store them in folders as part of her reading workflow. However, when she had ideas that concerned multiple papers simultaneously, she felt hesitant to write it down on any single paper. In contrast, with our system, she appreciated being able to lay out multiple papers side by side within a single workspace, read and compare them in parallel, scribble ideas on sticky notes, and place them wherever felt most appropriate: on specific pages, on entire papers, or anywhere between them on the 2D canvas (P1-B).

7.3.2 ***Switching (F9).*** On average, the participants spent 11.9%, 66.0%, and 22.1% of their time on finding, reading, and organizing papers, respectively (Table 4). During this process, they switched tasks an average of 304.6 times, spending an average of 54.8 seconds per task (Table 6, Figure 9). The participants who used both devices spent 69.5% of time on desktops and 30.5% of time on tablets (Table 6). On desktops, the participants spent 10.4% more time on finding, and on tablets, they spent 9.1% more time on reading (Table 6). The participants agreed that the system was helpful for transitioning between finding, reading, and organizing (Q9: 4.3/5.0) (Figure 10).

For example, P2, P5, and P12 found the ability to seamlessly switch between finding and organizing papers especially important at the beginning stages of new research projects. In their usual workflows, the web browsers they used to find papers quickly filled up with many tabs of papers and became disorganized. In contrast, our system provided a big enough workspace where they could lay out all the papers and transition back and forth between finding and organizing tasks as needed. While finding papers, they often gained insights that helped them better organize those papers, and while organizing papers, they often noticed gaps that led them to find further papers. Prior work found that reducing friction in such iterative cycles of information collection and reorganization can help manage situations where new information accumulates rapidly [38, 57] (P2-B, P5-B, P12-B).

P4 shared that she had purchased a tablet device at the beginning of graduate school. However, she rarely ended up using it, as she could not find any research tasks where the device made a meaningful difference for her. In contrast, with our system, the tablet form factor allowed her to continue literature review processes seamlessly even while moving between locations. She especially appreciated the ability to quickly zoom in and out using pinch gestures, which made it easy to see both the overall research landscape and specific handwritten notes on individual papers (P4-A).

7.4 Gardening Papers

During the long-term in-the-wild study, the participants leveraged the 9 confirmed usage patterns (F1–F9) to generate outcomes that helped them see the big picture of literature reviews (Q12: 4.8/5.0) and keep track of literature reviews (Q13: 4.8/5.0) (Figure 10). As a result, they found the system helpful for their research (Q14: 4.7/5.0) and expressed willingness to continue using it even after the study ended (Q15: 4.4/5.0) (Figure 10).

• **Visual thinking.** In their usual workflows, all participants managed their papers solely through folders or tables, which made it difficult to understand complex relationships between papers and construct a coherent narrative. As a result, they had to take extra steps to summarize the papers individually and then elaborate their connections on separate documents (P1, P7, P11, P12) or slides (P4, P9, P10). In contrast, P9 noted that our system helped him not only see each paper's place within the broader research landscape, but also how it contributed to that landscape (Figure 12h). P2 noted that, for her, the system transformed literature reviews into a process of visually shaping the "*mind map*" of research papers.

• **Spatial memory.** The participants noted that visual features of their canvases, such as visual appearances of papers and their annotations (P1-P12), spatial placement of paper nodes (P2, P3, P4, P9), and topological fingerprints of citation links (P2, P5, P7, P8), helped them keep track of their literature reviews. For example, P3, who often brought his tablet to meetings with his advisor, shared an instance in which a specific paper suddenly came to mind during a discussion. Although he could not recall the paper's title, he could quickly launch our system on his tablet and locate the paper based on where he had previously placed it on the canvas.

8 LIMITATIONS & FUTURE WORK

In this section, we discuss the limitations identified in our short-term lab study and long-term in-the-wild study, and suggest future research directions.

• **Support for broader sources.** Our system was designed primarily to help researchers query academic databases such as Google Scholar [15]. However, researchers often come across new and interesting papers through curated sources such as publisher websites, popular blogs, and social media platforms, which often provide timely highlights, contextual summaries, or expert commentary. Therefore, future work should accommodate the ways researchers discover papers across diverse and fragmented channels.

• **Support for intelligent assistance.** Our system provided basic annotation features comparable to those of commercial PDF viewers. However, there has been a growing trend in academia to actively leverage large language models (LLMs) for translating, summarizing, inquiring, and ideating, although such interactions are often limited to a single paper at a time. Therefore, future work should incorporate LLMs not only at the level of individual papers, but also at the level of graph topologies and clusters, enabling intelligent assistance grounded in a global research context.

• **Support for teams and organizations.** Our 4-week deployment provided insights into long-term use in a real-world setting. Nevertheless, researchers often spend several months or years managing multiple projects in collaboration with other researchers. As related papers accumulate, researchers may seek to consolidate them into a unified and evolving body of knowledge. Therefore, future work should explore supporting long-term, large-scale collaborative literature reviews for teams and organizations.

9 CONCLUSION

In this study, we present *Garden of Papers*, a visual, integrated, and flexible workspace for authoring a node-link diagram of research papers through a harmonious set of sketch-based interactions. In our system, users can search for papers through in situ graph-based queries, reducing the anxiety of missing important papers in the process of finding them; utilize the infinite 2D canvas as a workspace for browsing and note-taking, reducing feelings of being overwhelmed by a rapidly growing pile of papers to read; and progressively refine their knowledge graph of related work, reducing the frustration of wrestling with complex relationships among numerous papers to organize them into a coherent narrative.

We first developed a proof-of-concept system and conducted a 1-week lab study with 4 experienced researchers. Through this study, we identified both the system's potential and its limitations. Based on these insights, we upgraded the system and conducted a 4-week in-the-wild study with 12 experienced researchers from diverse fields. They used the system as part of their ongoing literature reviews for a total of 83 hours and 31 minutes, creating 21 canvases containing 503 paper nodes, 344 citation links, and 498 sticky notes.

From these studies, we quantitatively and qualitatively found that our system facilitates 9 usage patterns that enhance literature review practices. Researchers identified papers with important topological characteristics within citation graphs, such as foundational papers (F1: stem paper), converging papers (F2: fruit paper), and intermediate papers (F3: branch paper), and left placeholders for future searches (F4: fertilizing). Researchers also created meaningful spatial arrangements by clustering related papers together (F5: zoning), drawing new connections between disparate clusters (F6: grafting), and retaining selective information relevant to their research narrative (F7: pruning). While engaging multiple papers simultaneously, researchers captured spontaneous ideas whenever and wherever they emerged (F8: sprouting), and seamlessly transitioned between finding, reading, and organizing tasks as needed (F9: switching).

We describe these patterns using gardening metaphors, as they resemble the act of tending a garden, involving continuous cultivation, selective pruning, and thoughtful shaping over time, suggesting a conceptual framework for future literature review tools. We believe that such tools grounded in visual thinking and spatial memory will play a vital role in helping researchers not only make sense of, but also shape their evolving research landscapes.

Acknowledgments

This research was supported by the DRB-KAIST SketchTheFuture Research Center and the KAIST Convergence Research Institute Operation Program. We thank Jin Park, Hyelim Hwang, and Jeongche Yoon for their help with the figures and Chaeyeon Ahn for her help with the usage log analysis.

References

[1] Adobe. 2025. *Adobe Acrobat Reader.* Retrieved Jul 23, 2025 from https://get.adobe.com/reader
[2] Anand Agarawala and Ravin Balakrishnan. 2006. Keepin' it real: pushing the desktop metaphor with physics, piles and the pen. In *Proc. CHI '06.* 1283–1292. doi:10.1145/1124772.1124965
[3] American Association for the Advancement of Science. 2025. *Science.* Retrieved Jul 23, 2025 from https://www.science.org
[4] Joseph Chee Chang, Amy X. Zhang, Jonathan Bragg, Andrew Head, Kyle Lo, Doug Downey, and Daniel S. Weld. 2023. Citesee: augmenting citations in scientific papers with persistent and personalized historical context. In *Proc. CHI '23.* Article 737, 15 pages. doi:10.1145/3544548.3580847
[5] Duen Horng Chau, Aniket Kittur, Jason I. Hong, and Christos Faloutsos. 2011. Apolo: making sense of large network data by combining rich user interaction and machine learning. In *Proc. CHI '11.* 167–176. doi:10.1145/1978942.1978967
[6] Kiroong Choe, Seokweon Jung, Seokhyeon Park, Hwajung Hong, and Jinwook Seo. 2021. Papers101: supporting the discovery process in the literature review workflow for novice researchers. In *Proc. PacificVis '21.* 176–180. doi:10.1109/PacificVis52677.2021.00037
[7] Clarivate. 2025. *EndNote.* Retrieved Jul 23, 2025 from https://endnote.com
[8] Cold Spring Harbor Laboratory. 2025. *bioRxiv.* Retrieved Jul 23, 2025 from https://www.biorxiv.org
[9] Connected Papers. 2025. *Connected Papers.* Retrieved Jul 23, 2025 from https://www.connectedpapers.com
[10] Cornell University. 2025. *arXiv.* Retrieved Jul 23, 2025 from https://arxiv.org
[11] Niklas Elmqvist and Philippas Tsigas. 2007. CiteWiz: a tool for the visualization of scientific citation networks. *Inf. Vis.* 6, 3 (2007), 215–232. doi:10.1057/palgrave.ivs.9500156
[12] Elsevier. 2025. *Mendeley.* Retrieved Jul 23, 2025 from https://www.mendeley.com
[13] Allen Foster. 2004. A non-linear model of information seeking behaviour. *J. Am. Soc. Inf. Sci. Technol.* 55, 3 (2004), 228–237.
[14] Google. 2025. *Google Cloud.* Retrieved Jul 23, 2025 from https://cloud.google.com
[15] Google. 2025. *Google Scholar.* Retrieved Jul 23, 2025 from https://scholar.google.com
[16] Google. 2025. *Google Sheets.* Retrieved Jul 23, 2025 from https://docs.google.com/spreadsheets
[17] Google. 2025. *Google Slides.* Retrieved Jul 23, 2025 from https://docs.google.com/presentation
[18] Florian Haag, Steffen Lohmann, Stephan Siek, and Thomas Ertl. 2015. QueryVOWL: visual composition of SPARQL queries. *In The Semantic Web: ESWC 2015 Satellite Events* (2015), 62–66. doi:10.1007/978-3-319-25639-9_12
[19] Florian Heimerl, Qi Han, Steffen Koch, and Thomas Ertl. 2016. CiteRivers: visual analytics of citation patterns. *IEEE TVCG* 22, 1 (2016), 190–199. doi:10.1109/TVCG.2015.2467621

[20] Ken Hinckley, Koji Yatani, Michel Pahud, Nicole Coddington, Jenny Rodenhouse, Andy Wilson, Hrvoje Benko, and Bill Buxton. 2010. Pen + touch = new tools. In *Proc. UIST '10*. 27–36. doi:10.1145/1866029.1866036

[21] Hyunmo Kang, Catherine Plaisant, Bongshin Lee, and Benjamin B. Bederson. 2006. NetLens: iterative exploration of content-actor network data. In *Proc. VAST '06*. 91–98. doi:10.1109/VAST.2006.261426

[22] Hyeonsu B. Kang, Joseph Chee Chang, Yongsung Kim, and Aniket Kittur. 2022. Threddy: an interactive system for personalized thread-based exploration and organization of scientific literature. In *Proc. UIST '22*. Article 94, 15 pages. doi:10.1145/3526113.3545660

[23] Hyeonsu B. Kang, Tongshuang Wu, Joseph Chee Chang, and Aniket Kittur. 2023. Synergi: a mixed-initiative system for scholarly synthesis and sensemaking. In *Proc. UIST '23*. Article 43, 19 pages. doi:10.1145/3586183.3606759

[24] Rodney Michael Kinney, Chloe Anastasiades, Russell Authur, Iz Beltagy, Jonathan Bragg, Alexandra Buraczynski, Isabel Cachola, Stefan Candra, Yoganand Chandrasekhar, Arman Cohan, Miles Crawford, Doug Downey, Jason Dunkelberger, Oren Etzioni, Rob Evans, Sergey Feldman, Joseph Gorney, David W. Graham, F.Q. Hu, Regan Huff, Daniel King, Sebastian Kohlmeier, Bailey Kuehl, Michael Langan, Daniel Lin, Haokun Liu, Kyle Lo, Jaron Lochner, Kelsey MacMillan, Tyler Murray, Christopher Newell, Smita R Rao, Shaurya Rohatgi, Paul Sayre, Zejiang Shen, Amanpreet Singh, Luca Soldaini, Shivashankar Subramanian, A. Tanaka, Alex D Wade, Linda M. Wagner, Lucy Lu Wang, Christopher Wilhelm, Caroline Wu, Jiangjiang Yang, Angele Zamarron, Madeleine van Zuylen, and Daniel S. Weld. 2023. The semantic scholar open data platform. arXiv:2301.10140 [cs.DL]

[25] Jill H. Larkin and Herbert A. Simon. 1987. Why a diagram is (sometimes) worth ten thousand words. *Cogn. Sci.* 11, 1 (1987), 65–100. doi:10.1111/j.1551-6708.1987.tb00863.x

[26] Pengkai Liu, Xin Wang, Qiang Fu, Yajun Yang, Yuan-Fang Li, and Qingpeng Zhang. 2022. KGVQL: a knowledge graph visual query language with bidirectional transformations. *Knowl.-Based Syst.* 250, 108870 (2022). doi:10.1016/j.knosys.2022.108870

[27] Donghyeok Ma, Joon Hyub Lee, and Seok-Hyung Bae. 2024. Understanding visual, integrated, and flexible workspace for comprehensive literature reviews with SketchingRelatedWork. In *Proc. CHI EA '24*. Article 348, 7 pages. doi:10.1145/3613905.3650791

[28] Donghyeok Ma, Joon Hyub Lee, Junwoo Yoon, Taegyu Jin, and Seok-Hyung Bae. 2023. SketchingRelatedWork: finding and organizing papers through inking a node-link diagram. In *Proc. UIST '23 Adjunct*. Article 20, 3 pages. doi:10.1145/3586182.3616685

[29] Catherine C. Marshall and Russell A. Rogers. 1992. Two years before the mist: experiences with Aquanet. In *Proc. ECHT '92*. 53–62. doi:10.1145/168466.168490

[30] Justin Matejka, Tovi Grossman, and George Fitzmaurice. 2012. Citeology: visualizing paper genealogy. In *Proc. CHI EA '12*. 181–190. doi:10.1145/2212776.2212796

[31] Justin Matejka, Tovi Grossman, and George Fitzmaurice. 2021. Paper forager: supporting the rapid exploration of research document collections. In *Proc. GI '21*. 237–245. doi:10.20380/GI2021.27

[32] Miro. 2025. *Miro*. Retrieved Jul 23, 2025 from https://miro.com

[33] MongoDB. 2025. *MongoDB Atlas*. Retrieved Jul 23, 2025 from https://www.mongodb.com/products/platform/atlas-database

[34] Notion Labs. 2025. *Notion*. Retrieved Jul 23, 2025 from https://www.notion.so

[35] Obsidian. 2025. *Obsidian*. Retrieved Jul 23, 2025 from https://obsidian.md

[36] Srishti Palani, Aakanksha Naik, Doug Downey, Amy X. Zhang, Jonathan Bragg, and Joseph Chee Chang. 2023. Relatedly: scaffolding literature reviews with existing related work sections. In *Proc. CHI '23*. Article 742, 20 pages. doi:10.1145/3544548.3580841

[37] Robert Pienta, Acar Tamersoy, Alex Endert, Shamkant Navathe, Hanghang Tong, and Duen Horng Chau. 2016. VISAGE: interactive visual graph querying. In *Proc. AVI '16*. 272–279. doi:10.1145/2909132.2909246

[38] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *In Proc. Int. Conf. Intell. Anal.* 5, 1 (2005), 2–4.

[39] Antoine Ponsard, Francisco Escalona, and Tamara Munzner. 2016. PaperQuest: a visualization tool to support literature review. In *Proc. CHI EA '16*. 2264–2271. doi:10.1145/2851581.2892334

[40] Napol Rachatasumrit, Jonathan Bragg, Amy X. Zhang, and Daniel S. Weld. 2022. Citeread: integrating localized citation contexts into scientific paper reading. In *Proc. IUI '22*. 707–719. doi:10.1145/3490099.3511162

[41] ResearchRabbit. 2025. *ResearchRabbit*. Retrieved Jul 23, 2025 from https://www.researchrabbit.ai

[42] George Robertson, Mary Czerwinski, Kevin Larson, Daniel C. Robbins, David Thiel, and Maarten van Dantzich. 1998. Data Mountain: using spatial memory for document management. In *Proc. UIST '98*. 153–162. doi:10.1145/288392.288596

[43] scite. 2025. *scite*. Retrieved Jul 23, 2025 from https://scite.ai

[44] Selvas AI. 2025. *Selvy Pen SDK*. Retrieved Jul 23, 2025 from https://handwriting.selvasai.com

[45] SerpApi. 2025. *SerpApi*. Retrieved Jul 23, 2025 from https://serpapi.com

[46] Zeqian Shen, Michael Ogawa, Soon Tee Teoh, and Kwan-Liu Ma. 2006. BiblioViz: a system for visualizing bibliography information. In *Proc. APVIS '06*. 93–102.

[47] Frank M. Shipman and Catherine C. Marshall. 1999. Spatial hypertext: an alternative to navigational and semantic links. *ACM CSUR* 31, 4es (1999), 14–es. doi:10.1145/345966.346001

[48] Springer Nature. 2025. *Nature*. Retrieved Jul 23, 2025 from https://www.nature.com

[49] John Stasko, Carsten Görg, and Zhicheng Liu. 2008. Jigsaw: supporting investigative analysis through interactive visualization. *Inf. Vis.* 7, 2 (2008), 118–132. doi:10.1057/palgrave.ivs.9500180

[50] Hariharan Subramonyam, Colleen Seifert, Priti Shah, and Eytan Adar. 2020. texSketch: active diagramming through pen-and-ink annotations. In *Proc. CHI '20*. 1–13. doi:10.1145/3313831.3376155

[51] Nicole Sultanum, Christine Murad, and Daniel Wigdor. 2020. Understanding and supporting academic literature review workflows with litSense. In *Proc. AVI '20*. doi:10.1145/3399715.3399830

[52] Craig S. Tashman and W. Keith Edwards. 2011a. Active reading and its discontents: the situations, problems and ideas of readers. In *Proc. CHI '11*. 2927–2936. doi:10.1145/1978942.1979376

[53] Craig S. Tashman and W. Keith Edwards. 2011b. LiquidText: a flexible, multitouch environment to support active reading. In *Proc. CHI '11*. 3285–3294. doi:10.1145/1978942.1979430

[54] William Wright, David Schroh, Pascale Proulx, Alex Skaburskis, and Brian Cort. 2006. The sandbox for analysis: concepts and methods. In *Proc. CHI '06*. 801–810. doi:10.1145/1124772.1124890

[55] Haijun Xia, Ken Hinckley, Michel Pahud, Xiao Tu, and Bill Buxton. 2017. WritLarge: ink unleashed by unified scope, action, & zoom. In *Proc. CHI '17*. 3227–3240. doi:10.1145/3025453.3025664

[56] Vahan Yoghourdjian, Yalong Yang, Tim Dwyer, Lee Lawrence, Michael Wybrow, and Kim Marriott. 2021. Scalability of network visualisation from a cognitive load perspective. *IEEE TVCG* 27, 2 (2021), 1677–1687. doi:10.1109/TVCG.2020.3030459

[57] Pengyi Zhang and Dagobert Soergel. 2014. Towards a comprehensive model of the cognitive process and mechanisms of individual sensemaking. *J. Assoc. Inf. Sci. Technol.* 65, 9 (2014), 1733–1756. doi:10.1002/asi.23125

[58] Xiaolong Zhang, Yan Qu, C. Lee Giles, and Piyou Song. 2008. CiteSense: supporting sensemaking of research literature. In *Proc. CHI '08*. 677–680. doi:10.1145/1357054.1357161