# A Computational Inflection for Scientific Discovery

Tom Hope
tomh@allenai.org
The Allen Institute for AI
The Hebrew University of Jerusalem

Doug Downey
dougd@allenai.org
The Allen Institute for AI
Northwestern University

Oren Etzioni
oren@allenai.org
The Allen Institute for AI

Daniel S. Weld
danw@allenai.org
The Allen Institute for AI
The University of Washington

Eric Horvitz
horvitz@microsoft.com
Office of the Chief Scientific Officer
Microsoft



We stand at the foot of a significant inflection in the trajectory of scientific discovery. As society continues on its fast-paced digital transformation, so does humankind's collective scientific knowledge and discourse. We now read and write papers in digitized form, and a great deal of the formal and informal processes of science are captured digitally—including papers, preprints and books, code and datasets, conference presentations, and interactions in social networks and collaboration and communication platforms. The transition has led to the creation and growth of a tremendous amount of information—much of which is available for public access—opening exciting opportunities for computational models and systems that analyze and harness it. In parallel, exponential growth in data processing power has fueled remarkable advances in artificial intelligence, including self-supervised neural models capable of learning powerful representations from large-scale unstructured text without costly human supervision. Dramatic changes in scientific communication channels—such as the advent of the first scientific journal in the 17$^{\text{th}}$ century—have historically catalyzed revolutions in scientific thought and behavior. The confluence of societal and computational trends suggests that computer science is poised to ignite a revolution in the scientific process itself.

## Widening Gap

At the heart of the scientific process, a basic behavior has remained unchanged for hundreds of years: We build on *existing* ideas for forming *new* ideas. When faced with a new question or problem, we leverage knowledge from accumulated learnings and from external sources and perform synthesis and reasoning to generate insights, answers, and directions. But the last few decades have brought changes. The explosion of digital information and steep acceleration in the production of scientific data, results and publications [28, 38]—with over one million papers added every year to the PubMed biomedical index alone—stand in stark contrast to the constancy of human cognitive capacity. While scientific knowledge, discourse, and the larger scientific ecosystem are expanding with rapidity, our human minds have remained static, with severe limitations on the capacity for finding, assimilating and manipulating information. Herbert Simon's reflection that "…a wealth of information creates a poverty of attention," aptly describes the limited attention of researchers in the modern scientific ecosystem. Even within narrow areas of interest, there is a vast space of potential directions to explore—while the keyhole of cognition admits only a tiny fraction of the broad landscape of information and deliberates over small slices of possibility. The way we search through and reflect about information across the vast space—the areas we select to explore, and how we explore them—is hindered by cognitive biases [40] and lacks principled and scalable tools for guiding our attention [46]. "Unknowns" are not just holes in science, but important gaps in personal knowledge about the broader knowns across the sciences.

We thus face an imbalance between the treasure trove of scholarly information and our limited ability to reach into it. Despite technological advances, we require new paradigms and capabilities to address this widening gap. We see promise in developing new foundational capabilities that address the cognitive bottleneck, aimed at extending human performance on core tasks of research—e.g., keeping abreast with developments, forming and prioritizing ideas, conducting experiments, reading and understanding papers (see Table 1). We focus on a research agenda we call *task-guided scientific knowledge retrieval*, in which systems counter humans' bounded capacity by ingesting corpora of scientific knowledge and retrieving inspirations, explanations, solutions and evidence synthesized to directly serve task-specific utility. We present key concepts of task-guided scientific knowledge retrieval, including work on prototypes that highlight the promise of the direction and bring into focus concrete steps forward for novel representations, tools, and services. We review systems that help researchers discover novel perspectives and inspirations [16, 17, 19, 42], help guide the attention of researchers toward opportunity areas rife with uncertainties and unknowns [26, 46], and models that leverage retrieval and synthesis of scientific knowledge as part of machine learning and prediction [12, 37]. We conclude with a discussion of opportunities ahead with computational approaches that have the potential to revolutionize science.

## 1 HUMAN-CENTRIC PERSPECTIVE

Extraordinary developments at the convergence of AI and scientific discovery have emerged in specific areas, including new kinds of analytical tools, with the prominent example of AlphaFold, which harnesses deep neural models to dramatically improve the prediction of protein structure from amino acid sequence information [22]. While celebrating impressive achievements in modeling and simulation, we focus on a complementary set of opportunities for computing in scientific discovery, taking a *human-centered*, cognitive perspective: We pursue computational approaches that can augment the abilities of individual researchers, taking into account the diversity of tasks, contexts, and cognitive processes involved in consuming and producing scientific knowledge. Collectively, we

| Task/Activity | Description |
|---|---|
| *Attention to areas of interest* | A background process of keeping track of latest developments in relevant scientific communities. Involves applying selective attention, perceiving relevance and utility. |
| *Problem identification & prioritization* | Identifying new research questions and deciding on which ones to work. Involves factors such as subjective preferences and assessment of feasibility. |
| *Forming directions* | Given a problem/question, forming ideas to address it. Involves cognitive processes such as constructing mental models of a problem, problem reformulation, abstraction and decomposition, adaptation of relevant knowledge to new scenarios, and assessing likelihood of success. |
| *Literature search & review* | Accessing and ingesting knowledge in the literature. Involves many processes such as query formulation, skimming and assessing relevance, positioning ideas with relations and contrasts to existing work, and reading and summarization strategies. |
| *Learning, understanding, sense-making* | The cognitive processes and activities involved in assimilating new knowledge and concepts, and making sense of complex scientific information spaces. |
| *Experimentation, analysis, action* | A broad category referring to the many processes and activities involved in formulating and conducting experiments (e.g., planning data collection and measurements), performing analyses (e.g., understanding a set of data points, modeling and extrapolation, prediction, evaluation), and producing artifacts, techniques, theories, decisions, policies, actions. |
| *Research communication* | Writing research documents (papers, proposals, analyses), communicating with peers (feedback and review, collaboration, presentation). |

**Table 1: Research may be decomposed into salient tasks that are prime targets for computational augmentation (§ 2).**

group these under the *inner cognitive world* of a researcher[1] (see Figure 1). The researcher interacts with the scientific ecosystem—literature, resources, discussions—in order to inform decisions and actions. Researchers can have many different uses for scholarly information, depending on the task at hand and the stage of exploration (see Table 1 and discussion in Section 2). We pursue a research agenda around assisting researchers in their tasks, guided by two main desiderata:

(1) Systems for augmenting human capabilities in the sciences need to enhance the *effective flow of knowledge from the outer world of scientific information and discourse to the researcher's inner cognitive world*—countering humans' bounded capacity by retrieving and synthesizing information targeted to enhance performance on tasks. Achieving this goal requires methods that build and leverage rich representations of scientific content and that can align computational representations with human representations, in the context of specific tasks and backgrounds of researchers.

(2) Research on such systems should be rooted in *conceptual models of the inner cognitive world of a researcher*. Shining a spotlight on this inner world brings numerous factors and questions to the fore. How do researchers form ideas? How do they decide which problems to look into? How do they find and assimilate new information in the process of making decisions? What cognitive representations and bottlenecks are involved? What computing services would best augment these processes?

**Background and Related Themes.** In our research agenda, we leverage research in natural language processing, information retrieval, data mining and human-computer interaction and draw concepts from multiple disciplines. For example, efforts in *metascience* focus on sociological factors that influence the evolution of science [25], e.g., analyses of information silos that impede mutual understanding and interaction [53] and analyses of macro-scale ramifications of the rapid growth in scholarly publications [5] — work enabled by digitization of scholarly corpora (see Section 1.1). Metascience research makes important observations about human biases (desideratum 2) but generally does not engage in building computational interventions to augment researchers (desideratum 1). Conversely, work in *literature-based discovery* [48] mines information from literature to generate new predictions (e.g., functions of materials [52] or drug targets [36]) but is typically done in isolation from cognitive considerations; however, these techniques have great promise in being used as part of human-augmentation systems (see Sections 2-3). Other work uses machines to automate aspects of science. Pioneering work from Herbert Simon and Pat Langley [29] automated discovery of empirical laws from data, with models inspired by cognitive mechanisms of discovery (see Section 1.2). More recent work has focused on developing *robot scientists* [7, 23] that run certain biological experiments—not only formulating hypotheses but "closing the loop" by automated tests in a

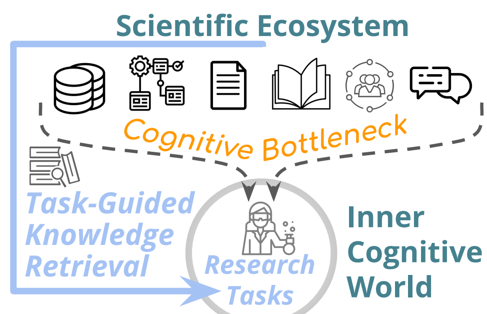


**Figure 1: Information flows from the outer world into the inner cognitive world of researchers, constrained by cognitive capacity and biases. We see opportunities to support researchers by retrieving knowledge that helps with tasks across multiple phases of the scientific process (Table 1).**

[1]We use the term *researcher* to include also practitioners in science-driven areas, such as medical doctors and technological engineers, who require deep scientific knowledge.

physical laboratory—where robots use narrow curated background knowledge (e.g., of a specific gene regulatory network [7]) and machine learning to guide new experiments. Related work explores automating scientific data analysis [12], which we discuss in Section 2 as a case of retrieval from scientific repositories to augment aspects of *experimentation and analysis* (see Table 1).

We now turn to a discussion of central concepts: the ecosystem of science, and the cognitive world. This presentation lays the foundations for our exposition of task-guided retrieval in Section 2 and the research opportunities in Section 3.

### 1.1 Outer World: Scientific Ecosystem

We collectively name the scientific ecosystem and the digital representations of scientific knowledge as the *outer world* (see Figure 1). The outer world is comprised of scientific communities, a complex and shifting web of peers, concepts, methodologies, problems and directions revolving around shared interests, understandings and paradigms [25]. This ecosystem generates a tremendous amount of digital information—digital "traces" of scientific thought and behavior—lying at the center of our attention as computer scientists interested in boosting human capacity to "reach into" the pool of scientific knowledge. This knowledge includes, first and foremost, *scholarly publications*, unstructured texts that appear in journals, conference proceedings, and online preprint repositories. Online publications can be seen as one main case of *digital research artifacts*; other examples of products of research include software, datasets, knowledge bases. Research artifacts are also associated typically with *signals of quality and interest*, such as citations to a specific paper or downloads of a dataset. The specific context of why a paper or resource was cited or used is often reflected in natural language descriptions. Different types of signals include *peer review* prior to publication (mostly not shared publicly), and *social media discussions* such as on Twitter, which has become a major virtual platform for academic dissemination and conversation [13]. Along with the trend in society, *private communication channels* among researchers are also digital—emails, online calls and messages. Similarly, *note taking and writing*—important activities across the scientific workflow—are done in digital form. This information is siloed in different platforms under privacy restrictions, yet represents a treasure trove for tools for the augmentation of scientific reasoning and exploration.

### 1.2 Inner World: Human Cognition in Science

The way researchers decide to interact with information in the outer world and the way they process and use this information is governed by a complex array of cognitive processes, personal knowledge and preferences, biases and limitations, which are only partially understood. We collectively name these the *inner world*, and briefly discuss several salient aspects.

Early work in AI by Herbert Simon and Alan Newell and later efforts by Pat Langley and Paul Thagard focused on cognitive and computational aspects of problem solving, creativity, decision making and scientific reasoning and discovery, seeking algorithmic representations to help understand and mimic human intelligence [29, 51]. Cognitive mechanisms that play important roles in scientific discovery include inductive and abductive reasoning, mental modeling of problems and situations, abstraction, decomposition, reformulation, analogical transfer and recombination [10, 51]; for example, in analogical transfer, given a situation or problem being considered in our working memory, we retrieve from our long-term memory prior analogous problems or situations [10].

This cognitive machinery powers humans' ingenuity. However, the human mind also has severe limitations—*bounded rationality* in the words of Simon—that impede these powerful mechanisms. Our limitations and capabilities have been studied for over a hundred years with the discipline of cognitive psychology. Our limitations manifest in bounded cognitive capacity and knowledge, and deeply-ingrained biases that govern our behaviors and subjective preferences. These limitations are all tightly interrelated. The ability to generate ideas, for instance, directly relies on prior knowledge and understandings; but, when a large volume of information from the outer world of science is met by insufficient cognitive capacity for processing and assimilating it, the result is information overload—a ubiquitous hindrance for researchers [28, 42]. Information overload in science strains the attentional resources of researchers, and forces researchers to allocate attention to increasingly narrow areas. This effect, in turn, amplifies a host of biases which researchers, just like all humans, suffer from [40, 46]. For example, scientists can be limited by confirmation bias [8], aversion to information from novel domains [24, 42], homophily [32], and fixation on specific directions and perspectives without consideration of alternative views [19, 40]. More broadly, selection of directions and areas to work on is a case of decision-making, and as such personal preference and subjective utility play fundamental roles. Our research decisions rely on subjective assessment of feasibility, long-term or short-term goals and interests, and even psychological factors (e.g., tendencies for risk aversion). These factors are of course also impacted by biases [40].

Clearly, the inner world of researchers is dauntingly complex. However, in the next section, we present encouraging results of applying computational methods to augment cognition in the sciences, helping to mitigate biases and limitations and enabling researchers to make better use of their powerful creative mechanisms.

## 2 AN APPROACH: TASK-GUIDED RETRIEVAL

How might we widen and deepen the connection between the fast-expanding outer world of science with researchers' limited cognitive worlds? We see a key bridge and research opportunity with developing tools for scientific *task-guided knowledge retrieval*. Drawing from discussions in literature on the process of scientific discovery, we enumerate in Table 1 salient scientific tasks and activities, such as *problem identification, forming directions, learning, literature search and review, experimentation.* These tasks could benefit from augmentation of human capabilities but remain underexplored in computer science.

Existing computational technologies for assisting humans in discovering scientific knowledge are underinvested in important aspects of the intricate cognitive processes and goal-oriented contexts involved in scholarly endeavors. The dominant approach to information retrieval research and systems can be summarized as "relevance first"—focusing on results that answer user queries as accurately as possible. Academic search engines assume users know what queries to explore and how to formulate them. For pinpointed literature search in familiar areas, this assumption *may* often suffice; but a broad array of other scholarly tasks, such as ideation or learning about a new topic, are very much underserved [17–19, 26, 42]. At the same time, many voices in the information retrieval community have discussed a different, broader view of *utility-driven* search, where the search is situated in a wider context of information seeking by users with specific intents and tasks [45]. Here, we adapt ideas and principles from this general paradigm.

We envision methods for task-guided scientific knowledge retrieval: systems that retrieve and synthesize outer knowledge in a manner that directly serves a task-guided utility of a researcher, while taking into consideration the researcher's goals, state of inner knowledge, and preferences. Consider the tasks in Table 1. For researchers engaged in *experimentation or analysis*, we envision systems that help users identify discussions of specific experiments and analyses in the literature to guide design choices and decisions. For researchers in early stages of *selecting problems* to work on, we picture systems that support this decision with information from literature and online discussions, synthesized and aggregated to obtain estimated impact and feasibility. As part of *forming directions* to address a problem, systems will help users find inspirations for solutions. Researchers who are *learning* about a new topic will be provided with retrieved texts and discussions that explain the topic in a manner personally tailored to personal knowledge. Importantly, task-guided knowledge retrieval follows the two desiderata introduced in Section 1; namely, systems should enable users to find knowledge that directly assists them in core research tasks by augmenting their cognitive capacity and mitigating their biases, and computational representations and services should align with salient cognitive aspects of the inner world of researchers.

We present work on initial steps and prototypes, including representative work that we have done and the work of others, framed in alignment with task-guided knowledge retrieval and tasks enumerated in Table 1. The main aim of this brief review is to stimulate discussion in the computer science community on tools for extending human capabilities in the sciences. Existing methods are far from able to realize our vision. For example, we see major challenges in representation and inferences about the inner world of knowledge and preferences, and aligning these with representations and inferences drawn from the outer world knowledge. Today's prototypes are limited examples of our vision, using very rough proxies of inner knowledge and interest based on papers and documents written or read by the user, or in some cases only a set of keywords. We discuss these research challenges and others in Section 3.

## 2.1 Prototypes of Task-Guided Retrieval

**Forming Directions.** We have developed methods for helping researchers generate new directions. A fundamental pattern in the cognitive process of creativity involves detecting *abstract connections* across ideas and transferring ideas from one problem to another [11]. Grounded in this cognitive understanding, we have pursued several approaches for stimulating creativity powered by retrieving outer knowledge. We developed and studied a system named Bridger that connects researchers to peers who inspire novel directions for research [42]. Bridger identifies matches among authors based on *commonalities and contrasts*, identifying peers who are both relevant and novel—working on similar problems but using very different methods, potentially inspiring new solutions. By doing so, Bridger helps mitigate the cognitive bias of *fixation* [19].

In this setting, inner knowledge is represented as mentions of problems and methods extracted automatically from a researcher's papers and weighted by term frequency. The outer knowledge being retrieved takes the form of other authors in computer science, following the same representation. For each retrieved author, the system displays salient problems, methods and papers, ranked by measures of relevance to the user. In studies with CS researchers, we found that Bridger dramatically boosted creative search and inspiration over state-of-art neural models employed by the Semantic Scholar search engine [6], surfacing useful connections across diverse areas; for example, one researcher drew novel connections between the mathematical area of graph theory and their own area of human-centered AI, by exploring a recommended author who applies graph theory to decision making.

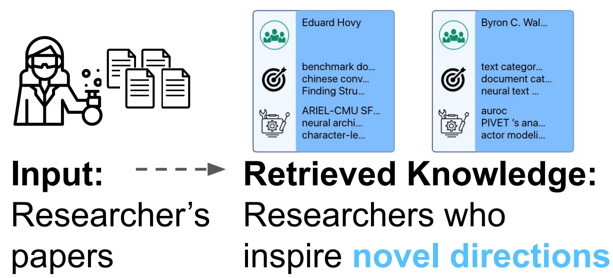


**Figure 2: Matching researchers to authors with whom they are unfamiliar, to help in generating directions. Author cards show key problems and methods extracted from their papers.**

We have also explored retrieving outer knowledge in the form of ideas that can explicitly enhance the human ability to find opportunities for analogical transfer [4, 16]. Extensive work in cognitive studies has highlighted the human knack for "analogical retrieval" as a central function in creativity—bringing together structurally related ideas and adapting them to a task at hand [9, 15]. We developed a search method that enables researchers to search through a database of technological inventions and find mechanisms that can be transferred from distant domains to solve a given problem. Given as input from the user a textual description of an invention, we retrieve ideas (inventions, papers) that have partial structural similarity to the input (e.g., inventions with similar mechanisms), to facilitate discovery of analogical transfer opportunities. We found that the method could significantly boost measures of human creativity in ideation experiments, in which users were asked to formulate new ideas after viewing inspirations retrieved with our approach versus baseline information retrieval methods. For example, a biomechanical engineering lab working on polymer stretching/folding for creating novel structures found useful inspiration in a civil engineering paper on web crippling in steel beams—abstractly related to stretching and folding.

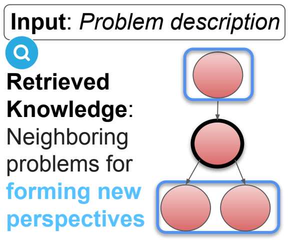


**Figure 3: Using an extracted hierarchy of problems to retrieve new perspectives on a focal problem of interest.**

Innovation may also involve traversing multiple levels of abstraction around a focal problem to "break out" of fixation on the details of a specific problem at hand by exploring novel perspectives. Given as input a problem description written by the user (as a proxy summary of the user's inner world of knowledge and purpose), we have pursued mechanisms that can retrieve diverse problem perspectives that are related to the focal problem, with the goal of inspiring new ideas for problem abstraction and reformulation [19] (see Figure 3). Using NLP models to extract mentions of problems, we mine a corpus of technological invention texts to discover problems that often appear together, and use this information to form a hierarchical problem graph that supports automatic traversal of neighboring problems around a focal problem, surfacing novel inspirations to users. In a study of the efficacy of the methods, over 60% of "inspirations" retrieved this way were found to be useful and novel—a relative boost of 50-60% over the best-performing baselines. For example, given an input problem of reminding patients to take medication, our system retrieves related problems such as in patient health tracking and alerting devices.

**Guiding attention and problem identification.** We see great opportunity in developing methods for guiding the attention of researchers to important areas in the space of ideas where there exists less knowledge or certainty [26, 46] (Figure 4). In one direction, we built a search engine that allows users to retrieve outer knowledge in the form of difficulties, uncertainties and initial hypotheses discussed in literature. The key goals of this mode of search are to bolster attention to rising and standing challenges of relevance to the end user so as to help overall with identification and selection of problems. We performed experiments with participants with diverse research backgrounds, including medical doctors working in a large hospital. Using query topics as a proxy for the inner world of participants' interests, we found the system could dramatically outperform PubMed search, the go-to biomedical search engine, at discovering important and interesting areas of challenges and directions. For example, while searching PubMed for the ACE2 receptor in the context of COVID-19 returns well-studied results, the prototype system by contrast focuses on finding statements of uncertainty, open questions and initial hypotheses, like a paper noting the *possibility* that ACE2 plays a role in liver damage in COVID-19 patients.

Another direction on biases and blindspots considers the long-term effort to identify protein-protein interactions (PPIs). A dataset of the growing graph of confirmed PPIs over decades was constructed and leveraged to identify patterns of scientific attention [46]. A temporal analysis revealed a significant "bias of locality," where explorations of PPIs are launched more frequently from those that were most recently studied, rather than following more general

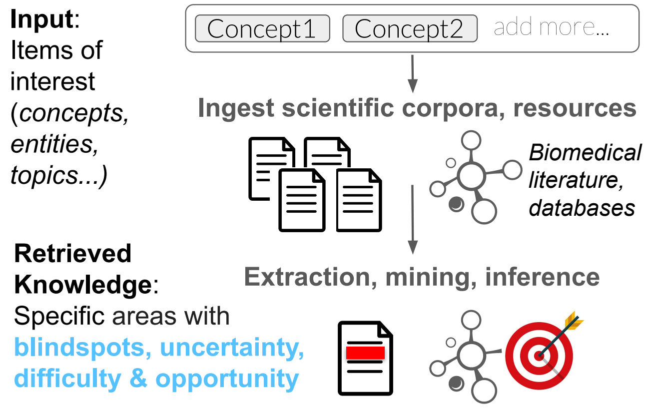


**Figure 4: Suggesting research opportunities for query concepts (e.g., medical topics) by identifying blindspots, gaps in collective knowledge and promising areas for exploration.**

prioritization of exploration. While locality reflects an understandable focus on adjacent and connected problems in the biosciences, the pattern of attention leads to systematic blindspots in large, widely used PPI databases that are likely unappreciated—further exacerbating attentional biases. The study further demonstrated mechanisms for reprioritizing candidate PPIs based on properties of proteins, and showed how earlier discoveries could be made with use of the debiasing methods. The findings underscore the promise of tools that retrieve existing outer world knowledge to guide attention to worthwhile directions. In this case, the outer knowledge source is a PPI *database*, and a user-selected sub-graph provides a proxy for inner world knowledge and interests.

**Literature search and review**. A great body of work on *literature search and review* has deep relevance to task-guided retrieval in the sciences. In particular, we see great opportunity with building on recent advances in information retrieval to (1) help biomedical researchers with domain-specific representations and (2) enhance scientific search by building new neural models.

Specialized search systems have been developed for the biomedical domain, with the overall vision of harnessing natural language understanding technologies to help researchers discover relevant evidence and expedite the costly process of systematic literature review [1, 41]. For example, Nye et al. [41] build a search and synthesis system based on automated extraction of biomedical treatment-outcome relations from clinical trial reports. The system is found to assist in identification of drug repurposing opportunities. As another recent example, the SPIKE system enables researchers to extract and retrieve facts from a corpus using an expressive query language with biomedical entity types and new term classes that the user can define interactively [49]. Together, this work underscores the importance of extracting a semantically meaningful representation of outer world knowledge that aligns with core aspects of inner world reasoning by researchers (see Section 3).

In separate work, neural language models built via self-supervision on large corpora of biomedical publications have recently led to performance boosts and new features in literature search systems [55], such as support for natural language queries that provide users with

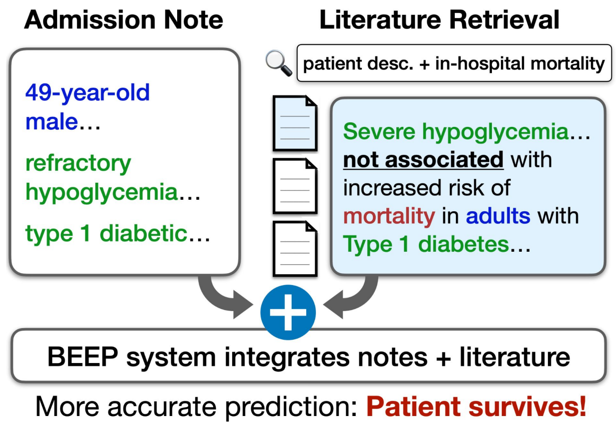


**Figure 5: Leveraging medical corpora to enhance the precision of AI models for inference about patient outcomes.**

a more natural way to formulate their informational goals. Neural models have also been trained to match abstract discourse aspects of pairs of papers (e.g., sentences referring to methodologies) and automatically retrieve documents that are aspectually similar [35]. By employing a representation that aligns with scientific reasoning across areas, this method achieves state-of-art results across biomedical and computer science literature.

**Experimentation, analysis, and action**. Beyond helping researchers via awareness and knowledge, we see great opportunities to use scientific corpora to construct task-centric inferential systems with automated models and tools for assisting with analysis, prediction and decisions. We demonstrate these ideas by casting two different lines of work as cases of task-guided retrieval.

(1) *Workflows* are multi-step computational pipelines used as part of scientific experimentation for data preparation, analysis and simulation [12]. Technically this includes execution of code scripts, services and tools, querying databases and submitting jobs to the cloud. In the life sciences, in areas such as genomics, there are specialized workflow management systems to help researchers find and use workflows, enabled by a community that creates and publicly shares repositories of workflows with standardised interfaces, metadata and functional annotations of tools and data [27]. As discussed in Gil [12], machine learning algorithms can potentially use these resources to automate workflow construction, learning to retrieve and synthesize data analysis pipelines. In this setting, outer world knowledge takes the form of workflow repositories, from which systems retrieve and synthesize modular building blocks; users' inner world is reflected via analysis objectives and constraints.

(2) In our work on clinical predictions [37], the goal is to enhance prediction of medical outcomes of patients hospitalized in the intensive care unit (ICU), such as in-hospital mortality or prolonged length of stay. Our system, named BEEP (biomedical evidence enhanced prediction), learns to perform predictions by retrieving medical papers that are relevant to each specific ICU patient, and to synthesize this outer knowledge in combination with internal EMR knowledge to form a final prediction. The primary envisaged user is a practice-oriented researcher—a medical doctor, whose inner knowledge is given by a rough proxy in the form of internal clinical notes from which we extract "queries" issued over medical papers. We find BEEP to provide large improvements over state-of-art models that do not use retrieval from the literature. BEEP's output can be aligned with inner world representations, e.g., matches between patient aspects and related cohorts in papers (see Figure 5).

**Learning and understanding**. We introduced a system [34] for helping users learn about new scientific concepts by showing definitions grounded in concepts the users know. For example, a new algorithm is explained as a variant of an algorithm the user is familiar with. Cognitive studies have asserted that effective descriptions of a new concept ground it within the network of known concepts [39]. Our system takes as input a list of concepts reflecting the user's inner knowledge as obtained from papers that they have written or read (drawn from Semantic Scholar logs). When the user seeks a definition of a new target concept, we retrieve outer knowledge in the form of definitions appearing in scientific papers, in which the *target concept* is explained in terms of *source concepts* that the user is familiar with. To further assist with reading complex scientific texts, we employ a neural text generation model to re-write the text in a structured, template form that relates the target concept to the source concepts.

## 3 OPPORTUNITIES AHEAD

The challenges of task-guided retrieval in support of researchers frames a host of problems and research opportunities. We discuss select challenges and directions (see also Table 2). We first motivate the discussion with an illustrative example, imagining a futuristic system for task-guided retrieval in science. We use this example to guide the discussion of future research needs.

### 3.1 Aspirations

We envision research tools with the capability of flowing outer world knowledge to researchers based on inferences about the state of the inner world— users' acute goals and difficulties, as well as users' knowledge, their past and present pursuits, and the tasks from Table 1 they are engaged in. Such systems would use multiple signals in these inferences, including users' papers, data and experiments, users' communication channels and documents, and would also engage in conversational interaction to understand users and their needs, suggesting solutions, hypotheses and experiments aiming to maximize utility on tasks captured in Table 1.

We foresee systems powered by rich representations of both inner and outer scientific knowledge. For a given concept, e.g., a certain algorithm or organism, an aspirational system would ingest all papers on the subject to form a multi-faceted representation of concepts as objects with associated properties and functions. Using this representation, the system could assist in *literature search and review*, enabling expressive queries to outer world information that target abstract aspects like functionalities, mechanisms, behaviors and designs in a manner that transcends field-specific jargon, abstracting away lexical differences that hindered historical search engines (e.g., Google Scholar). To help users *learn and understand new concepts* they encounter, the system would explain them in relation to other concepts the user already knows. A future system might also assist in automating *experimentation, analysis and action* and in *forming directions*, by forming compositions of concepts and predicting the affordances that would be formed as a result; for example, matching a certain algorithm with a suitable problem based on the algorithm's properties and the problem's requirements, matching an organism with a specific method of measurement or modification, or recombining parts of two devices to form a new device. The system could help identify related problems in the literature, synthesizing from them useful suggestions for problem reformulations. Considering the huge combinatorial space of potential suggestions, a system could assist in *prioritization* using estimated measures of interestingness, feasibility and value by synthesizing historical and current signals in literature, online discussions and knowledge bases.

Envisioned systems would be designed as human-centric, focusing on the individual researcher. The systems would enable users to convey preferences, goals and interests, and mediate the presentation of suggested directions and problem solutions based on personal prior knowledge, proposing concrete new directions grounded in representations that researchers can follow, and assisting users in *reading* complex retrieved texts by editing their language to conform with concepts that users are familiar with.

## 3.2 Research Directions

We are far from having machines with the capabilities described in the vision above. But imagining these capabilities can be valuable for guiding research efforts. We elaborate on challenges and directions, including limitations in representing scientific knowledge and inferring about researchers' inner worlds (see Table 2).

**Task-aligned representations and scientific NLP.** Paul Thagard writes: "thinking can best be understood in terms of representational structures in the mind and computational procedures that operate on those structures" [58]. We seek representations that can be aligned with human thinking—for insight-building, decision making and communication. How can we go beyond the representation of ideas in unstructured textual papers, toward representations that support such cognitive processes? The quest for formulating a *universal schema* that could represent scientific ideas and thinking goes back hundreds of years. Scholars such as Gottfried Leibniz and René Descartes were intrigued by the prospects of developing a universal codification of scientific knowledge. Leibniz proposed the *characteristica universalis*, a hypothesized formal language of ideas that would enable inferences with algebraic operators.

While such a representation is not within reach, envisioning its existence—and what may be needed to even roughly approximate it—points to important research directions. An exciting direction, for example, is obtaining representations that support a "computational algebra of ideas"—e.g., modeling novel compositions of concepts and the new affordances that would be formed as a result. Small steps toward this vision can be seen in work on representations of concepts as vectors in natural language embeddings [33] that support rudimentary forms of addition, subtraction, and analogy. More recently, models of language and images [2, 43] have made striking progress in generating new content and coherently combining concepts, albeit currently far from achieving full compositionality and relational reasoning [31]. We are excited by the prospect of future representations and operators developing and extending these capabilities to the complex world of scientific ideas.

One core limitation is that the underlying representations learned by these powerful models are currently far from understood and lack "hooks" for control and interpretability, critical in human-AI collaboration [56]. In our own work, in line with our focus on grounding representations of outer world knowledge with inner world cognitive aspects of researchers, we have pursued methods that "reverse engineer" scientific papers to automatically extract, using NLP models, *structured representations* that aim to balance three desiderata:

(1) Semantically meaningful representations, *aligned with a salient task* from the tasks in Table 1, grounded in cognitive research to guide us toward useful structures.
(2) Representations with *sufficient level of abstraction* to generalize across areas and topics.
(3) Representations *expressive enough for direct utility* in helping researchers as measured in human studies.

For example, we have extracted representations that capture *causal mechanisms* and *hierarchical graphs of functional relationships*. This kind of decomposition of ideas has enabled us, for example, to perform basic analogical inference in the space of complex technological and scientific ideas, helping researchers discover new inspirations (see Section 2). However, many forms of richer structures should be explored (e.g., of experimentation processes and methodologies to facilitate tasks in Table 1).

A central challenge is that current models' extraction accuracy is limited, and the diversity of scientific language leads to problems in generalization and normalization of terms and concepts. We have pursued construction of new datasets, models and evaluations for identifying similarity between concepts and aspects across papers [3, 35], with fundamental problems in resolving *diversity, ambiguity and hierarchy* of language. As our results have highlighted, neural models still tend to focus on surface-level lexical patterns, as opposed to deeper semantic relationships; this has also been echoed in recent work in the scientific NLP community [47]. More generally, substantial advances are needed to handle challenges posed by scientific documents. We require NLP models with full-document understanding, not only of text but of tables, equations, figures, and reference links [57]. Open access corpora (e.g., S2ORC [30], CORD-19 [54]) provide a foundation to address this challenge.

**New modes of writing and reading.** Perhaps the way we write can be dramatically different in the digital world, using machine-actionable representations? This theme is in its infancy [44] and faces massive cultural and technological barriers. Moreover, beyond mere reporting and documentation, the process of writing represents a direct channel between the inner and outer worlds, forcing us to communicate ideas in concrete language; this process often brings to light new questions which suggest new analyses and problem framings. Can systems accompany different phases of writing, suggesting new perspectives and shaping ideas?

In parallel, there is the task of *reading* what others have written; a new scientific reading interface has recently been built with interactive PDF documents that can, for example, provide customized concept definitions [14]. We imagine a future where every reader will see a different form of the same paper, with text re-written to

| Challenge | Description |
|---|---|
| *Task-aligned representations, scientific NLP* | How can we automatically and accurately extract conceptual representations of scientific knowledge, that are aligned with tasks that comprise the endeavor of science (Table 1)? How can we build NLP models that understand full scientific papers? |
| *Computational algebra of ideas* | Can we build representations of scientific knowledge that support composition of ideas? e.g., inferring the result of recombining two concepts. |
| *Identifying conceptual relationships across literature* | How do we detect important relationships across scientific ideas, across related discussions in different communities and areas? How can we resolve challenges of diversity, ambiguity, and multiple levels of detail in scientific language? |
| *Estimation of personal knowledge* | How can we estimate the knowledge of a given researcher? What are useful, practical models of this knowledge? What concepts does a researcher know, which of their aspects, and to what technical extent? How do we account for *latent* knowledge? |
| *Addressing gaps in knowledge* | Given an estimated model of a researcher's knowledge, and given a specific task in Table 1, what new knowledge would be useful for the task at hand? |
| *Estimation of preferences, goals, interests* | How can we estimate key latent preferences, interests and subjective utilities of researchers? Using information in papers and discussions to infer factors behind researchers' choices. |
| *Prediction and prioritization* | How might we identify promising sparse/unexplored areas in large "spaces of ideas" and prioritize directions that are novel, plausible and valuable? |
| *Developing new representations for learning and communicating* | Might the way we read and write papers change to be more effective? Might we communicate with machine-actionable, interlinked representations of scholarly knowledge. Might we create personalized "living" documents that tailor their content to readers' backgrounds. |

**Table 2: Directions with formulating and leveraging computational representations of scientific knowledge.**

align with readers' knowledge; e.g., our personalized concept definitions system [34] (§ 2) will insert new wording and explanations grounded in readers' knowledge.

**Internal world of researchers.** The latter point, of grounding new concepts in the previous knowledge of readers, suggests a wider and highly challenging problem. How can we estimate inner-world aspects of researchers? New methods are needed to enable researchers to specify their knowledge, preferences, and goals, in order to direct systems to carry out tasks. However, directly querying for these aspects places a burden on the researcher and may be prone to reporting biases. Digital scientific activity presents an opportunity for *automatically* approximating a researcher's knowledge, objectives, needs and interests—based on data. We are interested in using researchers' papers to estimate a *spectrum of knowledge*, what concepts users know and to what extent. We envision mixed-initiative interfaces [20] in which approximations of the inner world are presented to researchers and refined in human-machine collaboration, to identify personal gaps in knowledge in the context of a specific task and suggest new useful knowledge.

Representations of interest and preference are central in web commerce based on user activity histories. We are encouraged by results that highlight the feasibility of rich user models, including use in the personalization of general search engines [45, 50] and dynamically updated inferences [21]. Paul Samuelson wrote of "revealed preferences"— preferences that are revealed indirectly by the economic price people are willing to pay; while not directly equivalent, researchers' digital traces may reveal underlying choices, e.g., selecting to work on one problem and not another.

**Prediction and prioritization of directions.** Whenever we decide to work on a research direction, we are implicitly making a prediction about an area in "idea space" that is unknown to us. Can automated systems help make these predictions? This involves identifying promising areas and generating directions—hypotheses, ideas—in either natural or structured language, under constraints on a given user's background knowledge. It also involves ranking directions as function of estimated likelihood (feasibility, plausibility), utility (value) and novelty. Despite the great challenges involved, we are encouraged by advances in models trained on scientific datasets for predicting specific targets (e.g., protein structures [22]); we see potential in building on these advances as part of our *wider* agenda that considers the inner world of cognitive aspects and tasks, and the outer world outside the context of a narrow dataset.

## 4 SUMMARY

As the terrain of science widens at a fast pace, researchers are constrained by the limits of human cognition, and lack principled methods to follow developments, review literature, guide attention, and formulate and prioritize research directions. For the first time in the history of science, essentially all of scientific knowledge and discourse has moved into the digital space. This shift, coupled with dramatic advances in computational tools for analyzing and forming representations, presents tremendous opportunities for leveraging scientific corpora as databases from which knowledge, solutions, insights and inspirations can be gleaned to help scientists. We see great opportunity ahead for developing tools for researchers that can address the imbalance between the growing treasure trove of scholarly information and their limited ability to reach into it. Computational approaches have the potential to revolutionize the scientific process, harnessing humankind's collective knowledge and intelligence by performing syntheses of literature, databases, and discussions. Numerous challenges stand in the way of making progress on the path to achieving the vision we have laid out. However, even small steps forward will unlock vast opportunities for making advances at the frontiers of science.

## ACKNOWLEDGMENTS

We thank the members of the Semantic Scholar team for stimulating discussions. Projects were supported by NSF Convergence Accelerator Grant 2132318, NSF RAPID grant 2040196, and ONR grant N00014-18-1-2193.

## REFERENCES

[1] Gaelen P Adam, Byron C Wallace, and Thomas A Trikalinos. Semi-automated tools for systematic searches. In *Meta-Research*, pages 17–40. Springer, 2022.
[2] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.
[3] Arie Cattan, Sophie Johnson, Daniel Weld, Ido Dagan, Iz Beltagy, Doug Downey, and Tom Hope. Scico: Hierarchical cross-document coreference for scientific concepts. *Automated Knowledge Base Construction (AKBC) 2021*, 2021.
[4] Joel Chan, Joseph Chee Chang, Tom Hope, Dafna Shahaf, and Aniket Kittur. Solvent: A mixed initiative system for finding analogies between research papers. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–21, 2018.
[5] Johan SG Chu and James A Evans. Slowed canonical progress in large fields of science. *Proceedings of the National Academy of Sciences*, 118(41), 2021.
[6] Arman Cohan, Sergey Feldman, Iz Beltagy, Doug Downey, and Daniel Weld. SPECTER: Document-level Representation Learning using Citation-informed Transformers. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2270–2282, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.207. URL https://www.aclweb.org/anthology/2020.acl-main.207.
[7] Anthony Coutant, Katherine Roper, Daniel Trejo-Banos, Dominique Bouthinon, Martin Carpenter, Jacek Grzebyta, Guillaume Santini, Henry Soldano, Mohamed Elati, Jan Ramon, et al. Closed-loop cycles of experiment design, execution, and learning accelerate systems biology model development in yeast. *Proceedings of the National Academy of Sciences*, 116(36):18142–18147, 2019.
[8] Dieter Frey. Recent research on selective exposure to information. *Advances in experimental social psychology*, 19:41–80, 1986.
[9] Dedre Gentner. Structure-mapping: A theoretical framework for analogy. *Cognitive science*, 7(2):155–170, 1983.
[10] Dedre Gentner. Analogy in scientific discovery: The case of johannes kepler. In *Model-Based Reasoning*, pages 21–39. Springer, 2002.
[11] Dedre Gentner and Arthur B Markman. Structure mapping in analogy and similarity. *American psychologist*, 52(1):45, 1997.
[12] Yolanda Gil. Will ai write scientific papers in the future? *AI Magazine*, 2022.
[13] Stefanie Haustein. Scholarly twitter metrics. In *Springer handbook of science and technology indicators*, pages 729–760. Springer, 2019.
[14] Andrew Head, Kyle Lo, Dongyeop Kang, Raymond Fok, Sam Skjonsberg, Daniel S. Weld, and Marti A. Hearst. Augmenting scientific papers with just-in-time, position-sensitive definitions of terms and symbols. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021.
[15] Keith James Holyoak and Paul Thagard. *Mental leaps: Analogy in creative thought.* MIT press, 1996.
[16] Tom Hope, Joel Chan, Aniket Kittur, and Dafna Shahaf. Accelerating innovation through analogy mining. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017.
[17] Tom Hope, Jason Portenoy, Kishore Vasan, Jonathan Borchardt, Eric Horvitz, Daniel S Weld, Marti A Hearst, and Jevin West. Scisight: Combining faceted navigation and research group detection for covid-19 exploratory scientific search. In *EMNLP*, 2020.
[18] Tom Hope, Aida Amini, David Wadden, Madeleine van Zuylen, Sravanthi Parasa, Eric Horvitz, Daniel Weld, Roy Schwartz, and Hannaneh Hajishirzi. Extracting a Knowledge Base of Mechanisms from COVID-19 Papers . 2021.
[19] Tom Hope, Ronen Tamari, Hyeonsu Kang, Daniel Hershcovich, Joel Chan, Aniket Kittur, and Dafna Shahaf. Scaling creative inspiration with fine-grained functional facets of product ideas. *arXiv e-prints*, pages arXiv–2102, 2021.
[20] Eric Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 159–166, 1999.
[21] Eric J Horvitz, John S Breese, David Heckerman, David Hovel, and Koos Rommelse. The Lumiere project: Bayesian user modeling for inferring the goals and needs of software users. In *Proceedings of the Conference on Uncertainty in AI*, pages 256–265, 1998.
[22] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.
[23] Ross D King, Kenneth E Whelan, Ffion M Jones, Philip GK Reiser, Christopher H Bryant, Stephen H Muggleton, Douglas B Kell, and Stephen G Oliver. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature*, 427(6971):247–252, 2004.
[24] Aniket Kittur, Lixiu Yu, Tom Hope, Joel Chan, Hila Lifshitz-Assaf, Karni Gilon, Felicia Ng, Robert E. Kraut, and Dafna Shahaf. Scaling up analogical innovation with crowds and AI. 116. Publisher: National Academy of Sciences Section: Social Sciences.
[25] Thomas S Kuhn. *The structure of scientific revolutions*, volume 111. Chicago University of Chicago Press, 1970.
[26] D Lahav, JS Falcon, B Kuehl, S Johnson, S Parasa, N Shomron, DH Chau, D Yang, E Horvitz, DS Weld, et al. A search engine for discovery of scientific challenges and directions. 2021.
[27] Anna-Lena Lamprecht, Magnus Palmblad, Jon Ison, Veit Schwämmle, Mohammad Sadnan Al Manir, Ilkay Altintas, Christopher JO Baker, Ammar Ben Hadj Amor, Salvador Capella-Gutierrez, Paulos Charonyktakis, et al. Perspectives on automated composition of workflows in the life sciences. *F1000Research*, 10, 2021.
[28] Esther Landhuis. Scientific literature: Information overload. *Nature*, 535(7612): 457–458, 2016.
[29] Pat Langley, Herbert A Simon, Gary L Bradshaw, and Jan M Zytkow. *Scientific discovery: Computational explorations of the creative processes.* MIT press, 1987.
[30] Kyle Lo, Lucy Lu Wang, Mark Neumann, Rodney Kinney, and Daniel S. Weld. S2ORC: The Semantic Scholar Open Research Corpus. In *Proceedings of ACL*, 2020. URL https://arxiv.org/abs/1911.02782.
[31] Gary Marcus, Ernest Davis, and Scott Aaronson. A very preliminary analysis of dall-e 2, 2022. URL https://arxiv.org/abs/2204.13807.
[32] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444, 2001.
[33] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781 [cs]*, January 2013. URL http://arxiv.org/abs/1301.3781. arXiv: 1301.3781.
[34] Soni Murthy, Kyle Lo, Chandra Bhagavatula, Daniel King, Bailey Kuehl, Sophie Johnson, Jonathan Borchardt, Daniel S. Weld, Tom Hope, and Doug Downey. Accord: A multi-document approach to generating diverse descriptions of scientific concepts. 2022.
[35] Sheshera Mysore, Arman Cohan, and Tom Hope. Multi-vector models with textual guidance for fine-grained scientific document similarity. *NAACL*, 2022.
[36] Rahul Nadkarni, David Wadden, Iz Beltagy, Noah A Smith, Hannaneh Hajishirzi, and Tom Hope. Scientific language models for biomedical knowledge base completion: An empirical study. *Automated Knowledge Base Construction (AKBC) 2021*, 2021.
[37] Aakanksha Naik, Sravanthi Parasa, Sergey Feldman, Lucy Lu Wang, and Tom Hope. Literature-augmented clinical outcome prediction. *arXiv preprint arXiv:2111.08374*, 2021.
[38] Richard Van Noorden. Global scientific output doubles every nine years. In *Newsblog*. Nature, 2014. URL http://blogs.nature.com/news/2014/05/global-scientific-output-doubles-every-nine-years.html.
[39] NRC. How people learn: Brain, mind, experience, and school: Expanded edition. In *The National Academies Press*, Washington, DC, 2018. doi: https://doi.org/10.17226/9853.
[40] Regina Nuzzo et al. How scientists fool themselves–and how they can stop. *Nature*, 526(7572):182–185, 2015.
[41] Benjamin Nye, Jay DeYoung, Eric Lehman, Ani Nenkova, Iain J Marshall, and Byron C Wallace. Understanding clinical trial reports: Extracting medical entities and their relations. In *AMIA Annual Symposium Proceedings*, volume 2021, page 485. American Medical Informatics Association, 2021.
[42] Jason Portenoy, Marissa Radensky, Jevin West, Eric Horvitz, Daniel Weld, and Tom Hope. Bridger: Toward bursting scientific filter bubbles and boosting innovation via novel author discovery. *CHI*, 2022.
[43] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125*, 2022.
[44] Stuart Ritchie. The big idea: should we get rid of the scientific paper? *The Guardian*, 2022.
[45] Chirag Shah and Emily M Bender. Situating search. In *ACM SIGIR Conference on Human Information Interaction and Retrieval*, pages 221–232, 2022.
[46] Uriel Singer, Kira Radinsky, and Eric Horvitz. On biases of attention in scientific discovery. *Bioinformatics*, 12 2020. URL https://doi.org/10.1093/bioinformatics/btaa1036.
[47] Shruti Singh and Mayank Singh. The inefficiency of language models in scholarly retrieval: An experimental walk-through. *arXiv preprint arXiv:2203.15364*, 2022.
[48] D. R. Swanson. Fish oil, raynaud's syndrome, and undiscovered public knowledge. *Perspectives in Biology and Medicine*, 30(1):7–18, 1986.
[49] Hillel Taub Tabib, Micah Shlain, Shoval Sadde, Dan Lahav, Matan Eyal, Yaara Cohen, and Yoav Goldberg. Interactive extractive search over biomedical corpora. In *Proceedings of the 19th SIGBioMed Workshop on Biomedical Language Processing*, pages 28–37, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.bionlp-1.3. URL https://aclanthology.org/2020.bionlp-1.3.
[50] Jaime Teevan, Susan T Dumais, and Eric Horvitz. Personalizing search via automated analysis of interests and activities. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information*

*retrieval*, pages 449–456, 2005.
[51] Paul Thagard. *The cognitive science of science: Explanation, discovery, and conceptual change.* Mit Press, 2012.
[52] Vahe Tshitoyan, John Dagdelen, Leigh Weston, Alexander Dunn, Ziqin Rong, Olga Kononova, Kristin A Persson, Gerbrand Ceder, and Anubhav Jain. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature*, 571(7763):95–98, 2019.
[53] Daril A Vilhena, Jacob G Foster, Martin Rosvall, Jevin D West, James Evans, and Carl T Bergstrom. Finding cultural holes: How structure and culture diverge in networks of scholarly communication. *Sociological Science*, 1:221, 2014.
[54] Lucy Lu Wang, Kyle Lo, Yoganand Chandrasekhar, Russell Reas, Jiangjiang Yang, Doug Burdick, Darrin Eide, Kathryn Funk, Yannis Katsis, Rodney Michael Kinney, Yunyao Li, Ziyang Liu, William Merrill, Paul Mooney, Dewey A. Murdick, Devvret Rishi, Jerry Sheehan, Zhihong Shen, Brandon Stilson, Alex D. Wade, Kuansan Wang, Nancy Xin Ru Wang, Christopher Wilhelm, Boya Xie, Douglas M. Raymond, Daniel S. Weld, Oren Etzioni, and Sebastian Kohlmeier. CORD-19: The COVID-19 open research dataset. In *Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020*, Online, July 2020. Association for Computational Linguistics. URL https://aclanthology.org/2020.nlpcovid19-acl.1.
[55] Yu Wang, Jinchao Li, Tristan Naumann, Chenyan Xiong, Hao Cheng, Robert Tinn, Cliff Wong, Naoto Usuyama, Richard Rogahn, Zhihong Shen, et al. Domain-specific pretraining for vertical search: Case study on biomedical literature. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 3717–3725, 2021.
[56] Daniel S Weld and Gagan Bansal. The challenge of crafting intelligible intelligence. *Communications of the ACM*, 62(6):70–79, 2019.
[57] Michihiro Yasunaga, Jure Leskovec, and Percy Liang. Linkbert: Pretraining language models with document links. In *Association for Computational Linguistics (ACL)*, 2022.
[58] Edward N Zalta, Uri Nodelman, and Colin Allen. Stanford encyclopedia of philosophy, 1995.