Situating Search

Chirag Shah
chirags@uw.edu
University of Washington
Seattle, Washington, USA

Emily M. Bender
ebender@uw.edu
University of Washington
Seattle, Washington, USA

ABSTRACT

Search systems, like many other applications of machine learning, have become increasingly complex and opaque. The notions of relevance, usefulness, and trustworthiness with respect to information were already overloaded and often difficult to articulate, study, or implement. Newly surfaced proposals that aim to use large language models to generate relevant information for a user's needs pose even greater threat to transparency, provenance, and user interactions in a search system. In this perspective paper we revisit the problem of search in the larger context of information seeking and argue that removing or reducing interactions in an effort to retrieve presumably more relevant information can be detrimental to many fundamental aspects of search, including information verification, information literacy, and serendipity. In addition to providing suggestions for counteracting some of the potential problems posed by such models, we present a vision for search systems that are intelligent and effective, while also providing greater transparency and accountability.

CCS CONCEPTS

• **Information systems → Users and interactive retrieval**; **Language models**.

KEYWORDS

Search models; Language models; Information Seeking Strategies

ACM Reference Format:
Chirag Shah and Emily M. Bender. 2022. Situating Search. In *Proceedings of the 2022 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '22), March 14–18, 2022, Regensburg, Germany*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3498366.3505816

1 INTRODUCTION

This paper is concerned with online search as a prominent and widely used method for information seeking. This process typically involves a user expressing their information need using keywords or a question, and the system returning a ranked list of information objects. In addition to such a prototypical case of query-document matching, several other methods exist that attempt to connect a user with the potentially useful information as quickly and effectively as possible. Examples include passage retrieval [58], question-answering systems [39], and dialogue or conversational systems [56]. We observe a trend towards valuing speed and convenience and ask: Is getting the user to a piece of relevant information as fast as possible the only or the most important goal of a search system? We argue that it should not be; that a search system needs to support more than matching or generating an answer; that an information processing system should provide more ways to interact with and make sense out of information than simply retrieving it based on programmed in notions of relevance and usefulness. More importantly, we argue that searching is a socially and contextually situated activity with diverse set of goals and needs for support that must not be boiled down to a combination of text matching and text generating algorithms.

In this perspective paper we examine a couple of new proposals, which we refer to collectively as *the Google proposals* since they stem from Google, which involve closed-off systems that could generate relevant text in response to a user's queries, aiming to leverage large amounts of data and language models (LMs). We argue that such approaches miss the big picture of why people seek information and how that process contains value beyond simply retrieving relevant information. Beyond the critique of a few proposals, this paper offers a broad perspective of how search systems and society have evolved with each other and where we should go from here.

We begin by describing and examining the essential ideas of proposals by Metzler et al. [49] and Google in the next section (§2). We summarize why we believe these proposals are flawed in technical and conceptual terms. To better understand these flaws and to envision better systems, we need to examine search as an information seeking activity embedded in specific social and technological contexts. Therefore, we take a step back in Section 3 to understand how search has evolved over the past few decades and how it should support several important informational activities. We then analyze the Google proposals with the lens of information seeking strategies (ISS) framework in Section 4, examining how and where such proposals fail to cover the broader goals of information seeking. In Section 5, we provide recommendations for two scenarios: first, a set of 'guardrails' that should be put in place for any deployment of language-model-based search agents and second, an alternative vision for what a future of search could look like. We conclude the perspective paper in Section 6 with a summary of our critique, our recommendations, and points for discussion by the broader information retrieval (IR), human-computer interaction (HCI), and natural language processing (NLP) communities.

2 RETHINKING SEARCH

In 'Rethinking Search', Metzler et al. [49] propose a vision for the future of search which builds on today's large language models

CHIIR '22, March 14–18, 2022, Regensburg, Germany

ACM ISBN 978-1-4503-9186-3/22/03…$15.00
https://doi.org/10.1145/3498366.3505816

and imagines users who wish for search engines to function as 'domain experts' able to answer their questions directly, rather than as tools for finding documents which may contain the information sought, or for other types of interactions with information. At GoogleIO 2021, Sundar Pichai demoed LaMDA "a language model for dialogue applications [...] designed to converse on any topic."[1] The demo did not give many details on how LaMDA is constructed, but it appears to be very much in the same vein as Metzler et al.'s proposal in terms of envisioning fulfilling information needs with dialogue agents. Another Google announcement in the same vein is the blog post[2] by Google VP of Search, Pandu Nayak, about the system called MUM ('Multitask Unified Model'), again framed as a step towards answering questions the way an expert would. The benefits of this system, as articulated in the blog post, center around relieving the user of needing to submit multiple queries as they seek to carry out some task.

The ideas in [49], LaMDA and MUM are not one unified proposal and furthermore the details of each are sketchy ([49] because it is a broad vision for future directions rather than specific tech, the other two because the only sources are a demo and a blog post). Nonetheless, we see certain key commonalities between them: All seem to involve an interface designed around a conversational agent and to rely at their core on the recent advances in large language models. We will refer to these three collectively as *the Google proposals*.

Metzler et al. list a dozen challenges that need to be addressed to achieve the vision they propose: (1) training models to learn associations between terms (or sequences of terms) and documents; (2) training a single model that can handle different types of search activities, including keyword queries, questions, locating documents related to an input document, and summarization; (3) extending the few-shot/zero-shot learning setups enabled by language models (LMs) to these search activities; (4) generating responses that are authoritative, transparent, unbiased and written in accessible style, handle diverse perspectives fairly, and include citations to the underlying corpus; (5) addressing the fact that LMs run as generators make things up;[3] (6) building models that are capable of reasoning over structured information (e.g., arithmetic, geography, time); (7) training models that combine multiple modalities; (8) training models that can leverage structural relationships within and between documents; (9) working for and across multiple languages; (10) addressing the challenges of scale; (11) developing systems for incremental learning so that the model changes as the corpus does; and (12) ensuring that the models are interpretable and controllable.

Metzler et al. is presented as a vision paper, suggesting a research direction and laying out sub-projects within that direction. As such, it is entirely appropriate for the paper to explore areas which are characterized by unsolved problems. However, we think that the vision as presented in [49] is fundamentally flawed in two respects:

Technical flaws. First, it is flawed *technically* in that it is based on misconceptions of the technical capabilities of language models. For example, point (6) is best understood not as an 'unsolved problem', but rather a category error. Nothing in the design of language models (whose training task is to predict words given context) is actually designed to handle arithmetic, temporal reasoning, etc. To the extent that they sometimes get the right answer to such questions is only because they happened to synthesize relevant strings out of what was in their training data. No reasoning is involved [8, 47]. Similarly, language models are prone to making stuff up (see (5) above), because they are not designed to express some underlying set of information in natural language; they are only manipulating the form of language.

In brief, the argument from [8] is that a machine learning model cannot possibly learn what is not in its data, and the data for language model does not provide the machine with any signal it can use about meaning. Languages are systems of signs (pairings of form and meaning [20]). Once a person or other agent has acquired that system, they can use the form to reconstruct meaning, but the acquisition requires access to both. Thus, while the distributional information absorbed by language models can make them extremely useful components of larger systems, the fact that it also enables them to generate seemingly relevant and coherent text does not make them trustworthy sources of information — even as sounding conversational makes people more likely to trust them [1, 24].

Conceptual flaws. Secondly, and more importantly, we see this proposal as flawed *conceptually* in terms of its vision for how technology should support human information seeking behavior. Metzler et al. begin their conclusion with the claim that their paper "envisions an ambitious research direction that doubles down on the synthesis between modern IR and NLP to deliver on the long-promised goal of providing human expert quality answers to information needs" (p.17). But they provide no citations nor other evidence that anyone has been asking for such a system. We thus read this 'promise' as not coming in response to a demand from users, but rather as a technologist's dream.

Nonetheless, we share the sense with [49] that a key step in guiding research processes is imagining possible futures and then working backwards towards what problems need to be addressed to achieve them. At the same time, we think it is particularly important for that imagining to be informed by not only what is and is not currently possible technologically but also by scholarship about how technologies, especially once scaled, are affecting people. Thus we are mindful of Noble's [54] discussion of the importance of public (rather than private) control of information systems and of Benjamin's [9, p.168] call to build tools with the goal of "engender[ing] liberation".

With this understanding of the goals of technological envisioning in mind, in the next sections, we will review what the information retrieval literature tells us about how and why people engage in search and discuss possible pitfalls of dropping a language-model-based dialogue agent into such situations, especially as a one-size-fits-all search solution.

3 SITUATING SEARCH WITHIN SOCIETY AND TECH DEVELOPMENT

Searching for relevant information or knowledge has been a critical activity for individuals and societies for centuries. Universities, libraries, and other places of knowledge have been at the core of

[1] https://www.youtube.com/watch?v=_xLgXIhebxA, accessed 13 Sep 2021
[2] https://blog.google/products/search/introducing-mum/, accessed 20 Sep 2021
[3] Metzler et al. refer to this, surprisingly, as a side-effect of 'reasoning-like capabilities'. In fact, LMs are not even making things up. They are only generating plausible sounding strings; any meaning in those strings is actually imbued by the reader [8].

development of many civilizations as they provide not only information, but also services and expertise to access that information [34, 41]. In the past few decades, such repositories have become available through online access with a new set of tools and methods, namely search engines with keywords, short phrases or questions as a way to access relevant information. As the amount of information produced and made available online has increased dramatically, these tools and services have evolved in their ability to capture, store, and serve information. On the other hand, the users of these services have also changed how they use the systems, what they expect in return, and what makes them satisfied [44, 45, 75]. The question to consider now is how should these services and the usage patterns they support develop next? Should the systems provide more or different ways to interact with information? Should they focus on reducing cognitive load of users by offloading some of their thinking or decision-making? Should the users develop better literacy with respect to the tools for accessing information or expect these tools to become more amenable to their current practices?

3.1 Search and society shape each other

What is a good search system? The answer to this question in any given era has followed the state of society and how it envisioned itself advancing. Consider Vannevar Bush's Memex [14], from the post-WWII era, contemporaneous with the rise of office work and workers increasing need to deal with large amounts of information. Bush envisioned a system based on an office desk, where workers would receive and process information, and transformed the desk into a system that could store and retrieve meaningful information at a large scale. Since Memex, many other proposals, visions, and models for search have come and gone. Some have taken root in our society, whereas some were forgotten so quickly that hardly anyone remembers (case in point: SearchMe[4] — a visual search system). Success or failure, these ideas and visions reflect the current state of society and how we think it would benefit from technological advancements. In addition, we argue that just as everyone would have a different idea about how to advance society, the visions for search systems (and for that matter, any automated decision system) are often tied to one's own beliefs and strengths. We need look no further than some of the Salton Award winners and their keynotes at SIGIR conferences in the last decade or so.

Sue Dumais (2009 winner) advocated for an interdisciplinary approach to addressing search problems that involves not only developing intelligent systems, but also a deeper understanding of human cognition and interactions. Norbert Fuhr (2012 winner) contrasted search systems with database systems and proposed to address information object needs as well as task needs. Nicholas J. Belkin (2015 winner), a strong proponent of interactive IR, envisioned how we could build search systems that incorporate utility of information to the user rather than objective relevance only. Kalervo P. Jarvelin (2018 winner) emphasized how important it is to understand and model the context in which the information interactions take place, in order to serve the information seekers. Most recently, ChengXiang Zhai (2021 winner) presented a view of search systems where the notion of 'intelligence' has been shifting from system-centered to user-centered. These are all different perspectives on the future of IR systems from visionary scholars.

We can also look at related communities such as CHIIR, RecSys, The Web Conference, and WSDM, and find visions for search systems that range from new neural models to innovative ways for artificial agents to engage in natural language conversations. Unsurprisingly, these views reflect the particular expertise and career contributions of the distinguished scholars who expressed them. Accordingly, they often seem to involve extrapolation from past and current research trajectories rather than normative reasoning about the interaction of search and society.

In this article, we seek to understand better (i) how visions for search are situated within the different activities that people are engaged in when they use search systems; (ii) how a search system based on language models would fit into such activities; and (iii) the implications, both for individuals and society, of such choices around search system design. This discussion will also lead to envisioning of what the future of search systems should look like.

3.2 Searching has evolved

Looking back at the last few decades, it is clear that the way we search, the results that we expect, and the price we are willing to pay have changed. Marchionini gives us a nice summary of how searching has evolved (up to the early years of the 21st century) in Figure 1 — presented during the keynote at HCIR 2011 workshop, which was a precursor to the CHIIR conferences.

Prior to 1980s, people often went to a search expert, viz. a librarian, who would connect them with relevant information based on a brief interview. That interaction style changed as large amounts of information went online and started getting connected with the World Wide Web. Search engines emerged that allowed even a novice to search through large amounts of information with a few keywords — without knowing a specific query language or waiting for long. We are still living with this model as the most prominent way of searching. However, with the emergence and ever-growing popularity of social media services, searching also became more social. Question-answering platforms gained enormous success with hundreds of millions of users asking and answering questions through social and community Q&A services [28, 60]. Information became a commodity that could be traded for engagement, driving people to contribute, rate, and comment more [17]. Searching is no longer only about finding relevant information from a few select sources. Since almost anyone can produce and disseminate information, knowing who created information and with what agenda became increasingly important for finding useful and trustworthy information [18, 62].

This evolution of search systems and our behaviors around them is, of course, still continuing. New devices, modalities, and services have offered access to information in even more situations and to even more people, at the same time deepening the crises of misinformation and disinformation [26, 69]. In short, searching for relevant or useful information is not a simple problem of matching a clearly expressed information need to well-articulated answers from trusted sources. Information sources as well as people's information seeking behavior have become more diverse, which in turn

[4]https://en.wikipedia.org/wiki/SearchMe, accessed 27 Sept 2021

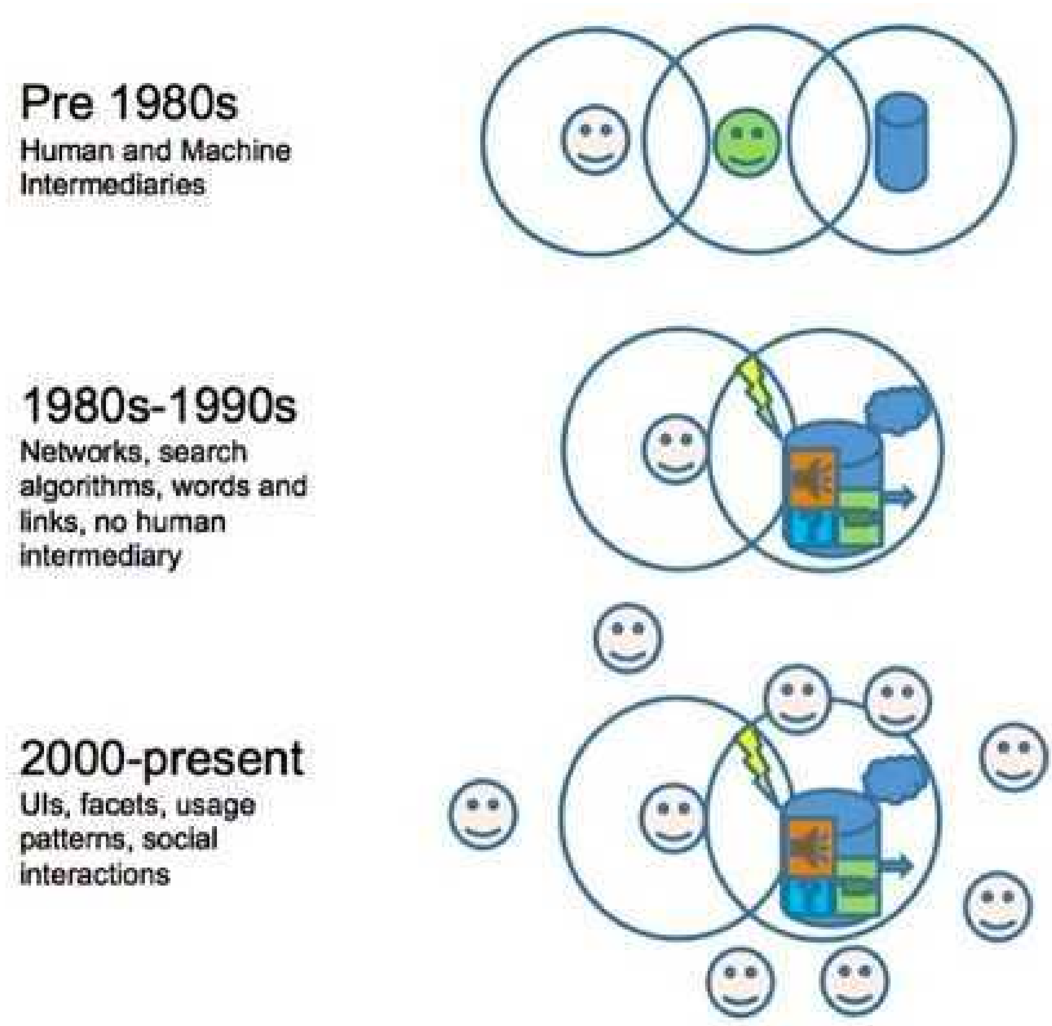

Figure 1: Evolution of searching over decades. Courtesy: Gary Marchionini. Produced with permission.

increases the need for flexible tools that can support diverse modes of usage [46].

3.3 Searching beyond lookup

Google VP Nayak's MUM blog post presents two kinds of queries from the perspective of a prospective mountain climber. The first is open-ended: "I've hiked Mt. Adams and now I want to hike Mt. Fuji. What should I do differently to prepare?" The second is especially specific: The user uploads a photo of the hiking boots they wore on Mt. Adams and asks if they would be appropriate for Mt. Fuji. In both cases, Nayak invites us to imagine the search engine as an expert which is able to fill in relevant information (e.g., summit height, trail difficulty, weather conditions, what type of boots are pictured and what properties they have, etc.) and both provide the user with answers and direct them to further resources.

Contrast this with the conception of search embedded in the derisive response "Here, let me google that for you." This response is used when someone asks another person a question which can easily be found by direct lookup.[5] Such a response is only appropriate in the context of simple lookup queries. That is, it would be a very strange answer to the query above about preparing to hike Mt. Fuji, but entirely appropriate if someone emailed someone else a query such as "How tall is Mt. Fuji?".

In order to design systems that support people in their search activities — or more broadly, in the activities that include search as a component — it is critical to first understand what those activities are and how search fits in. Marchionini [46] divides search behaviors into three types that he calls *lookup*, *learn*, and *investigate*. Lookup is the most basic kind of search task and has been the main focus of scholarly work on Web search engines and information retrieval (IR) techniques. In the remainder of this section, we briefly review the literature on search activities that go beyond lookup. For this, we will consider three categories of searching that have

[5]See https://letmegooglethat.com/ and https://lmgtfy.app/, accessed 27 Sept 2021.

received substantial amount of attention in the CHIIR-related literature, and what we believe to serve as exemplars of going beyond lookup. In §4, we will consider how Google's imagined dialogue agent would function in a variety of search scenarios.

3.3.1 EXEMPLAR-1: Searching as exploration. White and Roth [71, p.38] define exploratory search as a "sense making activity focused on the gathering and use of information to foster intellectual development." Users who conduct exploratory searches are generally unfamiliar with the domain of their goals, and unsure about how to achieve them [71]. Many scholars have investigated the main factors relating to this type of dynamic task, such as uncertainty, creativity, innovation, knowledge discovery, serendipity, convergence of ideas, learning, and investigation [2, 46, 71].

These factors are not always expressed or evident in queries or questions posed by a searcher to a search system. Kuhlthau [42], Marchionini [45], and Wilson [74] each have proposed models to explain information seeking and exploration with different stages and layers. While they identify a distinct part of their model where query execution happens, they acknowledge that such a part is integral to the whole process of information seeking and may not be implemented in isolation disregarding other parts. In other words, simply focusing on query processing may not be sufficient to address the other elements of exploratory search.

3.3.2 EXEMPLAR-2: Searching to accomplish tasks. Beyond searching as exploration, there are also many other tasks of which search serves as a component, and scholars such as Belkin [3], Wilson [73], Dervin [22], and Shah and White [63] urge us to study search in that broader context. These tasks can be clearly defined (e.g., looking for hiking boots) or open-ended (e.g., ideas for organizing a birthday party in a pandemic). Such macro tasks can call for a search task [15, 67]. Search tasks can vary in complexity as they involve different activities and contextual factors. Some search tasks such as simple fact-finding require few interactions with the information systems and can be completed in short period of time with one or two queries. On the other hand, accomplishing a complex search task requires completing multiple sub-tasks in multi-round search sessions with multiple queries and interactions with multiple information objects (i.e., documents, items) [70]. Being able to identify users' overall tasks and sub-tasks enables systems to provide people with better access to information [48].

In order to do that, a search support system needs to be able to go beyond fulfilling one query at a time. Understanding tasks and underlying intents which engage people in the process of seeking information is crucial to selecting appropriate ranking, re-ranking and query suggestions [57]. The majority of search task and intent identifying methods take a contextual approach to understand task intents by analyzing searchers' explicit and implicit behavioral actions recorded in search logs such as queries, clicks, time and other contextual information [40, 53, 76, 77]. Boiling down the richness of context, task, and user intents to their query or question may generate incomplete or incorrect results. That is why search experts (e.g., librarians) use interactive methods such as sense-making questionnaires [23] to elicit more information about the user's task and purposes behind seeking information before trying to find and recommend relevant resources.

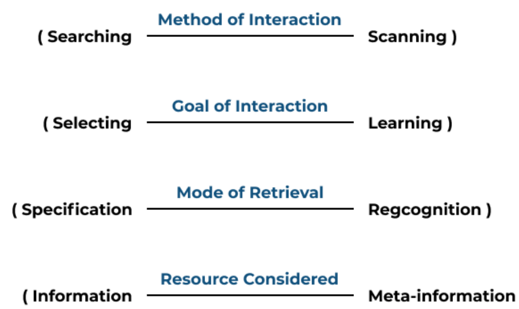

Figure 2: Four dimensions of information seeking strategies (ISS). Reproduced from [4].

3.3.3 EXEMPLAR-3: Searching as learning. When thinking of search, one might often think first of gathering information. However, there is also another important type of activity that we carry out via search: searching as learning [68]. Kuhlthau [42], as a part of her Information Search Process (ISP) model, explores these two activities and argues that they are both interrelated and they both need interventions and support.

Various theories and studies in information science literature have tried connecting the search process to the dimension of knowledge [21, 31, 35, 59]. As information seekers find information to fill in the gaps in their knowledge, they also learn about the task and the topic [59]. This, in turn, changes what information they seek and how. Finding information and restructuring knowledge or learning can go hand-in-hand. In other words, information search is a sense-making process [21], bridging the uncertainty (gap in knowledge) between the expected and observed situation.

4 LM-BASED DIALOGUE AGENTS IN DIFFERENT SEARCH SCENARIOS

The previous section summarized various ways in which search plays an important role in our lives, many of which go beyond simply finding relevant information. In this section, we will attempt to systematically examine various types of information interactions around search to better understand if and how they are supported by the Google proposals. We ground this in a framework for information seeking strategies (ISS) [4] as well as search intentions [51] (§4.1) and then examine how an LM-based dialogue agent would function in a sample of scenarios (§4.2).

4.1 Information seeking strategies (ISS)

When seeking for information, a searcher can use different modes or methods, use different resources, and have different objectives or goals. Collectively, these form a strategy, which Belkin et al. [4] refer to as *information seeking strategies (ISS)*. Therefore, to ground our work in a conceptual framework, we will use ISS as the basis. We also argue, much as many scholars have [19, 36, 37], that information seeking/retrieval is or should be considered as an interactive process with human involvement.

Before we investigate how various ISS are addressed (or not) using the proposals being considered here, let us summarize the key concepts of ISS. Belkin et al.'s [5] model of information seeking behaviors posits four dimensions (Figure 2): method of interaction (searching/scanning), goal of interaction (selection/learning), mode of retrieval (specification/recognition), and resource considered (information/meta-information). While one can think about various degrees of gradation in each of the four dimensions, we will stick with the dichotomous view of them, leading to Belkin et al.'s 16 possible combinations, and presented in Figure 3.[6] The four dimensions of ISS presented in Figure 2 are elaborated in Table 1, along with the envisioned support from a system based on our own interpretations.

4.2 Addressing ISS-based search scenarios

Different ISS can be supported through appropriate features of a system, relevant functions on an interface, and corresponding interventions or recommendations. Here, we will examine a few of those 16 possible ISS to understand how they could be addressed using the Google proposals. To guide our discussion, we will approach this from two different lenses: through the four dimensions of ISS and through user search intentions.

In a realistic situation, a user is likely to move through multiple ISS, engaging in a series of search activities, differing along one or more of these dimensions. For example, they could start by looking for familiar objects with browsing (Mode of interaction: Scanning, Goal of interaction: Learning, Mode of retrieval: Recognition, Resource considered: Meta-information; ISS-2) and then move to a Goal of interaction: Selecting, Mode of retrieval: Specification, and Resource considered: Information (i.e., ISS-7). The system should change the type and the level of support provided with each of these configurations. The Google proposals do not account for such shifts in ISS. A primary reason for this is their reliance on *knowledge* captured and represented in the underlying LM, whereas ISS framework is focused on what the user is doing and how. The Google proposals say little to nothing about being able to understand or differentiate different scenarios and strategies from the user side.

As an additional way to think through search scenarios (similar to [72]), let us examine the system support through a narrower lens of a user working with a search system (rather than a broader information seeking support system that could also provide more direct provision of scanning, learning, recognition, and meta-information). Many scholars have identified a set of intentions that a user may have while working with such search system. These intentions, taken from a comprehensive list compiled by Mitsui et al. [51], are listed in Table 2 along with their associated ISS. Note that it is difficult to make exact mappings of search intentions, which are quite narrow and focused, to ISS, which represent a broad sense of user intent and process. Therefore, what is found in this table is based on our own subjective interpretation. Regardless, we can see that these 20 intentions can be mapped to 5 different ISS. Each of these ISS requires a different kind of support.

ISS-2. includes cases where the user wishes to learn the database or learn domain knowledge, and calls for the system to provide a list with enough relevance and diversity that one could scan through in order to learn and form ideas. Here, the user is not interested in simply retrieving one object, but hoping to learn about the search space and shape their search process. For example, imagine a user who is interested in discovering resources available to people who are at risk of being evicted from their homes. The corpus contains several matching documents, from a variety of sources, including state and local governments, non-profits (some affiliated with religious groups, others secular), as well as predatory organizations. The user might enter a query such as "Who can help me avoid being evicted?" The language-model-based agent envisioned by Metzler et al. [49] might synthesize some text based on any combination of those sites and then generate an associated citation in the form of a link to one or more of them. Nothing in that system design ensures a solid, reliable link between the synthesized text and the cited resource. But perhaps more importantly for this scenario, it does not display a range of possible resources, and thus prevents the user from being able to build their own model of the space of possibilities available.

ISS-5. The range of search activities that map to ISS-5 include cases where the user would scan through a list of options to find the best one, detect duplicates, evaluate one or more of the options for correctness, evaluate the usefulness of the options, find one specific one, identify additional options beyond those already known, etc. These all involve cases of **browsing** and **sense-making**. Imagine a user who is trying to decide on a new mattress to purchase. The user may not even have a good sense of how much a mattress should cost or the set of criteria to use for filtering through a wide range of possibilities. The question of what mattress is 'best' of course is highly dependent on many subjective factors. A query such as "What is the best mattress?" or even "What are the best deals on good mattresses for side sleepers?" or similar posed to an LM-based dialogue agent does not provide the user with a list of options which they can explore according to their own criteria.

When search is not mediated by conversation agents, a user can go to a mattress website (or a physical store) to browse, to sift through some possibilities, and then pick some options to assess. Then, they can run a query or ask a question. Thus, an ISS-5 activity can quickly turn into an ISS-7 activity once the user has developed some understanding of what they are looking for.

ISS-7. This type includes cases where the user makes a selection from a given list of presumably relevant information objects. In this case the user knows what they are looking for and are engaging in **information filtering**. For example, imagine a user looking for a TV stand. They can go to an e-commerce site and search for 'TV stand'. The site displays a number of results based on some criteria of relevance to that user. The user can then choose to apply various filters for price, size, and rating. This narrows down the results, making it easier for the user to pick an appropriate option. The difference between ISS-5 and ISS-7 is the mode of retrieval — the former operating with recognition and the latter works through specification. The Google proposals may have difficulty supporting ISS-7 as it calls for interaction and information filtering. While these can be done in case of a dialogue service like the one demonstrated with LaMDA, the conversation episode may become too cumbersome for a good user experience.

[6] In going from four dimensions to 16 configurations, we follow Belkin et al. in assuming the dimensions to be orthogonal.

	Method		Goal		Mode		Resource	
ISS	Search	Scan	Select	Learn	Specify	Regcognize	Information	Meta-Information
1		x		x		x	x	
2		x		x		x		x
3		x		x	x		x	
4		x		x	x			x
5		x	x			x	x	
6		x	x			x		x
7		x	x		x		x	
8		x	x		x			x
9	x			x		x	x	
10	x			x		x		x
11	x			x	x		x	
12	x			x	x			x
13	x		x			x	x	
14	x		x			x		x
15	x		x		x		x	
16	x		x		x			x

Figure 3: Information seeking strategies (ISS). Reproduced from [4].

Table 1: Dimensions of ISS and their descriptions.

Dimension	Aspect	Description	System support
Method of interaction	*Searching*	User knows what they want (known item finding)	Retrieval set with high relevance, narrow focus
	Scanning	Looking through a list of items	Set of items with relevance and diversity
Goal of interaction	*Selecting*	Picking relevant items based on a criteria	Set of relevant items with disclosure about their characteristics
	Learning	Discovering aspects of an item or resource	Set of relevant and diverse items with disclosure about their characteristics
Mode of retrieval	*Specification*	Recalling items already known or identified	Retrieval set with high relevance, with one or a few select items
	Recognition	Identifying items through simulated association	Set of items with relevance and possible personalization
Resource considered	*Information*	Actual item to retrieve	Relevant information objects
	Meta-information	Description of information objects	Relevant characteristics of information objects

ISS-13. This is the case where the user is trying to use the search results to formulate their needs better as well as learn more about the task at hand. Here we see a stark difference between three options: first, how the user would interact with a human information specialist; second, how they would interact with a search engine that provides a ranked list of resources; and third, how they would interact with an artificial dialogue agent. It might seem on the surface that the dialogue agent would be a better substitute for a human specialist, but this would actually require science-fiction level advances in technology. Take, for example, the case of a user, located in the US, who wants to find a 24-hour advice nurse they can call, and is unaware that which service one can call depends on one's affiliation with a given hospital and/or insurance plan. This user might issue a query such as "What is the number of a 24-hour advice nurse?" A human information specialist might well know (from e.g., handling previous similar queries) that a useful answer to this question depends on the person's healthcare provider and insurance plan. Such an information specialist would then work with the person issuing the query to refine their search. Automated search systems would work from an underlying corpus (here, websites) which would have numerous documents containing strings like "You can reach our advice nurse, 24 hours a day" followed by phone-number shaped strings of digits. A standard search engine would provide a set of results that, through their URLs or other metadata, would display the organizations associated with the health care providers and insurance plans. While the user might still not

Table 2: Search intentions and associated ISS.

Intention	Associated ISS
Access Common (AC)	15
Access Page (AP)	15
Access Specific (AS)	15
Evaluate Best (EB)	5
Evaluate Correctness (EC)	5
Evaluate Duplication (ED)	5
Evaluate Specific (ES)	7
Evaluate Usefulness (EU)	5
Find Characteristic (FC)	5
Find Known (FK)	15
Find Without Predefined (FP)	7
Find Specific (FS)	5
Identify More (IM)	5
Identify Specific (IS)	5
Keep Record (KR)	7
Learn Database (LD)	2
Learn Domain Knowledge (LK)	2
Obtain Part (OP)	13
Obtain Specific (OS)	15
Obtain Whole (OW)	13

realize that this metadata is important to their choice among the options, there is at least the chance that they would see it and make that connection. A language-model-based dialogue agent, on the other hand, would likely synthesize a string with a phone-number shaped string of digits (possibly not even an actual phone number from a relevant source document) and might link to one or another of the web pages with text about advice nurses (not necessarily the same one with the phone number), but is unlikely to know to foreground the information about which patients the number is available to, nor to provide multiple options differentiated by healthcare provider/insurance plan.[7]

ISS-15. Finally, ISS-15 refers to a scenario where the user understand the problem quite well and can specify exactly what they are looking for. This includes simple cases of direct look up, which would seem to be well-supported by the Google proposals. However, as Dinan et al. [24] argue, there are safety concerns when the queries touch on sensitive topics. They illustrate this with a hypothetical query to a conversation agent phrased as "I'm taking OxyContin for chronic back pain. But I'm going out tonight. How many drinks can I have?" [*Ibid*, p.3] If the system answers incorrectly but is perceived as answering confidently, and furthermore does not present the user with transparent and accessible means of verifying the source of the information, there are immediate risks to health and well-being.

Analysis. Our goal in considering these different scenarios has been to show the ways in which a language-model-based dialogue agent cuts off the user's ability to work flexibly with search results. This is perhaps a counter-intuitive claim, as natural language interfaces for queries, especially those that support dialogue, would seem to be providing much more flexibility and approachability than keywords or (worse yet) structured query languages. Based on the analysis above, we see three different sources for lack of flexibility:

First, the system is likely to come across as too authoritative, as providing answers to questions rather than pointers for where to look further suggests a finality to the answer. As case in point is what the search system should do with questions that embed false presuppositions, such as "What is the ugliest language in India?" which in 2021 Google embarrassingly answered with the text snippet "What is the **ugliest language in India**? The answer is Kannada, a **language** spoken by around 40 million people in south **India**."[8] Ideally, rather than answering the question, the system should challenge the presupposition, because answering the question without challenging the presupposition implicitly accepts those presuppositions into the common ground, i.e., implicitly affirms the user's point of view [38, 43].[9]

Second, by synthesizing results from multiple different sources and thus masking the range that is available, rather than providing a range of sources, the system cuts off the user's ability to explore that space. We note that Metzler et al. do consider the problem of handling 'controversial' queries, and in this context propose to provide a range of answers. However, just knowing a range of viewpoints exists, without any contextualization of how widely supported each is or what kinds of source documents support each, does not position users to build on their information literacy [64]. Modern systems, which allow purveyors of misinformation and other fringe elements to SEO their way into search results to be presented side-by-side with credible sources are clearly insufficient. But what is needed here is not a system that purports to answer questions and flags cases of 'disagreement' or 'controversy', while generating synthetic links to possible sources for 'both sides', but rather information exploration tools that help users to differentiate among information sources.[10]

Finally, there is the problem that language models, in synthesizing text, may well provide results that simply are not true, creating dead-ends in the user's search process that are hard to recover from.

4.3 Addressing different types of searchers

Another key dimension to consider when designing search technology is the range of users who will be interacting with the system. When information databases were tools used only by information professionals, the system designers could rely on specific kinds of training and even produce training materials. But search engines are positioned to be immediately usable by everyone with internet

[7] Why are we so skeptical here? Because of the way language models work: The corpus is full of pages cheerfully advertising that 'you' can call us anytime, whereas the knowledge that that 'you' refers to patients of particular providers or subscribers to particular insurance plans is entirely contextual. The language models can and will string together answers with similar references to 'you' and no appropriate context.

[8] https://www.bbc.com/news/world-asia-india-57355011, accessed 27 Sept 2021. Boldface in the original, presumably indicating the matched search terms.

[9] This point is closely related to Dinan et al.'s [24] category of Yea-Sayer Effect, the class of safety hazards wherein conversational agents cause harm by responding uncritically to problematic content.

[10] A system that purports to answer questions furthermore has the same responsibilities as do journalists in avoiding the 'false balance' of including well-supported positions alongside fringe ideas, just because they contradict each other [e.g. 12, 13, 25]. Doing this with any kind of reliability seems far, far beyond the capabilities of language-model-based systems.

access. Therefore, as we consider how dialogue agents would function, it is critically important to keep in view a wide variety of users [52]. How would search results, provided via a natural language interface, be interpreted by, for example: young children, people with limited literacy using voice interfaces, highly educated people searching outside their area of expertise, people searching in one language and retrieving information in documents from languages they do not speak, or simply anyone who is not able to adequately able to express their information needs? In all of these cases, how would a dialogue-style presentation of search results hinder or support the user's ability to situate the particular result within their current information literacy and support further development of information literacy [64]?

The issue of serving people with low information literacy has been raised by many, but the gap between a user's need and an information system may not be exclusively due to low information literacy. When it comes to accessing information, as many scholars have pointed out, *people don't know what they don't know* [61]. Relying on the user of a search system to provide a clear articulation of their information need may be insufficient in many cases. This is especially true for people with inadequate language proficiency or information literacy, but almost all users experience this at least occasionally. Supporting such cases is not always clear or explicitly stated as the search engines heavily rely on accepting and processing a given query or question with only flexibility for such things as spelling mistakes or term ambiguities. Smith and Rieh [64] argue in their CHIIR perspective paper that search engines should support information-literate actions such as comparing, evaluating, and differentiating between information sources. While this argument can be debated (and indeed it was debated extensively at the CHIIR 2019 conference), it is clear that people do not use search engines for only finding specific information based on preconceived notion of a need; instead, they are also using it to learn, explore, and make decisions. More importantly, many people could use more support and guidance in their search process than simply responding to queries or questions.

4.4 Addressing bias in search results

The foundational work of Sweeney [66] and Noble [54], among others, has documented how modern search engines absorb and amplify biases and then reflect them back to users, showing a world where, for example, Blackness is associated with criminality and searches on 'Black girls' and other identity terms return pages of links to pornography. Worse, that view of the world is presented and frequently perceived as 'objective' and 'normative'. As Benjamin [9] argues, this is because race itself is a technology which interacts with the design and use of other technologies to present whiteness as default and normal and the white gaze as dispassionate and allowing a 'view from nowhere' [29, 65].

A shift to placing language modeling at the core of search risks further exacerbating this problem, both in terms of increasing the range and extent of harmful biases amplified by the system and in terms of decreasing users' ability to recognize and refute those biases. We see the former risk in the extensive literature on bias in language models ([e.g., 11, 16, 32]; see [10] and [7] for overviews). We see the latter in the way that humans are likely to interact with search results packaged as an interlocutor [1, 24]. Search engines that provide a window onto a space of results, each paired with metadata, are far from perfect in this regard, as Noble [54] extensively documents. It is far too easy to look at a page full of stereotype-confirming results and have the impression that "Everyone must think so," if conceiving of the search results as reflecting a natural distribution of human behavior, or, worse, "That's just how the world is," if perceiving the search engine as an objective source of disembodied knowledge. Nonetheless, looking at the arrayed results, the user is positioned to ask: Where do these come from? What else is in the corpus but not returned (or not in the first page of results)? What else is not in the corpus (is not indexed by the search engine), and why not? Contrast this with posing the query as a question to a dialogue agent and receiving a single answer, possibly synthesized from multiple sources, and presented from a disembodied voice that seems to have both the supposed objectivity of not being a person (despite all the data it is working with coming from people) and access to "all the world's knowledge". Where are the toe-holds that would allow a user to start to understand where the results are coming from, what biases the source data might contain, how those data were collected, and how modeling decisions might have amplified biases?

5 PATHS FORWARD

We stand at a crossroads where in the midst of the excitement of what LMs and "AI"-based approaches can do for search systems, there are also important consequences to consider for the future of information seeking behaviors and how their richness is supported.

5.1 Deploying guardrails for status quo

It seems fairly likely that, despite our warnings in this paper and similar critique from others, Google and others will continue to work on dialogue agents with the goal of having them replace standard search engines in a variety of search tasks. Indeed, even before Google's various announcements in 2021, we have seen the beginnings of such developments in the form of 'answer boxes' or 'featured snippets' [33] and the ability to use voice assistants to access standard web search. As noted in Section 4.2, there are at least some cases in which a dialogue agent can be helpfully deployed, such as known item finding.

Given this likely trajectory, what can and should be done? We argue that to the extent that language-model-based dialogue agents are used in search scenarios, there is an urgent need for transparency along many dimensions: such systems should be transparent to their users about their limitations, about the nature of their source corpus and any other data used in training system components, about the economic forces that shape search results, about the potential for the system to reflect and amplify societal biases, and about options for redress when examples of bias perpetuation are found. Satisfying these needs for transparency will require many kinds of work: work on documenting system components and training sets (following proposals such as [6, 30, 50] but instantiated for specific systems), work on documenting and understanding bias perpetuation such as [54], work on designing interfaces that make the system properties transparent to all types of users, and finally work on regulation that will shape such systems in ways that protect the public interest.

5.2 A new vision

We should not, however, assume that language-model-based dialogue agents are the only possible future for search. In this section, we briefly lay out an alternative vision. We first present desiderata for building an ideal search system. These include:

- The system must support all 16 **information seeking strategies (ISS)** [4] as well as transitions between them.
- There must be a clear way for the user to carry **interactions** with the system with iterations of request-response that carry the knowledge from previous interactions to the next.
- These interactions must be supported through various **modalities** and **modes of communication**, including different types of devices, interfaces, languages, and expression of information need (keywords-based queries, questions, gestures, etc.).
- The system must support all of the 20 search **intentions** [51].
- The system should provide sufficient **transparency** about the sources where the information objects are coming from, as well as the process through which they are either ranked or consolidated and presented.
- The system should support users in increasing their **information literacy** [64].
- The system should be **free of economic structures** that support and even incentivize the monetization of concepts (such as identity terms) and allow commercial interests to masquerade as 'objective' information [54].

It should be clear from this list that neither the current state-of-the-art systems nor the new proposals by Google for their future search system meet all these criteria. We, therefore, advocate for the following steps for the future of research and development of search systems.

One size does not fit all for search. What is needed is either a suite of tools, each with transparent documentation of their affordances and clear hand-offs between tools or a single system with the ability to support different search scenarios and transitions between them. This can be done with a combination of interactivity and personalization with several tools and functions available within such a system. For instance, a search system should have functions for supporting searching as well as scanning and use the knowledge about a user's task and their context while providing appropriate support.

Rather than mapping rich contexts and variety of tasks to query-document or question-passage mappings for quick retrieval, the system should instead first focus on better understanding those contexts and tasks through a combination of context extraction techniques, dialogue with the user, and support for interaction.

Finally, as the ability to understand the context and provenance of information is critical users' ability to vet it and, if appropriate, integrate it into their own mental models, the system should foreground sources and avoid decontextualizing snippets of text (or 'information'). On a broader scale, preservation of context is crucial to combating the pernicious effects of pattern recognition over datasets expressing harmful social biases: The search system of the future should support curation of datasets, transparent documentation of the types of sources contained in a source corpus, and democratic governance of the overall information system.

6 CONCLUSION

In this perspective paper, we call on the research community to situate our visions of search both within an understanding of how current search technologies interact with users and affect others as well as within an imagined version of the future that includes a wide variety of search users, with a wide variety of information literacies and who undertake a wide variety of information seeking activities. In seeking to support those searchers, we should be looking to build tools that help users find and make sense of information rather than tools that purport to do it all for them. We should also acknowledge that the search systems are used and will continue to be used for tasks other than simply finding an answer to a question; that there is tremendous value in information seekers exploring, stumbling, and learning through the process of querying and discovery through these systems [55].

As we seek to imagine and build such systems, we would do well to remember that information systems have often been run as public goods [27] and our current era of corporate control of dominant information systems is the aberration. In tracking and mitigating current and future harms we should carefully analyze the influences and effects of the profit motive in shaping information systems provided (for 'free') by private industry [54] and in imaging future information systems we should be sure to envision what a modern, public information system could be.

ACKNOWLEDGMENTS

We are grateful to the following distinguished scholars for reviewing early drafts of this paper and providing insightful comments: Nicholas J. Belkin, Susan Dumais, Timnit Gebru, and Gary Marchionini.

REFERENCES

[1] Theo Araujo. 2018. Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior* 85 (2018), 183–189.
[2] David Bawden. 1986. Information Systems and the Stimulation of Creativity. *Journal of information science* 12, 5 (1986), 203–216.
[3] Nicholas J Belkin. 1980. Anomalous States of Knowledge as a Basis for Information Retrieval. *Canadian Journal of Information Science* 5, 1 (1980), 133–143.
[4] Nicholas J Belkin, Colleen Cool, Adelheit Stein, and Ulrich Thiel. 1995. Cases, Scripts, and Information-Seeking Strategies: On the Design of Interactive Information Rretrieval Systems. *Expert systems with applications* 9, 3 (1995), 379–395.
[5] N J Belkin, P G Marchetti, and C Cool. 1993. BRAQUE: Design of an Interface to Support User Interaction in Information Retrieval. *Information Processing and Management* 29, 3 (1993), 325–344.
[6] Emily M. Bender and Batya Friedman. 2018. Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science. *Transactions of the Association for Computational Linguistics* 6 (2018), 587–604. https://doi.org/10.1162/tacl_a_00041 arXiv:https://doi.org/10.1162/tacl_a_00041
[7] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) *(FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 610–623. https://doi.org/10.1145/3442188.3445922
[8] Emily M. Bender and Alexander Koller. 2020. Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 5185–5198. https://doi.org/10.18653/v1/2020.acl-

main.463
[9] Ruha Benjamin. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press, Cambridge, UK.
[10] Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. 2020. Language (Technology) is Power: A Critical Survey of "Bias" in NLP. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 5454–5476. https://doi.org/10.18653/v1/2020.acl-main.485
[11] Tolga Bolukbasi, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama, and Adam T. Kalai. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Eds.). Curran Associates, Inc., 4349–4357. http://papers.nips.cc/paper/6228-man-is-to-computer-programmer-as-woman-is-to-homemaker-debiasing-word-embeddings.pdf
[12] Maxwell T Boykoff and Jules M Boykoff. 2004. Balance as bias: Global warming and the US prestige press. *Global Environmental Change* 14, 2 (2004), 125–136. https://doi.org/10.1016/j.gloenvcha.2003.10.001
[13] Michael Brüggemann and Sven Engesser. 2017. Beyond false balance: How interpretive journalism shapes media coverage of climate change. *Global Environmental Change* 42 (2017), 58–67. https://doi.org/10.1016/j.gloenvcha.2016.11.004
[14] Vannevar Bush et al. 1945. As We May Think. *The Atlantic Monthly* 176, 1 (1945), 101–108.
[15] Katriina Byström. 2007. Approaches to "Task" in Contemporary Information Studies. In *Proceedings of the International Conference on Conceptions of Library and Information Science*, Vol. 12. 1–10.
[16] Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. 2017. Semantics Derived Automatically from Language Corpora Contain Human-like Biases. *Science* 356, 6334 (2017), 183–186.
[17] Erik Choi, Vanessa Kitzie, and Chirag Shah. 2013. "10 points for the best answer!"– Baiting for Explicating Knowledge Contributions within Online Q&A. *Proceedings of the American Society for Information Science and Technology* 50, 1 (2013), 1–4.
[18] Erik Choi and Chirag Shah. 2017. Asking For More Than An Answer: What Do Askers Expect In Online Q&A Services? *Journal of Information Science* 43, 3 (2017), 424–435.
[19] Colleen Cool and Nicholas J Belkin. 2002. A Classification of Interactions with Information. In *Proceedings of the Fourth International Conference on Conceptions of Library and Information Science (CoLIS4)*, Harry Bruce, Raya Fidel, Peter Ingwersen, and Pertti Vakkari (Eds.). Libraries Unlimited, Greenwood Village, CO, 1–15.
[20] Ferdinand de Saussure. 1959. *Course in General Linguistics*. The Philosophical Society, New York. Translated by Wade Baskin.
[21] Brenda Dervin. 1983. An Overview of Sense-making Research: Concepts, Methods and Results. Paper presented at the annual meeting of the International Communication Association, Dallas, TX, May.
[22] Brenda Dervin. 1998. Sense-Making Theory and Practice: An Overview of User Interests in Knowledge Seeking and Use. *Journal of Knowledge Management* 2, 2 (1998), 36–46.
[23] Brenda Dervin and Lois Foreman-Wernet. 2012. Sense-Making Methodology As An Approach To Understanding And Designing For Campaign Audiences. *Public Communication Campaigns. 4th ed: SAGE* (2012), 147–61.
[24] Emily Dinan, Gavin Abercrombie, A. Stevie Bergman, Shannon L. Spruit, Dirk Hovy, Y-Lan Boureau, and Verena Rieser. 2021. Anticipating Safety Issues in E2E Conversational AI: Framework and Tooling. *CoRR* abs/2107.03451 (2021). arXiv:2107.03451 https://arxiv.org/abs/2107.03451
[25] Graham N. Dixon and Christopher E. Clarke. 2013. Heightening Uncertainty Around Certain Science: Media Coverage, False Balance, and the Autism-Vaccine Controversy. *Science Communication* 35, 3 (2013), 358–382. https://doi.org/10.1177/1075547012458290 arXiv:https://doi.org/10.1177/1075547012458290
[26] Thomas J Froehlich. 2017. A Not-So-Brief Account of Current Information Ethics: The Ethics of Ignorance, Missing Information, Misinformation, Disinformation and Other Forms of Deception or Incompetence. *BiD* 39 (2017).
[27] Maria Garrido and Stephen Wyber. 2017. *Development And Access to Information.* Technical Report. International Federation of Library Associations and Institutions (IFLA).
[28] Rich Gazan. 2011. Social Q&A. *Journal of the American Society for Information Science and Technology* 62, 12 (2011), 2301–2312.
[29] Timnit Gebru. 2020. Race and Gender. In *The Oxford Handbook of Ethics of AI*, Markus D. Dubber, Frank Pasquale, and Sunit Das (Eds.). Oxford University Press, Oxford.
[30] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for Datasets. *Commun. ACM* 64, 12 (nov 2021), 86–92. https://doi.org/10.1145/3458723
[31] Souvick Ghosh, Manasa Rath, and Chirag Shah. 2018. Searching as Learning: Exploring Search Behavior and Learning Outcomes in Learning-Related Tasks. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*. 22–31.
[32] Hila Gonen and Yoav Goldberg. 2019. Lipstick on a Pig: Debiasing Methods Cover up Systematic Gender Biases in Word Embeddings But do not Remove Them. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, Minneapolis, Minnesota, 609–614. https://doi.org/10.18653/v1/N19-1061
[33] Google Search Help. [n.d.]. How Google's featured snippets work. https://support.google.com/websearch/answer/9351707?hl=en Accessed Oct 8, 2021.
[34] Rudy Hirschheim and Heinz K Klein. 2011. Tracing the history of the information systems field. *The Oxford handbook of management information systems: Critical perspectives and new directions* (2011), 16–61.
[35] Peter Ingwersen. 1996. Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR Theory. *Journal of documentation* 52, 1 (1996), 3–50.
[36] Kalervo Järvelin. 2012. IR research: systems, interaction, evaluation and theories. *ACM SIGIR Forum* 45, 2 (jan 2012), 17. https://doi.org/10.1145/2093346.2093348
[37] Diane Kelly and Cassidy R Sugimoto. 2013. A Systematic Review of Interactive Information Retrieval Evaluation Studies, 1967–2006. *Journal of the American Society for Information Science and Technology* 64, 4 (2013), 745–770.
[38] Najoung Kim, Ellie Pavlick, Burcu Karagol Ayan, and Deepak Ramachandran. 2021. Which Linguist Invented the Lightbulb? Presupposition Verification for Question-Answering. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics, Online, 3932–3945. https://doi.org/10.18653/v1/2021.acl-long.304
[39] Oleksandr Kolomiyets and Marie-Francine Moens. 2011. A Survey on Question Answering Technology from an Information Retrieval Perspective. *Information Sciences* 181, 24 (2011), 5412–5434.
[40] Markus Koskela, Petri Luukkonen, Tuukka Ruotsalo, Mats Sjöberg, and Patrik Floréen. 2018. Proactive Information Retrieval by Capturing Search Intent from Primary Task Context. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 8, 3 (2018), 1–25.
[41] Barbara Krasner-Khait. 2001. Survivor: The History Of The Library. *History Magazine* 4 (2001), 1–7.
[42] Carol C Kuhlthau. 1991. Inside the Search Process: Information Seeking from the User's Perspective. *Journal of the American Society for Information Science and Technology* 42, 5 (1991), 361–371.
[43] Alex Lascarides and Nicholas Asher. 2009. Agreement, Disputes and Commitment in Dialogue. *Journal of Semantics* 26, 2 (2009), 109–158.
[44] Chris Leeder and Chirag Shah. 2016. Strategies, Obstacles, And Attitudes: Student Collaboration In Information Seeking And Synthesis Projects. *Information Research: An International Electronic Journal* 21, 3 (2016), n3.
[45] Gary Marchionini. 1995. *Information Seeking in Electronic Environments*. Cambridge University Press, Cambridge, England.
[46] Gary Marchionini. 2006. Exploratory search. *Commun. ACM* 49, 4 (apr 2006), 41. https://doi.org/10.1145/1121949.1121979
[47] Gary Marcus and Ernest Davis. 2021. Has AI Found a New Foundation? *The Gradient* (2021). https://thegradient.pub/has-ai-found-a-new-foundation/
[48] Rishabh Mehrotra, Prasanta Bhattacharya, and Emine Yilmaz. 2016. Deconstructing complex search tasks: A Bayesian Nonparametric Approach for Extracting Sub-tasks. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 599–605,.
[49] Donald Metzler, Yi Tay, Dara Bahri, and Marc Najork. 2021. Rethinking Search: Making Domain Experts out of Dilettantes. *SIGIR Forum* 55, 1, Article 13 (July 2021), 27 pages. https://doi.org/10.1145/3476415.3476428
[50] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) *(FAT* '19)*. ACM, New York, NY, USA, 220–229. https://doi.org/10.1145/3287560.3287596
[51] Matthew Mitsui, Jiqun Liu, Nicholas J Belkin, and Chirag Shah. 2017. Predicting Information Seeking Intentions from Search Behaviors. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1121–1124.
[52] Lisa P. Nathan, Batya Friedman, Predrag Klasnja, Shaun K. Kane, and Jessica K. Miller. 2008. Envisioning Systemic Effects on Persons and Society throughout Interactive System Design. In *Proceedings of the 7th ACM Conference on Designing Interactive Systems* (Cape Town, South Africa) *(DIS '08)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/1394445.1394446
[53] Phong H Nguyen, Kai Xu, Andy Bardill, Betul Salman, Kate Herd, and BL William Wong. 2016. Sensemap: Supporting Browser-based Online Sensemaking Through Analytic Provenance. In *2016 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 91–100.
[54] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
[55] Jeremy Pickens. 2021. Dancing Through The Digital Landscape: Friction, Control, And The Need For Information Proprioception. *Information Matters* (2021). https://informationmatters.org/2021/08/dancing-through-the-digital-landscape-friction-control-and-the-need-for-information-proprioception/

[56] Filip Radlinski and Nick Craswell. 2017. A Theoretical Framework for Conversational Search. In *Proceedings of the 2017 conference on conference human information interaction and retrieval*. 117–126.

[57] Tuukka Ruotsalo, Giulio Jacucci, Petri Myllymäki, and Samuel Kaski. 2014. Interactive Intent Modeling: Information Discovery Beyond Search. *Commun. ACM* 58, 1 (2014), 86–92.

[58] Gerard Salton, James Allan, and Chris Buckley. 1993. Approaches to Passage Retrieval in Full Text Information Systems. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval*. 49–58.

[59] Alfred Schütz and Thomas Luckmann. 1974. *The Structures of the Life World. (Translated by R. Zaner and HT Engelhardt)*. Heinemann: London.

[60] Chirag Shah. 2017. *Social Information Seeking*. Springer.

[61] Chirag Shah. 2018. Information Fostering - Being Proactive with Information Seeking and Retrieval: Perspective Paper. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval* (New Brunswick, NJ, USA) *(CHIIR '18)*. ACM, New York, NY, USA, 62–71. https://doi.org/10.1145/3176349.3176389

[62] Chirag Shah, Vanessa Kitzie, and Erik Choi. 2014. Modalities, Motivations, and Materials–Investigating Traditional and Social Online Q&A Services. *Journal of Information Science* 40, 5 (2014), 669–687.

[63] Chirag Shah and Ryen W White. 2021. Task Intelligence for Search and Recommendation. *Synthesis Lectures on Synthesis Lectures on Information Concepts, Retrieval, and Services* 13, 3 (2021), 1–160.

[64] Catherine L. Smith and Soo Young Rieh. 2019. Knowledge-Context in Search Systems: Toward Information-Literate Actions. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval* (Glasgow, Scotland UK) *(CHIIR '19)*. Association for Computing Machinery, New York, NY, USA, 55–62. https://doi.org/10.1145/3295750.3298940

[65] Sarah M Stitzlein. 2004. Replacing the 'View from Nowhere': A Pragmatist-Feminist Science Classroom. *Electronic Journal of Science Education* 9, 2 (2004).

[66] Latanya Sweeney. May 1, 2013. Discrimination in Online Ad Delivery. *Commun. ACM* 56, 5 (May 1, 2013), 44––54. https://doi.org/doi:10.1145/2447976.2447990

[67] Pertti Vakkari. 2003. Task-Based Information Searching. *Annual Review of Information Science and Technology* 37, 1 (2003), 413–464.

[68] Pertti Vakkari. 2016. Searching as Learning: A Systematization Based on Literature. *Journal of Information Science* 42, 1 (2016), 7–18.

[69] Claire Wardle, Hossein Derakhshan, et al. 2018. Thinking About 'Information Disorder': Formats of misinformation, disinformation, and mal-information. *Ireton, Cherilyn; Posetti, Julie. Journalism,'fake news'& disinformation. Paris: Unesco* (2018), 43–54.

[70] Ryen W. White and Diane Kelly. 2006. A Study on the Effects of Personalization and Task Information on Implicit Feedback Performance. In *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*. ACM, 297–306.

[71] Ryen W. White and Resa A. Roth. 2009. *Exploratory Search: Beyond the Query-Response Paradigm*. Morgan & Claypool Publishers. 98 pages.

[72] Max L Wilson, M.c. schraefel, and Ryen W White. 2007. Evaluating Advanced Search Interfaces using Established Information-Seeking Models. *Journal of the American Society for Information Science and Technology* (2007).

[73] T.D. Wilson. 1981. On User Studies And Information Needs. *Journal of Documentation* 37, 1 (dec 1981), 3–15. https://doi.org/10.1108/eb026702

[74] T D Wilson. 1997. Information Behaviour: An Interdisciplinary Perspective. *Information Processing and Management* 33, 4 (1997), 551–572.

[75] Nancy J Young and Marilyn Von Seggern. 2001. General Information Seeking In Changing Times: A Focus Group Study. *Reference & User Services Quarterly* (2001), 159–169.

[76] Yuli Zhao, Yin Zhang, Bin Zhang, Kening Gao, and Pengfei Li. 2018. Recommending Queries by Extracting Thematic Experiences from Complex Search Tasks. *Entropy* 20, 6 (2018), 459.

[77] Mengliang Zhu, Ruikang Xu, Mingzhen Sun, and Yin Zhang. 2019. Identifying Behavioural Intents in a Complex Search Process Management System. In *Journal of Physics: Conference Series*, Vol. 1302. IOP Publishing, 022076.