

Uncharted Territory: Understanding Exploratory Search Behaviours in Literature Reviews

AYAH SOUFAN, University of Strathclyde, UK

IAN RUTHVEN, University of Strathclyde, UK

LEIF AZZOPARDI, University of Strathclyde, UK

In the realm of *Information Seeking and Retrieval* (ISR), searching the literature for relevant references in the context of academic work, such as theses or publications, is widely recognised as an exploratory search task. This task becomes particularly challenging when searchers lack prior knowledge of the subject matter. To help address the growing need for supporting exploratory search endeavours, ISR researchers have developed exploratory search models and interfaces. However, while much attention has been given to conceptualising exploratory searches, little focus has been placed on understanding the specific approaches and behaviours that searchers employ during conducting literature reviews. This paper aims to bridge this gap by conducting semi-structured interviews with 30 Master's students at the end of a user study. This paper provides comprehensive definitions for the fundamental exploratory characteristics from the recent conceptual model, pinpoints potential factors that could influence these characteristics, introduces new exploratory dimensions, and extends our comprehension of existing ones in the academic context. It also uncovers a spectrum of approaches used in literature searches, shedding light on how individuals rely on specific paper sections to measure their relevance and highlighting essential facets of knowledge acquisition in the context of literature searches.

CCS Concepts: • **Human-centered computing** → **Interaction design theory, concepts and paradigms**.

Additional Key Words and Phrases: Exploratory Search; Information Seeking; User Study; Literature Search; Academic Search

ACM Reference Format:

Ayah Soufan, Ian Ruthven, and Leif Azzopardi. 2024. Uncharted Territory: Understanding Exploratory Search Behaviours in Literature Reviews. In *Proceedings of the 2024 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '24)*, March 10–14, 2024, Sheffield, United Kingdom. ACM, New York, NY, USA, 21 pages. <https://doi.org/10.1145/3627508.3638334>

1 INTRODUCTION

Exploring the literature to identify relevant references for reports, theses, or publications is commonly characterised as an exploratory search task [16, 18, 25, 35, 36]. Engaging in a literature review within a new domain entails the utilisation of diverse search tools, techniques, and methods to locate relevant information. This process typically unfolds over time and follows an ad-hoc approach, with searchers tentatively delving into the domain to acquire knowledge about their chosen topic [16, 18, 35, 36].

Researchers have proposed various ways to characterise exploratory search tasks, attempting to define their core dimensions and characteristics. According to White et al. [42], exploratory search typically occurs within information seeking contexts that are open-ended, persistent, multifaceted, involving opportunistic, iterative, and multi-tactical information seeking processes. For White et al. [41], exploratory search is adopted by individuals interested in a specific area but lacking the requisite expertise to formulate precise queries. In contrast, in some segments of the literature,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

literature review tasks are often perceived as more formal, well-defined, and structured processes, typically characterised by a clear and specific goal [8, 28, 38].

The understanding that conducting literature reviews is a structured process contradicts the perspective of researchers in the information seeking and retrieval community who widely view literature review tasks as a tool for exploring and understanding exploratory searches [16, 18, 25, 35]. This apparent contradiction underscores the need for a more nuanced exploration of the approaches, behaviours, and practices involved in conducting literature reviews within the broader framework of exploratory search. Therefore, this paper addresses the following research questions:

- RQ1. What are the key exploratory characteristics that come into play when users search for literature in a new domain?
- RQ2. What are the main approaches that users use when conducting literature searches in unfamiliar domains?
- RQ3. How do users evaluate the relevance of documents when they are searching for literature in a new domain?
- RQ4. What are the processes users employ in acquiring knowledge during literature searches?

We conducted a user study involving 30 participants. These participants were tasked with using some interfaces to generate outlines on topics of their choice, with the condition that the topic should be something they had no prior experience with. Subsequently, we conducted semi-structured interviews to gain deeper insights into their experiences, behaviours, and practices while conducting literature reviews in general. Our analysis, involving multiple rounds of coding and thematic analysis, enriches our understanding of exploratory searches, behaviours, approaches, characteristics, decision-making patterns, and knowledge acquired throughout the literature search.

This paper investigates the behaviours and search approaches associated with exploratory searches during literature reviews. Our primary goal is to support existing theories with empirical evidence, providing insights into the distinctive behaviours observed during exploratory searches in the context of literature reviews. Additionally, we explore the criteria used by searchers to determine the relevance of papers and examine the knowledge acquisition process employed in generating literature reviews.

2 BACKGROUND

In this section, we provide an overview of three pertinent lines of research. First, we delve into the exploration of exploratory search. Second, we explore the landscape of literature review tasks. Third, we examine the existing models and theories related to information seeking behaviour.

2.1 Exploratory Search

In the context of classifying search tasks, it often simplifies into a binary classification: either it is a known-item search or an exploratory one [22]. The latter can be particularly challenging. Searchers frequently employ exploratory search strategies when tackling complex tasks such as literature searches or delving into unfamiliar domains [16, 36, 42]. These strategies typically involve searchers initiating tentative queries to locate relevant documents and then exploring the information environment to gather cues guiding their subsequent actions [42].

White and Roth [43] identified vital user characteristics of exploratory searches, including a lack of prior knowledge in a specific domain, uncertain goals, and an unclear path to fulfilling their information needs. Kules and Capra [21] introduced operational characteristics for exploratory search tasks, emphasising that answers are not typically found on the initial interaction, leading users to interact with results, reformulate queries, and search for multiple items. In their review, Wildemuth and Freund [44] consolidated the main features of exploratory search tasks as general, open-ended,

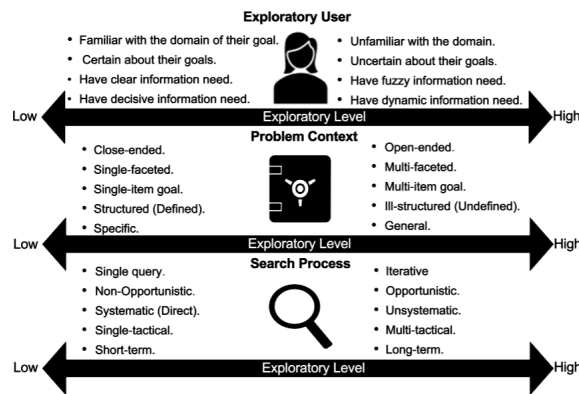


Fig. 1. The recent conceptual model of the exploratory search based on Soufan et al. [36].

focused on learning and investigative goals, often involving multiple documents or items, and motivated by ill-defined or ill-structured problems.

For Hassan et al. [15], exploratory sessions occur when individuals confront open-ended and multifaceted information seeking tasks. Athukorala et al. [2] argue that searchers' knowledge, interests, goals, and information needs are in constant flux throughout the exploratory search process. The degree to which a task is considered "exploratory" is dependant upon the user's expertise and experience [10, 17]. Soufan et al. [36] have recently proposed a conceptual model of exploratory search, which builds upon a thorough examination of foundational papers and studies in the field. Their recent model comprises three primary dimensions, as shown in Figure (1). Each dimension has distinct characteristics, and their research findings indicate the potential existence of a fourth dimension that has to do with knowledge gain/change.

For Soufan et al. [36], the main exploratory dimensions are 1) the user dimension, which includes users who are unfamiliar with the domain, uncertain about their goals, have vague information needs, and exhibit dynamic information needs, 2) the problem context dimension which describes the problem context as open-ended, multi-faceted, and often ill-structured, 3) the search process dimension which characterises this process as iterative, opportunistic, unsystematic, and multi-tactical, 4) the knowledge gain/change dimension which encompasses elements such as encountering surprising or unexpected information during the search, learning about and encountering new concepts or keywords, and choosing to delve deeper into these newfound areas of interest.

The recent conceptual model of the exploratory search of Soufan et al. [36] has indeed expanded our understanding of this multifaceted concept. However, it is important to acknowledge that there may still be unexplored aspects related to particular approaches, behaviours, and practices, especially in the context of conducting a literature review in a new domain. Previous exploratory search models have indeed significantly contributed to our understanding of exploratory search. However, it is important to note that empirical studies in this domain remain limited. Specifically, there is a distinct lack of empirical research focusing on the methods and information seeking behaviors associated with exploratory searches during literature reviews in unfamiliar domains. There is also a shortage of research addressing search behavior and knowledge acquisition during literature reviews in new domains.

2.2 Literature Review Task

Literature reviews provide an overview of previously published works on a topic, encompassing various scholarly materials like books, theses, dissertations, and journal articles [3]. They summarise and evaluate a literature collection related to a specific theme, theory, or method to enhance the existing knowledge base [19, 29]. They require various skills, including finding and evaluating relevant materials, synthesising information from diverse sources, employing critical thinking, and effectively summarising content [7].

Conducting an effective literature review with a logical flow demands various skills, including locating and evaluating pertinent materials, synthesising information from various sources, employing critical thinking, and effectively paraphrasing content [7, 32]. Exploratory search strategies are crucial in conducting literature reviews, where individuals aim to gain insights into specific topics. Soufan et al.'s work [36] revealed that around 80% of nearly 400 participants viewed literature reviews as exploratory search tasks. It is important to note that literature reviews are regarded as processes of knowledge construction [37, 46]. This involves collecting research literature, comprehending its content, establishing connections among gathered data, and synthesising insights [37, 46]. The term “mental model” is frequently used in sense-making literature [31, 37, 46] to describe this evolving comprehension as searchers encounter new information during exploratory searches [5].

Several models exist to describe information seeking behaviours among academics, including Ellis's model [13], Bates' berry-picking model [4], and Vakkari's model [39]. However, the current literature lacks comprehensive insights into the search behaviours and practices involved in the exploratory search, specifically in academic contexts and when searching for literature. For instance, Hoeber et al. [16] primarily investigated the information seeking models utilised in search scenarios they had hypothesised. Athukorala et al. [1] classified five primary motivations driving computer scientists' engagement in literature reviews. Prior research about academics' information seeking behaviour mostly centred around their responses to emerging electronic information tools and the influence of factors like demographics, environmental contexts, and academic disciplines on their information behaviour, as evident in works by Brown [6] and Niu and Hemminger [27].

Further studies explored different facets, including changes in academics' information seeking strategies, challenges when using new electronic tools, and preferences for resources and channels [34]. Moreover, some studies focused on specific academic groups, such as computer scientists [1], data scientists [24], social scientists [13, 23, 34], or graduate students [14]. However, while these studies have shed light on aspects of academic information seeking, they do not provide a comprehensive understanding of searchers' behaviours, approaches, decision-making processes, and knowledge acquisition when they engage in exploratory searches during academic tasks.

In some previous studies, the examination of the actual information seeking behavior of students in a digital scholarly environment was primarily based on system logs. For instance, Nicholas et al. [26]'s study revealed a distinctive form of information seeking behaviour specific to students. It highlighted differences between their behaviour and that of other academic community members. However, it is worth noting that this study focused on academics utilising various systems to read scholarly papers, and the system logs were collected independently of specific search tasks. In another study, Yu et al. [45] sought to identify students' strengths and weaknesses in locating, retrieving, and citing information effectively. The study revealed that different student groups cited varying numbers of items in total, with some groups citing more books and journal articles than others.

However, it is important to note that the data were collected independently of specific tasks, and the study did not provide insights into the exact approaches, experiences, and step-by-step processes students follow when conducting a

literature review. Consequently, this research aims to bridge this knowledge gap, specifically in the context of conducting exploratory searches in academic settings, particularly when searching for scholarly literature. In the upcoming section, we will provide further details into how various models of information seeking behaviours can also be applied to describe searching the literature in a new domain.

2.3 Information Seeking Behaviour Models & Theories

Numerous models have been developed in information seeking behaviour to describe various facets of information seeking activities, the underlying causes, the consequences of such activities, and the intricate relationships underpinning these behaviours. For example, Dervin's Sense-Making theory [11] introduces a triangular framework comprising situational factors, a gap or bridge, and outcomes. Within this construct, situational factors represent the contextual scenarios giving rise to information-related challenges. The gap serves as a conceptual marker delineating the disparity between the existing situational context and the desired state while the bridge signifies a mechanism employed to traverse this gap. In their quest to bridge this gap, individuals engage in information seeking activities to generate novel insights. We view the process of conducting a literature review in an unfamiliar domain as a form of sense-making, where researchers seek to understand and make sense of the information space and the resources they encounter.

The Information Foraging Theory [33] underscores two distinct foraging strategies: 1) specialists who concentrate their efforts on a single high-density "patch" of sources that they encounter through informal communication channels and heavily rely on sources within their collections, and 2) generalists who, in contrast, adopt a broader approach by gathering sources from a diverse array of "patches". This theory offers insights into information seeking behaviour, which aligns with much of what individuals do when searching the literature. However, we lack a comprehensive understanding of how people respond to specific signals during this process, as we view queries and documents as parts of patches.

Kuhlthau's *Information Search Process* (ISP) model [20] provides valuable insights into the emotional and cognitive aspects involved in information searching. Although this model was studied in the context of students conducting literature reviews over a term, we still lack detailed information on students' actual approaches, resource assessment methods, and the key exploratory characteristics they employ during this task. Ellis's Behavioural Model of Information Searching Strategies [9] examined the various activities of social scientists. However, it still does not comprehensively understand the specific activities involved in literature review tasks. It is worth noting that Ellis's model was developed before the advent of digital libraries and online environments.

In summary, there is a need to study and understand the behaviours, and approaches used in literature review tasks. This paper focuses explicitly on individuals conducting literature reviews in unfamiliar domains, with the aim of comprehensively understanding their information seeking behaviours in this context.

3 METHODOLOGY

To investigate the primary behaviours and approaches individuals employ when conducting literature searches in unfamiliar domains, explore the key exploratory characteristics relevant to such searches, understand how users assess the relevance of documents during literature searches and delve into the knowledge acquisition process during these searches, we conducted semi-structured interviews with 30 participants after they engaged in a literature review task. This data collection was part of a broader user study evaluating various exploratory interfaces. The following section summarises the study setup.

3.1 User Study Setup

We conducted a mixed-methods laboratory-based user study involving 30 Master’s students (between-subjects). The study employed two user search interfaces, with the user interface serving as an independent variable. Each participant had access to only one interface. The interfaces utilised in the user study were specifically designed to support academic search, encompassing around 200k papers in Computer Science and related fields.

Participants were tasked with searching the literature on a topic of personal interest that they were unfamiliar with. Participants were given as much time as they desired, and no time limit was set for their engagement with the search task. The objective was for participants to produce the most comprehensive and well-structured literature review outline possible. Completion times for this task varied, with some participants finishing in less than half an hour while others dedicated up to one and a half hours. On average, participants spent approximately 45 minutes on this task.

The only condition specified in the study was that the chosen topic should be unfamiliar, reflecting a subject of personal interest they would like to explore further. Participants were encouraged to focus on their Master’s project, ensuring they had not previously conducted a literature review on the chosen topic. At the conclusion of the search session, they provided us with a literature review outline related to their selected topic. The main author observed them while using the system, searching, and conducting the task.

After completing the task, we conducted thorough semi-structured interviews with each participant. The semi-structured interviews were the final phase of the user study. These interviews focused not only on participants’ experiences in conducting literature reviews using the provided systems but also extensively delved into participants’ overall experiences and approaches to conducting literature reviews in general.

3.2 Data Collection

We gathered quantitative and qualitative data for the user study through system logs, questionnaires, and semi-structured interviews. However, this paper concentrates solely on the qualitative data obtained from the semi-structured interviews and our observations. Observing participants’ interactions with the provided systems and how they conducted literature reviews during the search session offered valuable context and insights. These observations guided the formulation of specific and relevant questions during the interviews.

During the interviews, we inquired about participants’ experiences using our interfaces for literature reviews and whether their approach aligned with their general literature review practices. Additionally, they were probed about their broader and general experiences and approaches in conducting literature reviews in real-life scenarios. Participants were asked to walk through their thought processes and to provide a detailed overview of their typical approaches when conducting literature reviews in real-life scenarios.

The interviews aimed to uncover participants’ criteria for determining the relevance of research papers, factors influencing their choices in selecting and reading specific papers, resources they rely on, and tools used for literature reviews in general. The semi-structured interviews primarily focused on prompting users to recollect their experiences when conducting literature reviews from memory [40]. The user study sessions were recorded via Zoom technology, enabling the generation of transcripts from the audio recordings.

Subsequently, the transcripts underwent a thorough review involving careful listening to the recorded interviews to ensure the accuracy of the transcripts for subsequent data analysis. The interviews’ duration varied between 10 to 20 minutes for each participant. We opted for a multi-stage coding process to code the interviews effectively, including deductive and inductive coding. More details are in the following sections.

3.3 Deductive Coding: Key Exploratory Characteristics when Searching for Literature

To address our first research question (RQ1), we conducted a deductive coding process that involved constructing a coding scheme rooted in the 14 characteristics proposed by Soufan et al. [36] in their conceptual exploratory search model. As illustrated in Figure (1), this model lays the foundation for understanding exploratory search. Their model encompassed all the constructs proposed in the literature about exploratory search, and they employed a questionnaire instrument to empirically validate it. While Soufan and colleagues' conceptual model primarily focuses on exploratory search in a broad context, we tailored it to align with the specific demands of literature review tasks.

For every characteristic of the 14 exploratory characteristics, we provided a customised version of its definition to make it contextually relevant to literature review tasks. This adaptation involved the transformation of the 14 characteristics into a codebook, a critical tool during the deductive thematic analysis of the interview transcripts. These codes underwent multiple iterations and were reviewed by two senior researchers with experience in thematic analysis, and literature reviews to ensure the clarity and accuracy of the provided examples. This validation process was instrumental during the deductive coding phase. We utilised the NVivo tool to apply this codebook, facilitating the further refinement of these characteristics to suit the literature review context better.

The outcomes of this customisation are elaborated upon in the results section. The appendix has the codebooks we used during the deductive coding stage. The primary author of this paper reviewed the transcripts of the semi-structured interviews line by line. During this process, the author assigned codes based on the existing codebook, drawing connections between the textual data and predefined codes. This detailed coding approach aimed to systematically analyse and categorise the content of the interviews, ensuring a comprehensive exploration of participants' responses.

3.4 Inductive Coding: Exploring Literature Review Approaches

To answer our second, third, and fourth research questions (RQ2, RQ3, and RQ4), we employed an inductive thematic analysis approach. The primary objective was to unveil underlying patterns and establish codes within our data. Through semi-structured interviews, we gained valuable insights into participants' approaches for conducting literature reviews in unfamiliar domains. This involved exploring the factors guiding their paper selection and how they structured their literature reviews by navigating the complex information landscape.

It was imperative to delve deeper into these aspects as they substantially contribute to our holistic comprehension of the exploratory approaches, behaviours, and decision-making processes inherent in the literature search process for constructing a literature review. To facilitate this, we formulated a set of dimensions and accompanying examples. These dimensions underwent a review process, including modifications made over several iterations in consultation with two senior researchers. Here are the steps we followed during the inductive coding:

- (1) We identified and coded all text pieces related to participants' behaviours and approaches when conducting literature reviews.
- (2) We identified and coded all text pieces related to the criteria that influenced participants' paper selection.
- (3) We identified and coded all text pieces related to the process of knowledge acquisition during literature reviews.

We conducted further iterations on each of these codes. This allowed us to delve into the text to understand the various approaches participants took while conducting literature reviews, including whether they prepared an outline or structure before searching for literature. We also examined how they progressed through the task, read and investigated them, and how they decided which papers to read. Furthermore, we explored how participants used papers they

considered relevant and their methods for acquiring and utilising knowledge during the literature review process. The findings of this analysis are elaborated upon in the results section.

3.5 Ethics and Recruitment

We received ethics approval for our user study from our department’s Ethics Committee (Application ID 2035). The interfaces we used cater to academic searches in Computer Science and related fields. Therefore, we aimed to recruit Master’s students with backgrounds in Computer Science or related disciplines, all of whom had completed at least one literature review. To select our participants, we devised a screening questionnaire with essential questions. These questions covered whether they were currently enrolled as Master’s students, their field of study, whether they were required to conduct a literature review for their Master’s project, details about their Master’s dissertation/ thesis topic (either the title or a broad idea), their progress on their Master’s projects (measured on a scale from 0 to 100), and how many times they had previously produced a literature review.

The screening survey served the purpose of ensuring that we recruited Master’s students primarily from the field of computer science or related disciplines. This criterion was crucial to ensure a relatively uniform educational background among our participants and to confirm that they lacked prior experience with their Master’s projects. We disseminated this screening questionnaire through various channels, including university student mailing lists and social media platforms like Twitter, Facebook, and LinkedIn. As a token of appreciation for their time and participation, we offered online shopping vouchers worth approximately 35 US dollars.

3.6 Participants

Out of 94 individuals who completed the screening questionnaire, roughly half of them met our eligibility criteria and were deemed suitable for participation. Ultimately, only 30 participants actively responded to our emails and took part in the user study. All thirty participants were pursuing Master’s degrees, representing various academic fields such as Artificial Intelligence, Computer Science, Information and Library Studies, Software Engineering, Data Science, and more. These students were on the verge of embarking on their Master’s projects in the upcoming semester.

4 RESULTS

This section unveils the outcomes derived from the inductive and deductive coding applied to the semi-structured interviews we gathered, offering insights that address our research questions.

4.1 RQ1: Key Exploratory Characteristics

The findings presented in this section address our first research question (RQ1): “What are the key exploratory characteristics that come into play when users search for literature in a new domain?”. Here, we elucidate the information behaviours and search processes within the context of Soufan et al. [36]’s exploratory search model. As previously mentioned, we adapted their conceptual model to suit the specific demands of literature review tasks. This involved translating their 14 exploratory characteristics into a codebook, which we utilised during the coding process. We structured our presentation of findings for the deductive thematic analysis by focusing on each exploratory dimension and its associated characteristics.

Moreover, we forged connections between these characteristics and the experiences, behaviours, and approaches articulated by users throughout their literature searches, as revealed in the insights shared during the semi-structured interviews. This approach offered an exploration of how these dimensions and characteristics were reflected in our

participants' experiences. Furthermore, we identify and highlight the key exploratory characteristics of the user dimension based on the insights derived from the qualitative data we collected and analysed.

4.1.1 User Dimension: The exploratory conceptual model developed by Soufan et al. [36] outlined four characteristics within the User dimension. The subsequent insights pertain to this dimension.

Domain Familiarity/Unfamiliarity: This attribute assesses users' familiarity with the domain in which users are searching and reflects their existing familiarity or lack of it in terms of concepts, keywords, authors, sources, topics, knowledge, expertise, understanding of the domain and the topic. Through our data analysis, a prevailing theme emerged: the majority of participants openly admitted to their limited expertise in the field they were investigating. For instance, one participant stated, "*I am not an expert in this field in AI*" (P1), while another mentioned, "*The topic was completely new*" (P2). Some participants revealed that while they were familiar with certain terms and concepts in their chosen field, they were entirely unfamiliar with other aspects they encountered during their searches. One participant commented, "*I knew what the topic is about, but not how it was being done*" (P3). This outcome was expected since we deliberately tasked our participants with exploring a topic they had no prior knowledge about. This scenario mirrors real-life situations where individuals often find themselves conducting searches in entirely new domains.

Goal Certainty/Uncertainty: This attribute pertains to users' certainty level regarding their search objectives, the content they seek, the search structure, and the expected outcomes. Users might exhibit varying degrees of goal certainty. Some may possess uncertain or ambiguous goals, be unsure about the outcomes they anticipate, or deal with the challenge of pursuing answers that lack clarity. From our study findings, some participants revealed that they initially grappled with uncertainty regarding how to approach the task and doubted their ability to make meaningful progress. Throughout their search process, they encountered episodes of uncertainty at different stages. As one participant put it, "*At the beginning, I was a bit, what do I use now? I didn't think I was going to make so much progress*" (P4). Additionally, participants noted that it was easy to become sidetracked by the numerous subtopics they encountered during their literature reviews. One participant explained, "*...when you're writing a literature review, it's easy to get sidetracked among the many various subtopics that happen.*" (P5). Conversely, some participants demonstrated a higher level of certainty when defining their search objectives. They were more precise in specifying the content they were seeking and exhibited a structured approach to their searches. It is worth noting that several of these participants mentioned receiving training in conducting literature reviews and had prior experience performing them. This raises questions about additional factors that may influence goal certainty, which we will explore further in the discussion section.

Information Need Clarity/Fuzzy: This attribute relates to the extent of clarity users have regarding their information requirements – specifically, how clearly they can define the information or data they seek during a particular step of their search. Exploratory users often contend with vague, unclear, or imprecise information needs, making it challenging to articulate the information necessary to fulfill their goals precisely. Our data analysis unveiled a recurring theme: Participants frequently needed to invest time in exploration to grasp the available information fully. Many had to sift through multiple papers to learn and determine how to proceed to the next phase. Participants often encountered situations where, after reviewing specific papers, they found themselves uncertain about their subsequent steps. Some even described struggling to address their information needs effectively. For instance, one participant articulated this challenge as, "*...You know what you want, but you don't know how you're supposed to address it*" (P4). Similar to the previous attribute, certain participants displayed a distinct sense of direction and exhibited confidence in their information requirements. They demonstrated a clear understanding of how to progress from one step to the next in their search process.

Information Need Decisive/Dynamic: This attribute pertains to the dynamic nature of users' information needs as they progress through the search process. It encompasses how their search objectives, understanding, and answers may evolve and transform during exploration. Users invest time refining their comprehension and knowledge as they search, leading to a shifting focus on their information needs. Additionally, users may adapt and modify their search queries, including keywords, as they delve deeper into their search process. Our findings highlight two key facets of this characteristic including 1) dynamic information gathering: participants indicated that when they lacked prior knowledge about a topic, they searched for relevant keywords or picked up terminology from the papers they read. Consequently, they often moved from one paper to another, with each new source altering their understanding and information needs; 2) evolution of documentation: participants mentioned that their notes, summaries, or the structure they were developing for their literature review changed as they progressed in their search. For instance, one participant stated, "...after going through those papers, I did get an idea of where I needed to go next" (P4). Another participant explained, "...I would say it's a lot of updating of this structure..., every time I find something interesting then I say, oh, where can I write this in my structure." (P6). These two aspects reflect how participants adapted and refined their information seeking and documentation approaches while navigating unfamiliar topics during their literature reviews.

When discussing users conducting literature reviews in a new domain, it is clear that their familiarity with the topic is limited, a fundamental aspect of exploratory search. Regarding the nature of information needs, it is evident that this is a central characteristic of exploratory search. Participants continuously engaged in exploration, learning, investigation, and relearning throughout the search sessions. However, the level of goal certainty and information need clarity varied among participants despite their shared unfamiliarity with the subject. This variance could be linked to participants' prior experiences in conducting literature reviews, raising questions about how these experiences and training in conducting literature reviews influence this characteristic.

4.1.2 Results of Problem Context Dimension: The exploratory conceptual model developed by Soufan et al. [36] outlined five characteristics within the Problem Context dimension. The subsequent insights pertain to this dimension.

Open/Close-Ended: This attribute pertains to the nature of the problem and the desired output. In the context of a literature review task, there can be multiple potential answers or outcomes. Additionally, it highlights the continuous evolution of literature in the field being researched, the extensive volume of available information, the inherent limitations in searches, and the impossibility of encompassing all aspects of a topic. Based on our findings, it was evident that some participants perceive their fields of interest as highly dynamic and subject to continuous change. For instance, one participant emphasised the rapid evolution of their field by stating, "*Because natural language processing definitely is a cutting-edge technology, so every day it is changing.*" (P7). None of our participants viewed literature review as a closed-end problem with only one answer.

Multi/Single Faceted: This attribute characterises outcomes that involve multiple sub-tasks or cover various aspects or concepts within the explored domain. It relates to conducting a comprehensive exploration that spans different facets of the topic. According to our findings, participants generally viewed the literature review task as multifaceted. They discussed two primary approaches for tackling this complexity: 1) Some participants mentioned breaking the task down into smaller sub-tasks, 2) While others emphasised the identification of various topics and sub-topics relevant to their research area. They then assembled these relevant topics and concepts to construct their literature review. One participant said, "*It was like trying to piece together the relevant kinds of topics and concepts, and then taking it from there to understand how it all fits together*" (P8). It was uncommon to encounter participants who described literature reviews as single-faceted tasks.

Multiple/Single Item: This attribute refers to outputs that involve multiple documents or resources. It also describes an information goal that is typically achieved by combining information from various sources encountered during the search. Our research findings reveal a common pattern in the literature review process. Participants often initiate their review by reading a single paper. Subsequently, they frequently expand their exploration by delving into additional papers, usually prompted by citations or when searching for specific keywords and concepts that they initially encountered in the first paper. One participant mentioned bringing eight books from the library related to their literature review's domain. As one participant put it, "*I think it is like the Rabbit Hole technique where you find one paper, and you find like seven different terms, and these lead to seven different papers*" (P8). However, it is important to note that participants read many papers, some of which may not have been directly related to their final literature review. One participant mentioned, "*I spent time reading about 55 papers, but in the end, I used only around 26 to 30 of them.*" (P9). We did not encounter participants who mentioned relying on only one paper or resource to construct a literature review on a particular topic. However, participants may come across an "inspirational paper" closely related to their topic, prompting them to follow its structure or citations for guidance. This characteristic was one that many participants reflected on.

Structured/Ill-Structured: This attribute pertains to the requirements of the literature review task, which can often be unclear, imprecise, or subject to significant changes during the search session. Consequently, participants frequently encounter challenges in comprehending the task and often seek support in navigating its inherent unclarity and ambiguity. Our findings show that some participants found it challenging to initiate the task, primarily because the requirements were not specific. They perceived it as receiving a broad topic and being tasked with creating an outline, but they were uncertain about how to proceed. One participant expressed this by saying, "*At the beginning, I was a bit unsure about what to use, ..., it felt like I was given a topic and asked to create an outline*" (P10). It is important to highlight that most participants found the task instructions for this specific exercise clear. We addressed their questions and ensured their understanding of the task, which involved creating an outline for a topic of interest in an unfamiliar domain using the provided interface. However, in real-life tasks, they might not find the same support.

General/Specific: This attribute characterises a task that is presented with a broad and unclear description, resulting in general and vague answers. This situation often relates to the literature review task, where the information needed lacks specificity and clarity. Additionally, there is ambiguity in defining the scope of the search and identifying the specific aspects of the domain that should be the primary focus. Some participants shared that when conducting a literature review in an unfamiliar domain, they begin with the most general information and gradually narrow down their focus. For example, one participant mentioned, "*I'd start from the most generic stuff.*" (P4). What we observed here is the potential for participants to take various routes or approaches when addressing the same specific topic. This indicates the open-ended nature of the task, highlighting that multiple solutions or answers can exist. This might hint at a new characteristic related to the possibility of pursuing different directions in the search process to arrive at the final output, indicating a multi-directional search process. More on this will be discussed in the results of the search process.

We consider the aforementioned characteristics to be the primary exploratory elements within the problem context. Nonetheless, our results provide further depth and establish connections that enhance our comprehensive understanding, particularly regarding the multi-item and multi-faceted characteristics.

4.1.3 Results of Search Process. The recent exploratory conceptual model by Soufan et al. [36] delineated five characteristics within the Search Process dimension. The subsequent insights pertain to this dimension.

Iterative/ Not-Iterative: This attribute describes the iterative and evolving nature of exploratory search. Users typically begin with a tentative or imprecise query and then gradually refine their search through multiple iterations. User may follow this iterative approach, progressively narrowing down the scope of their information needs to obtain more relevant and useful results. Our findings revealed that many participants described their approach to conducting a literature review as starting with an initial query, reading some initial paper(s) to gain preliminary knowledge, and then actively searching for related concepts or topics mentioned in those papers. This often led to what could be described as a “rabbit hole” effect, where they moved from one paper to another, gradually deepening their understanding. Participants mentioned that they would conduct additional searches to clarify their understanding when encountering something they did not fully comprehend, such as an unfamiliar term or concept. One participant explained, *“Whatever I do not understand in that paper, what I do is I try to find those phrases and words and topics and papers around that.”* (P11). Additionally, participants noted that they iteratively refined their queries to find relevant information. Furthermore, some participants mentioned that they occasionally identified gaps in their knowledge during the actual writing process. In response, they paused their writing to conduct additional searches to fill these gaps. One participant stated, *“There are points where I am writing, and I realise I do not have enough information on this topic... I will pause the writing to do another quick search just to fill that out sometimes.”* (P8). Our analysis reveals another aspect of this characteristic: the multi-query approach. Participants described using multiple queries to gather information on various facets of the topic. This approach differs from issuing different versions of the same query to align more closely with their information needs. The multi-query approach is related to the multi-directional search, as participants, while reading and learning about the topic, have the option to steer their research in different directions by issuing new queries, either more general or more specific, to gather relevant information for the final output they are working on. It is worth noting that this characteristic, with its new facets, was frequently mentioned, and many participants reflected on it.

Opportunistic/ Not Opportunistic: This attribute describes users’ willingness to take greater risks with the expectation of potentially obtaining valuable information from reading or searching for documents. Our findings show that some participants described a somewhat spontaneous approach to their literature review process. They mentioned that they would randomly select a paper to start with and then explore further based on what they found in that initial paper and its references. In this way, they allow their exploration to be guided by the content they encounter. Additionally, participants noted that they often focus on specific sections of papers, particularly those they deem relevant to their research interests. For instance, one participant explained, *“There have been papers on..., so it’s completely off to me, not something that I want to use, but out of curiosity I go and read about it so that I know there is something out of the box approach as well”* (P9). Another facet of this characteristic is following a speculative approach while searching for relevant information. Some participants mentioned starting the search to see what the interface would show them, and based on the results, they decided what to read and where to go.

Systematic/Unsystematic: This attribute is related to the presence or absence of a systematic pattern or structured approach during the search process. Exploratory users may follow unpredictable and non-linear paths in their searches, often deviating from predefined search strategies. Our findings demonstrate that some participants described a more systematic method while searching the literature. They would begin with one paper, explore others related to it, and then return to the original paper. This systematic approach often involved tracing the development of ideas over time, resembling a study of the history and progression of a concept. As one participant put it, *“I start with the original paper, I will branch out, based on what I find, but I keep coming back to the original until I am finished.”* (P8). In contrast, others followed a more unsystematic approach. They might jump from one paper to another without necessarily completing any of them. One participant expressed this approach, saying, *“I’m trying to read a paper, but then... I just lose track of*

Manuscript submitted to ACM

what I was actually looking for..., sometimes I finish a paper, sometimes (usually) I don't." (P12). Our analysis highlights the presence of both systematic and unsystematic approaches to the search process. Regardless of the chosen approach, participants actively engaged in exploration activities, including finding, investigating, learning, and navigating between different information sources. Therefore, for this specific exploratory search task, this characteristic might not be suitable to differentiate exploratory search processes. This is because exploratory users engaged in various activities that showed different patterns of systematic and unsystematic search approaches.

Multi/Single tactical: This attribute pertains to a search approach where users use different systems, strategies (methods of finding documents), or tools to access relevant information and adapting search tactics in response to evolving information needs. Our findings reveal that the majority of participants utilise various tactics during their literature reviews. They frequently discover papers by tracing citations from one paper to another, searching for keywords and concepts, and conducting quick Google searches. This multifaceted approach helps them gather a comprehensive set of resources for their literature reviews. Furthermore, they leverage diverse tools and resources, including Google, Google Scholar, Wikipedia, their university library, digital libraries, and reference management software like Mendeley. Some participants incorporate books and YouTube videos into their information-gathering process. One participant mentioned, *"I go from one paper to another from like the citations of one to another."* (P6). Another one said, *"If I don't understand something, I will just read Wikipedia and then go back."* (P13).

Long/Short term: This attribute describes the duration of the literature review task that may take place over multiple sessions, and it can be as long as hours, days, or even months. Our participants emphasised that conducting a literature review is a time-consuming task, often requiring more time than initially anticipated. One participant mentioned, *"If I had more time, of course I would have spent more time"* (P10).

In literature review searches, users utilise iterative, evolving, and multi-query processes to cover various aspects of their topics, including different, subtopics, and sometimes unconventional themes. The multi-query approach is frequently overlooked in discussions of exploratory search processes. Also, contrary to common descriptions of search processes as unsystematic, our findings indicate that users in literature searches may employ both systematic and unsystematic approaches while exploring the literature.

4.2 RQ2, RQ3, & RQ4: Literature Review Strategies

As mentioned earlier, in the inductive thematic coding, we focused on understanding how participants conduct literature reviews, acquire knowledge, determine paper relevance, and explore their behaviours, approaches, and strategies in detail.

4.2.1 Literature reviews approaches: Here we try to address our second research question (RQ2): "What are the main approaches that users use when conducting literature searches in unfamiliar domains?" Our analysis revealed two main approaches that participants utilise when conducting literature reviews:

- **Predefined Structure:** Some participants begin with a preliminary outline, structure, or framework before starting their literature search. Depending on their familiarity with the domain, these structures can be quite generic, outlining main sections such as introduction, background, methods, approaches, etc. They might also structure topics and sub-topics if they possess some initial knowledge. These structures are dynamic and evolve as they delve deeper into their search and acquire new insights.
- **Evolving Structures:** Some other participants initiate their literature review without a predefined structure or outline. They start their search and reading process without a rigid framework in mind. However, as they

Manuscript submitted to ACM

read various papers, they gradually identify patterns and sometimes discover foundational concepts or core methodologies. It is not uncommon for them to encounter what they refer to as an “inspirational paper”, one that resonates with their goals, and they might use this paper’s structure as a guide when they begin crafting their literature review.

4.2.2 Paper Relevance: In response to our third research question (RQ3), which investigates “How do users evaluate the relevance of documents when they are searching for literature in a new domain?” Our findings emphasise the critical role of the abstract. For many participants, the abstract serves as the primary focal point when deciding a paper’s relevance. They carefully assess the abstract to ascertain whether it contains the points or concepts they seek. If the abstract aligns with their research needs, they proceed to explore the paper further. Some participants have a specific process: they begin with the abstract, proceed to the conclusion, and, based on these sections, decide on the paper’s relevance. Another group mentioned reading both the abstract and the introduction before deciding. A few participants base their decision on the paper’s title while others rely on the paper’s keywords. A minority mentioned looking at the results section mainly. Even when participants decide that a paper is relevant, many often employ a strategy of skimming the paper, and searching for specific information. They might revisit the abstract, introduction, and conclusion during this process. Unless a paper is particularly “inspirational” or highly pertinent to their research, they typically do not read the entire paper. Instead, they selectively focus on specific sections such as methods, models, approaches, or other targeted information that aligns with their research objectives.

4.2.3 Learning and Knowledge Gain: In response to our fourth research question (RQ4): “What are the processes users employ in acquiring knowledge during literature searches?” Our findings highlight several key aspects:

- **Keyword Acquisition:** Participants noted that the papers they read often provide them with keywords, terms, phrases, and concepts to search for in subsequent steps. These terms serve as entry points for deeper exploration.
- **Intermittent Searching:** During the reading process, participants frequently pause to conduct searches related to keywords or concepts they encounter. These searches may range from quick Google searches for basic information to more comprehensive searches involving other papers or resources. In some cases, they explore papers referenced in the initial paper.
- **Rabbit Hole Strategy:** Some participants employ a “rabbit hole” strategy, moving from one paper to another or from one resource to another, sometimes without returning to the original source. This approach allows them to follow threads of information to gain a deeper understanding.
- **Repetition for Understanding:** Participants may encounter certain concepts across different papers or resources multiple times. It often takes repeated exposure to fully comprehend these concepts. They search for new topics and concepts, hoping to grasp them, but understanding often comes after encountering them several times.
- **Note-taking and Summarising:** Many participants take notes or create summaries of the papers they read. These notes are dynamic and subject to frequent updates and changes as they continue their search. These summaries serve as valuable resources for the eventual literature review, although sometimes participants need to revisit the original paper if their notes are insufficient.
- **Selective Citation:** Participants may read a substantial number of papers but end up citing only a subset, typically those most relevant to their research objectives. They might read papers out of curiosity or to gain a broader understanding of the general topic before delving into more specific sources.

Overall, the process of acquiring knowledge during literature reviews is dynamic, involving continuous searching, reading, note-taking, and revisiting sources. Participants adapt their strategies as they progress through the research process, seeking to build a comprehensive understanding of their chosen topic.

5 DISCUSSION AND CONCLUSION

While previous research has predominantly focused on comprehending and supporting exploratory searches through frameworks, search systems, and interfaces tailored to this task, this paper takes a deeper dive into a specific aspect of exploratory search: literature reviews. This section highlights some key findings and concludes with the research implications and potential avenues for future work.

First, regarding the characteristics of the user dimensions; in terms of the problem context dimension, we found that participants often adopt various routes or approaches as they progress in their search and gain more insights into their topic of interest. This implies the existence of multiple directions towards the relevant information that contributes to the final output. This expands the understanding of the open-ended nature of the problem context. Regarding the search process dimension, the findings indicate that adopting a less methodological or systematic approach does not necessarily make the search more exploratory when searching the scholarly literature. Some participants followed a systematic approach, such as exploring references in a paper and conducting keyword searches. On the other hand, other participants pursued an unsystematic, unstructured approach as they explored the information space, moving between papers. Nevertheless, both groups were still engaged in exploration. This implies that the systematic/unsystematic search process might be a relatively minor characteristic in the context of exploratory searches. Furthermore, prior literature has discussed how users often commence their search with a tentative or imprecise query, subsequently honing their search through multiple iterations. However, what is often overlooked is the multi-query nature of the search process. Participants employ a range of queries to gather information that encompasses the multifaceted and diverse aspects of the problem. This enhances the understanding and expands the definition of the exploratory search process, and suggests the emergence of a new characteristic within the search process, which relates to using multiple queries.

Second, regarding the literature review strategies we have identified various approaches to conducting literature reviews. Furthermore, this research revealed that individuals primarily rely on specific parts of papers to determine their relevance. We also highlighted some key knowledge acquisition aspects. This underscores the importance of designing support interfaces tailored to assist individuals engaged in tasks like literature reviews.

Third, regarding the information seeking behaviours when conducting literature reviews, we consider conducting literature reviews as a sense-making problem, and the key exploratory characteristics elucidate the bridge proposed by Dervin [12]. The results provide a detailed account of users' information seeking activities when searching the scholarly literature to make sense of the surrounding information space for their tasks. Also, the findings empirically describe the Information Foraging Theory [30] as it applies to information seeking. We illustrate how participants interact with patches of information (papers and resources) and employ these patches to navigate to other sources of information. Additionally, we delve into how they leverage these patches of information to acquire knowledge and construct outlines and structures for their literature reviews. Furthermore, the research contributes to a better understanding of Ellis's Behavioural Model of Information Searching Strategies [13]. This model categorises information seeking patterns among social scientists into six major categories: starting, chaining, browsing, differentiating, monitoring, and extracting. We contextualise these activities within the literature review task and provide a comprehensive account of the activities themselves and how they are utilised for knowledge acquisition and decision-making patterns.

Overall, the above findings significantly contribute to our understanding of the information seeking approaches and behaviours in exploratory search, particularly within the context of conducting literature reviews. The results provide detailed insights into the actual approaches, the key exploratory characteristics employed during literature reviews, resource assessment methods, and the patterns of knowledge acquisition. This richness in exploratory search concepts and its multifaceted nature presents challenges. However, it also underscores the potential for effective support. Some implications of our findings pertain to the design of user support interfaces for exploratory searches. This paper encourages the research and design community to carefully consider the discussed findings when developing user interfaces intended to support users in conducting exploratory searches, especially literature review tasks. Given the diverse approaches and experiences revealed by the findings, interfaces should offer various support options to cater to users who follow different approaches or have different levels of experience.

Many of the current academic search engines and digital libraries typically display lists of papers with basic information like the title, authors, publication year, and brief snippets of the abstract. Based on the research findings, it would be beneficial to provide users with the option to access the full abstract, introduction, and conclusion with a single click before opening a document. This would allow users to quickly glance at these key sections and make informed decisions about whether they want to explore the paper further or not.

User interfaces should also aid users in handling the multi-item and multi-faceted nature of problems by presenting a visual map of various facets, topics, and sub-topics within a specific domain alongside a list of relevant papers. Additionally, support systems could enhance the iterative and systematic nature of the search process by introducing features that facilitate the exploration of additional papers or concepts related to the initial paper and allow users to track their progress effectively within the initial papers. Support interfaces also could incorporate features that cater to individuals' different approaches during their literature review process. However, since all researchers eventually aim to produce a written literature review, there could be interfaces specifically designed to facilitate note-taking and the construction of the review's structure and outline.

In conclusion, our main contribution of this chapter lies in conducting a comprehensive examination of the behaviours and information seeking patterns exhibited by 30 Master's students while engaging in a specific exploratory search task: literature review. One limitation of this work is the potential for different interpretations of the collected data and various perspectives on the findings. However, we adapted the recent conceptual model of exploratory search to the new context, establishing it as the core framework for understanding how people approach, behave, and interact during literature searches. We believe this framework provides a foundational basis for comprehending and studying exploratory searches in this specific context.

While this study is tailored to literature review tasks, further investigations should encompass a broader spectrum of exploratory search tasks. This would facilitate an assessment of the generalisability of the present results to various tasks. It is worth noting that while we believe the findings in this paper could be generalised to traditional or narrative literature reviews that are part of essays, dissertations, theses, or reports, this may not necessarily apply to other types of literature reviews in the medical domain or systematic reviews in general.

There is still a room to understand users' affections while working on conducting literature review tasks. Studying users' emotions during the search was beyond the scope of this study. However, it would be interesting to explore and connect these emotions to Kuhlthau's ISP model, particularly in the context of literature review tasks conducted in an online space. Additionally, we believe various factors may influence exploratory characteristics, such as experience and previous training in conducting literature reviews. Future studies should investigate other user-related factors that may impact exploratory behaviours.

Manuscript submitted to ACM

ACKNOWLEDGMENTS

We express our gratitude to all the participants who dedicated their time to take part in the user study. Special thanks to everyone involved in the development of the CRUISE-Literature application during the DoSSIER CRUISE-dev workshop, with a special mention to Wojciech Kusa and Óscar E. Mendoza. The work reported here is funded by the DOSSIER project under European Union's Horizon 2020 research and innovation program, Marie Skłodowska-Curie grant agreement No 860721.

REFERENCES

- [1] Kumaripaba Athukorala, Eve Hoggan, Anu Lehtiö, Tuukka Ruotsalo, and Giulio Jacucci. 2013. Information-seeking behaviors of computer scientists: Challenges for electronic literature search tools. *Proceedings of the American Society for Information Science and Technology* 50, 1 (2013), 1–11.
- [2] Kumaripaba Athukorala, Antti Oulasvirta, Dorota Glowacka, Jilles Vreeken, and Giulio Jacucci. 2014. Narrow or broad? Estimating subjective specificity in exploratory search. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. 819–828.
- [3] Lisa A Baglione. 2018. *Writing a research paper in political science: A practical guide to inquiry, structure, and methods*. Cq Press.
- [4] Marcia J Bates. 1989. The design of browsing and berrypicking techniques for the online search interface. *Online review* (1989).
- [5] Nicholas J Belkin. 1980. Anomalous states of knowledge as a basis for information retrieval. *Canadian journal of information science* 5, 1 (1980), 133–143.
- [6] Cecelia M Brown. 1999. Information seeking behavior of scientists in the electronic information age: Astronomers, chemists, mathematicians, and physicists. *Journal of the American Society for Information Science* 50, 10 (1999), 929–943.
- [7] David Budgen and Pearl Brereton. 2006. Performing systematic literature reviews in software engineering. In *Proceedings of the 28th international conference on Software engineering*. 1051–1052.
- [8] Harris M Cooper. 1988. Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in society* 1, 1 (1988), 104–126.
- [9] ELLIS David. 1989. A behavioural approach to information retrieval design. *Journal of Documentation* 45, 3 (1989), 171–212.
- [10] Shelda Debowski. 2001. Wrong way: Go back! An exploration of novice search behaviours while conducting an information search. *The Electronic Library* (2001).
- [11] Brenda Dervin. 1983. An overview of sense-making research: Concepts, methods, and results to date. (1983).
- [12] Brenda Dervin. 1998. Sense-making theory and practice: an overview of user interests in knowledge seeking and use. *Journal of knowledge management* (1998).
- [13] David Ellis. 1989. A behavioural approach to information retrieval system design. *Journal of documentation* (1989).
- [14] Carole George, Alice Bright, Terry Hurlbert, Erika C Linke, Gloriana St Clair, and Joan Stein. 2006. Scholarly use of information: graduate students' information seeking behaviour. *Information Research: An International Electronic Journal* 11, 4 (2006), n4.
- [15] Ahmed Hassan, Ryen W White, Susan T Dumais, and Yi-Min Wang. 2014. Struggling or exploring? Disambiguating long search sessions. In *Proceedings of the 7th ACM international conference on Web search and data mining*. 53–62.
- [16] Orland Hoerber, Dolinkumar Patel, and Dale Storie. 2019. A Study of Academic Search Scenarios and Information Seeking Behaviour. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*. 231–235.
- [17] Ching-Ting Hsin, Ying-Hsueh Cheng, and Chin-Chung Tsai. 2016. Searching and sourcing online academic literature: Comparisons of doctoral students and junior faculty in education. *Online Information Review* (2016).
- [18] Taraneh Khazaei and Orland Hoerber. 2017. Supporting academic search tasks through citation visualization and exploration. *International Journal on Digital Libraries* 18, 1 (2017), 59–72.
- [19] Jeffrey W Knopf. 2006. Doing a literature review. *PS: Political Science & Politics* 39, 1 (2006), 127–132.
- [20] Carol C Kuhlthau. 1991. Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science* 42, 5 (1991), 361–371.
- [21] Bill Kules and Robert Capra. 2008. Creating exploratory tasks for a faceted search interface. *Proc. of HCIR* (2008), 18–21.
- [22] Gary Marchionini. 2006. Exploratory search: from finding to understanding. *Commun. ACM* 49, 4 (2006), 41–46.
- [23] Lokman I Meho and Helen R Tibbo. 2003. Modeling the information-seeking behavior of social scientists: Ellis's study revisited. *Journal of the American society for Information Science and Technology* 54, 6 (2003), 570–587.
- [24] Sheshera Mysore, Mahmood Jasim, Haoru Song, Sarah Akbar, Andre Kenneth Chase Randall, and Narges Mahyar. 2023. How Data Scientists Review the Scholarly Literature. In *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*. 137–152.
- [25] Ya R Nedumov and Sergei D Kuznetsov. 2019. Exploratory Search for Scientific Articles. *Programming and Computer Software* 45, 7 (2019), 405–416.
- [26] David Nicholas, Paul Huntington, Hamid R Jamali, Ian Rowlands, and Maggie Fieldhouse. 2009. Student digital information-seeking behaviour in context. *Journal of Documentation* 65, 1 (2009), 106–132.
- [27] Xi Niu and Bradley M Hemminger. 2012. A study of factors that affect the information-seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology* 63, 2 (2012), 336–353.

- [28] Guy Paré, Marie-Claude Trudel, Mirou Jaana, and Spyros Kitsiou. 2015. Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management* 52, 2 (2015), 183–199.
- [29] Justin Paul and Alex Rialp Criado. 2020. The art of writing literature review: What do we know and what do we need to know? *International business review* 29, 4 (2020), 101717.
- [30] Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological review* 106, 4 (1999), 643.
- [31] Peter Pirolli and Stuart Card. 2005. 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. McLean, VA, USA, 2–4.
- [32] Diana Ridley. 2012. The literature review: A step-by-step guide for students. (2012).
- [33] Pamela Effrein Sandstrom. 1994. An optimal foraging approach to information seeking and use. *The library quarterly* 64, 4 (1994), 414–449.
- [34] Yi Shen. 2007. Information seeking in academic research: A study of the sociology faculty at the University of Wisconsin-Madison. *Information technology and libraries* 26, 1 (2007), 4–13.
- [35] Soumya Shukla and Orland Hoeber. 2021. Visually Linked Keywords to Support Exploratory Browsing. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*. 273–277.
- [36] Ayah Soufan, Ian Ruthven, and Leif Azzopardi. 2022. Searching the literature: an analysis of an exploratory search task. In *ACM SIGIR Conference on Human Information Interaction and Retrieval*. 146–157.
- [37] Nicole Sultanum, Christine Murad, and Daniel Wigdor. 2020. Understanding and supporting academic literature review workflows with litsense. In *Proceedings of the International Conference on Advanced Visual Interfaces*. 1–5.
- [38] Richard J Torraco. 2016. Writing integrative literature reviews: Using the past and present to explore the future. *Human resource development review* 15, 4 (2016), 404–428.
- [39] Pertti Vakkari. 2001. A theory of the task-based information retrieval process: A summary and generalisation of a longitudinal study. *Journal of documentation* (2001).
- [40] Doménique Van Gennip, Elise Van Den Hoven, and Panos Markopoulos. 2015. Things that make us reminisce: Everyday memory cues as opportunities for interaction design. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 3443–3452.
- [41] Ryen W White, Bill Kules, Steven M Drucker, et al. 2006. Supporting exploratory search, introduction, special issue, communications of the ACM. *Commun. ACM* 49, 4 (2006), 36–39.
- [42] Ryen W White, Gary Marchionini, and Gheorghe Muresan. 2008. Evaluating exploratory search systems. *Information Processing and Management* 44, 2 (2008), 433.
- [43] Ryen W White and Resa A Roth. 2009. Exploratory search: Beyond the query-response paradigm. *Synthesis lectures on information concepts, retrieval, and services* 1, 1 (2009), 1–98.
- [44] Barbara M Wildemuth and Luanne Freund. 2012. Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the symposium on human-computer interaction and information retrieval*. 1–10.
- [45] Fei Yu, Jan Sullivan, and Leith Woodall. 2006. What can students’ bibliographies tell us?—Evidence based information skills teaching for engineering students. (2006).
- [46] Xiaolong Zhang, Yan Qu, C Lee Giles, and Piyou Song. 2008. CiteSense: supporting sensemaking of research literature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 677–680.

A APPENDIX

Table 1. Codebook of the User Dimension

Soufan et. al's Explanation	Updated Explanation with Examples
<p>Unfamiliarity with the domain: Users are likely to employ exploratory search to learn about a new topic. Previous researchers claim that exploratory users might be unfamiliar or new to the domain they are searching in, have little knowledge about it, have insufficient expertise, or poorly understand the problem context's domain, resulting in more exploratory behaviors. Additionally, some researchers suggest that exploratory users might not know the right keywords beforehand, and they might gather information and study the topic of interest immediately in the process of search.</p>	<p>The "Domain_familiarity_unfamiliarity" code is applied to participants' comments on their level of familiarity with the domain in which they are searching. They may either comment on their lack of familiarity or their existing familiarity. This familiarity or lack of it may be expressed in terms of concepts, keywords, authors, sources, topics, etc. They may mention their knowledge, schol Example: "I have no knowledge on some of these concepts."</p>
<p>Uncertain about the goal: Users who engage in exploratory search might not have a precise goal in mind when they start, and the answer of the exploratory search may not be immediately apparent. Therefore, exploratory searches are typified by uncertainty about the space they search in and the nature of the problem that motivates the search. Additionally, exploratory users might seek different opinions on a topic and explore various aspects to ascertain an overview of a topic.</p>	<p>The "Goal_certainty_uncertainty" code is applied to participants' comments regarding the certainty of their search goals (answers, contents, structure, end results). They may indicate that they have uncertain or vague goals, are unsure about their end results, or that the answer they are seeking is not clear-cut. Additionally, they may comment on having clear goals and answers and being sure about their end results. Examples: "I don't know what I want, exactly".</p>
<p>Have fuzzy information need: Users' information needs might be unclear, ambiguous, or imprecise as exploratory users might not know the domain that they search in well or have a general interest but not specific knowledge of that domain. Therefore, users' keywords are a-priori unknown, vague, and keep evolving. Moreover, the exploratory problem context may be ill-structured, and users' search goal may not be apparent. Therefore, exploratory users might require additional information to clarify their goals.</p>	<p>The "Information_need_clarity_fuzzy" code is assigned to participants' comments when they discuss the clarity of their information needs (information/data they are seeking at some point of their search). This code is used when they express what they look for is vague, unclear, or imprecise. They may mention spending time exploring to understand what information is available in the domain. When users' comments indicate that they can't clearly describe the information that would fulfil their goal, it signifies a fuzzy information need. Conversely, some participants may indicate having a clear and precise information need, accompanied by knowing the right keywords to employ. Examples: "I just go through at least 10 papers before realizing what I need".</p>
<p>Have dynamic information need: Exploratory information seeking has a dynamic nature. For some researchers, the exploratory search process starts with inadequately explained search goals; therefore, exploratory users might submit tentative queries, selectively exploring and passively gathering cues about the next steps. As a result, users' knowledge and information need constantly change throughout the iterative search process. During the exploratory searches, users are likely to understand the problem context better; as a result, they tend to make more informed decisions about interaction or information use.</p>	<p>The "Information_need_decisive_dynamic" code is used for participants' comments about the nature of their information needs. It is applied when participants mention that what they seek evolves and changes as they engage in the search process. They may mention spending time refining their understanding, knowledge, and answers during the search. Participants may also mention using evolving keywords during their search queries to explore available information or that their output is developing as they search. Conversely, some participants may state that their information needs remain constant and precise throughout the search process. Examples: "As the research goes on, the main headings will change to cater to the research accordingly".</p>

Table 2. Codebook of the Problem Context Dimension

Soufan et. al's Explanation	Updated Explanation with Examples
<p>Open-ended (persistent): Open-endedness relates to uncertainty over the information available or incomplete information on the nature of the search task. The information need, the search goal, the problem context, the search process, and the search result of exploratory search are open-ended. The exploratory problem context, goal, and information need do not stay the same from the start to the end of the exploratory search. Additionally, having an open-ended problem context might lead to having fuzzy and dynamic information needs.</p>	<p>The "Open_close_ended" code is used for participants' comments on the problem and the answer. They may also highlight the existence of numerous potential answers. Also, they could point out the ongoing development of literature in their field, the vast expanse of information, the inherent incompleteness of searches, and the impossibility of containing all aspects of a topic. Conversely, some participants might indicate that they've thoroughly addressed all dimensions and obtained a comprehensive answer. Example: "There is always new papers in my topic, it keeps changing".</p>
<p>Multi-faceted: The exploratory problem context and the end result might include various aspects and different concepts. Besides, the exploratory problem context might incorporate multiple sub-tasks. Therefore, exploratory users might end up searching for information related to various aspects of the domain they are looking for in.</p>	<p>The "Multi_single_faceted" code is used for the text describing results that encompass multiple sub-tasks or various aspects/concepts of the explored domain. It is used for comments related to comprehensive exploration spanning various facets of the topic. Conversely, it may be used when participants mention a single-faceted answer, indicating a narrower perspective or isolated sub-task. Example: "I just start the search by breaking the topic into subtopics".</p>
<p>Multiple-item goal: Because the exploratory problem context might be open-ended, and multifaceted, a single target answer may not exist, and the target of the search is multiple items/documents. The final result may be an integration of different aspects of the domain. Moreover, the information goal is likely to be satisfied with a combination of information encountered during the search using multiple queries dedicated to addressing different aspects of a topic.</p>	<p>The "Multi_single_item" code is used for text describing results that encompass multiple documents/resources. It is also used for text describing an information goal that is likely to be satisfied through a combination of information encountered during the search, using multiple documents/resources. Conversely, it may be used when participants mention a single item answer using a single resource. Example: "I literature review results from reading different papers."</p>
<p>Ill-structured (Ill-defined): The problem context has imprecise task requirements. It also may remain undefined or in a significant flux for much of the search session. Therefore, users require additional information from external sources to clarify their goals and actions.</p>	<p>The "Structured_unstructured" code is used for text describing the LR task's requirements as unclear, imprecise, or subject to significant changes during the search session. They may indicate encountering challenges in understanding the task and seek support to navigate the unclarity and ambiguity. Thus, they may follow a search process that is on fly instead of a structured one. Conversely, the opposite is true for comments about structured problem context with clear requirements. Example: "I was not sure what to do or from where to start"</p>
<p>General rather than specific: The problem context is general, with a vague and under-specified description. Exploratory search tasks provide a low specificity about the information necessary for their search, finding the required information, and recognizing the needed information. The exploratory problem context might be widespread among different areas of the domain and consists of various aspects of the domain.</p>	<p>The "Generic_specific" code is applied to text that describes a task with a general and vague description and general answers. This code is used when participants' comments suggest that the LR task they are working on lacks specificity and clarity regarding the information needed. They may indicate encountering challenges in defining the scope of their search and identifying the specific aspects of the domain they should focus on. The opposite is true for comments about specific tasks with clear and well-defined information needs. Example: "My topic is very abroad, I can go to any direction"</p>

Table 3. Codebook of the Search Process Dimension

Soufan et al.'s Explanation	Updated Explanation with Examples
<p>Iterative: The exploratory search process starts with submitting tentative queries, selectively seeking and passively obtaining leads about the following steps, and iteratively searching with evolving information needs. The search process begins with an imprecise query, and then through several successive iterations of exploring the retrieved information and reformulating queries, the scope of the information need narrows down</p>	<p>The "Iterative_not_iterative_query" code is applied to text that reflects the iterative and evolving nature of the search. This code is used when participants' comments describe starting with a tentative or imprecise query and progressively refining their search through multiple iterations. Participants may mention following this iterative approach and gradually narrowing down the scope of their information needs to obtain more relevant and useful results. Conversely, the opposite is true for comments indicating a more direct and non-iterative search approach. Example: "I start with initial search, from what I get I search for specific topics"</p>
<p>Opportunistic: It can be defined as taking a greater risk on the premise of a bigger payoff. Since users don't have complete control over what to expect and do not precisely know what they are looking for, users' search process might be less direct. Users tend to select results that might have a higher opportunity to be beneficial for them. Moreover, users do not plan their next steps in advance but rather decide their next steps at each search stage.</p>	<p>The "Opportunistic_not_opportunistic" code describes text related to participants' willingness to take greater risks with the expectation of potentially obtaining valuable information from reading or searching for documents. Conversely, this code applies to comments indicating a search approach unwilling to take risks even if there might be valuable opportunities. Examples: "I just spent a good amount of time just reading through it, see what I could get from it, and then I use that to search more."</p>
<p>Unsystematic: Users who engage in the exploratory search process are unsure how to achieve their goals (either the technology or the process). While searching and browsing, users encounter new information and concepts of interest, generating additional needs and guiding the search to new directions. Users' exploratory search process might follow an unpredictable non-linear path during the search.</p>	<p>The "Systematic_unsystematic" code describes text indicating a lack of a systematic pattern or structured approach while searching. Participants may mention following an unpredictable non-linear path during the search and not adhering to predefined search strategies. Conversely, this code applies to comments reflecting a more systematic, structured, and methodical search approach. Example: "I start with the original paper. I'll branch out, but I keep coming back to the original."</p>
<p>Multi-tactical: Users might employ multiple search approaches and consult different sources throughout the exploratory search process. The exploratory search process is characteristic of the alternation and iteration of querying and browsing moods. Users might use other ways and systems to reach the wanted information. They also might use these systems in different ways to find more relevant information.</p>	<p>The "Multi_single_tactical" code describes text indicating that participants employ multiple search approaches, tactics and strategies to find the information they need. It applies when they mention the use of different systems, strategies (ways of finding documents), or tools to access relevant information and modify their search tactics based on emerging information needs. Conversely, comments suggesting a more singular and fixed approach to the search process. Examples: "I'm improving how I'm querying, sometimes I search for things that are not papers."</p>
<p>Long-term: Exploratory searches can take place over multiple sessions, and it can be long as hours, days, or even months.</p>	<p>The "Short_long_term" code is used to describe text related to the duration of the task that may take place over multiple sessions, and it can be long as hours, days, or even months. Examples: "Conducting literature reviews take a lot more than this."</p>