Falling Down the Rabbit Hole: Exploratory Search and The Case for Personal Information Environments

AMY RAE FOX, ROBERT KAUFMAN, and JAMES D. HOLLAN, University of California, San Diego

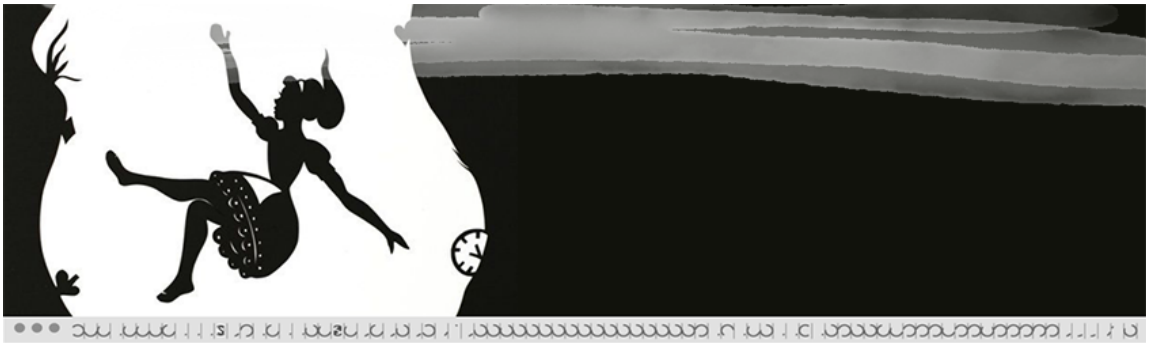

Fig. 1. New Yorker columnist Katheryn Schulz writes that, “In its most purely Carrollian sense...to fall down a rabbit hole means to stumble into a bizarre and disorienting alternate reality”[46] [illustration by anonymized]

Some describe the web browser as a window onto the world. In practice, it is more often a *rabbit hole* through which we fall. Despite twenty years of research on exploratory search yielding innovative algorithms, visualizations, and domain-specific tools, this pervasive information activity frequently remains frustrating and ineffective. To understand the troubling gap between research effort and user experience, we conducted a six-month participatory design engagement. Through survey, synthesis, activity-recording, and browser-logging, we observed how a common exploratory activity—literature review—is distributed through bespoke ecosystems of tools, time, physical and digital spaces. In turn, we offer a conceptual reformulation of “the problem of exploratory search” based on how modern web browsers are used as (impoverished) personal information environments. We contribute a descriptive dataset, restructured problem space, and argument for how understanding the situated context of exploratory search can yield progress on this old but persistent problem.

CCS Concepts: • **Human-centered computing** → *HCI theory, concepts and models*; **Empirical studies in HCI**; • **Information systems** → *Search interfaces.*

Additional Key Words and Phrases: Information Seeking & Search, Tasks/Interruptions/Notification, Information Spaces, Literature Review, Ethnography, Participatory Design

ACM Reference Format:
Amy Rae Fox, Robert Kaufman, and James D. Hollan. 2020. Falling Down the Rabbit Hole: Exploratory Search and The Case for Personal Information Environments. 1, 1 (December 2020), 19 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Authors’ address: Amy Rae Fox; Robert Kaufman; James D. Hollan, University of California, San Diego.

1 INTRODUCTION

In 2008, Yan Qu and George Furnas published an article in a special issue section of *Information Process and Management* devoted to the evaluation of Exploratory Search Systems (ESSs) [40]. At the time, their understated contribution joined a growing chorus of calls for a paradigm shift in the study of Information Retrieval: from focus on the search algorithm to the user's interactive processes. Starting from an established behavioural model [42], they performed formative evaluations to elaborate a space of design possibilities, concluding with three refinements to the model and corresponding design recommendations. First, they exposed the extraordinarily tight-coupling between a user's observable search behaviour and the construction of conceptual representations of what is found. Second, they found that as searchers develop representations to structure their growing knowledge, these ideas are sourced from both prior knowledge and newly-found information. Finally, they proposed that when a user encounters a structural-information need (i.e., a need for organization, rather than facts), strategies emerge to derive the necessary structure by conducting additional searches. In sum, Qu & Furnas contributed a plausible, empirically-grounded update to Russell, Stefik, Pirolli & Card's seminal model of sensemaking [42], and concluded that to support the *situated reality* of exploratory search, tools need to provide better support for: (1) expressing structural needs, (2) finding useful structures in the world, and (3) supporting the overarching context of task management.

Similar recommendations have been echoed in hundreds of exploratory search publications in the intervening years. So why do we draw your attention to Qu & Furnas? Because twelve years on, the most common domain-general exploratory search tools—the web browser and search engine—provide scarcely better support for this pervasive complex cognitive activity. The paradigm in Information Retrieval has shifted. Ample attention is being paid to both growing adaptation in search algorithms, and the situated behaviour of users. Nevertheless, users of the world's most popular web browsers and search engines are not reaping the rewards.

To complement the substantial body of empirical research in this area—particularly laboratory studies, formative and summative assessments of domain-specific tools—we set out to investigate the cognitive and affective experience of exploratory search in the broader context of activity in distributed information environments. We elected to concentrate on a personally-relevant genre of search: academic literature review. We convened an interdisciplinary group of graduate and undergraduate students, and over the course of six months, collectively examined our lived experience of *falling down the rabbit hole*. What emerged was two perspectives on this predominantly frustrating activity: (1) divergences—a set of factors implicating individual differences in behaviour, and (2) convergences—a set of common challenges we universally face.

We begin with a brief review of theoretical perspectives on exploratory search and sensemaking (2.1), and the distribution of information processing during complex cognitive activities (2.2). In (2.3), we examine recent proposals for *personal information spaces* and describe how the modern web browser serves as a de facto personal information environment (3). In Section (4) we introduce our research methodology, followed by the divergent (5) and convergent (6) factors that emerged from our data. We conclude with a discussion (7) of how our observations can contribute to the design and realization of personal information environments that provide more holistic support for these complex, interleaved information activities.

2 BACKGROUND

2.1 Exploratory Search and Sensemaking

The term *exploratory search* is nearly as complex, multifaceted and ill-defined as the activity to which it refers. Rather than a single action, it is a class of information-seeking behaviour, and as a research area grew out of the Information Science and Information Retrieval communities in the mid-2000s as a foil to traditional lookup-based models of search [57]. Perhaps most easily understood as 'not simple lookup', we perform exploratory searches when we have an information need, but also a gap in knowledge about the topic of interest [40]. This 'problem context' is ill-defined or open-ended [56], and is typically re-structured over a period of time, often spanning numerous possibly lengthy sessions. The adjective *exploratory* describes both the process by which the search occurs, as well as the context that motivates the user's action [33], although the term *search* does not necessarily characterize the user's navigation. In practice, exploratory search is realized through a combination of querying and browsing interactions. The most important factor in distinguishing the exploratory from other information-seeking activities is a user's evolving understanding of their own information needs, which is why models of exploratory search are largely built upon theories of *sensemaking*[1].

Sensemaking is a canonical complex cognitive activity [37], a 'meta-activity' that refers to combinations of numerous non-linear sub-activities, that we can only endeavour to precisely operationalize in the context of some task, in some situation, for some individual. As a construct it is *more than just* making a decision or forming a judgement or comprehending [29]. Its non-linear nature can be unravelled with methods like cognitive task analysis [42], and described in terms of information processing and the construction of representations [40]. In models of exploratory search, sensemaking fills the gap between making a query, and making *the next* query. It's how we evaluate the results of a search, decide what's relevant, retrieve some information, *construct knowledge*, and iterate, not necessarily in that order. White & Roth [57] argue that exploratory searches always involve some degree of sensemaking. It is the need for sensemaking that in part distinguishes a search as exploratory.

2.2 Supporting Complex Cognitive Activities

Russell & colleagues argue that "representation design is central to the sensemaking enterprise" [42, pg.8], and we add that this hinges on the complex interplay between internal and external forms [36, 44, 60], enacted in the body [28] distributed in the environment [19]. Sedig & Parsons [37, 47, 48] offer a conceptual model for exploring how computation is distributed through this environment. They describe five (metaphorical) spaces that together form a "human-information interaction epistemic cycle" [47, pg.93]. The *information space* consists of the set of information with which users might interact, and the *computing space* its storage and manipulation (i.e. machine computation). In the *representation space*, information is encoded and made available to the user's perception. [2] (The representation space is an abstraction, but is reified in computers as 'the interface'.) The *interaction space* affords exchange of information via action and perception: where the user performs actions and receives reactions. In the *mental space* the user constructs knowledge. We cannot overemphasize the importance of conceptualizing these spaces as metaphorical, and not simultaneously reifying the layers as physical systems with linear exchanges of information. In practice, processing emerges dynamically, simultaneously across the spaces, through time. This framework serves as tool for thinking about

[1]White & Roth [57] contextualize exploratory search in the greater landscape of research on Information Behaviour as a *type* of information seeking and *type* of sensemaking. They characterize information foraging [38] and berrypicking [4] as *classes* of Information Behaviour that describe how users find information and adapt to their information environments. Along with sub-disciplines like information visualization and interactive information retrieval, they position information foraging and berrypicking as relevant but not superordinate to exploratory search.
[2]Sedig & Parsons describe this in the context of visualization and visual perception, but is equally applicable to other sensory modalities.

how information and computation might be distributed across the cognitive system, in support of *complex cognitive activities* [37] (i.e. higher order processing).

2.3 Human-Centered Information Spaces

The term *information space* has also been used non-metaphorically to describe a particular conception of computational systems. In [20] Hollan & Satyanarayan offer a scenario for how the kind of 'personal information space' they envision would transform the workflow of a research scholar. They describe a multi-scale, zoomable interface where the researcher can access, "all the information (e.g., files, emails, messages, papers, web pages, notes, sketches, analyses, and visualizations) of her computer-mediated activities" [20, pg.2]. Their idealized system implies multiple dimensions of metadata, in that it affords interaction with the information *itself*, as well as the user's history of *interaction with* their information. They describe a symbiotic combination of a user with an intelligent-machine that responds to the user's past interaction and context of current activity to make recommendations or filter information, empowered with the ability to dynamically select appropriate representations. This vision is grounded in the claim that despite decades of computational evolution, our interaction with computers is, "too often difficult, awkward, and frustrating" [20, pg.1].

Fox & colleagues elaborate on this vision, arguing for "human-centered information spaces" as the catalyst for a paradigm shift in personal computing. They suggest that our interaction with computers is not as natural as it could be or should be, drawing on Ivan Illiach's notion of *conviviality* as an incitation of what is possible [22] . Their argument joins with the historical chorus questioning the prevailing flat, disconnected [8] static and depersonalized [26, 27, 53–55] presuppositions of information. In [12] they propose three architectural requirements for the design of such human-centered spaces. First, an escape from the silos of documents and applications. They argue that users must be able to access their information across applications, implying the need for integration, and/or instrumental techniques [5, 6, 30]. Secondly, they require that users be able to flexibly choose or construct their own (external) representations of information, arguing that to support higher order thinking, the user needs the representations and contextually-appropriate subset of interactions that suit the structure of their thinking at any given time. Finally, they propose that in order to support the user's integration of tasks across activities and working sessions, any such system must be responsive to the context of the user's past and present activity.

3 PERSONAL INFORMATION ENVIRONMENTS, THE BROWSER, AND EXPLORATORY SEARCH

Building on this vision, and applying Sedig & Parsons' model[3] of distributed information processing [48], we reconstrue the "human-centered information space" as a personal information *environment*. This shift in terminology is deliberate: both to prevent further overloading the spatial metaphor, and acknowledge that any such system is endeavouring to integrate across document and application boundaries, between devices and through time. To be certain, we are referring to yet another metaphorical space—a sort of 'desktop' for intellectual activity—but one likely to be realized via a series of interfaces, instruments, and integration services.

One can imagine performing a literature review in such an environment. You'd be relived from the cognitive load of remembering which of the thirty-five right-click open-in-new tabs contained promising papers. Your search engine of choice might be aware of your history of past searches, and know that when you query on 'learning' you mean (human) learning and be clever enough to filter out (machine) learning resources. As you browsed the search results, the way you interacted with the interface would capture, infer, and take action on the tacit judgements and decisions you were

[3]Although their model is presented in the context of visual analytics, we argue that Sedig & Parsons' ideas apply more broadly to design of any system for human-information interaction.

making in each moment. Not just 'this paper looks relevant open in new tab'—but how relevant (giving it a priority), to which part of your goal (giving it structure), and how do you intend to act on it (adding it to the *top* of the queue).

We aspire for the texture of this user experience, but are grounded in the reality of our present tools. Through this participatory research project we've learned that—like us—the Exploratory Search System leveraged by the average student is no system at all. Rather, a bespoke ecosystem of PDF-readers, note-taking apps, notebooks, post-its and to-do's on the backbone of the browser and search engine. In this play the web browser takes center stage, providing the interfaces and functional representations for perceiving, processing, organizing and acting on search results. In this sense, we argue that the modern web browser is serving as a personal information environment for exploratory (literature) search. It is *through* the browser that we fall down the rabbit hole. But as we've found, our web browsers, sitting aside partner programs in their application silos, are not collaborating on the challenging tasks we face.

A Brief Aside On Rabbit Holes. This couldn't be a paper about rabbit holes without a brief aside about rabbit holes. New Yorker columnist Katheryn Schulz writes that, "In its most purely Carrollian sense...to fall down a rabbit hole means to stumble into a bizarre and disorienting alternate reality"[46]. Of course that's not the way this turn of phrase has found its resurgence in the modern lexicon. Shulz suggests that our online rabbit holes ("which breed rabbit holes the way rabbits breed rabbits") come in three essential forms. The first is an incremental chain of progressive distraction that starts with an intention, interrupted by a simple distraction (like finding a reference to a paper on tabbed browsing) and ends some time later in wonder at 'where has the time gone' (for us with two hundred and fourty-five papers on tabbed browsing). Schulz describes the second sort as 'exhaustive', when you start out with the intention of learning about some topic, but end up with a compedium of information much greater than intended that you're not entirely able to make sense of. These are *exactly* the rabbit holes we are referring to.

Her third account is 'associative': when you look up one thing, leading to something distantly related, and then something further afield, and so on and so forth. These rabbit holes, often generated by instigating websites (e.g., Wikipedia), have an analogous construct in the Information Science literature: exploratory browsing. Exploratory browsing is differentiated from exploratory search in that it involves an *intentional* dive into uncertainty, with some overarching intention; perhaps seeking inspiration, or in the case of Wikipedia, general knowledge.

At the bottom of this particular rabbit hole is the insight that Carol's geologic metaphor has been adopted by the masses to refer to both deliberate and accidental descents into both captivating and confusing states of being. At present we are concerned with the *latter*: when the needs of the searcher exceed the support provided by the search ecosystem, and the result is an unsupported feeling of free fall, ripe with cognitive overload, fatigue, disorientation, and confusion leading to inefficient and often ineffective exploration.

4 RESEARCH METHODS

Our project was structured as a participatory design engagement where, alongside the research team, a group of participants collectively constructed a dataset of individual search experiences, analytical reflections, and design ideations. We drew upon methods from multiple research traditions, including cognitive ethnography [19], action research [17], and participatory design [9]. In such qualitative contexts the relationship between *the researcher* and *the researched* is both consequential and complex. As our object of study (the literature review) is one in which we are deeply implicated, we judged it impractical to maintain the impartiality of observers, and instead adopted the participant-observer duality of action research [17] with student-scholar activity as the site of inquiry.

Our objective was to better understand the process of scholarly literature reviews. We assembled a group of participants based on their interest in this topic, and proceeded with the work in three phases (Figure 2). The first phase centered on understanding the nature of literature review as an exploratory search activity, and how it occurs embedded within information work. The second centered on documenting the problems faced when performing literature reviews, with attention to the factors that seem to drive individual differences in experience. In the third phase, we turned to prioritizing and structuring these problems into a design space of potential solutions, with a focus on common challenges. Throughout the engagement, the group completed a series of individual task-assignments[4] designed to elicit search behaviours, and convened on a weekly basis for discussion, synthesis and reflection.

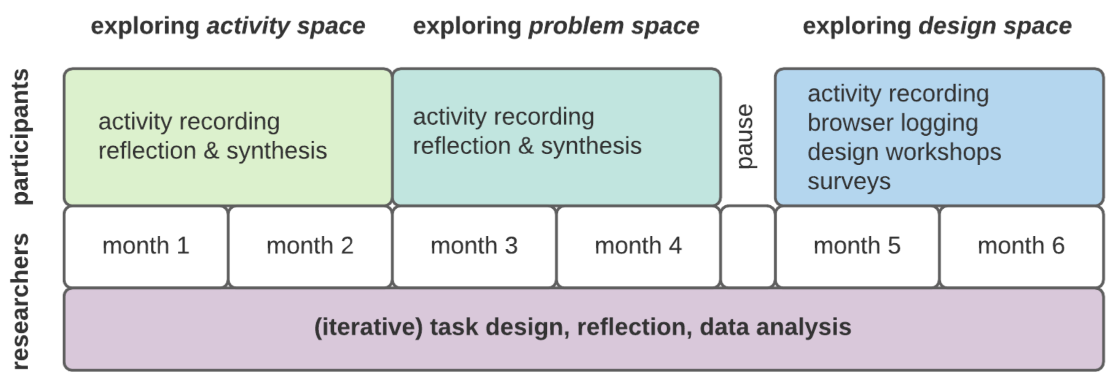

Fig. 2. Our engagement included three phases over a six month period.

4.1 Participants

Over the six month engagement, five students from our university collaborated with the authors in a series of tasks ranging from auto-ethnographic reflection to collaborative solution design. By convention we refer to these individuals as *participants*. The group was comprised of three PhD students, one masters student, and three upper division undergraduates, ranging in age from 19 to 38 (3 female, 4 male). Student topics of study included psychology, neuroscience, computer science, and human-computer interaction.

4.2 Tasks

4.2.1 Activity Recording. We leveraged the affordances of screen-recording software to capture the temporal and spatial dynamics of exploratory search activities "in the wild", and exploited the rich context of this corpus of activity data in multiple ways.

As an orienting exercise, in Phase One we asked participants to complete literature reviews on assigned topics relevant to our work (e.g., tabbed browsing behaviour, information overload, reinstating mental context). The goal was to generate 2-3 page research summaries for the group. The task took between 2-5 hours to complete over the course of one week, and participants recorded all of the working sessions it required, noting interruptions. The following week, participants reviewed their recordings and reflected in an auto-ethnographic fashion before convening to discuss

[4]Task instructions and data analysis guidelines can be found in our Supplemental Materials.

and derive insights as a group. After viewing our own recordings, each participant performed a qualitative content analysis of the sessions for someone else in the group (in a round-robin fashion). This analysis was guided by a set of orienting questions and coding scheme (developed by the researchers) [5] centering on three components: tools, activities, and strategies. Given the perspective of performing and reviewing our own activity, collective reflection, followed by analysis of another's activity, we spent the remainder of Phase One examining the similarities and differences between individual searchers and between sessions within a task, culminating in a set of factors we believe influence individual differences in the exploratory search experience. (Section 5).

In Phases Two and Three, participants completed ad hoc literature reviews on topics of their choice. Everyone was encouraged to explore the literature and tools relevant to the thread of our collective inquiry at the given time, and to record these sessions whenever possible. A subset of these recordings were coded by one participant and one researcher according to a finer-grained coding scheme. The results of these analyses informed a deeper understanding of the temporal dynamics of tasks within exploratory search, and are discussed in Section 5.3.

4.2.2 Browser Logging. In our collective synthesis of activity recordings, we reflected on a tension experienced when reviewing search results: to continue exploring the result set, or begin exploiting individual results. A preference for when to stop exploring and start exploiting seemed to represent a stable diverging factor between individuals across search sessions. It proved onerous to explore this hypothesis via activity recording data which requires manual annotation of browser events, so we developed a browser logging utility implemented as a Chrome extension.[6] This tool logs all tab, window, and webNavigation API events, and was used in conjunction with activity recording during Phases Two and Three to shed light on organizing behaviours during search, especially how participants use windows and tabs to organize search results and indicate activity status (Section 5.3).

4.2.3 Collaborative Synthesis. The most important and enjoyable component of our project was our working sessions. These were structured as reflections on our individual activities from the prior week. Our approach is firmly situated in the theoretical stance that knowledge is actively constructed through the representation and re-representation of ideas [28, 36, 44, 52, 59, 60]. Throughout our research we leveraged collaborative digital white-boarding software [7] to capture and share the external representations of our collective inquiry. The crux of this work was deconstructing metaphors like "tab hell", and "rabbit hole" into specific challenges with potential causes and effects. This processing transpired *in situ*, via lively discourse, with participants brainstorming ideas on the board, and others structuring and continuously re-structuring them into categories, themes, stages and dimensions as we progressively elaborated the (vast) landscape of our chosen problem. The structure of this paper is grounded in these organizing systems.

4.2.4 Design Workshops. Our engagement culminated in two, half-day participatory design workshops. These sessions were separated by 1-month and took place remotely via video conference[8]. The purpose of the workshops was to summarize our "common challenges" and iterate potential design solutions. During each workshop, we cycled between group, individual and breakout sessions aimed at diagramming and structuring. After prioritization, we turned to creative ideation and solution design. Although discussion of design solutions is beyond the scope of this paper, we share our common challenges in Section (6).

[5] See Supplemental Materials
[6] Project [anonymized] can be found on GitHub at [anonymized] and in our Supplementary Materials
[7] Miro: https://miro.com
[8] due to COVID-19

5 DIVERGENCE: INDIVIDUAL DIFFERENCES IN SEARCH BEHAVIOUR

To the best of our knowledge (and exploratory search ability) the study of individual differences in exploratory search is in its infancy [35]. Certainly much can be learned from the related study of individual differences in Personal Information Management (ie. [7, 15, 23]), where there is emphasis on how attitudes and preferences contribute to the selection of tools and strategies. We believe these preferences, especially organizational tendencies like 'piling vs. filing' or 'browsing vs. searching' are likely to influence exploratory search in similar ways. In the initial weeks of investigation, our participants were united in frustrated fascination with the phenomenon of lit review rabbit holes, sharing stories of 'tab hell' and browser crashes. But as we scratched beneath the shared context of our activity-recordings, differences emerged: in our strategies, tools, attitudes and feelings toward our experiences of search. In the sections that follow, we describe three salient dimensions along which behaviours diverged.

5.1 Attitudes, Affect, and Maladaptive Behaviour

We used surveys to collect baseline data on two constructs we believed were likely to to drive individual differences in exploratory search behaviour. To assess *information literacy* we used the 28-item Information Literacy Self-Efficacy Scale [49] (reliability $\alpha = 0.92$). Participants were asked to rate how often a statement was true (on a 7-point Likert scale): "I feel competent and confident... [performing information skill]". Individual scores ranged from 4.7 to 6.4, with a sample mean of 5.34 (SD = 1.60, n = 7), corresponding to an average response between "often true" and "usually true". Interestingly, the rank ordering of individual scores *did not* conform to the age, years of research experience, or program level of the participants; some undergraduates expressed higher self-efficacy than graduate students. To assess *the degree of distress associated with feelings of too much information*, we used the 15-item Information Overload Scale [58] (reliability $\alpha = 0.90$). Participants were asked to rate how strongly they agreed with a statement: "I feel ...[emotion] when [circumstance]" on a 5-point Likert scale. Individual scores ranged from 2.8 to 4.7, with sample mean 3.96 (SD = 1.26, n = 7) indicating a moderate experience of information overload [9].

We found greater variance between individuals in our group with respect to *information overload* (an affective experience) than *information literacy* (a subjective assessment of skill). These data comport with the narrative descriptions of skill and affect early in our engagement. Participants rarely expressed doubt or negativity toward their own abilities to manage information. Yet our conversations were ripe with lively discussion about our emotional experience while searching. Some participants described their frustrations or struggles with search with a detached affect: focusing attention on the tool, or design, or analytical justification for why something seemed to work or not work. Others described a highly emotional experience, with references to "information anxiety", feelings of being overwhelmed and even distressed. We spent much time discussing the behavioural responses that resulted from these feelings. Asked how they might feel about losing tabs in a crashed browser session, one participant expressed horror, others dismay, some relief, and others indifference. Two participants described frequently querying and evaluating search results, but not *processing* the found resources. One student described organizing search results into windows based on search topic, and using a session management browser extension to "save" the results, in essence *squirrelling the information* for another time. The same student described feeling like they were *hoarding information*, in that it was being saved but not consumed. Three students described routinely avoiding either their devices, or web browser, for feelings of guilt or anxiety at the number of open but un-visited tabs. As researchers, we started to refer to these narratives as *maladaptive information behaviours* drawing analogy to the concept of maladaptive behaviour in clinical psychology,

[9] Both instruments can be found in Supplemental Materials.

where an individual responds to distressing emotions in a way that is self-defeating. We propose that there are likely information-behavioural equivalents of many maladapative responses, making this a promising area for interdisciplinary research with potential impactful applied interventions.

5.2 Ecosystems

We use the term *ecosystem* to describe the distribution of tools (digital and physical) through spaces (physical and digital) in the user's environment. With the privileges of choice and ready access to computational technology, we intentionally assemble our ecosystems for information work, often in highly individualized ways. Though anyone who's struggled through a tool-choice appreciates many of the decisions we make are *satisficing* ones, and the tools assembled to support the whole are not necessarily the most suitable for any particular part. Tool choice is a complex problem, and it is reasonable to expect that choices are both influenced *by* individual histories and differences (such as attitudes and organizational preferences) and exert influence *on* individual differences (such as strategies).

When performing literature review, all seven in our group relied primarily on personal laptops, running either MacOS or Windows. All activity recordings were made on laptops, though participants discussed occasionally reading papers or 'managing' activities (to-do lists, scheduling) on their phones, particularly while riding public transportation. Some routinely used both laptop and tablet devices when searching at home or school; using the laptop for querying and organizing, and the tablet for reading and note-taking.

Participants tended to prefer Google Chrome as their primary web browser, and we all turned to Google Scholar as the literature search engine of choice.[10] In only one of 25 recordings did a participant use an alternate search engine (the PsychInfo database). While Chrome's popularity may not be surprising, we did find it interesting that we universally turned to Google Scholar as our single point of access, despite frustration with its limited metadata display and features for query refinement. Some were aware of more literature-specific search engines such as Web of Science, and the ACM digital library. Two students reported that they would use the (new) ACM Digital Library if looking for literature in a specific conference series. Despite offering advanced options for query-refinement and filtering, no one was willing to use Web of Science. In comparison to Google Scholar, the interface was described as "clunky", and the page-to-page response times not sufficiently performant.

While we were consistent in our choice of browser and search engine—the foundation of the search experience—we saw variation in the supporting cast of tools used for reading and note-taking, moving beyond note-taking and 'synthesizing' or integrating ideas and information, tools for organizing documents, and planning or managing our tasks. Through discussion, we learned that the choice of these tools was highly personal, and influenced by both what tools were already a part of the individual's 'standard suite' of software, as well as their openness to investigating new tools and processes for Personal Information Management. Tool choices were influenced by convenience (which applications came standard with the OS), and requirements for classes or recommendations from friends and colleagues, in addition to the more obvious requirements for features. Although we do not draw casual inferences from our observations, we do hypothesize a correspondence between variance in tools and the level of abstraction of the cognitive activities a tool was used to accomplish. We used one search engine for posing queries, and one browser for inspecting results, five different tools for reading, but ten different tools for *synthesizing*—thinking about what had been read and constructing new ideas. Individuals held strong preferences, sometimes articulated as 'compulsions', toward certain representational

[10]Possibly much to the chagrin of research librarians.

features for this activity, such as mind or concept mapping, free-form drawing, canvases for pasting and organizing, or hierarchical text.

We believe a relationship between tool diversity and complexity or abstraction of cognitive activity is a plausible one. The backbone of sensemaking is the construction of (mental) representations—the knowledge that is derived from the information [40, 42]. But this knowledge comes about both through the scanning, evaluation and reading of the returned information resources *as well* as the tacit behaviours the user employs: such as externalizing words or diagrams that specify structure, or implicitly indicating structure through the organization of information resources. As a cognitive activity scales in complexity (contrasting say, *reading* with *designing*) it is reasonable for the number of different representations that might be used to support that activity to scale as well. The tools the user chooses for the task drive not only the causal outcomes of their task performance (i.e. by measures of information retrieval), but more importantly, both influence *and* are influenced by how the individual prefers to *think with representations.*

5.3 Strategies

Our coding and subsequent discussions of activity-recording data revealed differences in our deliberate strategies for search and exploration, as well variance in the behaviours employed to achieve them. We use the term *strategy* to refer to a number of choices an individual makes, including what tools to use, how much time to allocate, how 'broadly' and 'deeply' to explore the search space, and when to stop searching. We expected to see differences in strategies based on the context of the search task (e.g., what is the required outcome?) and prior knowledge of the task domain. This proved to be the case, as participants made fewer queries and seemed to explore the results less systematically when doing ad-hoc literature reviews, compared to recordings when they were tasked with preparing a summary on a specific topic. For some participants, this difference influenced the tools they chose to accomplish the task. When the perceived demand for systematic understanding was high, additional tools such as concept maps and reference management software were employed.

A question of substantial interest was the explore-exploit tradeoff: How does one decide when to stop browsing results, and start consuming the (found) resources? The answer to this question depended first on the quality of any individual query. If after a simple scan the results seemed not relevant, another query would be executed. When the query returned some relevant results, however, the answer was less straightforward. Some participants reported that they, 'typically browse at least [x] pages of results' regardless of the context or query. These folks seemed driven by some perceived minimum requirement of exhaustiveness, even with the awareness that the relevance of results does not scale linearly across the number of results pages, either within or across queries. Some participants were not able to introspect on this question, instead reporting that they thought they would browse until the results stopped seeming relevant (a strategy predicted by [38]).

In coding our activity data, we sought to identify and characterize the temporal dynamics of discrete components of the search activity that were readily observable. We used a directed approach to content analysis [21] that allowed us to design a coding scheme first by drawing on the canonical stages of information seeking in electronic environments identified by Marchionini [32], and then update the scheme, grounded in properties of the empirical data. This hybrid approach yielded a taxonomy of sub-tasks we observed exploratory literature search, which is fully described in our Supplemental Materials.

Our content analysis of these data are still ongoing, but one interesting pattern we've observed is divergence in the sequence of three particular sub-tasks: reviewing search results (i.e. looking at the results page(s), evaluating resources (i.e. visiting the url for a particular result) and consuming them (i.e. reading). We've identified three patterns

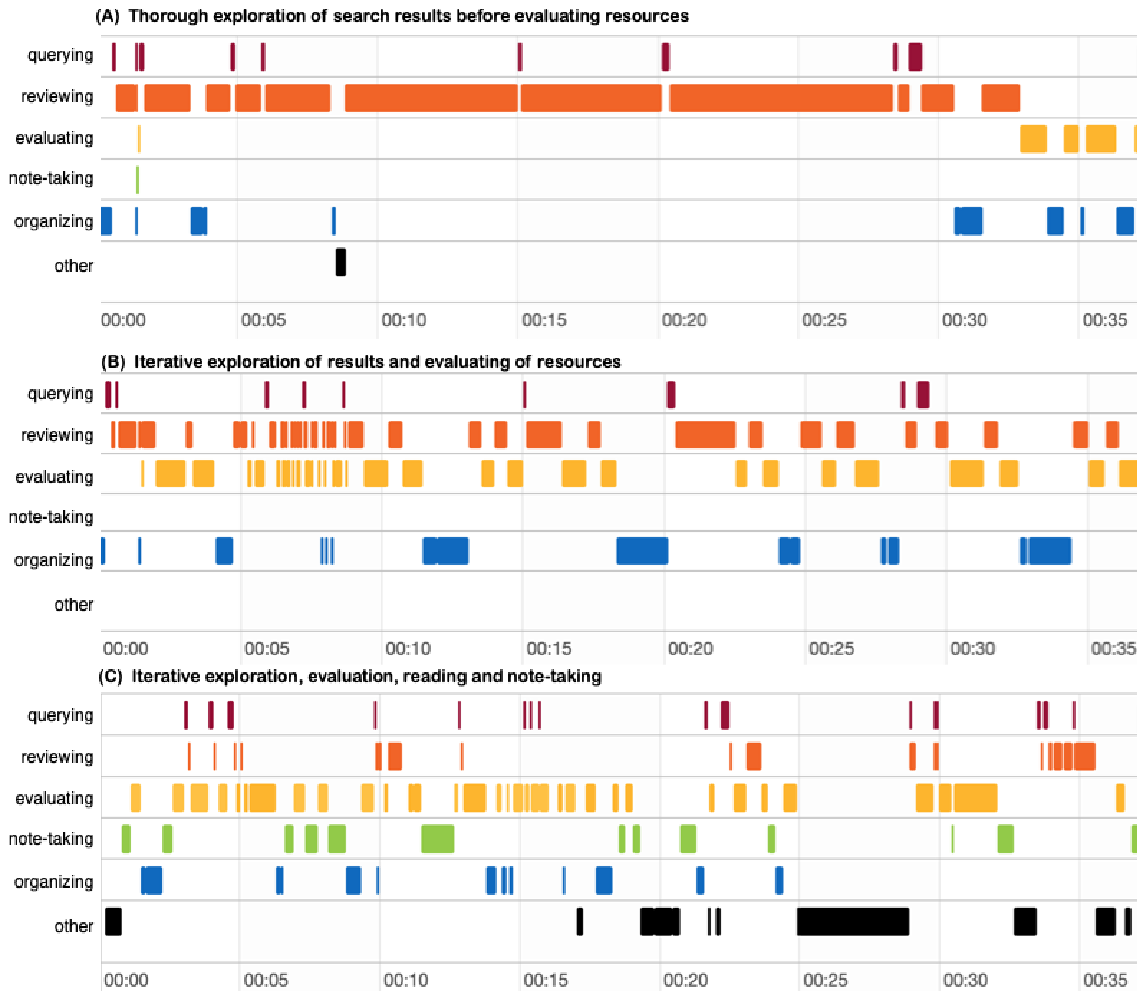

Fig. 3. The content analysis of sub-tasks revealed three patterns of sequential activity. These are examples drawn from participants' data.

in these activities that correspond to different locations on an explore-exploit continuum. In the first (Figure 3A) the searcher spends long spans of time reviewing the results page. This participant reported the most salient experience of information overload, as well as a desire to thoroughly explore the results of every query. In the recording, we saw this as activity on the results page (scrolling and opening resources in new tabs) for nearly 30 minutes before the first resource was visited. The second pattern (3B) corresponds to a more iterative dynamic between reviewing individual results and evaluating the corresponding resources. Tabs were visited shortly after they were opened. The third signature represents a further expansion of this cycle, where before proceeding to review the next result, the resource is read, and perhaps notes are taken. In future work, we plan to use our logging extension to explore this activity at scale to determine variablity and frequency.

The final substantial strategy difference we noted was the number of browser tabs and windows participants had open at any given time. Although everyone described challenges with managing tabs, analysis of the logging data showed very different patterns of tab and window use. All participants reported using windows as task-semantic organizing structures. A specific activity—like literature search for our project—might start in its own browser window, and split into multiple windows in order to organize the results the user wants to visit during their next working session, or save for a future date. Figure (4) depicts the structure of windows and tabs for 5 participants over one 24-hour period that included at least one exploratory search session. The high variance in number of tabs per window is consistent with the self-report that windows are often used as semantic organizers, though the total number of windows and tabs are highly variant.

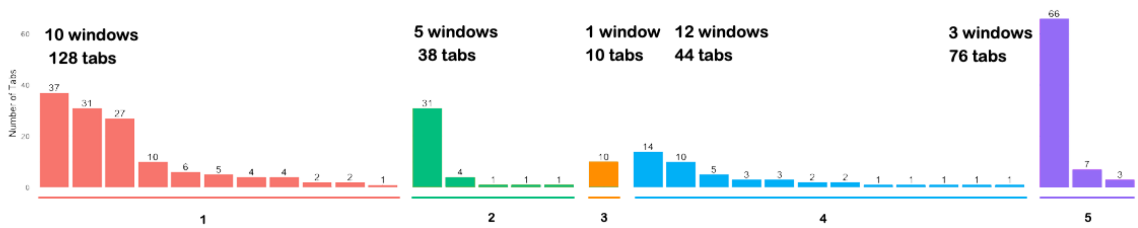

Fig. 4. Browser logging data revealed a large distribution in number of windows and tabs across 5 users in a 24 hour period. Colours differentiate the five participants. Each bar represents a unique window, and its height, the maximum number of tabs in that window.

6 CONVERGENCE: COMMON CHALLENGES IN SEARCH EXPERIENCE

Although we were at times surprised by the ways our search experiences diverged, we did share common search challenges. Through analysis and discussion of activity recordings, we enumerated the common challenges we experienced performing literature reviews (Figure 5). When exploring the structure of this space, we discovered a multi-faceted classification problem: some challenges could be categorized by when in the (iterative) process they occur, while others transcended all stages. Some pertained to the skill or knowledge of the searcher, while others centered on 'lower level' but equally disruptive concerns like navigation, or moving information between applications. Here we take a crosscutting approach, discussing: (1) challenges arising from the inherent difficulty of the exploratory search process, (2) challenges associated with sensemaking, (3) difficulties with tools, (4) difficulties integrating search in the context of other activities, and (5) our common affective experience.

6.1 Challenges Inherent to Search

Many challenges we shared concerned the specific gaps in knowledge that make exploratory search such a rich area of research. Graduates and undergraduates in our group alike were plagued by not knowing: (1) what we don't know, (2) what search terms to use, (3) what to do next, and (4) when to stop searching. These challenges were evident in our conversations, and also in our activity data, demonstrated by frequent 'shots in the dark' for search terms that yielded few relevant results, and simultaneous parallel queries that revealed inadequate organizational structures for the desired information. At times this manifested as not knowing what to do next, followed by taking the easiest available path rather than decomposing the topic into the more manageable parts that would be evident to someone with greater prior knowledge. Together, we believe these challenges constitute the intrinsic cognitive load [14] of exploratory search: the things that make search hard, no matter what tools you have at your disposal.

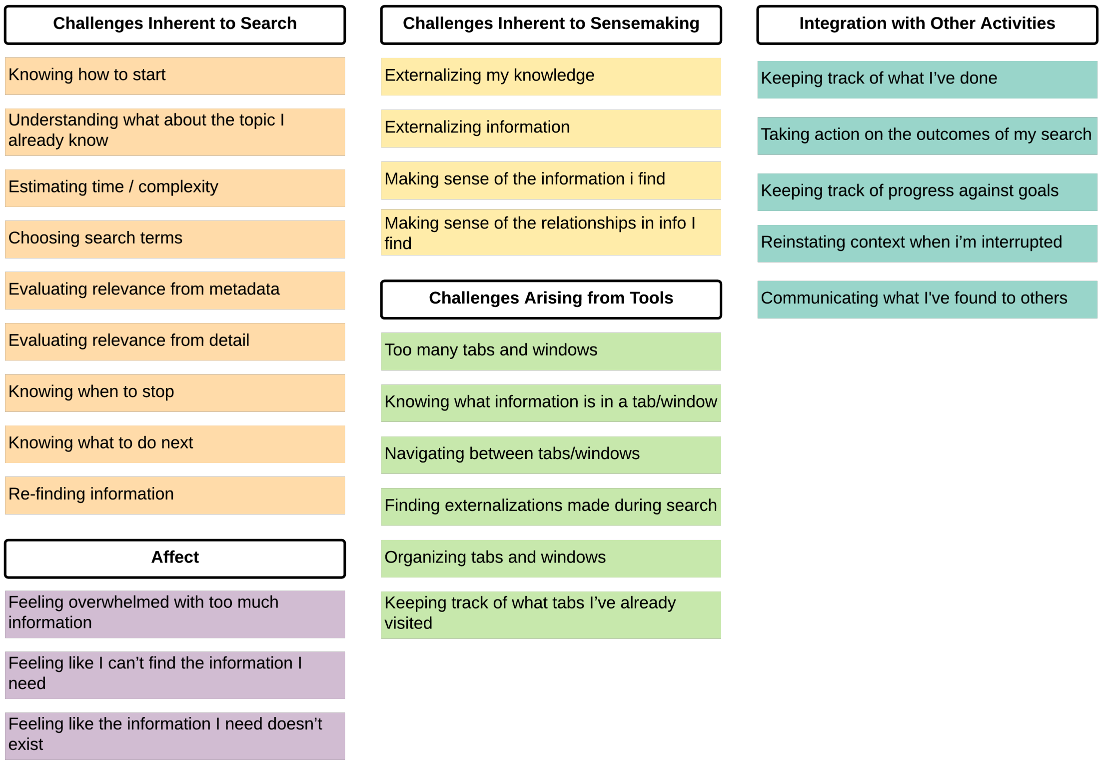

Fig. 5. Specific challenges we face during literature review.

In prior work researchers have endeavored to reduce this intrinsic load with recommendation algorithms and alternative representations like faceted and aspect browsing. Most relevant to the problem of literature review, Zhang et al. contributed the exploratory search system Citesense, addressing the common issue of searchers not having an initial structure for their exploration. Citesense helped searchers locate relevant work by clustering resources into hierarchical categories. These categories could be refined as the searcher learned more about a topic, and new documents would be suggested on the basis of these refinements [61]. Similarly, the Apolo system used machine learning to generate and cluster relevant resources, then identifying further potential relations to any particular cluster via its citation network [10]. Systems like Citesense and Apolo are especially helpful after the first few documents are deemed relevant, although they do not provide any direct scaffolding for determining appropriate initial search terms. Recommendation on the basis of citation analysis is also a topic of interest in the bibliometrics and scientometrics communities, and some support for reference-driven recommendation has found its way into Reference Management Software like ReadCube Papers[11].

[11] www.papersapp.com

6.2 Challenges With Sensemaking

While there is limited utility in differentiating between what is and is not sensemaking during exploratory search, we draw a line here to suggest what challenges might be addressed primarily through innovation in interfaces and visualization. It was no surprise that participants experienced challenges integrating newly found information with prior knowledge: sensemaking is hard. But we did develop a sense of exasperation with the fact that participants' earnest attempts at integrating information and constructing knowledge were too-often thwarted by inadequate representational tools. Individuals in our group expressed universal difficulty in tracking their sense-making processes, including goals, current and desired states, so that as explorations evolved, they could be reminded of and build upon their own ideas. The first route of externalization was note-taking, but many reported dissatisfaction with note-taking infrastructure, as application silos that often precluded taking notes within the application in which an insight originated. Further difficulty finding and navigating to notes meant they were often out of date, or had organizational structures that required manual revision. In general, participants struggled with externalizing information in efficient ways, and reported that until they "saw the information in a particular way," they weren't able to organize it or make sense of it.

Considerable prior work has focused on helping individuals organize and make sense of found information. Dumais, Jones, and colleagues' early work on keeping found information found is a seminal example [24, 25]. Searchers often express a desire to organize information into meaningful groups to help with the sensemaking process [18]. The Scatter/Gather system created by Pirolli et al. clusters documents based on summaries that can be high level or more specific, allowing searchers to traverse from big picture to details to help understand the relational structure of the information space [39]. Similarly, Jigsaw shows the connections between informational entities within and between particular documents in order to aid seeing new perspectives and sensemaking [51]. Network visualizations in particular show promise to assist with sensemaking, and can be leveraged to represent multiple dimensions of structure in tasks such as literature review [11, 13, 44, 45].

6.3 Challenges Arising From Tools

Participants frequently complained of having too many windows and tabs. Despite all it's power, the design of the browser itself can challenge exploratory search efficiency. Participants had difficulty knowing what was in tabs, with organizing them, locating the 'right' one at any particular moment, and at times found themselves aimlessly cycling between tabs having forgotten what they were trying to return to in the first place. One participant referred to bouts of 'tab hell': having dozens of tabs with potentially-relevant resources; their location, evaluation status, and future intent held *only* in fragile working memory. Participants also remarked that the visual representation of browser tabs offered insufficient cues to what they contained, and wanted features for organizing or even labelling. Moving findings and notes between applications was also particularly troublesome, contributing to disorientation, and loss of information or insights. We argue that this class of challenges introduce *extraneous* cognitive load to the search process: load that is not strictly required by the task, but serves to interfere with its effective completion. Our design efforts should be aimed at minimizing these sources of extraneous load so that any available processing capacity can be devoted to the constructive crux of sensemaking.

Challenges associated with windows and tabs have been well documented. In a survey of browser users conducted by Kulkarni et al., 50.7% of respondents "considered tab clutter a problem" [31]. The Vivaldi web browser is unique in offering configurable layout and navigation options for tabbed browsing, including the ability to stack tabs, manage on a sidebar instead of at the top of the window, search tabs, and view active tabs side-by-side in a split screen [2]. Yet, it hasn't become popular.

6.4 Integration Across Activities

Participants reported several challenges related to the meta-processes of tracking and managing (1) where they've been, (2) what they've done, (3) why they performed an action, and (4) what they feel they should do next. Participants described losing track of the past activities they had performed, such as how deeply a particular search result had been examined, what relevance determinations they had made, and changes to their organizational structure, hindering their ability to see their own progress, derive meaning from the changes, reflect, refine, justify, and verify. Losing track often resulted in a loss of progress and the need to re-do past work. Participants also felt unsupported in bridging the gap between their current organizational structure (what they know) and the structure of information in the world (what they need to find) while keeping track of overall goals. This was especially challenging when trying to reconcile information at different levels of detail.

There have been multiple attempts to use tools and visualizations to help searchers know where they've been and to save documents for later use. The most popular methods are browser-based search histories and bookmark functions. However, Kulkarni et al. found that 50.3% and 61.3% did not find history and bookmarks, respectively, useful for refinding information [31]. To make history more useful, Mayer and Bederson created Browsing Icons to break a search into user-defined tasks and visualize a searcher's history of websites previously visited and paths from a repeatedly visited site [34]. The search results page of the Bento Browser also allows users to save and keep track of their progress on any given search, enabling history within tasks and sub-tasks to be viewed [16]. Finally, Shrinivasan et al.'s visualization system included a history tree that allowed users to visualize the exploratory search process itself. Their system provides multiple phased views for additional control and the opportunity to fine-tune the features of each phase based on specific needs [50].

Reinstating context and getting back on track after following a tangent, taking a break, or being interrupted were difficult for participants. A lack of support for saving context requires holding information in short-term memory, and participants expressed difficulty remembering search progress, goals, prior activities, future intentions, and, perhaps most important, their insights after even minor distractions. This resulted in needing to re-do work and somtimes to giving up on a search. Rapid shifts between high-level overviews and low-level details, such as when moving from entities in an organizational structure to reading a particular document, was disorienting and especially troublesome when paired with task-switching. Re-finding previously seen information and remembering if a specific piece of information had already been found often resulted in wasted time and the need to re-do past work.

To address the challenges related to reinstating context, Rajamanickam et al. created TabFour, an example of a tool to reinstate an interrupted task by allowing users to annotate tasks and associated found documents [41]. Workona [3] and Session Buddy [1], two commercially-available Chrome extensions, help users recontextualize by allowing them to make workspaces for task-based browsing, save tabs, and search through the tabs they've saved. This allows users to reopen a previous browser environment and maintain context between search sessions. In addition, there is functionality to automatically save open tabs and windows in case of a technical interruption. Sarrafzadeh details work on Hierarchical Knowledge Graphs, an alternative representation for exploratory search results designed to help users stay oriented while moving between levels of detail [43]. Scatter/Gather and Jigsaw, mentioned above, both help searchers understand relational structure by showing them overviews of information with the ability to see details [39, 51]. This type of functionality can also help users remain oriented while traversing levels of abstraction.

6.5 Affect

Staying focused during search was challenging for participants as they often needed to divide attention between activities, tasks, and managing browser infrastructures. They reported feeling overwhelmed managing search tasks (e.g., keeping track of tabs, remembering goals and notes) while trying to process new search results. Some expressed feeling a need to process *all* the information they found immediately rather than focusing on the most salient information. This was described by one participant as stemming from "information FOMO" (fear of missing out) and the worry they'd neglect a particular aspect of a topic and thus might need to re-do work later. Participants often created a "backlog" of documents that was increasingly overwhelming to process. The one universally agreed on sentiment was the feeling that the exploratory literature review was much harder than it needed to be.

7 A WAY FORWARD

> *"Regardless of where the exploration takes place, it is clear that more computational resources will be devoted to exploratory search and the next search engine behemoths will be the ones that provide easy to apply exploratory search tools that help information seekers get beyond finding to understanding and use of information resources." —Marchionini, 2006 [33, pg.45]*

In 2009, following a series of workshops and conferences reflecting growing attention to exploratory search, White & Roth shared a set of unequivocal requirements for exploratory search systems [57] :

(1) support querying and query refinement
(2) offer facets and metadata-based result filtering
(3) leverage search context
(4) offer visualizations to support insight and decision-making
(5) support learning and understanding
(6) facilitate collaboration
(7) offer histories, workspaces and progress updates
(8) support task management: allow users to store, retrieve, and share search tasks in support of multisession and multiuser exploratory search scenarios

Twelve years later, our data suggest that the tools commonly used to support our participants' scholarly activity, and we would argue that of most researchers, offer scant if any support for these requirements beyond queries and query refinement. Exploratory search remains primitive but why is this the case?

Are we using the wrong tools? Why are we using Google Scholar for literature review, when more powerful tools exist? Is the feature-light simplicity of the interface the reason it is the first option to which we turn? Perhaps Google Scholar offers *just enough* functionality (e.g., 'cited by' and search within 'cited by') to be useful. Still it is perplexing that there are such glaring oversights as the inability to search by a custom date range and then sort by date. Is the future one of just adding small features to augment individual tools, or are we building the wrong tools?

Are we building the wrong tools? Marchionini's 2006 prediction of "search engine behemoths" with exploratory search tools has certainly not arrived. Why, given the pervasiveness and importance of exporatory search, are there not widely used domain general Exploratory Search Systems?

Perhaps these challenges are not for the web browser or search engine to solve. We argue that the underlying issue is more fundamental: the unquestioned presupposition of a document/applications-centered design paradigm. This focuses development on separate applications rather than on developing integrative personal information environments.[12] The central argument for personal information environments is that in order to adequately support the complex cognitive processing required of modern information work—in the human-centered convivial fashion we deserve—a system must afford the user access to information integrated across applications, beyond documents, with flexible re-representation in the context of current activity. Exploratory search is exactly the type of activity that is likely to remain unreasonably unruly and difficult in the absence of such an environment.

The genesis of our investigation was an ongoing long-term ethnographic study of scholarly information workflows. While interviewing interdisciplinary graduate students, we observed a common theme of frustration and dissatisfaction when scholars discussed their processes for literature review. In stark contrast to the well-reasoned, sophisticated, and highly-structured workflows that were described for data collection and analyses, scholars reported haphazard, chaotic and cluttered procedures for literature review. Scholars seemed universally dissatisfied, some even embarrassed, at the disorganization of their document collections, with reading and note-taking distributed across various tools and devices in varying states of completion. This was the catalyst for our own descent into the exploratory search literature, and the realization that innovations in recommendation algorithms and alternative visualizations in the Information Retrieval community were not making their way into the domain-general tools used by information-seekers on a daily basis. We began with the intuition that the most impactful gap in support for the average exploratory search was in its *situated context*—that is, the meta-level activity in which the search is embedded. *We end, convinced more than ever that the way forward is designing overarching personal information environments.*

REFERENCES

[1] 2020. Session Buddy. https://sessionbuddy.com/
[2] 2020. Vivaldi Browser. https://vivaldi.com/
[3] 2020. Workona. https://workona.com/
[4] Marcia J. Bates. 1989. The design of browsing and berrypicking techniques for the online search interface. *Online Review* 13, 5 (May 1989), 407–424. https://doi.org/10.1108/eb024320
[5] Michel Beaudouin-Lafon. 2000. Instrumental Interaction: An Interaction Model for Designing post-WIMP User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. ACM, New York, NY, USA, 446–453. https://doi.org/10.1145/332040.332473
[6] Michel Beaudouin-Lafon. 2017. Towards Unified Principles of Interaction. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter (CHItaly '17)*. ACM, 1:1–1:2.
[7] Ofer Bergman and Steve Whittaker. 2016. *The Science of Managing Our Digital Stuff.* MIT Press.
[8] Tim Berners-Lee, James Hendler, and Ora Lassila. 2001. The Semantic Web. *Scientific American* 284, 5 (2001), 34–43. http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21
[9] S. Bodker and M. Kyng. 2018. Participatory design that matters—Facing the big issues. *ACM Transactions on Computer-Human Interaction* (2018), 4:0–4:31.
[10] Duen Horng Chau, Aniket Kittur, Jason I Hong, and Christos Faloutsos. 2011. Apolo: interactive large graph sensemaking by combining machine learning and visualization. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. 739–742.
[11] Richard Cox. 1999. Representation construction, externalised cognition and individual differences. *Learning and instruction* 9, 4 (1999), 343–363.
[12] Amy Rae Fox, Philip Guo, Clemens Nylandsted Klokmose, Peter Dalsgaard, Arvind Satyanarayan, Haijun Xia, and James D. Hollan. 2020. Towards a dynamic multiscale personal information space: beyond application and document centered views of information. In *Conference Companion of the 4th International Conference on Art, Science, and Engineering of Programming ({Programming} '20)*. Association for Computing Machinery, New York, NY, USA, 136–143. https://doi.org/10.1145/3397537.3397542
[13] Nitesh Goyal, Gilly Leshed, and Susan R Fussell. 2013. Effects of visualization and note-taking on sensemaking and analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2721–2724.

[12] Of course there are commercial tool suites that provide some integration across component applications but most information typically remains encapsulated within applications and documents. In addition, we often need to use applications outside of commercial tool suites or collaborate with people who use different tool suites. Even the more open web environment evidences acceptance of the unquestioned presupposition of a document/applications-centered design paradigm.

[14] Jacek Gwizdka. 2010. Distribution of cognitive load in web search. *Journal of the American Society for Information Science and Technology* 61, 11 (2010), 2167–2187.

[15] Jacek Gwizdka and Mark Chignell. [n.d.]. Individual Differences. In *Personal Information Management*. University of Washington Press, Seattle and London, 206–220.

[16] Nathan Hahn, Joseph Chee Chang, and Aniket Kittur. 2018. Bento browser: complex mobile search without tabs. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.

[17] Gillian R. Hayes. 2014. Knowing by Doing: Action Research as an Approach to HCI. In *Ways of Knowing in HCI*, Judith S. Olson and Wendy A. Kellogg (Eds.). Springer, New York, NY, 49–68. https://doi.org/10.1007/978-1-4939-0378-8_3

[18] Marti A Hearst. 2006. Clustering versus faceted categories for information exploration. *Commun. ACM* 49, 4 (2006), 59–61.

[19] James D. Hollan and Edwin L. Hutchins. 2009. Opportunities and Challenges for Augmented Environments: A Distributed Cognition Perspective. In *Designing User Friendly Augmented Work Environments: From Meeting Rooms to Digital Collaborative Spaces*, Saadi Lahlou (Ed.). Springer.

[20] James D Hollan and Arvind Satyanarayan. 2018. Designing Cognitively Convivial Physics for Dynamic Visual Information Substrates. In *Rethinking Interaction: From Instrumental Interaction to Human-Computer Partnerships. CHI 2018 workshop.*

[21] Hsiu-Fang Hsieh and Sarah E Shannon. 2005. Three approaches to qualitative content analysis. *Qualitative health research* 15, 9 (Nov. 2005), 1277–88. https://doi.org/10.1177/1049732305276687

[22] Ivan Illich and Anne Lang. 1973. Tools for conviviality. (1973).

[23] William Jones. [n.d.]. *Building a Better World with Our Information: The Future of Personal Information Management, Part 3.* Morgan & Claypool Publishers.

[24] William Jones. 2007. *Keeping Found Things Found: The Study and Practice of Personal Information Management: The Study and Practice of Personal Information Management.*

[25] William Jones, Harry Bruce, and Susan Dumais. 2001. Keeping Found Things Found on the Web. In *CIKM '01: Proceedings of the tenth international conference on Information and knowledge*. 119–126.

[26] Alan Kay and Adele Goldberg. 1977. Personal Dynamic Media. *Computer* 10, 3 (1977), 31–41.

[27] Alan Kay and Adele Goldberg. 1988. A History of Personal Workstations. New York, NY, USA, Chapter The Dynabook: Past, Present, and Future, 249–264.

[28] David Kirsh. 2010. Thinking with external representations. *AI and Society* 25, 4 (2010), 441–454. https://doi.org/10.1007/s00146-010-0272-8

[29] Gary Klein, Brian Moon, and Robert R. Hoffman. 2006. Making Sense of Sensemaking 1: Alternative Perspectives. *IEEE Intelligent Systems* 21, 4 (July 2006), 70–73. https://doi.org/10.1109/MIS.2006.75

[30] Clemens Nylandsted Klokmose and Michel Beaudouin-Lafon. 2009. VIGO: Instrumental Interaction in Multi-surface Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, 869–878.

[31] Minchu Kulkarni, Kshitij Kapoor, Deva Surya Vivek Madala, Sanchit Bansal, and Sudheendra Hangal. 2019. Compartmentalizing web browsing with Sailboat. In *Proceedings of the 10th Indian Conference on Human-Computer Interaction*. 1–8.

[32] Gary Marchionini. 1995. *Information seeking in electronic environments.* Cambridge University Press, New York, NY, US.

[33] Gary Marchionini. 2006. Exploratory Search: From Finding to Understanding. *Commun. ACM* 49, 4 (April 2006), 41–46. https://doi.org/10.1145/1121949.1121979

[34] Matthias Mayer and Benjamin B Bederson. 2003. *Browsing icons: a task-based approach for a visual Web history.* Technical Report.

[35] Heather L. O'Brien, Rebecca Dickinson, and Nicole Askin. 2017. A scoping review of individual differences in information seeking behavior and retrieval research between 2000 and 2015. *Library & Information Science Research* 39, 3 (July 2017). https://doi.org/10.1016/j.lisr.2017.07.007

[36] Stephen E. Palmer. 1978. Fundamental Aspects of Cognitive Representation. In *Cognition and Categorization*, E. Rosch and B. B. Lloyd (Eds.). Erlbaum, Hillsdale, NJ, 259–302.

[37] Paul Parsons and Kamran Sedig. 2014. Distribution of Information Processing While Performing Complex Cognitive Activities with Visualization Tools. In *Handbook of Human Centric Visualization*. 671–691. http://link.springer.com/10.1007/978-1-4614-7485-2

[38] Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological Review* 106, 4 (1999), 643–675. https://doi.org/10.1037/0033-295X.106.4.643

[39] Peter Pirolli, Patricia Schank, Marti Hearst, and Christine Diehl. 1996. Scatter/gather browsing communicates the topic structure of a very large text collection. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 213–220.

[40] Yan Qu and George W. Furnas. 2008. Model-driven formative evaluation of exploratory search: A study under a sensemaking framework. 44, 2 (2008), 534–555. https://doi.org/10.1016/j.ipm.2007.09.006

[41] Mohan Raj Rajamanickam, Russell MacKenzie, Billy Lam, and Tao Su. 2010. A task-focused approach to support sharing and interruption recovery in web browsers. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. 4345–4350.

[42] Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. 1993. The Cost Structure of Sensemaking. In *Proceedings of InterCHI.*

[43] Bahareh Sarrafzadeh. 2020. Supporting Exploratory Search Tasks Through Alternative Representations of Information. (2020).

[44] M Scaife and Y Rogers. 1996. External cognition: how do graphical representations work? *International Journal of Human-Computer Studies* (1996), 185–213. http://www.sciencedirect.com/science/article/pii/S1071581996900488

[45] Mike Scaife and Yvonne Rogers. 2005. External cognition, innovative technologies, and effective learning. *Cognition, education and communication technology* (2005), 181–202.

[46] Kathryn Schulz. 2015. The Rabbit-Hole Rabbit Hole. https://www.newyorker.com/culture/cultural-comment/the-rabbit-hole-rabbit-hole

[47] Kamran Sedig and Paul Parsons. 2013. Interaction Design for Complex Cognitive Activities with Visual Representations: A Pattern-Based Approach. 5, 2 (2013), 84–133.

[48] Kamran Sedig, Paul Parsons, and Alex Babanski. 2012. Towards a Characterization of Interactivity in Visual Analytics. 3, 1 (2012), 12–28. https://doi.org/10.1145/0000000.0000000

[49] S. Serap Kurbanoglu, Buket Akkoyunlu, and Aysun Umay. 2006. Developing the information literacy self-efficacy scale. *Journal of Documentation* 62, 6 (Jan. 2006), 730–743. https://doi.org/10.1108/00220410610714949

[50] Yedendra Babu Shrinivasan and Jarke J van Wijk. 2008. Supporting the analytical reasoning process in information visualization. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1237–1246.

[51] John Stasko, Carsten Görg, and Zhicheng Liu. 2008. Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization* 7, 2 (2008), 118–132.

[52] Barbara Tversky. 2011. Visualizing Thought. *Topics in Cognitive Science* 3, 3 (2011), 499–535. https://doi.org/10.1111/j.1756-8765.2010.01113.x

[53] Bret Victor. 2011. A Brief Rant on the Future of Interaction Design. http://worrydream.com/ABriefRantOnTheFutureOfInteractionDesign/

[54] Bret Victor. 2016. Stop Drawing Dead Fish. https://vimeo.com/64895205

[55] Bret Victor. 2019. Dynamicland. http://dynamicland.org//

[56] Ryen W. White, B Kules, S Drucker, and m.c. schraefel. 2006. Supporting Exploratory Search: Introduction to special section. *Commun. ACM* 49, 4 (2006), 36–39.

[57] 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. https://doi.org/10.2200/S00174ED1V01Y200901ICR003

[58] Jeanine Williamson, P. E. Christopher Eaker, and John Lounsbury. 2012. The Information Overload Scale. *Proceedings of the American Society for Information Science and Technology* 49, 1 (2012), 1–3. https://doi.org/10.1002/meet.14504901254

[59] Jiajie Zhang. 1997. The Nature of External Representations in Problem Solving. *Cognitive Science* 21, 2 (1997), 179–217. https://doi.org/10.1207/s15516709cog2102_3

[60] Jiaje Zhang and DA Norman. 1994. Representations in Distributed Cognitive Tasks. *Cognitive science* 18, 1 (Jan. 1994), 87–122. https://doi.org/10.1207/s15516709cog1801_3

[61] 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.